Development of a multi-epitope peptide vaccine against human leishmaniases

Joana da Silva Pissarra

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Biologie Santé

École Doctorale Sciences Chimiques et Biologiques pour la Santé

Unité de recherche UMR177 INTERTRYP – Institut de Recherche pour le Développement (IRD)

Development of a multi-epitope peptide vaccine against human leishmaniases

Présentée par Joana PISSARRA

Le 26 Juin 2019

Sous la direction du Dr. Jean-Loup LEMESRE

Devant le jury composé de

Bernard MAILLÈRE, Dr., CEA-Saclay	Président / Examinateur
Claude LECLERC, Pr., Institut Pasteur Paris	Rapporteur
Sylviane PIED, Dr., Institut Pasteur Lille	Rapporteur
Nicolas BLANCHARD, Dr., Université de Toulouse	Examinateur
Rachel BRAS-GONÇALVES, Dr., IRD Montpellier	Co-encadrant
Philippe HOLZMULLER, Dr., CIRAD Montpellier	Co-encadrant
Amel GARNAOUI, Dr., Institut Pasteur Tunis	Co-encadrant
Stéphane DELBECQ, Pr., Faculté de Pharmacie Montpellier	Invité
Jean-Loup LEMESRE, Dr., IRD Montpellier	Directeur de thèse

UNIVERSITÉ DE MONTPELLIER

ACKNOWLEDGEMENTS

Undertaking this PhD has been a eventful and exciting experience and it would not have been possible without the support and guidance that I received.

This project has received funding from the **European Union**'s **Horizon 2020 research and innovation programme** under the Marie Sklodowska-Curie International Training Network ***grant agreemen*t No 642609.** I gratefully acknowledge the additional funding received towards my PhD from *Fondation des Treilles* and SATT AxLR.

I would like to express my sincere gratitude to my supervisor Dr. Jean-Loup Lemesre for his guidance and help throughout my PhD. I am grateful for his continuous support, encouragement and optimism which have made my PhD experience productive and stimulating.

I would like to thank the thesis committee and all the PhD defense jury members. A special thanks to Dr. Rachel Bras-Gonçalves and Dr Phillippe Holzmuller for their insightful comments and expertise, but also for the fruitful discussions and their contribution to this work. I would like to thank Dr. Amel Garnaoui for the scientific collaboration, her support and assistance of my thesis work during my stays in Tunis.

My sincere thanks also goes to Dr. Bernard Maillere who provided me an opportunity to visit the laboratory and research facilities. His precious support was invaluable for this research project and my PhD thesis.

I would like to thank Dr. Etienne Loire and Dr. Vincent Bonhomme for their significant aid and support, determinant to this thesis work. My thanks also go out to the support I received from the collaborative work I undertook with the *Plateforme de Protéomique Fonctionnelle* de Montpellier.

I would also like to thank Elodie Petitdidier and Julie Pagniez, great labmates who have helped and motivated me, and made the lab a friendly place.

I am grateful to my sister, mother and father for the support they provided me my entire life. A special and heartfelt thanks to Nuno. I am also grateful to my other family members and friends who have supported me along the way.

ABSTRACT

Leishmaniasis is a vector-borne neglected tropical disease endemic to 98 countries worldwide. Twenty *Leishmania* species are capable of establishing intracellular infection within human macrophages, causing different clinical presentations. Vaccine development against leishmaniases is supported by evidence of natural immunity against infection, mediated by a dominant cellular Th1 response and production of IFN-γ, IL-2 and TNF-α by polyfunctional TCD4+ and TCD8+ cells, ultimately leading to macrophage activation and parasite killing.

Excreted-secreted proteins are important virulence factors present throughout *Leishmania* life stages and are able to induce durable protection in dogs, a good model for human infection. We aim to develop a second generation vaccine from the *Leishmania* secretome, with the potential for large scale dissemination in a cost-effective, reproducible approach.

The secretome of six main pathogenic species (plus *L. tarentolae*) was analysed by Mass-Spectrometry and conserved candidate antigens were searched in the complete dataset. A total of 52 vaccine antigen candidates were selected, including 28 previously described vaccine candidates, and an additional 24 new candidates discovered through a reverse vaccinology approach.

In silico HLA-I and –II epitope binding prediction analysis was performed on all selected vaccine antigens, with world coverage regarding HLA restriction. To select the best epitopes, an automated R script was developed in-house, according to strict rational criteria. From thousands of potential epitopes, the automated script, in combination with optimal IC50, homology to host and solubility properties, allowed us to select 50 class I and 24 class II epitopes, synthesized as individual peptides. *In vitro* toxicity assays showed these selected peptides are non-toxic to cells.

The peptides' immunogenicity was evaluated using immunoscreening assays with immune cells from human donors, allowing for the validation of *in silico* epitope predictions and selection, and the assessment of the peptide's immunogenicity and prophylactic potential. Healed individuals, which had active infection and received treatment, possess *Leishmania*-specific memory responses and are resistant to reinfection, being considered the gold standard of protective immunity. On the other hand, the naive population is extremely important to include in the experimental validation step since it is the target population to vaccinate with a prophylactic vaccine. Importantly, a minimum specific T-cell precursor frequency is needed to induce long-lasting memory protective responses. Furthermore, there is also a positive correlation between immunodominant epitopes and a high frequency of specific T-cell precursors. Peptides able to induce Th1 and/or cytotoxic immune responses in both background are promising candidates for a vaccine formulation. Altogether,

experimental validation exclusively in human samples will provide us a very strong base for a vaccine formulation and allow to accelerate translation to the field.

Results show *Leishmania*-specific peptides successfully induce IFN-γ production by total PBMC from healed donors, and by specific T cells amplified from the naïve repertoire. Preliminary evidence exists for peptides which are immunogenic in both immune backgrounds (eight HLA-class I 9-mer peptides and five class II 15-mer peptides) which are, for now, the most promising candidates to advance for the multi-epitope peptide design.

Through the combination of proteomic analysis and *in silico* tools, promising peptide candidates were swiftly identified and the secretome was further established as an optimal starting point for vaccine development. The proposed vaccine preclinical development pipeline delivered a rapid selection of immunogenic peptides, providing a powerful approach to fast-track the deployment of an effective pan-specific vaccine against leishmaniases.

RESUMÉ

La leishmaniose est une maladie tropicale négligée à transmission vectorielle qui est endémique dans 98 pays dont les plus pauvres. Vingt espèces de *Leishmania* sont capables d'établir une infection intracellulaire au sein des macrophages humains, provoquant différentes manifestations cliniques. Le développement d'un vaccin contre les leishmanioses est étayé par des preuves d'immunité naturelle contre l'infection, induite par une réponse à médiation cellulaire de type Th1 dominante associée à la production d'IFN-γ, d'IL-2 et de TNF-α par des cellules T polyfonctionnelles $TCD4^+$ et $TCD8^+$, conduisant à l'activation classique des macrophages entrainant la destruction des parasites. Induire une protection robuste et durable et déterminer les épitopes immunodominants responsables de la protection naturelle représente un véritable défi.

Les protéines sécrétées sont des facteurs de virulence jouant un rôle important dans le cycle de vie des leishmanies et sont capables d'induire une protection durable chez le chien, un bon modèle pour l'infection humaine. Notre objectif est de développer, à partir du sécrétome de *Leishmania*, un vaccin de seconde génération reproductible et facile à produire à bas prix dans les zones d'endémie, avec des rendements de production rendant possible son utilisation à grande échelle.

Les sécrétomes des six espèces les plus pathogènes de leishmanie (plus *L. tarentolae*) ont été analysés et comparées par spectrométrie de masse. Les antigènes candidats ont été recherchés dans l'ensemble des données protéomiques disponibles. 52 antigènes candidats vaccin ont ainsi été sélectionnés, dont 28 avaient déjà été décrits dans la littérature et 24 sont nouveaux et découverts grâce à une approche de vaccinologie réverse.

Une analyse de la prédiction de liaison des épitopes *in silico* HLA-I et –II a été réalisée sur tous les antigènes candidats vaccin, prenant ainsi en compte le polymorphisme HLA de la population mondiale. Pour sélectionner les meilleurs épitopes parmi des milliers d'épitopes potentiels, un script R automatisé a été développé en interne, selon des critères rationnels stricts. Ainsi, 50 épitopes de classe I et 24 épitopes de classe II ont été sélectionnés et synthétisés sous forme de peptides individuels. Des essais de toxicité in vitro ont montré l'absence de toxicité cellulaire de ces peptides.

Les individus guéris par chimiothérapie généralement développent des réponses immunitaires protectrices à *Leishmania*. Des tests de stimulation des PBMC ont donc été réalisés avec des échantillons biologiques provenant de donneurs guéris de Tunisie et la production d'IFN-γ a été évaluée par ELISpot. De plus, il était important d'inclure dans l'étape de validation expérimentale des peptides des échantillons provenant d'individus naïfs, population cible à vacciner avec un vaccin prophylactique. Les résultats montrent que des peptides spécifiques de *Leishmania* induisent avec

succès la production d'IFN-γ par les PBMC totaux provenant de donneurs guéris et par les lymphocytes T spécifiques amplifiés à partir du répertoire naïf.

Globalement, la validation expérimentale des peptides réalisée exclusivement sur des échantillons humains nous fournira une base préclinique très solide pour développer un vaccin efficace capable de protéger les populations touchées par ces maladies. Elle constituera un moyen sûr et rentable de mieux sélectionner les candidats retenus pour le vaccin et d'éliminer ceux qui présentent un risque d'échec élevé au tout début du processus de développement du vaccin.

Grâce à la combinaison de l'analyse protéomique et d'outils in silico, des candidats peptidiques prometteurs ont été rapidement identifiés pour le développement d'un vaccin. Le « pipeline » de développement préclinique du vaccin proposé fournit une sélection rapide de peptides immunogènes, offrant une approche puissante pour accélérer le déploiement d'un vaccin pan-spécifique efficace contre les leishmanioses.

RESUMÉ DETAILLÉ

Les leishmanioses sont des maladies parasitaires à transmission vectorielle liées à l'infection par plus de 20 espèces de parasites protozoaires flagellés appartenant au genre *Leishmania*. Parmi les parasitoses, la leishmaniose est le deuxième plus grand tueur dans le monde après le paludisme, avec 50 000 décès estimés par an. Signalées dans 98 pays (dont 72 sont des pays en voie de développement), elles exposent 1 milliard de personnes au risque d'être infecté et de développer une des 4 formes cliniques de la maladie : la leishmaniose viscérale (LV), forme la plus sévère aussi connue sous le nom de kala-azar ; la leishmaniose cutanée (LC), forme la plus fréquente ; la leishmaniose cutanéo-muqueuse (LCM), forme la plus mutilante et défigurante; ou la leishmaniose cutanée post-kala-azar (LDPKA), pouvant présenter des complications graves après une LV.

Le contrôle de la leishmaniose repose principalement sur la lutte anti-vectorielle et le traitement, qui présentent plusieurs inconvénients, notamment la toxicité, le prix et le manque d'efficacité. A l'heure actuelle, aucun vaccin contre les leishmanioses humaines n'est disponible sur le marché. La vaccination est pourtant le moyen le plus adapté pour interrompre la transmission des leishmanies et contribuer à l'élimination des leishmanioses.

Des recherches menées par mon laboratoire d'acceuil sur les antigènes d'excrétion-sécrétion (AES) purifiés de leishmanies, est né CaniLeish® : le premier vaccin antiparasitaire Européen contre la leishmaniose viscérale canine, commercialisé par la société Virbac depuis 2011. C'est une innovation majeure en immunologie parasitaire et un atout essentiel dans la prévention de la leishmaniose canine mais aussi humaine, le chien étant le principal réservoir de parasites potentiellement transmissibles à l'homme. CaniLeish® est capable de déclencher une réponse à médiation cellulaire protectrice de type Th1 par immunisation avec des AES purifiés de cultures de promastigotes de *Leishmania infantum*. La résistance à *Leishmania* chez l'Homme est également basée sur une réponse Th1 et des réponses cytotoxiques (production d'IFN-γ, d'IL-2 et de TNF-α par les lymphocytes T CD4 + et CD8 + polyfonctionnels), conduisant à l'activation des macrophages et à la destruction des parasites intracellulaires. Par contre, la progression de la maladie est associée à des réponses cellulaires de type Th2 (prédominance d'IL-10 et d'IL-4).

L'objectif principal de mon projet de thèse est de développer un vaccin de deuxième génération contre les leishmanioses humaines. Fort du succès du vaccin CaniLeish®, les AES de leishmanies ont été choisis comme source d'antigènes pour le développement d'une stratégie de vaccination prophylactique à visée humaine, à base de peptides multi-epitopiques, réunissant "les meilleures parties des meilleurs antigènes" en un seul candidat vaccin polyvalent. Notre objectif est

de concevoir et de synthétiser des peptides avec des épitopes immunodominants multiples (poly-épitopiques) et appropriés dérivés de protéines excrétées / sécrétées identifiées à partir de données protéomiques provenant de 6 espèces pathogènes de leishmanies: *L. infantum*, *L. major*, *L. tropica*, *L. amazonensis*, *L. braziliensis* et *L. donovani*). Des études de phase préclinique ont été réalisées exclusivement sur des cellules humaines afin d'évaluer l'immunogénicité des peptides sélectionnés. La validation expérimentale des peptides synthétiques a consisté à évaluer les profils immunitaires après stimulation peptidique de cellules d'individus exposés ayant développé une immunité à l'infection par *Leishmania* (individus guéris) par rapport à des sujets naïfs. D'autre part, il est extrêmement important d'inclure dans l'étape de validation expérimentale de l'immunogénicité des peptides des échantillons provenant d'individus naïfs car il s'agit de la population cible à vacciner avec un vaccin prophylactique. De manière importante, une fréquence minimale de précurseur spécifique de cellules T est nécessaire pour induire des réponses protectrices mémoire.

D'autres résultats étaient attendus des données protéomiques comme la caractérisation détaillée du sécrétome et la découverte d'éventuelles corrélations entre la distribution géographique, les manifestations cliniques, l'immunomodulation et la pathogenèse. Nous avons également inclus dans notre étude les données du sécrétome d'une espèce de *Leishmania* non pathogène : *L. tarentolae*. Nous espérons aussi contribuer à accroître nos connaissances sur la variabilité interspécifique des leishmanies et éventuellement à identifier et caractériser de nouveaux facteurs de virulence pouvant contribuer au diagnostic de ces maladies et/ou à la mise au point de médicaments contre les leishmanioses.

Les tâches accomplies dans ce projet de recherche comprennent :

- l'identification exhaustive et la caractérisation des AES présents dans le sécrétome de six espèces pathogènes de *Leishmania* par spectrométrie de masse (*L. infantum*, l'espèce utilisée pour la production de CaniLeish ®, et aussi *L. major*, *L. tropica*, *L. amazonensis*, *L. braziliensis* et *L. donovani*) ainsi qu'une espèce non pathogène pour l'homme, *L. tarentolae*;
- la sélection de 52 antigènes vaccinaux les plus pertinents sur l'ensemble des données protéomiques générées. Ceux-ci comprennent 28 protéines déjà décrites dans la littérature scientifique comme des candidats vaccins (Set A), et 24 nouvelles protéines sélectionnées, en utilisant une approche de vaccinologie reverse (Set B) selon un critère de « non-homologie de séquences protéiques avec celles de l'hôte » ;
- l'identification d'épitopes T (séquences peptidiques qui se lient spécifiquement aux molécules HLA-I et -II, et qui activent le système immunitaire adaptatif) les plus affins pour l'ensemble des molécules HLA majoritairement représentées dans les populations humaines (36 allèles

HLA-I, correspondant à 11 supertypes et 98% de la population mondiale, et 21 allèles HLA-II, correspondant à 95% de la population mondiale), à l'aide de plusieurs serveurs de prédiction d'épitopes hautement performants;

- la seléction des épitopes le plus immunogènes, avec l'aide d'un script R automatisé, que j'ai développé pour ce projet (Pissarra J et al, publication en cours de révision), permettant de sélectionner facilement les épitopes les plus pertinents en intégrant différents critères prédéfinis. Les critères utilisés pour la hiérarchisation des épitopes sont : la forte affinité de liaison, la conservation entre les espèces, la prédiction par au moins 2 algorithmes différents, et la faible homologie des séquences peptidiques avec celles de l'hôte. Le script R applique ces critères aux données brutes collectées à partir des algorithmes de prédiction pour filtrer rapidement un nombre extrêmement important d'épitopes potentiels (stratégie "best of"). Le script est très polyvalent et applicable à d'autres prédicteurs, à d'autres protéines, mais aussi à d'autres pathogènes. Ainsi, nous avons sélectionné 50 épitopes HLA-I (9-mer) et 24 épitopes HLA-II (15-mer). Les peptides HLA-I proviennent de 23 protéines différentes (11 du Set A et12 du Set B), et les peptides HLA-II proviennent de 15 protéines (7 du Set A et 8 du Set B).

L'originalité et la pertinence de notre stratégie vaccinale réside aussi dans l'utilisation de tests fonctionnels précliniques comme voie d'exploration de l'efficacité de nos candidats vaccins : l'utilisation *ex vivo* de cellules humaines pour développer un vaccin humain. En effet, les modèles murins sont inappropriés pour étudier les réponses immunitaires humaines contre *Leishmania*. Il est difficile et dangereux d'extrapoler des résultats obtenus chez la souris à des hôtes naturels de l'infection. Globalement, la validation expérimentale des peptides réalisée exclusivement sur des cellules humaines devait fournir une base préclinique très solide pour développer un vaccin humain efficace capable de protéger les populations touchées par ces maladies. Elle constituait un moyen sûr et rentable de mieux sélectionner les candidats retenus pour le vaccin et d'éliminer ceux qui présentaient un risque d'échec élevé au tout début du processus de développement du vaccin.

Grâce à la combinaison de l'analyse protéomique et d'outils in silico, des candidats peptidiques prometteurs ont été rapidement identifiés pour le développement d'un vaccin. Le « pipeline » de développement préclinique du vaccin proposé fournit une sélection rapide de peptides immunogènes, offrant une approche puissante pour accélérer le déploiement d'un vaccin pan-spécifique efficace contre les leishmanioses. Les individus guéris d'une infection leishmanienne possèdent des réponses immunitaires mémoires contre les parasites qui les rendent résistants à la réinfection, et sont considérés comme le gold standard de l'immunité protectrice. La validation

expérimentale des peptides doit également être effectuée sur des échantillons provenant de sujets ayant d'autres statuts immunitaires : individus asymptomatiques et naïfs. Les individus asymptomatiques sont infectés par le parasite mais ne développent pas la maladie, ce qui signifie que leur système immunitaire parvient à contrôler l'infection sans pour autant éliminer le parasite. De plus, la population naïve est extrêmement importante à inclure dans la validation expérimentale, car elle représente la population cible pour un vaccin prophylactique. En effet, une fréquence minimale de précurseurs de cellules T spécifiques est nécessaire pour induire une réponse protectrice de longue durée. Il existe également une corrélation positive entre les épitopes immunodominants et la fréquence élevée de précurseurs de cellules T spécifiques, ce que nous chercherons dans notre étude. Les peptides sont validés pour leur capacité à induire des réponses mémoires préexistantes spécifiques à *Leishmania* sur des échantillons provenant d'individus guéris (en collaboration avec l'Institut Pasteur de Tunis). L'efficacité des peptides à stimuler *ex vivo* des cellules mononuclées du sang périphérique humain (PBMC) et à produire des cytokines de type Th1 comme l'IFN-γ (cytokine associée à la protection) a été évaluée.

Pour tester l'efficacité des peptides à activer le répertoire naïf d'individus sains, des essais de co-culture de cellules T sont réalisés avec plusieurs cycles d'amplification cellulaire (en raison de la rareté des cellules spécifiques), et la production d'IFN-y spécifique est recherchée par la technique ELISpot. Le typage HLA des donneurs naïfs permet une sélection « sur mesure » pour l'optimisation des immuno-essais, et aussi la conclusion effective sur les résultats de prédiction d'épitopes et la restriction HLA associée.

Ce projet aura aussi des répercussions importantes sur la connaissance de la biologie du parasite, grâce notamment au jeu de données protéomiques et à l'identification des régions d'immunogénicité des antigènes choisis (*epitope-mapping*). Ainsi, cela permettra d'augmenter nos connaissances sur la variabilité inter-espèces, et de révéler potentiellement de nouveaux facteurs de virulence importants et utiles pour le diagnostic de la leishmaniose ou le développement de nouveaux médicaments.

Nous pensons que ce projet contribuera à la découverte de peptides immunogènes pouvant, sous forme de peptides multi-épitopes, entrer dans la composition d'un vaccin efficace contre les leishmanioses humaines. Notre stratégie devrait concourir à minimiser les risques d'échec à un stade précoce du développement du vaccin et lors de la réalisation de futurs essais cliniques. Enfin, elle offre aussi les bases méthodologiques nécessaires aux suivis immunologiques des individus vaccinés lors d'essais cliniques de vaccination en zone d'endémie.

THESIS OUTLINE

The present thesis describes the preclinical development of a peptide-based vaccine against human leishmaniasis and preliminary experimental validation on the proposed peptide candidates.

Chapter I provides a contextualization of *Leishmania* biology, the problematic of leishmaniasis control and the importance of vaccine development. Particularly, it reviews current epidemiologic information on *Leishmania* spp. parasites, host-pathogen interactions and host immune responses against the parasite, as well as current vaccine pipeline and peptide-based vaccines.

Chapter II describes the proteomic analysis of the secretome of seven *Leishmania* species, responsible for the main clinical forms of leishmaniasis, and characterization of the secretome as an important source of virulence factors and of vaccine antigen candidates. Also, the results of this section provide the dataset used for vaccine antigen selection.

Chapter III exposes the strategies selected for vaccine antigen selection. A total of 52 protein candidates were selected from the secretome proteomic datasets through two parallel approaches: searching peptide candidates previously described in the literature (set A) and through a reverse vaccinology approach (set B).

Chapter IV describes the immuno-informatic tools available and used in this study, *in silico* epitope human leukocyte antigen (HLA)-binding predictions and epitope selection, including the development of a semi-automated R script to select the best epitopes from the vast HLA-binding prediction data corresponding to the selected 52 protein antigens. The selected HLA class-I and -II epitopes were synthesized as 9- and 15-mer peptides, respectively, for experimental validation.

In Chapter V, current methods to assess T-cell mediated peptide immunogenicity are reviewed, and cellular immune responses against the peptide candidates are evaluated through immunoassays with samples from humans with different immune status regarding *Leishmania* infection (naive and healed individuals).

Finally, **Chapter VI** comprehends a synthesis of main findings of this thesis work, general discussion and conclusions, as well as future perspectives.

TABLE OF CONTENTS

CHAPTER II - THE LEISHMANIA SECRETOME

CHAPTER III - VACCINE ANTIGEN SELECTION

CHAPTER IV - *IN SILICO* EPITOPE DISCOVERY AND SELECTION

CHAPTER V - EXPERIMENTAL VALIDATION OF SYNTHETIC PEPTIDES

LIST OF FIGURES

CHAPTER I – GENERAL INTRODUCTION

CHAPTER II – THE LEISHMANIA SECRETOME

CHAPTER III – VACCINE ANTIGEN SELECTION

CHAPTER IV - *IN SILICO* EPITOPE DISCOVERY AND SELECTION

CHAPTER V - EXPERIMENTAL VALIDATION OF SYNTHETIC PEPTIDES

CHAPTER VI – GENERAL DISCUSSION AND FUTURE PERSPECTIVES

LIST OF TABLES

CHAPTER I – GENERAL INTRODUCTION

CHAPTER II – THE LEISHMANIA SECRETOME

CHAPTER III – VACCINE ANTIGEN SELECTION

CHAPTER IV - *IN SILICO* EPITOPE DISCOVERY AND SELECTION

CHAPTER V - EXPERIMENTAL VALIDATION OF SYNTHETIC PEPTIDES

ABBREVIATIONS

a.a.	Amino-acid
Abs	Antibodies
APC	Antigen-Presenting Cells
CCL	Chemokine (C-C motif) ligand
CCR	C-C chemokine receptor type
CD40L	ligand of CD154
CL	Cutaneous leishmaniasis
CMI	cellular-mediated immune response
CTL	Cytotoxic T lymphocytes
DAMP	Damage-Associated Molecular Pattern
DAT	Direct Agglutination Test
DC	Dendritic cells
DCL	Diffuse Cutaneous Leishmaniasis
DMSO	Dimethyl sulfoxide
ELISpot	Enzyme-linked Immunospot assay
ELISA	enzyme-linked immunosorbent assay
ESA	Excreted-secreted antigens
ESP	Excreted-secreted proteins
FcR	Fc receptor
GM-CSF	Granulocyte-macrophage colony-stimulating factor
HLA	Human Leukocyte Antigen
IDRI	Infectious Disease Research Institute
IL	Interleukin
ILC	Innate Lymphoid Cell
IgG	Immunoglobulin G
IFN-	Interferon
L.	*Leishmania*
LCL	Local cutaneous leishmaniasis
LPS	Lipopolysaccharide
LST	Leishmanin Skin Test
LZ	Leishmanization
MCL	Mucocutaneous leishmaniasis
MHC	Major histocompatibility complex
MoDC	Monocyte-derived dendritic cells
MPL-SE	monophosphoryl lipid A
NF-κB	Nuclear factor-kappa B
NK cells	Natural killer cells
NO	Nitric Oxide
NTD	Neglected Tropical Disease
PAMP	Pathogen-Associated Molecular pattern
PBMC	Peripheral Blood Mononuclear cells
PBS	Phosphate-buffered saline
PHA	Phytohemagglutinin-L
PKDL	Post Kala-Azar Dermal Leishmaniasis
pMHC	peptide:MHC complex
PMN	Polymorphonuclear Neutrophils
PRR	Pattern Recognition Receptor
PSA	Promastigote Surface Antigen
R&D	Research and Development
SLA	Soluble Leishmania Antigen
SSG	Sodium StiboGluconate
TCM	Central Memory T cell
TCR	T-cell receptor
TEM	Effector Memory T cell
TFH	Follicular T Helper
TGF	Transforming Growth Factor
Th1	Type 1 T helper
Th2	Type 2 T helper
Th17	Type 17 T helper
TLR	Toll-Like Receptor
TNF	Tumor Necrosis Factor
Tregs	Regulatory T cells
TRM	Resident memory T cell
TSCM	Stem cell memory T cell
VL	Visceral leishmaniasis

LIST OF PUBLICATIONS AND PRESENTATIONS

Peer-reviewed articles:

- **Geiger et al 2016 (1)**

Geiger A, Bossard G, Sereno D, **Pissarra J**, Lemesre J-L, Vincendeau P, et al. Escaping Deleterious Immune Response in Their Hosts: Lessons from Trypanosomatids. Front Immunol. 2016;7(May):1–21

- **Holzmuller et al 2018 (2)**

Holzmuller P, Geiger A, Nzoumbou-Boko R, **Pissarra J**, Hamrouni S, Rodrigues V, et al. Trypanosomatid Infections: How Do Parasites and Their Excreted–Secreted Factors Modulate the Inducible Metabolism of l-Arginine in Macrophages? Front Immunol. Frontiers; 2018 Apr 20;9:778

- **Pissarra et al (under review)**

Pissarra J, Bonhomme V, Loire E, Lemesre J-L, Holzmuller P. Exploring -omics datasets for epitope-based development of vaccines and therapeutics – SILVI, an open-source pipeline script for T-cell epitope selection. (submitted to Plos One).

- **Petitdidier E et al (submitted to npj Vaccines)**

Petitdidier E, Pagniez J, Pissarra J, Holzmuller P, Papierok G, Vincendeau P, Lemesre J-L, Bras-Goncalves R. Peptide-based vaccine successfully induced protective immunity against canine visceral leishmaniasis.

- **Pissarra et al (under preparation)**

Provisional title: Proteomic analysis of the total secretome of seven *Leishmania* species.

Poster presentations:

- **11th ENII EFIS EJI Summer School on Advanced Immunology**, Sardinia, Italy (May 7-14th 2016)

Pissarra J, Holzmuller P, Guizani-Tabbane L, Lemesre J-L. "Development of a multiepitope peptide-based vaccine against human leishmaniasis"

- **2nd International ParaFrap Conference** (French Parasitology Alliance for Health Care, ParaFrap), Ile des Embiez, France (October 2-5th 2016)

Pissarra J, Holzmuller P, Bras-Gonçalves R, Guizani-Tabbane L, Lemesre J-L. " Development of a multi-epitope peptide vaccine against human leishmaniasis"

- **CBS2 day 2016** (Journée de l'Ecole Doctorale CBS2), Montpellier, France (May 26th 2016)

Pissarra J, Holzmuller P, Bras-Gonçalves R, Garnaoui A, Lemesre J-L. "A new peptide-based vaccine against human leishmaniasis – optimised preclinical development to prevent late-stage failure".

- **10th Meeting Club de Vaccinologie,** Lyon, France (March 20-21st 2017)

Pissarra J, Holzmuller P, Guizani-Tabbane L, Bras-Gonçalves R, Lemesre J-L. "A new peptide-based vaccine against human leishmaniases – optimised preclinical development to prevent late-stage failure".

- **6th World Congress on Leishmaniasis** (Worldleish), Toledo, Spain (May 16-20th 2017)

Pissarra J, Holzmuller P, Bras-Gonçalves R, Garnaoui A, Lemesre J-L. "A new peptide-based vaccine against human leishmaniasis – optimised preclinical development to prevent late-stage failure".

- **TwinntoInfect Spring School in Infection and Immunity 2018** by the TwinnToInfect Network (Instituto de Medicina Molecular, Francis Crick Institute, Institut Pasteur Paris)

Pissarra J, Holzmuller P, Bras-Gonçalves R, Garnaoui A, Lemesre J-L. "Preclinical development of a new peptide-based vaccine against human leishmaniases".

Oral presentations:

- **CBS2 day 2017** (Journée de l'Ecole Doctorale CBS2), Montpellier, France (May 24th 2017)

Oral presentation (*awarded prize for best oral presentation*): Pissarra J, Holzmuller P, Bras-Goncalves R, Garnaoui A, Lemesre JL. "A new peptide-based vaccine against human leishmaniasis – optimized preclinical development to prevent late-stage failure".

- **10th European Congress on Tropical Medicine and International Health Institute of Tropical Medicine** (ECTMIH) during Session 1S5 – Vaccinology (October 16-20th 2018)

Oral presentation: Pissarra J, Holzmuller P, Bras-Gonçalves R, Garnaoui A, Lemesre J-L. "Preclinical development of a new peptide-based vaccine against human leishmaniasis".

The abstracts of the 10th ECTMIH were published at Oral Presentation Sessions. Tropical Medicine and International Health 2017 Oct;22:2–114 (3).

- **1st International Caparica Congress on Leishmaniasis 2018**, Lisbon, Portugal (October 29-31st 2018)

Shotgun Presentation (*awarded prize for best shotgun presentation*): Pissarra J, Holzmuller P, Bras-Gonçalves R, Garnaoui A, Lemesre J-L. "A new epitope-based peptide vaccine against human leishmaniasis".

- **Final EUROLEISH.net meeting**, Barcelona, Spain (November 12-13th 2018)

Oral presentation (*awarded prize for best oral presentation*): Pissarra J. "Preclinical development of a new peptide-based vaccine against human leishmaniasis" .

CHAPTER I

GENERAL INTRODUCTION

1. Leishmaniasis is a Neglected Tropical Disease

Leishmaniasis is a term that refers to any form of a complex group of diseases caused by protozoan parasites of the genus *Leishmania* (belonging to the Trypanosomatida order*: Trypanosomatidae* family*)* and are transmitted by sand flies, phlebotomine vectors. Over 20 different *Leishmania* (*L.*) species are known to cause disease in humans and other mammals. Despite being closely related and sharing a common lifecycle and an invertebrate host, different *Leishmania* species are transmitted by different vector species, have different epidemiological features, namely, zoonotic or anthroponotic transmission, and cause quite different clinical presentations. The main four clinical forms of leishmaniasis are: cutaneous leishmaniasis (CL), mucocutaneous leishmaniasis (MCL), visceral leishmaniasis (VL) also known as kala-azar, and post-kala-azar dermal leishmaniasis (PKDL).

Neglected Tropical Diseases (NTDs), also known as Neglected Infectious Diseases (NIDs), englobe several communicable diseases caused by diverse infectious organisms. They prevail in 149 tropical and subtropical countries causing a massive economic and development burden to the affected societies (http://www.who.int/neglected_diseases/diseases/en/). Poor and/or rural populations are the most vulnerable to these infections, and the least likely to have access to healthcare services. NIDs share features that advocate both for Public Health initiatives that successfully detect, prevent disease and treat patients, but also for the development of new and more effective control tools.

Leishmaniasis is one of the most neglected tropical diseases as diagnostics and treatment tools are often ineffective or toxic, and improvement and development are impaired by the lack of funding and R&D (4–6). Leishmaniasis is endemic to the poorest countries in the world, and evidence shows that epidemics (or increases in incidence) are closely associated with socio-economic conditions, war and conflicts, malnutrition and food insecurity, and access to healthcare (7–10). Leishmaniasis surveillance is seldom based on passive case detection which further contributes to the underestimation of its burden and impairment of control efforts, whilst contributing to active parasite transmission (11,12).

In the last decades, efforts towards controlling NIDs have increased, notably since the World Health Organization's 2012 Roadmap on NTDs, which calls for enhanced control, prevention, elimination and eradication of NTDs, namely the regional elimination of VL in the Indian subcontinent by 2020 (13). Shortly after, the London Declaration on NTDs was signed (www.who.int/neglected_diseases/London_Declaration_NTDs.pdf), wherein several pharmaceutical

companies, donors, endemic countries and non-government organisations declare their commitment to contribute to NIDs elimination through R&D and control programme implementation. The latest Report of the WHO's Strategic and Technical Advisory Group for Neglected Tropical Diseases (www.who.int/neglected_diseases/NTD_STAG_report_2017.pdf) further cements the commitment towards NIDs elimination and shows the progress achieved so far. Regarding leishmaniasis specifically, the Resolution WHA60.13 was adopted by the 60th World Health Assembly in 2007 to promote the awareness of the burden of leishmaniasis, and to monitor of progress of leishmaniasis control programmes (11,14).

Leishmaniasis prevention needs an integrated approach targeting both human and animal hosts (One Health approach). Measures that aim at reducing the incidence of leishmaniasis are directed: i) to people, e.g. diagnosis and treatment of cases (15); ii) to the reservoir, e.g. applying protective insecticide treatment to dogs (16); and iii) to vector, e.g. insecticide spraying. Since no effective vaccine against human leishmaniasis exists, the most effective method of controlling *Leishmania* transmission to date is vector control (17,18).

A large VL elimination campaign was launched in 2005 in South Asia, relying heavily on indoor residual spraying, long-lasting insecticidal bed nets, and environment management, has contributed to the reduction of reported cases (12,19,20). However, success longevity depends on continuous application of control measures, which may not be assured once targets have been achieved, and/or given the cyclical and geographical shifts in leishmaniasis transmission (12,20). As most sand fly species bite mostly outdoors, there is no strong argument for insecticide spraying – in this scenario, and for *L. infantum*- and *L. major*-endemic areas, reservoir control may prove to be a more useful tool.

The main milestones concerning leishmaniasis, since the implementation of the objectives set in 2012's Roadmap for NTDs, are i) the improvement of surveillance and case management ('District Health Information System' platforms), and ii) the introduction of standardized tools for the collection of indicators from all member states, some accessible through the Global Health Observatory website (http://apps.who.int/gho/data/node.main.NTDLEISH), and others limited to high-burden countries (11,12). Regarding reservoir control, three canine vaccines are licensed and currently available in Europe and Brazil, Canileish® and Letifend®, and Leish-Tec®, respectively.

Overall, current existing tools can greatly contribute to decrease *Leishmania* transmission, however, the need for investment in new diagnostics, treatment and prevention tools still stands (21–23). Alternative approaches such as immunochemotherapy or immunotherapy should also be further explored (12,21), and beyond innovation, product accessibility must also be taken into account (24,25).

Particularly, the introduction of vaccines in endemic areas remains a primary objective in the context of leishmaniasis control (26–28), either a prophylactic vaccine preventing infection, disease progression and transmission, and/or an immunotherapeutic vaccine (12).

A few cost-effectiveness studies on the impact of the introduction of human leishmaniasis vaccines were performed, and these have demonstrated that vaccines remain the most cost-effective tool for leishmaniasis control programmes. A vaccine against VL in the endemic Bihar state in India with conferring only 50% protection during 5 years is highly cost-effective compared with current treatments, as well as a vaccine against CL deployed in American countries with providing 70% protection during 10 years (29,30).

2. *Leishmania* parasites and leishmaniases

2.1. *Leishmania* spp. life cycle

The general life cycle is common to all *Leishmania* species (Figure I.1 panel A), and the vertebrate host stage begins when an infected female sand fly takes a blood meal from a naive host. The main reservoirs for *Leishmania* spp. are dogs, rodents and humans.

Sand flies from the genus *Phlebotomus* (Old World) or from the genus *Lutzomyia* (New World) are modified pool feeders, meaning they bite superficially multiple times and feed on pooled blood. Less competent flyers than mosquitoes, sand flies breed in walls, rubbish or rubble, or rodent burrows (31–33). Other transmission routes, which remain exceptional, include congenital transmission, blood transfusion, sharing of infected needles, or (rarely) sexual transmission (31,34–36).

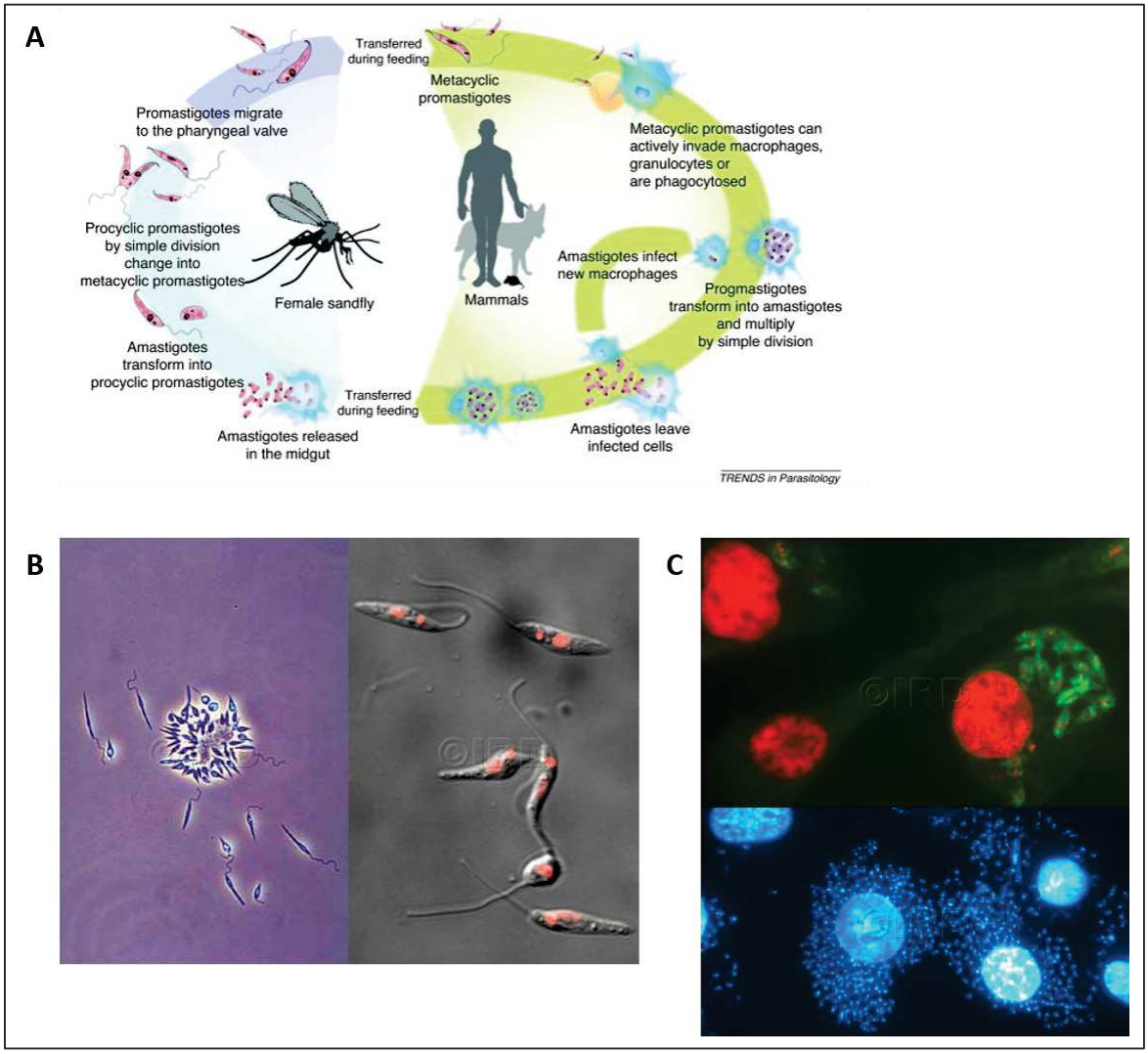


Figure I. 1* Leishmania *parasites and lifecycle. *A)* Leishmania *parasites general lifecycle* (37). *B) Microscopy images of* Leishmania *promastigotes cultured in vitro. Left image, contrast microscopy (© IRD Lemesre, Jean-Loup); right image, parasite DNA stained with red fluorescence dye (© IRD Vergnes, Baptiste). C) Microscopy images of* Leishmania *amastigotes inside human macrophages (© IRD Vergnes, Baptiste). Top image, overlap between green-labeled parasites and DNA stained with red fluorescence dye.*

Once inside the sand fly midgut, the parasite differentiates into a motile extracellular promastigote, firstly to a proliferative procyclic promastigote, and subsequently, to a non-proliferative infectious metacyclic promastigote within approximately one week (Figure I.1 panel B). The infected sandfly takes a bloodmeal from a naive host, injecting metacyclic promastigotes that invade phagocytes (mostly macrophages and neutrophils), where they differentiate into intracellular amastigotes and establish infection (Figure I.1 panel C). The sand fly saliva enhances promastigote infectivity as it contains vasodilator and immunomodulatory molecules (38,39).

Once in the vertebrate host, these parasites have developed key strategies that allow them to thrive in drastic conditions destined to kill them, inside acidic phagolysosomes. Phagolysosomes are cytoplasmic organelles formed after the fusion of the phagosome with one or more lysosomes, that become acidic and contain antimicrobial peptides and hydrolytic enzymes, killing intracellular pathogens (40). The complex *Leishmania* lifecycle is highly adapted to the host's immune system, which is actively manipulated to the parasite's benefit.

2.2. Leishmaniasis distribution

Leishmaniases are distributed worldwide across the tropical, subtropical, and temperate regions in 98 countries, 72 of which are in developing areas of the world (Figure I.2) (12). 350 million people are at risk worlwide and an estimated 12 million people suffer from leishmaniasis (11). There are an estimated half a million new VL cases per year, and 1 to 1.5 million new CL cases per year, with 2.4 million disability-adjusted life-years (DALYs), but these numbers are likely underestimated. Over 90% of all cases of VL are found in seven countries (Brazil, Ethiopia, India, Kenya, Somalia, South Sudan, Sudan) (11,12). Approximately 90% of all CL cases occur in Afghanistan in Central Asia; Iran, Saudi Arabia, and Syria in the Middle East; and in Brazil and Peru in Latin America. CL is also a substantial issue for travellers, and military personnel visiting endemic areas (31). Finally, 90% of the cases of MCL occur in three South American countries: Bolivia, Brazil, and Peru.

Co-infection with HIV has emerged as an important public health threat in areas in southern Europe and other regions where the two diseases coexist, as well as for immunocompromised individuals, as is the case of organ transplants and other conditions affecting cell-mediated immunity (41).

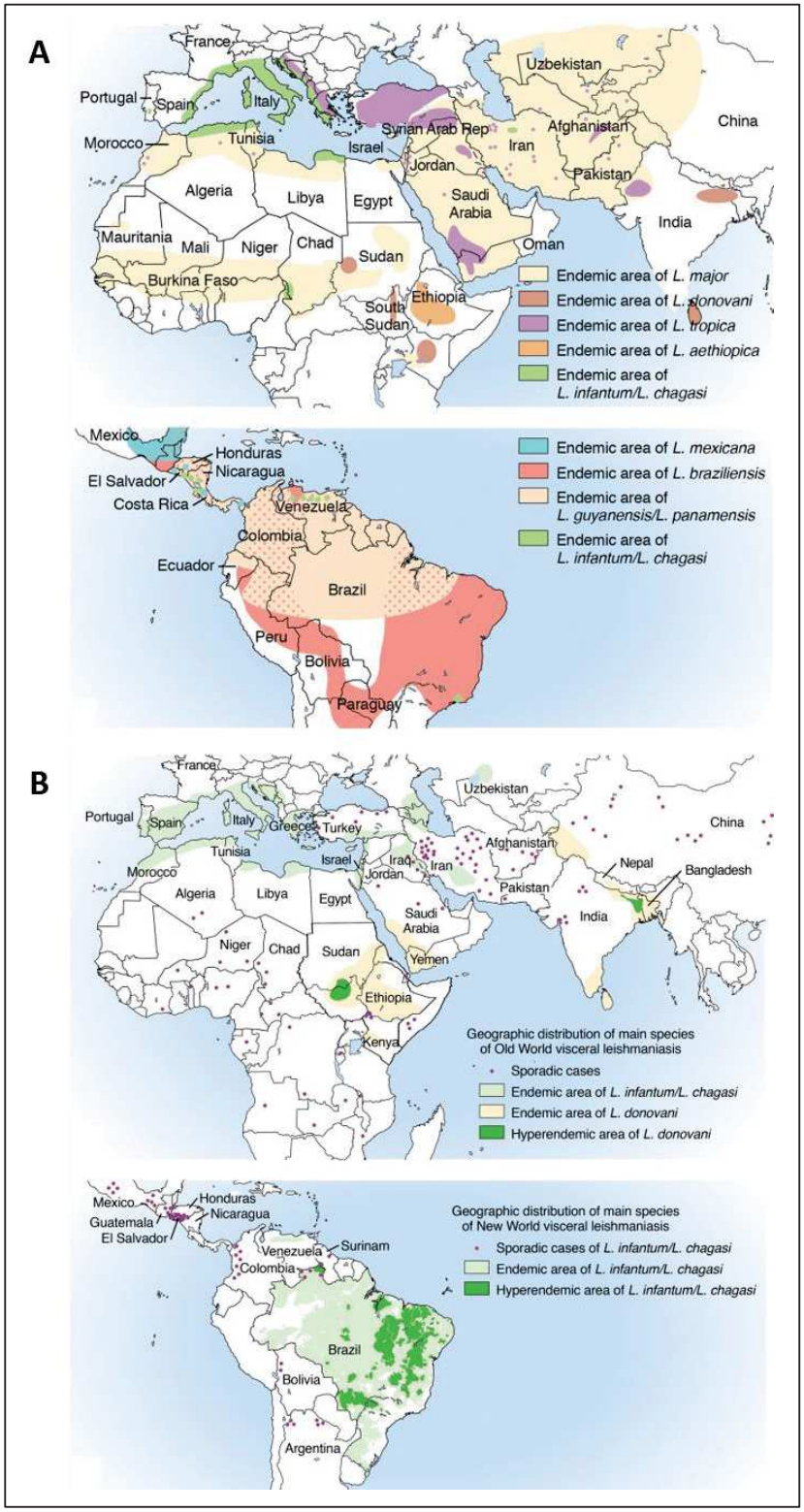


Figure I. 2* Leishmania *world distribution (31,42)***. A) World distribution of CL-causing*** Leishmania ***species. B) World distribution of VL-causing*** Leishmania ***species.***

2.3. Clinical syndromes

The main disease-causing *Leishmania* species and respective clinical syndromes are described in Table I.1.

Table I. 1* Leishmania *parasites and major clinical syndromes. *Adapted from Burza S. et al 2018 and Magill A. 2015 (12,31). ©WHO, World Health Organization campaign photos.*

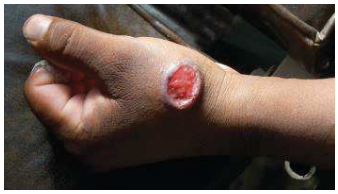
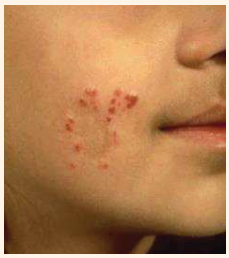
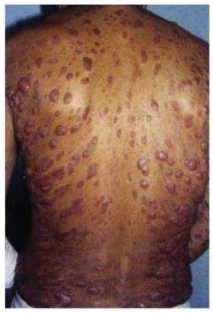
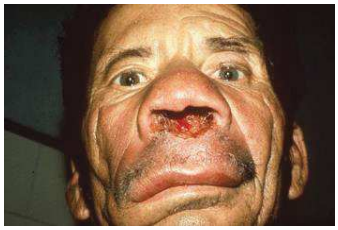
Clinical Syndromes	*Leishmania* spp. and location	Natural Progression	
Visceral leishmaniasis (VL), also known as kala-azar: generalized involvement of the reticuloendothelial system (spleen, bone marrow, liver, lymph nodes)	*L. (L.) donovani* causes classic VL in Asia; *L. (L.) infantum* causes infantile VL in the Old World. *L. (L.) chagasi=L. (L.) infantum* causes VL in the Americas; *L. (L.) donovani* and *L. (L.) infantum* in East Africa (Ethiopia, Kenya, Somalia, Sudan, Uganda); *L. (L.)* amazonensis is an uncommon cause of atypical VL in the Americas; *L. (L.) tropica* is rarely associated with VL syndrome, often atypical.	VL is fatal within 2 years (natural progression)	 © WHO
Post–kala-azar dermal leishmaniasis (PKDL)	*L. (L.) donovani* (Indian subcontinent) *L. (L.) donovani*, *L. (L.) infantum* (East Africa)	PKDL develops in apparently cured VL individuals (5-10% in India, 50-60% in Sudan); PKDL lesions self-heal in up to 85% of cases in Africa but rarely in India	 © WHO
Old World cutaneous leishmaniasis (CL): single or limited number of skin lesions	*L. (L.) major* (also known as moist or rural oriental sore) *L. (L.) tropica* (also known as dry or urban oriental sore) *L. (L.) aethiopica* *L. (L.) infantum=L. (L.) chagasi* (rare) *L. (L.) donovani, L. (L.) infantum*	Self-healing in over 50% of cases within 8 or 12 months (different ulcer morphology and scarring/species); *L. (L.) infantum* healed lesions confer individual immunity	 © WHO/C.Black
New World cutaneous leishmaniasis (CL): single or limited number of skin lesions	*L. (L.) mexicana* (chiclero's ulcer) (Central and South America) *L. (L.) amazonensis* (Brazil and Amazon Basin) *L. (V.) braziliensis* (Central and South America) *L. (V.) guyanensis* (Guyana, Surinam, Amazon basin) *L. (V.) peruviana* (uta) (Western Andes) *L. (V.) panamensis* (Central America) *L. (V.) colombiensis* (Central America) *L. (L.) infantum*/*L. (L.) chagasi* (Central and South America)	Ulcerating lesions; often self-healing lesions within 3-4 months (*L. mexicana*) or within 6 months (*L. guyanensis*)	 (31)

Table I.1 (continued)

Clinical Syndromes	***Leishmania* spp. and location**	**Natural Progression**	
Leishmaniasis recidivans (LR)	*L. (L.) tropica* (North Africa, Middle East)	May last for many years (Tuberculosis-like presentation)	 (31)
Diffuse cutaneous leishmaniasis (DCL)	*L. (L.) amazonensis* (Brazil, Amazon basin) *L. (L.) mexicana* (Central and South America) *L. (L.) aethiopica* (East Africa)	Slow to heal lepra-like lesions (within 2 to 5 years); DCL and HIV co-infection seldom reported	 (43)
Disseminated leishmaniasis	*L. (V.) braziliensis* and *L (V.) amazonensis* (Brazil)	Not well described	
Mucocutaneous leishmaniasis (MCL)	*L. (V.) braziliensis* (espundia) (Central and South America) Other *Leishmania (V.)* spp. (*guyanensis, panamensis*) are rare	Ulcerating lesions (palpable lymph nodes before and early on in the onset of lesions; Possible self-healing) 2-5% of patients infected by these species develop MCL	 (31)

Depending on the infecting species and host immunity, different clinical syndromes develop. However, most infected individuals remain asymptomatic, indicative of the development of an effective immune response, even if incomplete. Evidently, the ratio between asymptomatic:active infection varies widely according to the infecting species and endemic region (Table I.2).

Table I. 2 Ratio between number of asymptomatic and active disease cases, according to geographical region. *Adapted from Singh OP et al 2014* (44).

Country	Ratio (asymptomatic:active infection)	Reference
Sudan	1:2,4	(45)
Kenya	4:1	(46)
Ethiopia	5,6:1	(47)
Brazil	18:1	(48)
Spain	50:1	(49)
Bangladesh	4:1	(50)
India and Nepal	8,9:1	(51,52)

In patients where immune responses are inadequate to control parasite proliferation, the infection progresses to active disease. Leishmaniasis encompasses a spectrum of clinical syndromes (Table I.1 and Figure I.3), from self-healing CL, to chronic, or disseminated visceral disease, indicative of both parasite diversity and variable host responses (31,53,54).

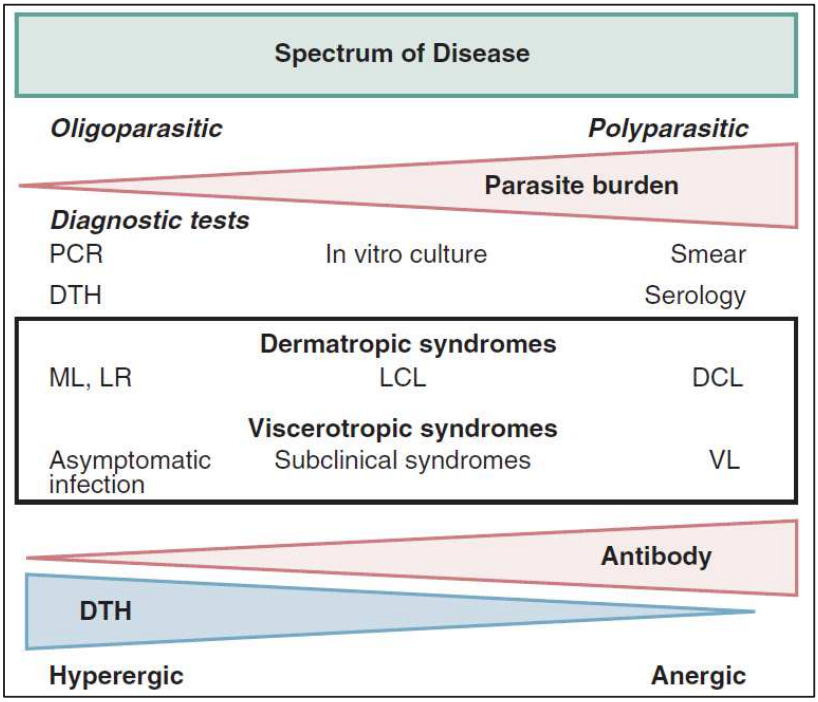


Figure I. 3 Spectrum of* Leishmania *infection and disease (31)***.*** *DCL, diffuse cutaneous leishmaniasis; DTH, delayed-type hypersensitivity; LCL, localised cutaneous leishmaniasis; LR, leishmaniasis recidivans; ML, mucosal leishmaniasis; PCR, polymerase chain reaction; VL, visceral leishmaniasis.*

Polyparasitic active leishmaniasis (DCL, PKDL or VL) presents with heavily parasitized macrophages in the dermis, with few lymphocytes present, and peripheral PBMC that do not proliferate upon antigenic stimulus, do not produce IFN-γ or IL-2, and patients do not produce any DTH reaction and show high levels of IL-10 (31). Oligoparasitic active leishmaniasis is characterised by the slow progression of chronic lesions, with high PBMC recruitment and infiltration, and can progress to granuloma formation as in the case of *L. donovani* liver infection, a site of chronic inflammation usually triggered by persistent infectious agents which have a central area of macrophages, often fused into multinucleate giant cells, surrounded by T lymphocytes (31,55).

Asymptomatic and subclinical infections are not yet well defined. Cases can be detected by a positive serological test, PCR or Leishmanin skin test (LST) in individuals otherwise healthy. However, currently available serological tests were developed to detect active VL, and PCR positivity fluctuates greatly over time (low parasitemia, different target sequences, and short DNA half-life in the body). The LST is useful to detect cellular-mediated responses in patients or exposed individuals, but it is seldom negative in active VL patients. Recovered / healed individuals can usually be identified by a positive LST, as well as positive responses against *Leishmania* antigens *in vitro* stimulation,

although with lower sensitivity (56). Recovered and self-resolving individuals present long-term protection against disease, in the absence of immunosuppression (57,58). Despite no technique is really adapted and developed for asymptomatic case detection, it is common to use two or more markers for asymptomatic infection detection, seldom serology and PCR (59).

It is still not possible to known which patients will progress to active disease, remain asymptomatic carriers or achieve parasite clearance (44,60). Factors that contribute to the spectrum of *Leishmania* infection and disease are associated to complex interactions between environmental factors, parasite and host-related factors (Figure I.4). The main risk factors for developing human leishmaniasis include environmental or behavioural risks (migration, urbanisation, deforestation, irrigation, lack of bed nets, open houses, house dampness), vector-associated factors (proximity to sand fly breeding sites, proportion of infected vectors, preferred sand fly feeding behaviour), host factors (malnutrition, immune status including HIV infection, other co-infections, age, genetic background), and parasite-associated factors (infecting species, tropism, virulence, co-infections with *Leishmania* RNA virus) (23,41).

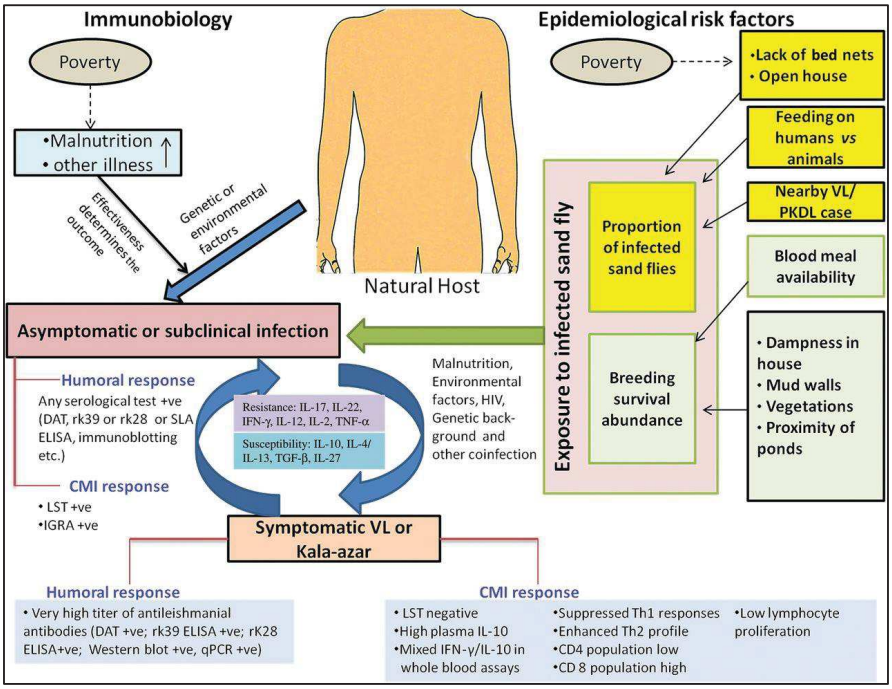


Figure I. 4 Risk factors for the development of active or asymptomatic leishmaniasis (44)***.*** *DAT, direct agglutination test; SLA, soluble leishmania antigen; ELISA, enzyme-linked immunosorbent assay; CMI, cellular-mediated immune response; LST, leishmanin skin test; IGRA, Interferon-gamma release assay.*

3. *Leishmania* spp. pathogenesis and host immunity

3.1. Innate immunity against *Leishmania* spp. infection

The first immune barriers for the parasite to overcome are the complement system and innate immune cells (neutrophils, innate lymphoid cells, antigen-presenting cells) (Figure I.5). *Leishmania* parasites initially interact with skin-resident cells – dermal macrophages, Langerhans cells, and keratinocytes (61). Interestingly, keratinocytes provide a mechanical obstacle against infection, but may also have an important role in the immune response, since they secrete effector cytokines, namely IL-6 which is associated to protection in mice (61–63).

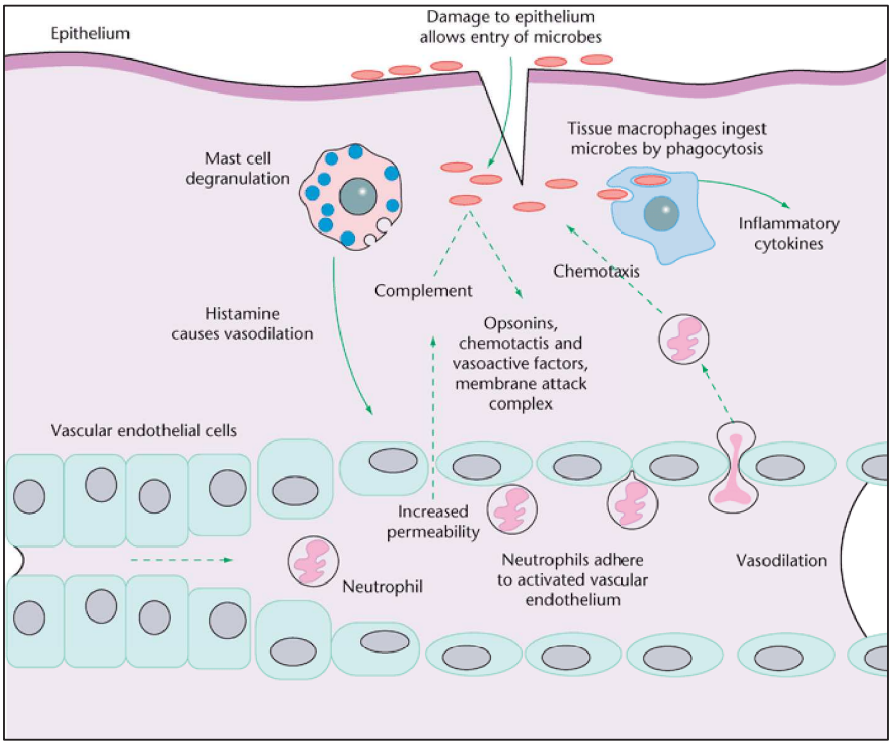


Figure I. 5 Overview of the acute phase of inflammation (64)***.*** *Tissue damage causes mast cell degranulation in the tissues thereby releasing histamine and chemotactic factors. This increases the expression of adhesion molecules, enabling phagocytic neutrophils to adhere and cross into the tissue. Innate immune responses are triggered first. Neutrophils, guided by chemotactic factors, ingest microorganisms by phagocytosis. Increased permeability allows complement components to enter, generating a variety of antimicrobial and pro-inflammatory molecules. Meanwhile, tissue macrophages ingest any microorganisms, releasing simultaneously inflammatory cytokines. These cause vasodilation, increased permeability and expression of adhesion molecules. Inflammatory cytokines also cause Langerhans cells (not represented) to migrate to draining lymph nodes where they activate T lymphocytes to initiate an adaptive immune response.*

a) First obstacles: the complement system

The complement system comprises a large number of plasma proteins which interact among themselves and react with other components of the immune system, enhancing antibody and phagocyte function and inducing inflammatory responses against infection (65). There are three complement activation routes – alternative, lectin, or classical pathways (65). All pathways culminate with the activation of C3 convertase which generates C3b and C3a from C3 (complement component 3). C3b is deposited on the surface of the pathogens allowing for parasite killing through opsonisation and subsequent phagocytosis by neutrophils and macrophages, or through the formation of the membrane attack complex (MAC) complex and cell lysis (66) (Figure I.6).

All three pathways seem to be important for *Leishmania* clearance, and undoubtedly the alternative complement activation is critical for parasite elimination (67). Between 85% to all promastigotes are killed by complement within a few minutes in human blood (*L. donovani*, *L. infantum*, *L. major*, and *L. amazonensis*), so parasites must quickly establish infection inside phagocytes, namely inside macrophages, their definitive host cell (67,68).

Leishmania parasites found elegant solutions to not only overcome targeting by the complement system, but also to exploit host opsonins to invade host cells and modulate their response in its favour, mostly through the interaction with glycocalyx molecules and other virulence factors (62,69,70). The major parasite surface proteins involved in initial first host-pathogen molecular interactions are: lipophosphoglycan (LPG), a glycoconjugate; the GP63 metalloprotease, also known as promastigote surface protease (PSP) or leishmanolysin, present only in promastigotes; and glycosyl-inositol phospholipids (GPI), present in large numbers in both promastigotes and amastigotes (63,69).

Both LPG and GIPLs bind to the mannan-binding serum protein (MBP), which can activate the complement system in an antibody-dependent manner through the lectin activation pathway (63). LPG is the major acceptor of C3b. Simultaneously, LPG and GP63 are involved in promastigote complement resistance. The LPG on the parasite surface prevents the complement membrane attack complex insertion: when the C5b-MAC complex binds, it does not affect the parasite membrane, and is released as soluble C5b-9 complex (63,71). GP63 can inactivate C3b and totally inhibit the formation of the C5b convertase complex (63,68). In addition, *Leishmania* promastigote membrane kinases (LPK-1 and potentially others) phosphorylate the C3, C5 and C9 components of the complement, inhibiting complement activation (63,72) (Figure I.6).

Moreover, GP63 cleaves C3, and resulting iC3b opsonises promastigotes promoting macrophage through the interaction with surface C3 receptor, also known as CR3 or MAC-1 (68,73,74). The CR3-dependent entry route is important for establishing infection inside macrophages, favouring parasite development and persistence (73,75) due to associated IL-12 downregulation (63,76). GP63 also cleaves host tyrosine phosphatase SHP1 preventing IFN-γ-mediated classical macrophage activation (63). LPG interacts with the early inflammatory C-reactive protein, triggering phagocytosis (77) (Figure I. 6).

C3b-coated parasites can be internalized by macrophages via the CR1 receptor (complement receptor 1) (73). Promastigotes can also invade macrophages through interactions with host macrophage Fc-γ, mannose or fibronectin receptors (68,74). FcγR-dependent *Leishmania* phagocytosis is required for clearance and protection from disease (78) (Figure I. 6).

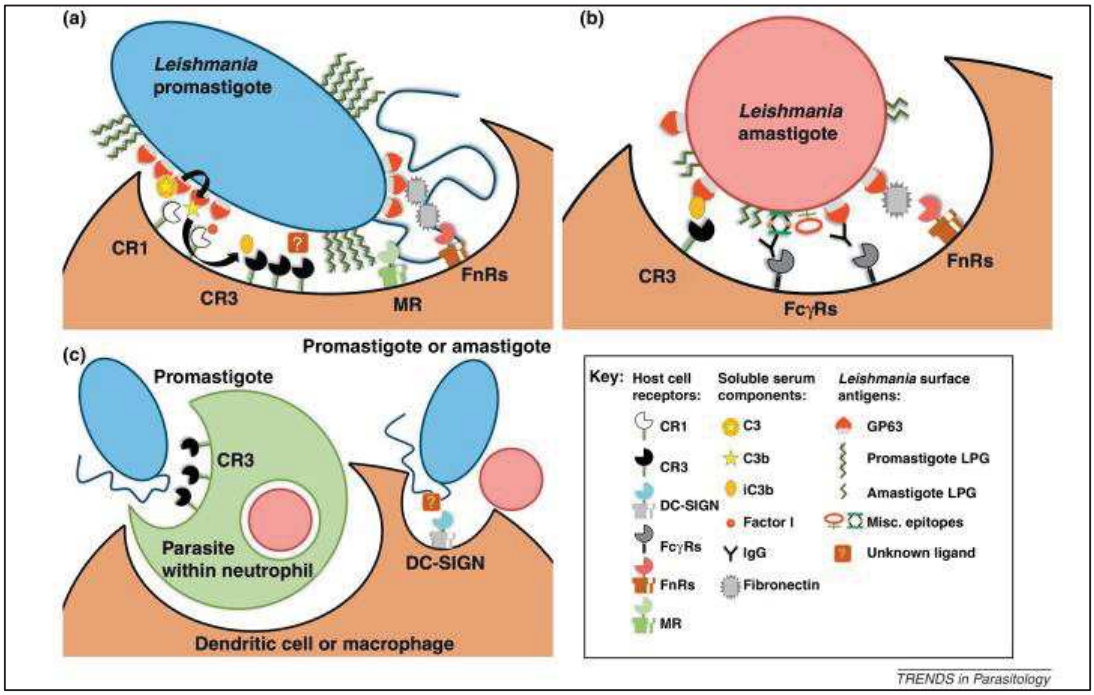


Figure I. 6 Receptor-mediated phagocytosis of* Leishmania *parasites (73). *A) GP63-mediated entry via interaction with CR3 receptor. CR3 may also mediate direct binding to promastigotes via a yet unknown surface epitope on promastigotes. GP63 also binds fibronectin, which then bridges the parasite to fibronectin receptors (FnRs). B) LPG expression on amastigotes is absent, possibly allowing the low levels of GP63 to become opsonized with iC3b protein and subsequently ligate CR3. Antibody and fibronectin detection of amastigotes leads to ligation of Fc gamma receptors (FcγRs) and FnRs, respectively. B) Immediately following inoculation by the sand fly, promastigotes parasitize predominantly Polymorphonuclear Neutrophils (PMN). Promastigotes enter PMN via CR3 and then enter macrophages or DC while contained in the short-lived granulocyte. Promastigotes and amastigotes may also directly enter DC via DC-specific DC-SIGN.*

b) Neutrophils and macrophages, evasion and exploitation of host innate immune responses

Shortly after parasites are injected in the skin, neutrophils are the first cells to be recruited to the inflammation site (79). Because of the sand fly's natural pool feeding behaviour, the tissue damage caused and the injection of saliva proteins together with the promastigote gel and *Leishmania*-produced exosomes, create a highly inflammatory environment through the release of alarmins (signal for tissue damage), cytokines, and chemokines (63,80). Neutrophils can act against the intracellular *Leishmania* parasites through reactive oxygen species (ROS), neutrophil elastase (NE), and neutrophil extracellular traps (NETs) (79,81–87). Nevertheless, neutrophils become infected by promastigotes during the first 18 hours post-infection, providing a transient shelter allowing for parasite development and immunomodulation of subsequent response cascades (79,88).

The parasite has evolved to both evade and exploit host innate immune responses, reviewed in Geiger et al 2016 (1). There is evidence that *Leishmania* infection prolongs neutrophil survival *in vitro* and *in vivo*, as parasites delay the onset of apoptosis in infected neutrophils (68,89). This delay provides the time needed for the recruitment of antigen-presenting cells (macrophages and DC) to the infection site. As neutrophils undergo apoptosis, the parasites take advantage of **a Trojan horse strategy**, where phagocytosis of apoptotic neutrophils by dermal macrophages or dendritic cells triggers anti-inflammatory signal pathways (high TGF-β, IL-10 and low IL-12), providing an optimal route for intracellular development, while simultaneously impairing subsequent adaptive immune responses (90–93).

Furthermore, *Leishmania* disease development depends on the presence of apoptotic promastigotes in the virulent inoculum, which expose phosphatidylserine (PS) also enabling a silent cell invasion and inducing the production of anti-inflammatory cytokine TGF-β (68,94,95). It is interesting to note that neutrophils readily phagocytized promastigotes, but amastigotes are not uptaken by these cells, rather by macrophages through endocytosis (CR3- or clathrin- and caveolae-mediated) (63,94). 48 hours after infection, the majority of infected cells are macrophages (96). Parasite are also capable of escaping pre-apoptotic macrophages in membrane blebs that are uptaken by bystander macrophages (68).

During promastigote internalization, the parasite transforms the phagolysosomal microenvironment and survives inside an adapted parasitophorous vacuole (PV) allowing a successful differentiation to amastigote form, by either delaying phagolysosomal fusion (*L. major*, *L. infantum*, *L. donovani*, with or without LPG), or producing large vacuoles that dilute and impair hydrolytic enzymes (*L. mexicana* and *L. amazonensis*) (97,98).

The production of reactive nitrogen and oxygen species (RNS, ROS) by macrophages is tightly controlled by the balance between energy metabolism and cell redox state (2). The enzyme arginase produces either ornithine, used for polyamine synthesis, or it produces RNS and ROS (namely, O_2^-, NO, H_2O_2, and peroxynitrite $ONOO^-$), responsible for intracellular parasite killing. *Leishmania* parasites actively modulate this balance, skewing arginase function towards glycolysis, and ensuring parasite intracellular survival. These mechanisms are reviewed in Holzmuller et al 2018 (2).

Neutrophils and macrophages have concerted activities: neutrophils augment macrophage function to quickly neutralise early inflammatory stimulus, and apoptotic neutrophils signal macrophages to decrease inflammation and promote healing (99,100). In the context of non-healing lesions, neutrophils contribute to the development of chronic infection through the impairment of the recruitment of inflammatory monocytes and monocyte-derived DC and of subsequent Th1 response (92,99,101–104).

c) Sensing danger – TLR signalling

In addition to receptor-mediated phagocytosis, macrophages and other antigen-presenting cells (APC) possess receptors responsible for sensing danger and reacting accordingly through the production of cytokines and chemokines. In this context, danger is perceived as any molecule able to harm the host, from pathogens (Pathogen-Associated Molecular patterns, PAMPs) or from cell stress and injury (Damage-Associated Molecular Patterns, DAMPs). The main PAMPs are microbial nucleic acid, lipoproteins, surface glycoproteins, and other membrane components, and are recognised by Pattern Recognition Receptors (PRRs), including toll-like receptors (TLRs) (66).

Humans have ten different membrane-bound TLRs, most use MyD88 for intracellular signal transduction, except TLR3 that uses MyD88-independent mechanisms (TRIF, TIR-domain-containing adapter-inducing interferon-β; TIRAP, TIR domain containing adaptor protein; or TRAM, TRIF-related adaptor molecule), and TLR4 that uses both. Among them, TLR2 and TLR4 can recognize extracellular promastigotes, TLR3 and TLR9 can recognize parasites in intracellular vacuoles (66,105).

Studies have shown a major role for TLR activation in anti-*Leishmania* immunity, namely TLR4 and TLR9 (106). TLR9 activation has been linked to protective adaptive responses (107). TLR recognition results in the activation of transcription factors (NF-κb and IRFs) and the production of inflammatory cytokines and promotion of NO production (68). TLR4 inhibition results in a M2b macrophage phenotype, high IL-10 production contributing to Th2 responses, whereas TLR4 ligands

are shown to induce IL-12 and NO production, with low concomitant IL-10 production, suggesting a protective effect by the TLR4/MyD88/IL-12 pathway (108,109).

TLRs are crucial for the coordination of the type and magnitude of the innate response, with a dual role in protection against *Leishmania* parasites. Albeit a non-decisive role in infection outcome, TLRs contribute towards an early Th2 bias, promoting parasite survival (108). TLR signalling and control of early innate responses becomes particularly relevant for *Leishmania* immunity since PAMPs are conserved and expressed constitutively across pathogenic species, and they are key virulence factors, suggesting their inclusion when designing vaccine formulations against *Leishmania* (108).

d) Innate lymphoid cells – NK cells

Natural killer (NK) cells belong to the group of Innate lymphoid cells (ILCs), immune cells that mirror T cell phenotypes and functions, yet ILCs do not express antigen receptors not do they undergo clonal selection and expansion upon activation (110). According to cytokine production patterns and associated transcription factors, ILCs can be divided in 3 groups. NK cells belong to group 1 ILCs (ILC1s) that predominantly express interferon-γ (IFN-γ) and depend on the transcription factor T-bet, associated with Th1 T-cell development (111).

NK cells are an important source of IFN-γ in early host parasite interactions, driving $TCD4^+$ cell differentiation into a Th1 phenotype. This secretion can be dependent on IL-2 produced by specific $TCD4^+$ cells and IL-12 secretion by DC, as in the case of *L. infantum* infection, where NK cell activation depends on IL-12 production and TLR9 expression by myeloid DC, emphasising the complex interactions that will eventually lead to either infection control or progression (61,112). LPG interacts with TLR2 and activates NK cells, triggering IFN-γ and TNF-α production, contributing for host protection in the mouse model, possibly through TLR9 activation (108,113). In *L. major* infection, NK cells produce IFN-γ, which can amplify IL-12 production by DC and induce Th1 differentiation (protective role). Pathology-associated functions have also been described – NK cells (NKp46+ CD49b+) are recruited to the spleen and hepatic granulomas where they secrete IL-10 and impair protective immunity in experimental VL models (114,115). During *L. donovani* infection, extensive activation of NK cells induces IL-10 secretion (negative immunoregulatory role of cell-mediated immunity) (114).

3.2. Bridging the gap – innate and adaptive immune response coordination

The early innate immune effector mechanisms described above, namely phagocytosis, cytokine and chemokine production, expression of co-stimulatory molecules on APC, are essential for **T cell differentiation**. Early *Leishmania* infection is marked by a 'silent phase' during which parasites replicate inside host cells without any associated symptoms for about 4 to 5 weeks (116,117). The development of a protective immune response requires the coordinated action of cells of the innate and adaptive immune response (63,117,118) (Figure I.7). The complex early innate immune responses and induced adaptive immune responses are summarised in Figure I.7.

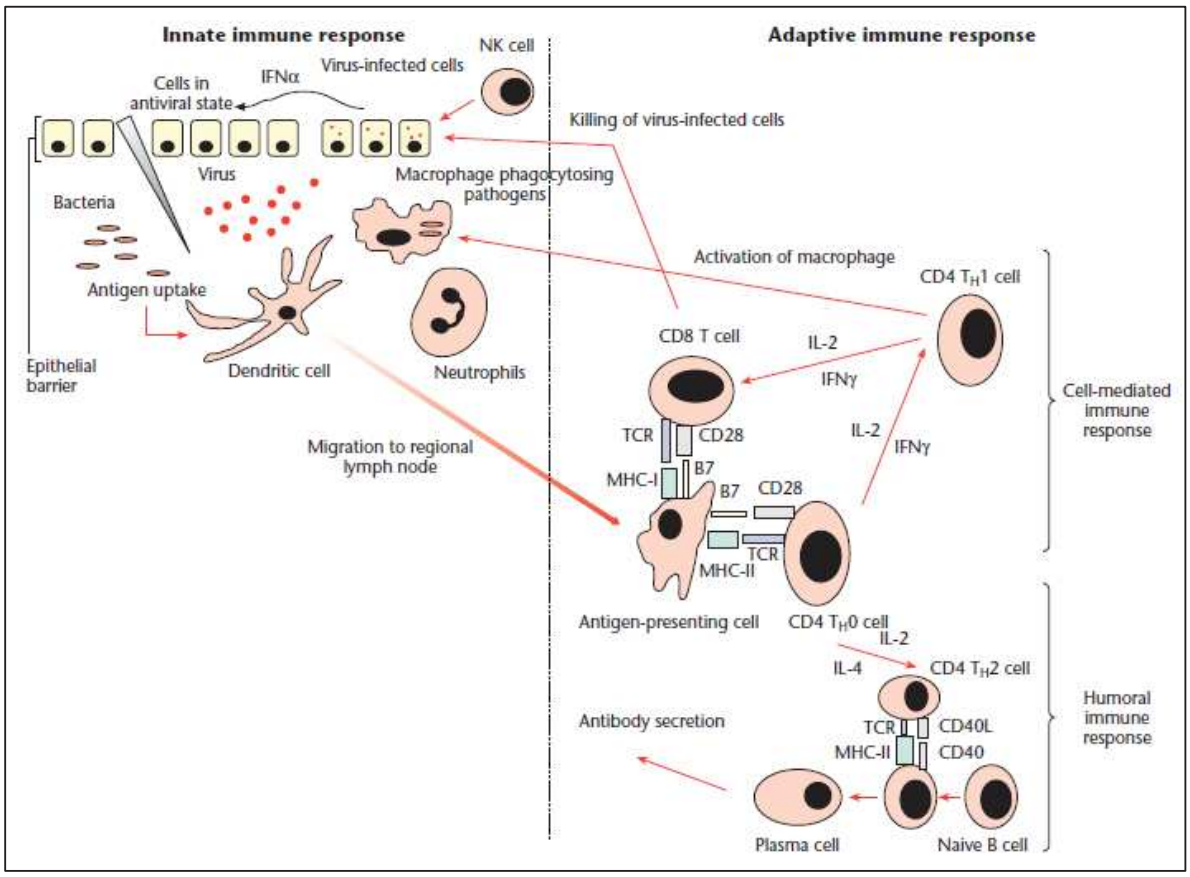


Figure I. 7 Crosstalk between innate and adaptive immune responses (119). *Macrophages, neutrophils and NK cells are attracted to the site of pathogen entry. Dendritic cells take up antigen and migrate to regional lymph nodes. Here antigen presentation to T cells takes place, and these differentiate into helper T cells (Th1 or Th2 T CD4+ cells) and cytotoxic T cells (T CD8+ cells). Activation of B cells and immunoglobulin production are also initiated.*

Skin-resident DC, Langerhans cells and dermal DC can efficiently uptake and present parasite antigens in a pro-inflammatory environment, migrate to the lymph nodes where they can activate

specific naive T cells (120). Most DC become infected after contacting with infected neutrophils. However, some weeks post-infection, the number of DC in the lesion site is higher due to increased recruitment, and these infected DC can prime naive T cells locally, an effector function essential for acquired resistance against *Leishmania* (120) (Figure I.7).

After antigen recognition, naive T cell precursors differentiate into effector cells (**priming**) (Figure I. 8). Naive TCD8$^+$ cells recognise specific peptides presented by major histocompatibility complex (MHC)-class I molecules on the surface of nucleated cells, and differentiate into cytotoxic CD8$^+$ T lymphocytes which kill infected cells. On the other hand, CD4$^+$ T lymphocytes are activated by APC and differentiate into different effector subsets – T helper 1 (Th1), Th2, Th17 or follicular T_{FH} which activate the target cells; or regulatory T cells (Tregs), which inhibit cell activation (121). Naive T cell activation depends on i) antigen recognition of specific pMHC by the T-cell receptor (TCR); ii) expression and interaction of co-stimulatory molecules and, iii) the cytokine environment which determines the differentiation pathway (121). T cell activation will be further discussed below (3.2 b Immunological synapse between APC and T cells).

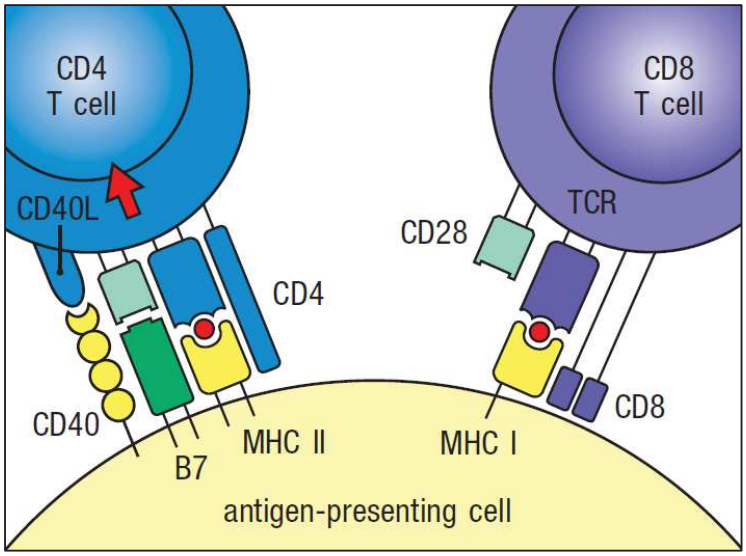


Figure I. 8 Naive T cell priming. *Interaction between antigen-presenting cells and naive T cells (represented here: TCR/pMHC interaction and costimulatory molecules) leading to the differentiation into effector cells* (122)*.*

a) Antigen presentation and the Major Histocompatibility Complex (MHC)

Two classes of MHC molecules, membrane-associated glycoproteins, present antigens to $TCD4^+$ and $TCD8^+$ cells:

- **MHC-class I** exist on the surface of all nucleated cells, and present 8-12 amino-acid peptide epitopes to $TCD8^+$ cells, acting against intracellular pathogens. Briefly, short peptides from intracellular antigens are processed in the proteasome, transported to the endoplasmatic reticulum by the transporter associated with antigen processing (TAP protein), where they are trimmed and loaded onto nascent MHC-class I molecules (123,124).
- **MHC-class II** exist only on the surface of APC (B cells, macrophages, DC, epithelial thymus cells), and they present 13-25 amino-acid epitopes to $TCD4^+$ cells, acting against extracellular pathogens as well as inducing differentiation of other T helper phenotypes thereby regulating the immune response. Briefly, endocytosed extracellular antigens or antigens from intra-vesicular pathogens inside macrophages are degraded in endocytic vesicles and presented by MHC class II molecules to $TCD4^+$ cells resulting in APC activation (123).

Soluble antigenic peptides of the correct length and sequence can directly bind to MHC-class I molecules. Although artificial, this direct binding can be exploited in the context of immunoassays, when peptides cannot be presented via endogenous class I presentation (123–125).

Exogenous peptides can be presented by MHC class I molecules through cross-presentation (126), and it has been described to occur during *Leishmania* infection (127,128).

In humans, MHC molecules are named **Human Leukocyte Antigen (HLA)** (Figure I.9).

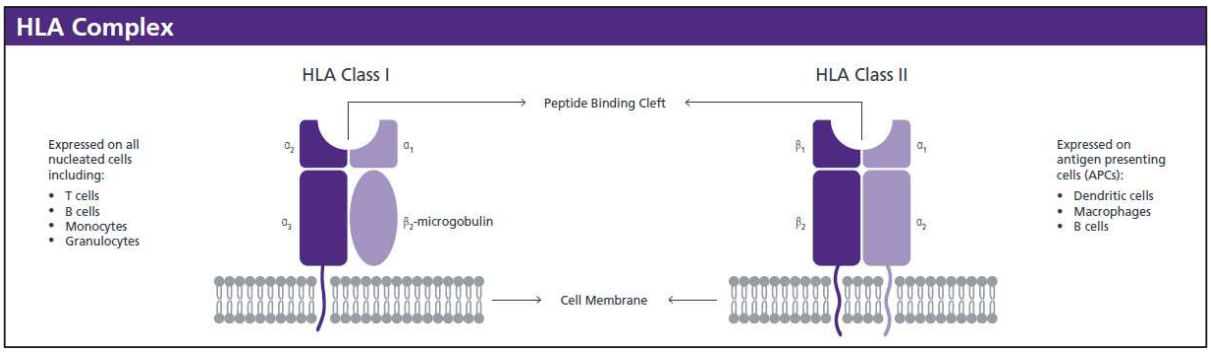


Figure I. 9 HLA class I and class II molecules *(©STEMCELL Technologies Inc. 2017).*

The HLA molecules are encoded in a cluster of closely linked genes at chromosome 6 comprising 6 class I loci and 13 class II loci (Figure I.10). There are two components to HLA diversity: i) gene families, encoding HLA-class I heavy chains and -class II alpha and beta chains; and ii) genetic polymorphism, responsible for multiple alternative forms of a given gene (alleles), and respective encoded protein (allotype). HLA gene polymorphism corresponds to variability at a gene locus in which all variants occur at a frequency greater than 1% (55). HLA gene expression is codominant, the two alleles are expressed in roughly equal amounts in the heterozygote, and each individual expresses a unique set of HLA alleles, specific to a unique set of antigenic peptides (123,124).

A given isoform is the product of an individual HLA allele (123). There are monomorphic genes and oligomorphic genes, for which different alleles have been described (Figure I.10). Six HLA-class I isotypes: **A, B, C**, highly polymorphic genes, and with a role in antigen presentation to $TCD8^+$ cells; **F**, monomorphic gene, intracellular; **E** and **G** genes, oligomorphic and important for NK-mediated responses. Also, there are five HLA-class II isotypes: **DM, DO**, oligomorphic genes encoding proteins involved in the loading of other class II molecules; **DR, DQ, DP**, highly polymorphic genes, encoding proteins responsible for presentation to $TCD4^+$ cells (123).

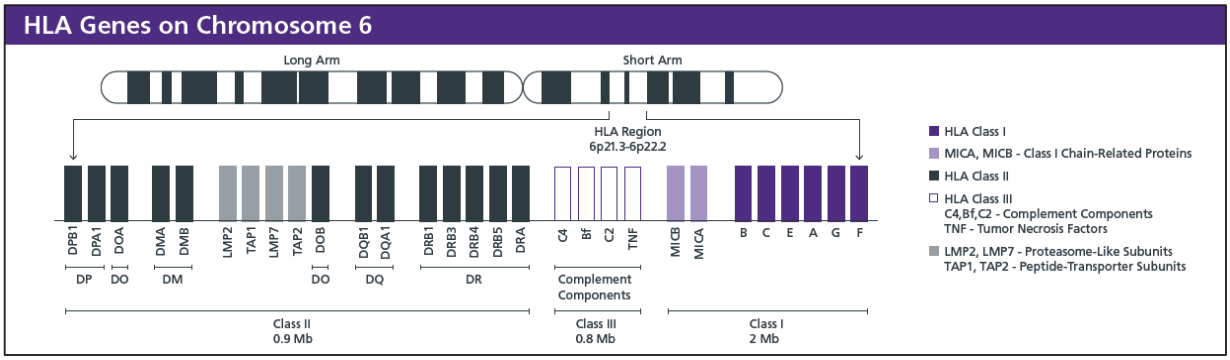
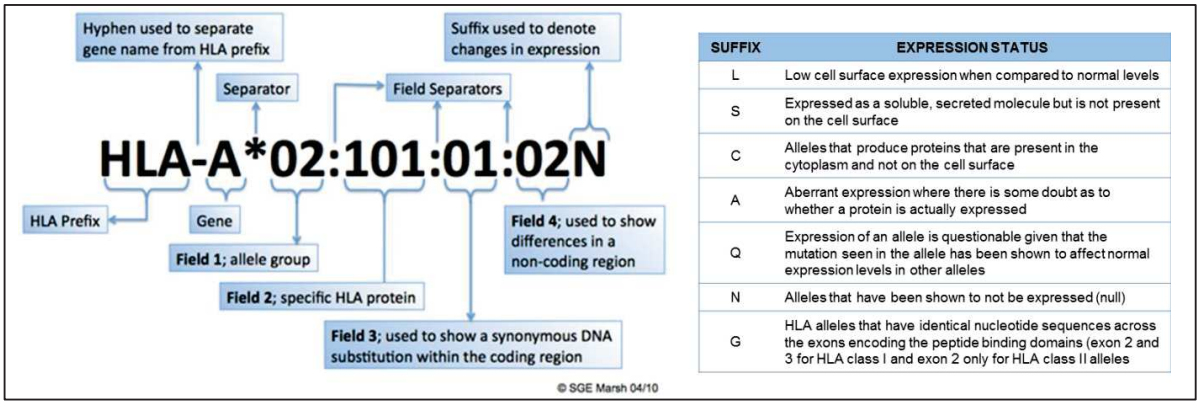


Figure I. 10 Genetic loci in chromosome 6 encoding for all HLA proteins *(©STEMCELL Technologies Inc. 2017).*

HLA gene nomenclature describes this diversity (Figure I.11), e.g. **HLA-A*02:101:01:02N** corresponds to: "**HLA**" prefix for human genes; "**A**" genomic locus; "**02**" allele group; "**02:101**" is the encoded protein, the specific HLA protein numbered by order of discovery; "**01**" is the allele variant, a silent/synonymous polymorphism within the coding region; "**02**" is a SNP number, denotes differences in a non-coding region; "**N**" is a suffix to denote additional information. For example, alleles that have been shown not to be expressed - '**N**ull' alleles - have been given the suffix '**N**'. Alleles that have been shown to be alternatively expressed may have the suffix '**L**', '**S**', or '**Q**'. (Figure I.11) (129).



SUFFIX	EXPRESSION STATUS
L	Low cell surface expression when compared to normal levels
S	Expressed as a soluble, secreted molecule but is not present on the cell surface
C	Alleles that produce proteins that are present in the cytoplasm and not on the cell surface
A	Aberrant expression where there is some doubt as to whether a protein is actually expressed
Q	Expression of an allele is questionable given that the mutation seen in the allele has been shown to affect normal expression levels in other alleles
N	Alleles that have been shown to not be expressed (null)
G	HLA alleles that have identical nucleotide sequences across the exons encoding the peptide binding domains (exon 2 and 3 for HLA class I and exon 2 only for HLA class II alleles

Figure I. 11 HLA gene nomenclature (130)**.**

The **peptide-binding cleft or groove** is the longitudinal cleft in the top surface of an MHC molecule into which the antigenic peptide is bound (55). Each allele has a unique binding groove that defines its specificity, and many of the MHC polymorphisms change the pockets affecting peptide binding specificities (123). The bonds between MHC molecules and the peptidic epitope are determined by their molecular structure and main anchor binding positions, respectively. The peptides insert their amino-acid side chains (anchor residues) into pockets in the peptide binding groove. For MHC-class I molecules, the peptide binding groove is formed by the α1 and α2 domains, and the 9-mer peptide anchor residues are often determined by the B and F pockets (peptide's P2 and P9 residues) (124) (Figure I.12). On the other hand, the peptide binding groove of MHC-class II molecules is formed by the α1 and β1 domains and, contrarily to MHC-class I, they are open at both ends, presenting longer peptides of variable length (123). Also, MHC-class II molecules present higher binding pocket variability, which hinders the identification of anchor residues, although four main pockets were identified which interact with peptide's P1, P4, P6 and P9 residues (131) (Figure I.12). The MHC-class II binding pockets accommodate a 9-mer core region which determines binding affinity and specificity (131) (Figure I.12).

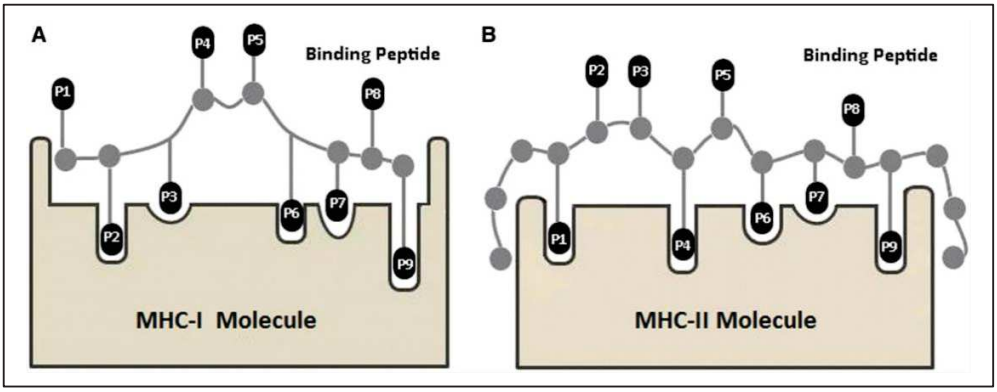


Figure I. 12 MHC binding cleft and interacting residues along the peptide sequence (124)*. A) MHC-class I binding groove. B) MHC-class II binding groove.*

HLA **supertypes** are allele groups with shared binding specificities (supermotifs), meaning that if a peptide is able to bind to an allele within a supertype, it will bind to other alleles in the same supertype (132,133). From binding data, structural analysis and motif identification, nine HLA-class I supertypes were described - HLA-A1, -A2, -A3, -A24, -B7, -B27, -B44, -B58, -B62 (134). Lund et al further defined three new HLA-class I supertypes (HLA-A26, -B8, and -B39) (135). Sidney et al proposed a revised version of these groups, and found that 80% of the 945 alleles evaluated were classified into one of the nine supertypes previously identified (133). However, some alleles were found to have binding specificities pertaining to different supertypes (9 alleles A01A03, and 10 alleles A01A024) and some alleles were 'unclassified' (133). MHC-binding motif information is readily accessible (www.iedb.org) and MHC sequence data are also available in the IMGT, the international ImMunoGeneTics information system database (www.imgt.org). Allele classification in supertypes allows to reduce their inherent complexity, assisting experimental design and validation of potential epitopes (132). Evidently, despite the practical advantages, oversimplification of the HLA allele complexity may introduce unwanted bias.

Moreover, 62% of all EBV- and HIV-derived peptides have motifs associated to two or more supertypes and include 21% of peptides with supermotifs bound with an affinity of maximum 100nM (133). 95% of all epitopes were recognized by individuals expressing different alleles and even other supertypes (133), and this feature is referred to as epitope **promiscuity** (136).

Although evidence is scarce, some **HLA polymorphisms** have been associated with leishmaniasis susceptibility. Genome-wide association studies in Indian and Brazilian populations found polymorphisms in the HLA-DRB1–HLA-DQA1 class II region contributes to visceral leishmaniasis susceptibility, suggesting that genetic risk factors for visceral leishmaniasis are shared

despite differences in geography and infecting parasite species (137). Another study in the Sri Lankan population, found a high heterogeneity in detected alleles, but provides preliminary evidence for some alleles associated with protection (the B*07 allele and the DRB1*15 DQB1*06 haplotype which are over represented in controls; and the haplotypes DRB1*04 DQB1*02 and DRB1*07 DQB1*02 which were absent in patients) or susceptibility (DRB1*15 allele which is over represented in patients), suggesting a role for certain class I and class II HLA genes in LCL susceptibility. However, another study with 110 Peruvian individuals, which also focussed on HLA class II loci, did not find any association between the different alleles and haplotypes detected, and susceptibility to CL or MCL (138). Another study focusing in HLA class I genes showed a statistical significance between HLA-A26 expression and susceptibility to disease in VL patients (139).

b) Immunological synapse between APC and T cells

Naive T cells recognise specific peptide:MHC complexes (pMHC) in the APC surface, in combination with co-stimulatory signals, which triggers T cell proliferation (IL-2 mediated) and differentiation (122). Briefly, naive T cell priming depends on 3 consecutive signals (Figure I.13):

i) stable interaction between the pMHC and the TCR and CD4 or CD8 (activation signal);

ii) the interaction between costimulatory molecules CD28 on the T cell with B7.1/B7.2 (also known as CD80/CD86), and between CD40L on T cells and CD40 on the APC surface (survival or costimulatory signal); the lack of costimulary molecules inhibits cell maturation and activation;

iii) cytokine signalling, namely IL-2, IL-6, IL-12, IL-4, TGF-β (differentiation signal).

Once cells are in contact, when T-cell receptors bind to an antigen and cell-adhesion molecules bind to their counterparts on the two cells, the immunological synapse is formed, a highly organised interface between a T-cell and the target cell, and the adaptive response ensues.

Specifically, macrophages and dendritic cells need two signals for activation and differentiation from Th1 cells – IFN-γ signalling, and CD40 ligand (CD40L) interaction on the T cell surface. These signals further increase expression of CD40 and TNF-α receptors, produced by APC themselves, functioning in synergy with IFN-γ to raise and maintain activation (Figure I.13). Secreted or receptor-mediated **costimulatory signals** are essential for cell activation and for the generation of protective immunity (Figure I.13).

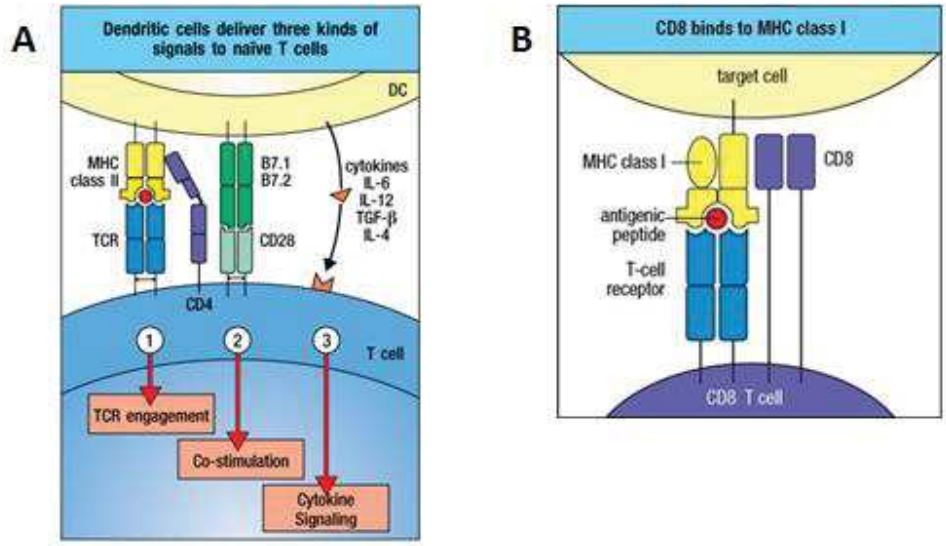


Figure I. 13 Interactions between T cells and antigen-presenting cells. *A) TCD4 cell interaction with MHC-class II molecules on the surface of a dendritic cells and intracellular signalling. B) TCD8 cell interaction with MHC-class I on the target cell surface* (122).

After cell proliferation, the interaction with infected macrophages or dendritic cells, results in the differentiation and acquisition of Th1 or Th2 effector functions: naive TCD4+ cells activated in presence of IL-12 and IFN-γ commit to differentiate into Th1 cells, whereas naive TCD4+ cells activated in the presence of IL-4 commit to a Th2 differentiation (Figure I.14). The main factors that influence the differentiation fate are the cytokine environment, the type of APC, the abundance of pMHC and the binding affinity with the TCR receptor. Molecules produced by activated macrophages are toxic to the host as well, so Th1 responses are tightly controlled by Th2 responses (IL-4, TGF-β, IL-10, IL-13) and regulatory T cells (122) (Figure I.14).

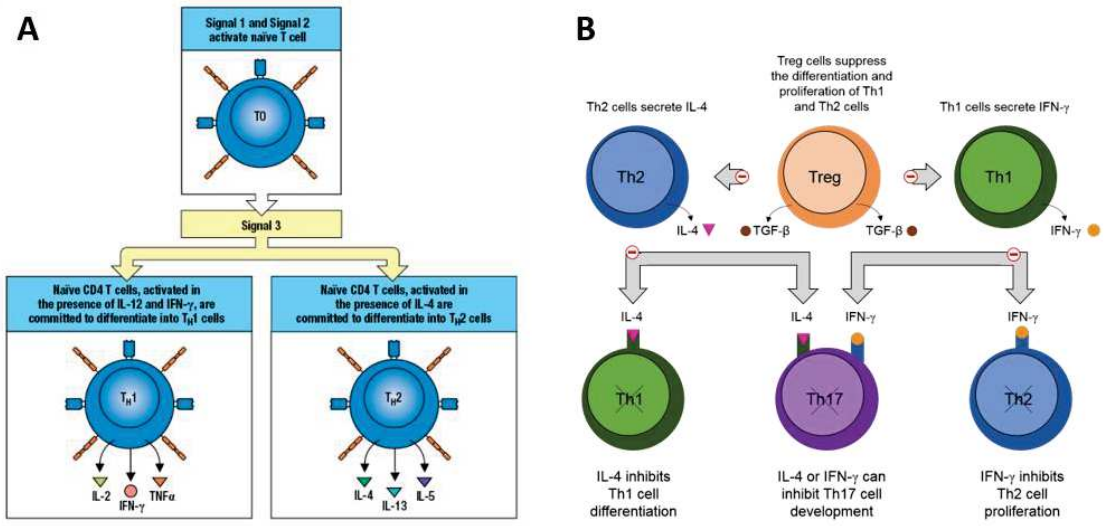


Figure I. 14 Naive T cell differentiation in Th1 or Th2 effector cells and cross-regulation*, adapted from* (121)*. A) Th0 cells differentiate into Th1 or Th2 effector cells according to cytokine environment. B) Th1/Th2/Treg cross-regulation.*

Leishmania parasites actively prevent macrophage activation and interfere with the development of adaptive immune responses through several mechanisms (69,140,141):

i) impairing cell function, by decreasing MHC class-I and -II expression through direct parasite internalization or targeted vacuolar fusion; or by changing TCR interaction through increased membrane fluidity, or by interfering with adhesion molecules and co-receptors;

ii) *Leishmania* cathepsin-B-like protease induces TGF-β production, activating macrophage arginase, and resulting in high ornithine, low NO levels, which favour parasite survival;

iii) anergy induction, in active VL TCD8+ cells typically do not produce IFN-γ, and express high levels of CTLA-4 and PD-1, negative regulators of T cell activity associated with T cell anergy and exhaustion.

3.3. Adaptive immunity mechanisms

a) Primary and secondary immune responses

The human adaptive immune system includes both **naive T cells** (with randomly assigned TCR specificity by genetic recombination during thymic development) and **memory T cells** (selected and differentiated from effector cells after initial antigen encounter into long-lasting specific T cells).

Each TCR is uniquely arranged and randomly selected from over 10^{10} possible combinations. TCR expression is clonal, meaning each mature lymphocyte has a single type of receptor of unique specificity (Figure I.15). The overall diversity of T cell clones in an individual comprises at least 10^8 different lymphocyte specificities, providing the ability to recognise virtually all possible epitopes from a given pathogen and the basis for the complexity of the adaptive immune system. Upon antigen recognition, specific T cells become activated, undergo clonal expansion and become effector cells (Figure I.15). However, lymphocyte diversity and the unique TCR specificity per clone also imply that few antigen-specific T cell clones for a given epitope are present, which poses a problem for response detection.

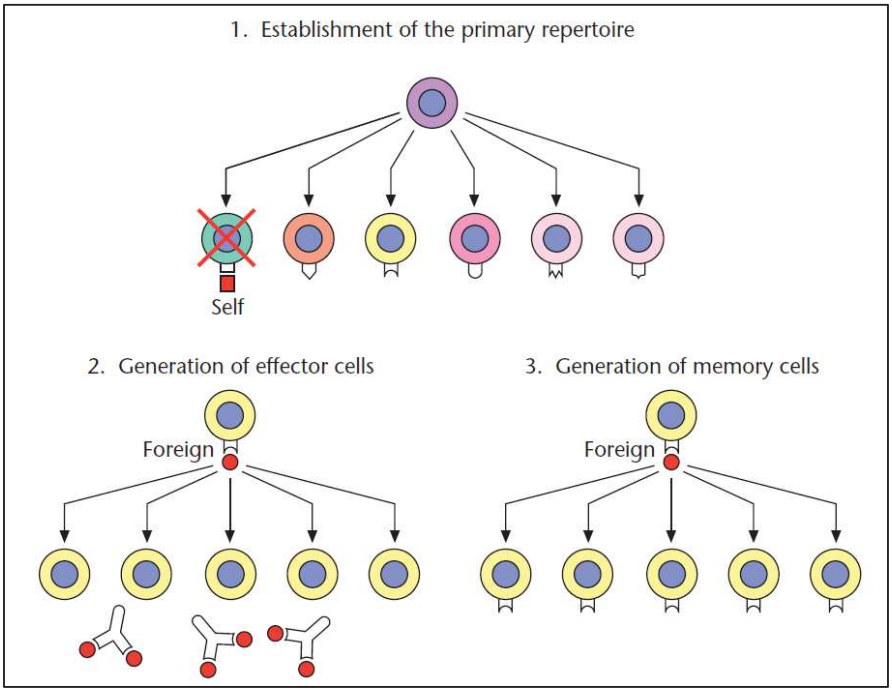


Figure I. 15 T cell clonal selection model (142)*. 1) during development, thousands of cell clones bearing unique TCR are generated in central lymphoid organs. Clones that bind self-antigens with high affinity are deleted and the remainder colonise the secondary lymphoid tissues (primary immune repertoire). 2) The introduction of foreign antigen induces activation, proliferation and differentiation of the specific T cell clones.3) a proportion of these antigen-specific T cells differentiates into memory T cells responsible for stronger and faster responses upon re-exposure.*

During a primary response against infection, specific naive T cells proliferate and differentiate into effector T cells that eliminate the pathogen or infected cells. As the immune response enters the resolution phase, some T cell clones differentiate into long-lasting memory cells and remain in circulation (Figure I.14). Activated memory T cells will proliferate more rapidly and produce higher amounts of effector cytokines. Re-stimulation of these cells results in another proliferation cycle, followed by a new resting phase (Figure I.16). The expansion and maintenance of the antigen-reactive cell populations provide the basis for immunological memory, which translates into enhanced and prolonged secondary responses (Figure I.16). The T-cell clonal selection model applies to both humans and mice.

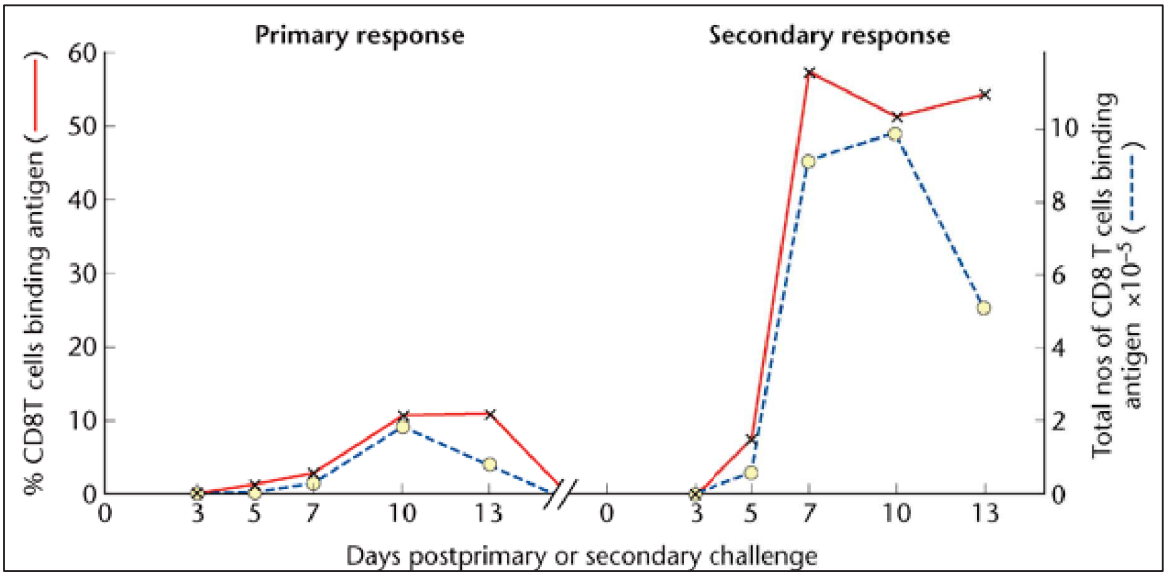


Figure I. 16 Characteristics of primary and secondary cellular responses (143)***.*** *The figure represents the cellular responses (TCD8 lymphocytes) in mice against an Influenza A virus-specific peptide. Mice were immunized twice intranasally, 8 months apart; rechallenge 8 months after initial challenge induces a faster and greater cytotoxic TCD8 response in the lungs.*

b) T cell activation - from peptide presentation to T cell effector functions

Besides peptide- and MHC-related features, T-cell activation depends TCR signalling and TCR-pMHC binding parameters, namely binding affinity and avidity (144).

The TCR-peptide-loaded MHC (pMHC) binding strength is referred to as **affinity**, the sharing of a monovalent interaction, or, the binding strength of one molecule to its ligand at a single site (55,125). Antigen-presentation dictates the fate of developing T cells. In the periphery, unlike in the thymus or bone marrow, T cells are activated if they recognise antigens with high affinity. The T cells

recognising antigens with low affinity generally become nonresponsive or anergic (145). Chronic infections, such as leishmaniasis, can induce high or chronic doses of antigen, in which case, T cells can become anergic, exhausted or apoptotic.

Affinity by itself cannot directly describe the activation extent of a given TCR-pMHC interaction. A high TCR-pMHC dissociation time and serial TCR triggering are more stimulatory than a strong affinity (125,144).

Binding **avidity** refers to all multivalent interactions involved in cell-cell interaction, i.e. the sum of the strength of binding of two molecules or cells to another at multiple sites (55). Avidity-based mechanisms mediate antigen sensitivity in T cells, which results in different binding and activation parameters in naive versus memory cells with the same TCR specificity (146). Throughout the immune response, **primed T cells** acquire greater sensitivity to pMHC signalling than naive T cells, undergoing **avidity maturation** (147). Avidity maturation may be related to decreasing antigen levels during the resolution phase, in preparation for the generation of specific memory T cells, more effective against secondary infection (148).

Although $TCD8^+$ cells are undoubtedly important for host immune responses against *Leishmania* parasites (see below), some controversial results were observed in animal models, namely, CD8-knockout mice which are resistant to infection (149). These differences may be attributed to the antigen dose, as evidenced by the fact that $TCD8^+$ cells are effective against low parasite doses, but not against high parasite loads (150). In turn, these observations further strengthen the need for optimal T cell priming for the generation of long-lasting T-cell memory.

Requirements for the successful induction of an effective adaptive immune response:

- Magnitude of response, a sufficient number of specific T cells is needed;
- Breadth of response, broader diversity of T cell specificity is more effective than responses targeting one epitope;
- Avidity, infection control mediated by high avidity T cells is more effective;
- Cell function, the induction of adequate effector phenotypes is needed (e.g. Th1 and cytotoxic responses)

c) Adaptive immunity against *Leishmania* spp. infection – lessons from mice

Two pivotal aspects regarding adaptive immunity make it an intrinsic mechanism for immunization – response specificity, through the use of highly variable and specific receptors, and memory generation, allowing for a stronger secondary immune response (151).

Generally, protective immunity against leishmaniasis is associated with a predominant inflammatory Th1 response, while disease is associated with a predominant anti-inflammatory Th2 response (63,152). During the initial 'silent phase', immune cells in the infection site secrete IL-4, promoting a Th2 response and thereby an environment that favours parasite development (116).

Interestingly, the Th1/Th2 dichotomy observed in *L. major* mouse infection was incorporated in the first description of Th1 and Th2 helper T cell populations by Mosmann TR et al 1988 (153).

The correlation of Th1 responses and protection against *Leishmania* originates from landmark discoveries from experimental infection studies in mice:

i) *L. major* infection in mice produces different outcome according to genetic background – BALB/c mice develop a Th2 immune response (IL-4-mediated) and are susceptible; C57BL/6 mice develop a Th1 response (with IFN-γ and IL-12 production) and are resistant (121,154);
ii) *Leishmania* spp. immunomodulation of IL-12 production – the decreased IL-12 production by infected macrophages, prevents IFN-γ production by NK cells, and Th1 differentiation and function favouring parasite development (121,155,156);
iii) *Leishmania* spp. induces IL-10 production by regulatory T cells, preventing parasite clearance – the increased IL-10 downregulates MHC-II expression, TNF-α and NO production, leading to reduced parasite clearance and suppressed activation of Th1 cells (157–159);
iv) Passive T cell transfer from resistant C57BL/6 mice to irradiated BALB/c susceptible mice generates protection against *L. major* infection (160).

More recent studies in *L. major* mouse infection (in resistant C57BL/6), CCL2/MCP-1 production resulted in the recruitment of $Ly6C^+$ inflammatory monocytes, capable of capturing and killing parasites by oxidative burst and migration to lymph nodes where they differentiate into specialized DC subsets (161). These $CCR2^+$ monocytes capture *L major* parasites, produce IL-12 and differentiate into iNOS-expressing DC in the lymph nodes, promoting Th1-mediated protection, making them an interesting target cell population for vaccine formulations (161). Also, studies have found that in experimental VL and CL (mice), resistance needs anti-parasitic Tbet+ IFN-γ producing TCD4+ cells (161,162).

d) Adaptive immunity – pathogenesis and human-specific mechanisms against *Leishmania* spp. infection

Human leishmaniases often exhibit a mixed type 1 and type 2 cytokine profiles, so the Th1/Th2 dichotomy observed in mouse models does not apply (163–165). Although mouse models are extremely useful to detail the molecular mechanisms involved in immunopathology, these only offer an incomplete assessment human-specific mechanisms (163,166,167). Notwithstanding, the complex immune responses observed in dogs against canine leishmaniases bear a much closer resemblance to human responses (168).

During active VL and DCL (diffuse cutaneous leishmaniasis), the host immune response is dysfunctional, and in human CL and MCL immune-mediated tissue pathology is observed. Active VL patients possess weak T-cell mediated responses unable to control infection – Th1 responses are balanced with immunosuppressive mechanisms, i.e. specific Th1 cells are activated but not enough to prevent disease development.

In **human CL**, the cell-mediated responses at the lesion site are crucial for disease outcome. Human CL caused by *L. major* induces high levels of IFN-γ, IL-10, IL-12 mRNA in lesions, indicative of mixed TCD4+ responses. IFN-γ-producing cells are definitely predominant in healing cutaneous lesions, but in chronic cutaneous lesions and mucocutaneous lesions we observe both Th1 and Th2 cytokines, with high levels of IL-4 and IL-10 (169).

In **human VL**, no clear correlation exists with increased IL-4 levels (169). However, IL-10 has been implicated with the development of active disease in infected patients (169,170). In active VL, high levels of IL-4 and IFN-γ are detected in the spleen, but these levels decline after cure, and the same profile is observed upon antigenic *in vitro* PBMC stimulation (165,169). IL-10 is produced by different cell types (Tregs, Th2, Th1 and other cells) and differences between CL and VL are still not clear.

The anti-inflammatory cytokine **IL-10** plays a key role in the regulation of host immune responses (121). There is a consensus regarding the suppressive effect of IL-10 on the immune response in VL and its association with disease severity, since this cytokine is an important immunosuppressant and inhibitor of macrophage microbicidal activity in both mice and humans with VL (63,152,171). In human VL, IL-10 is increased in plasma, as well as increased mRNA levels in the spleen, bone marrow and lymph nodes (170). Total PBMC from healed VL patients stimulated with *Leishmania* soluble antigens co-produce IFN-γ and IL-10 (172). On the other hand, IL-10 blockade in VL patient sera or splenic aspirate cultures results in suppressed parasite replication in macrophages

and enhanced Th1 responses (170,173). IL-10 immunomodulation may also involve the induction of T cell exhaustion and anergy (174). It is interesting to note that while in VL patients IL-10 is detected in the plasma (170), this is not the case for patients infected with *L. major* and *L. aethiopica* (175). However, IL-10 mRNA is detected in lesions caused by *L. major* and *L. tropica* (176). This is most likely due to the immune response compartmentalisation, and possibly due to different sources of IL-10 production.

The mechanisms behind Th1/immunosuppression regulation during *Leishmania* infection are not fully understood, however, the cytokine IL-10 has a crucial role in these processes. T cell-derived IL-10 is mostly produced by IFN-γ-producing $TCD4^+$ cells, Type 1 regulatory cells (Tr1, conventional T cells that convert to FoxP3+ regulatory cells) and, to a lesser extent, by thymus-derived Tregs ($CD4^+CD25^+FoxP3^+$) (177). The Treg cell subpopulation has an important dual role in immunosuppression and promotion of concomitant immunity in CL. Treg-produced IL-10 modulates APC functions and suppresses TCD4+ effector functions against *L. major* and *L. braziliensis* and helps maintain a stable infected macrophage reservoir. Tregs bind IL-2 (high levels of CD25 = IL2R) and deprive other cells from this growth factor, causing apoptosis (176,178). Ambiguous and little evidence for Treg involvement in human or experimental VL needs further development (173,179,180).

PKDL pathogenesis remains mostly unknown, there is consensus that immune suppression allows the multiplication of latent parasites from the viscera or residing in the skin. PKDL is characterised by increased IFN-γ and TNF-α, and IL-10 and TGF-β at lesion sites (181). In Indian PKDL, a low expression of receptors is observed, a high number of $TCD8^+$ cells in circulation and in lesions, as well as increased antigen-induced IL-10 production by TCD8+ cells, impaired antigen-induced proliferation and increased Th17 responses (IL-17A, IL-23, RORγT) (182). The presence of Tregs at the skin ($FoxP3^+CD25^+CTLA4^+$) is correlated with parasite burden in Indian PKDL (183). In Sudanese PKDL, some susceptibility-associated polymorphisms in the IFN-γ receptor have been described, and immune responses similar to cured VL patients (184). PBMC from Sudanese PKDL patients proliferate in response to *Leishmania* antigens and $TCD4^+$ cells secrete IL-10 and IFN-γ, similarly to responses observed in cured VL patients. Th17 cells may have a role in parasite clearance in PKDL: PBMC from Indian PKDL patients stimulated with *Leishmania* antigens produce IL-17A and IL-23, and PBMC stimulated with IL-17A show enhanced production of TNF-α and NO (182).

MCL is characterised by the chronic inflammation of the nasal mucosa and a hyperactive T cell response (high levels of IFN-γ and TNF-α, decreased levels of IL-10 and TGF-β). In this case, pathogenesis is caused by a poorly regulated T cell response (174,185). In human MCL patients, the number of $TCD8^+$ cells in the lesion site increases as disease progresses. These cells express high levels of granzymes and perforin (cytolytic activity), contributing to inflammation and disease pathology via perforin-mediated cytotoxicity (186). Th17 cells is also associated with pathogenesis in MCL patients: MCL lesions have increased IL-17A mRNA levels, and TGF-β, RORγT, IL-23 (associated with Th17 differentiation), and IL-17 is produced by not only $TCD4^+$ but also $TCD8^+$, $CD14^+$ and $CCR6^+$ cells (187). High IL-17 increases neutrophil recruitment, suggesting it promotes inflammatory response making it a potential target for immunotherapy (187).

DCL is a severe manifestation of CL, characterised by high antibody titers, TGF-β in circulation, and a defective cellular immune response against *Leishmania* antigens which is restricted to anti-parasitic responses, as responses to unrelated antigens rest intact (188). Human DCL patients have increased parasite numbers in skin lesions, as well as low levels IFN-γ, IL-2 mRNA, and high levels of IL-10, IL-4, IL-5. Therapeutic cure enhances IFN-γ production with low IL-10, further suggesting the need for Th1 responses. In DCL, high antigen exposure may be responsible for T cell unresponsiveness, alternatively, T cell responses may promote localized parasite growth in the skin (189). DCL patients respond poorly to drug treatment and better results are obtained in combination with immunomodulation (IFN-γ and BCG plus antimonial treatment) (189).

3.4. Generation of immunological memory is essential for vaccine development

a) Human immunological memory

Immunological memory is defined as the ability of the immune system to respond more rapidly and more effectively on a second encounter with an antigen; immunological memory is specific for a particular pathogen and is long-lived (55).

Memory T cells are identified by the expression of CD45RO isoform, and by the lack of CD45RA isoform expression ($CD45RO^+CD45RA^-$) (190) (Figure I.17). CD45RA is considered a marker for naive T cells ($CD45RO^-CD45RA^+$). They are also categorized according to the expression of lymph node-homing molecules, such as CCR7 – central memory T cells (T_{CM}) express CCR7 and traffic to lymphoid tissues, and effector memory T cells (T_{EM}) do not express CCR7 and migrate to peripheral tissues

(191). Both subsets produce effector cytokines in response to infection and other stimuli, yet T_{CM} cells produce more IL-2, and T_{EM} have higher proliferative capacity (191). Stem cell memory T cells (T_{SCM}) are $CD45RO^{-}CD45RA^{+}$, similarly to naïve T cells, they also express co-receptors CD27 and CD28, IL-7Rα, CD62L and CCR7, and most importantly, can differentiate into other subsets, including T_{CM} and T_{EM}, hence the designation. Current consensus lies in a differentiation lineage between the different T cell subsets, based on signal strength and degree of activation, from precursor to effector functions. T_{CM}, T_{EM} and T_{SCM} all produce IFN-γ, TNF-α and IL-2 but vary in the proportion of cells producing these effector cytokines (191,192) (Figure I.17).

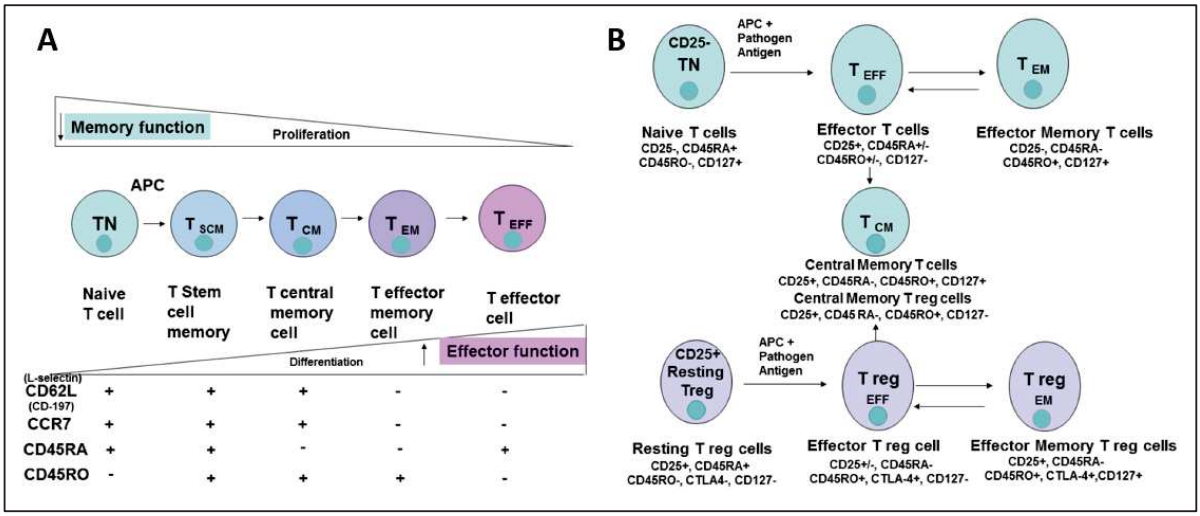


Figure I. 17 T cell subset populations and differentiation (193)*. A) Memory TCD8+ cell differentiation. B) Memory TCD4+ cell differentiation.*

Overall, memory $TCD4^{+}$ or $TCD8^{+}$ memory cells in blood circulation rapidly produce IFN- γ and IL-2 upon non-specific stimulation with phorbol 12-myristate 13-acetate (PMA)/ionomycin and can produce multiple cytokines in effector responses. The presence of polyfunctional memory T cells is associated to strong recall and protective responses (191,194).

Initial T cell recruitment site following an infection dictates the generation of memory cells and even a compartmentalization of pathogen-specific memory T cell responses, as observed by studies with canine leishmaniasis models (195), other viral infections (196,197), or *Mycobacterium tuberculosis* PPD intradermal administration (191) .

The distribution is an important aspect regarding memory T cell populations, as most subsets were described in peripheral blood but there is evidence for memory T cell presence in tissues in mice and humans (191,192). Tissue-resident memory T cells (T_{RM}), described initially in mice, have been shown to be present throughout the human body, with $TCD4^{+}$ subsets persisting in either the mucosa

(CCR7$^-$) or lymph nodes (CCR7$^+$), and TCD8$^+$ cells persisting mostly as CCR7$^-$ cells in peripheral tissue. Most of these cells express a putative marker CD69, absent in circulating memory T cells. Tissue-specific properties have also been described – skin T_{RM} express cutaneous lymphocyte antigen (CLA) and skin-associated homing chemokine receptors CCR4 and CCR10; T_{RM} in the intestinal mucosa express gut-homing chemokines CCR9 and integrin α4β7 – suggesting tissue-specific mechanisms for memory generation (191,192). Most human T_{RM} cells (bone-marrow, lung, mucosal) are polyfunctional for effector and cytolytic cytokines and produce IL-17, whereas skin-specific T_{RM} produce IL-22.

Immunological memory responses against *Leishmania* after drug treatment or self-healing involves central and effector memory TCD4$^+$ and TCD8$^+$ cells (57,80,198–200).

b) Cell populations involved in immune responses against *Leishmania* parasites

The Leishmanin skin test (LST), or Montenegro test, evaluates delayed-type hypersensitivity (DTH) reactions, a form of cell-mediated immunity elicited by an antigen in the skin stimulating sensitized Th1 CD4 and CD8 lymphocytes are methods used to evaluate infected individuals' immune response to *Leishmania* (55). To detect *Leishmania*-specific responses, *Leishmania* antigen (Leishmanin) is injected intradermally in the skin, and induration is measured 48 to 72 hours after (generally indurations>5 mm are considered positive). While active VL patients are generally LST-negative due to cell anergy, six months after successful chemotherapy, around 80% of patients become LST-positive. Hence, while the LST has little value for diagnostic purposes, it may prove useful to detect asymptomatic infections and/or previous exposure in epidemiological studies. Nevertheless, LST is rarely performed nowadays due to supply issues and lack of standardization in antigen production (12,201).

In individuals with history of CL (*L. major* or *L. tropica*), studies have shown the presence of both T_{EM} cells (CD4$^+$CD45RO$^+$CD45RA$^-$CCR7$^-$) that produce IFN-γ, and T_{CM} cells (CD4$^+$CD45RO$^+$CD45RA$^-$CCR7$^+$) that produce IL-2, and their role in protective recall responses (202). An important role was also found for memory CD8+ T_{EM} cell subsets (CD8$^+$CD45RO$^+$CD45RA$^-$CCR7$^-$) in recall responses in healed CL individuals after *L. major* or *L. tropica* infection (200).

Similarly, it was shown that individuals who recovered from *L. infantum* infection, or who are asymptomatic, possess effective and specific memory responses against *Leishmania* (57). Stimulated PBMC from healed individuals showed increased CD69$^+$ expression, a T cell activation marker, in both TCD4$^+$ and TCD8$^+$ cells, increased CD25 expressing in TCD8$^+$ cells, and increased percentage of TCD4$^+$ cells expressing memory marker CD45RO when compared with controls and individuals with

symptomatic disease (57). Asymptomatic individuals can control *L. infantum* infection through the presence of IFN-γ producing TCD8$^+$ cells (53,203).

Furthermore, studies with cured individuals after *L. braziliensis* infection have shown that while both CD8$^+$ and CD4$^+$ specific T cells are present, with the number of TCD8$^+$ cells increasing after cure, CD4$^+$ T_{EM} cells are the main source of IFN-γ produced after antigenic stimulation (204,205). Also, cured CL patients show increased lymphoproliferative potential and higher IFN-γ production, while asymptomatic individuals infected with *L. braziliensis* have higher levels of IL-10, probably involved in the balance between immunoregulatory and effector responses responsible for parasite control without associated tissue damage (205).

After therapy, parasite burden decreases along with *Leishmania*-specific immune responses, an equilibrium that appears to be essential for long-term protection – constant antigenic stimuli (due to reinfection or parasite persistence) can induce and maintain specific memory T cells but also contribute to a chronic activation of effector cells and T cell exhaustion.

There is evidence that time after cure may influence host responses since a reduction in TCD4$^+$ and TCD8$^+$ cell activation is observed two years after the initial *L. braziliensis* infection and cure, and this is again consistent with increased regulatory responses (206). Differences between early lesions and late lesions (*L. braziliensis*) positively correlate the duration of illness with increased CD4$^+$CD69$^+$ T cells, and negatively with CD4$^+$CD25$^+$ T cells (207). Furthermore, again during active CL caused by *L. braziliensis*, both CD4$^+$ and CD8$^+$ T_{EM} cells (CD45RO$^+$CCR7$^-$) are present in lesion sites, in much higher frequency than T_{CM} cells, and this enrichment in effector cells can be associated with immunopathology and tissue damage (207). Nevertheless, the recall responses present in healed individuals persist, and CD4$^+$ and CD8$^+$ T_{EM} cells proliferate and produce IFN-γ after secondary stimulation, whilst having regulatory mechanisms that prevent immunopathology without the loss of protective immunity (206).

Humans immunized with a whole crude vaccine (prepared from promastigote antigens from New World *Leishmania* species) are resistant to reinfection and mount similar responses to those found in recovered patients after *L. braziliensis* infection, showing IFN-γ production without IL-4 as well as a higher proportion of TCD8+ cells, consistent with a long-lasting protective role of TCD8+ cells (53,175,208,209). IL-4 production is only detected during active MCL infection (175).

It is important to note, most human studies are performed with peripheral blood, and additional tissue-specific memory responses cannot be excluded.

c) Memory maintenance and concomitant immunity

Memory T cell maintenance mechanisms remain unclear, and no successful human vaccine targeting cellular-mediated immunity exists, but there is evidence of long-lasting memory T cells. Even for antibody-based vaccines, as is the case of individuals vaccinated against the vaccinia virus, there are specific memory T cells in circulation 25 to 70 years after immunization (191,210). Moreover, a study has found a direct correlation between influenza-specific memory $TCD4^+$ cells and reduced disease severity after challenge (211), and circulating memory $TCD4^+$ cells are associated with non-progression of HIV infection (194).

In mice, the cytokines IL-17 and IL-15 have been implicated in the maintenance of $TCD8^+$ memory T cells, whereas $TCD4^+$ memory cells need TCR and/or MHC-II signalling for their functional maintenance (191,192). In humans, these differences may be associated with tissue distribution where memory cells in the bone marrow, lymphoid tissue, spleen and blood can be maintained by IL-17 and IL-15, and memory T cells in the mucosa are maintained by TCR cross-reactions due to the high antigen density (191,212–214).

Interestingly, healed and asymptomatic individuals continue to carry viable parasites, in roughly constant numbers throughout their life (215,216). This is a very important aspect for infection transmission (217) and for potential relapse in immunosuppressive conditions (44), but also for memory generation and maintenance. Persistent parasites may be necessary for memory maintenance, as revealed by the fact that sterile cure of persistently infected mice makes them susceptible to secondary infection, and the prominent role of T_{EM} cells in anti-*Leishmania* recall responses (218). As previously referred in 2.3, asymptomatic infection is much more common than active disease and even more than sterile cure, and these infections show low parasite levels, no pathology, and protective immunity against reinfection and/or disease severity.

Persistent parasites appear to include two subpopulations: i) quiescent non-replicative parasites which can resist host immune responses; and ii) persistent replicative parasites, similar to acute infection parasites, maintained through a numbers game between active parasite replication and immune clearance (215,218) (Figure I.18). HIV-*Leishmania* co-infection case studies, wherein an underlying *Leishmania* infection becomes active upon immunosuppression (219–221), and the reactivation of latent infection after administering iNOS or IFN-γ inhibitors in mice (222), further support this model. Also, constant parasite replication and killing help maintain the effector memory T cells, and agrees with the loss of immune protection after sterile cure (57,198,215,223,224).

This process of memory maintenance is named **concomitant immunity**. Immune memory against *Leishmania* is mediated through continuous cell priming that maintains memory cells in

circulation and in the skin, allowing for a small number of parasites to develop and replicate inside APC and maintain the memory cell populations responsible for protective immunity (53,215) (Figure I. 18). In mouse models, protection against *L. major* reinfection is associated to parasite persistence mediated by $CD4^+CD25^+$ Treg cells that produce IL-10 (218,225,226).

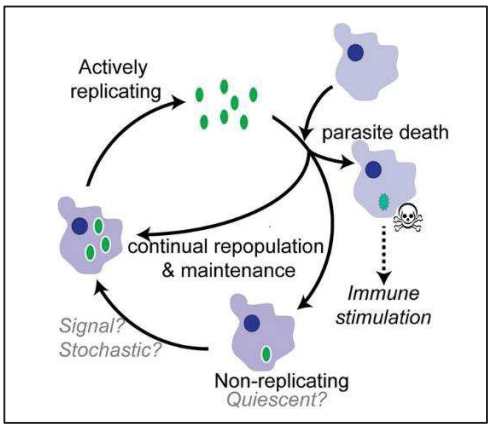


Figure I. 18 Model for the maintenance of concomitant immunity against* Leishmania *parasites (227).

In humans, there is evidence for T_{EM} and T_{CM} cells participation in maintaining immunity against CL, and only T_{EM} require persistent parasite antigens (80,202). After treatment, parasites persist in original site of infection, partly due to IL-10 mediated mechanisms. Persistent parasites maintain CD4+ T_{EM} cells that protect against reinfection (161,218). Alternatively, long-lasting central memory TCD4+ cells (T_{CM}) develop in the absence of persisting antigens and acquire effector functions upon reinfection (223). T_{CM} require additional IL-12 signals to fully develop into functional Th1 cells. In the absence of IL-12, they can convert into IL-4 producing cells. These cells are generated early in infection and are responsible for clearance of primary infection and control of secondary infections (174). Lastly, skin-resident memory TCD8+ cells (T_{RM}) are also maintained in the absence of persistent parasites and provide protection against *L. major* in the mouse model (198,228). In mice, T_{RM} induce the recruitment of inflammatory monocytes, which produce ROS and NO, contributing to the control of parasite growth (228). $CD69^+CCR4^+CCR10^+CLA^+$ skin T_{RM} cells may be important targets in anti-*Leishmania* vaccination (198).

- Early immune responses greatly impact subsequent adaptive response;
- Both TCD8+ and TCD4+ cells mediate protection or susceptibility to *Leishmania* infection;
- IFN-γ is absolutely required to control the parasites sufficiently and control the disease;
- IL-4 is associated with the development of a non-protective Th2 phenotype;
- Human memory subsets involved in long-lasting protection are not very well described, but an important role for T_{EM} and T_{CM} was described;
- Memory maintenance against *Leishmania* is associated with parasite/antigen persistence.

3.5. Leishmaniasis is a vaccine-preventable disease

Overall, the main observations supporting vaccine feasibility are: i) protection against reinfection following spontaneous or drug-induced recovery; ii) protection can be induced through leishmanization; iii) Th1 cells are essential for resistance; iv) $TCD4^+$ and $TCD8^+$ cells are involved in immune memory.

The strongest evidence supporting the idea of inducing protection against *Leishmania* infection comes from leishmanization (LZ), the deliberate inoculation with live, virulent *L major* parasites (229,230). The intradermal inoculation of naive individuals with live promastigotes mimics a natural cutaneous infection, inducing self-resolving lesions, typically smaller than by natural infection, which in turn protect the individual from secondary infection (229,230). The outcomes of LZ reflect the natural host response diversity observed in natural infection, with most of the individuals developing a self-healing lesion and some being able to control infection and never develop a lesion. LZ was used in large scale control programs in endemic countries in Asia and Middle East (namely in Uzbekistan, Iran, Iraq) until the 80‘s, and has proven to be efficacious against Old World CL (231).

LZ-induced protection is associated with parasite persistence and development of concomitant immunity (see 3.4.c). This protection is T cell-mediated: IFN-γ-producing $TCD4^+$ cells are recruited to dermal site of infection, where they perform effector functions including activating microbicidal mechanisms in infected macrophages (232–234).

This protection depends on high parasite virulence and infectivity, as low virulence parasites stimulate DTH responses but do not provide protection (174,232). LZ has also been proposed as an assay to measure vaccine candidates’ efficacy, as it stands as the gold-standard regarding correlates of protection (231). LZ was abandoned due to logistical and safety issues related to dose control,

strain and virulence standardization, and also the occasional development of persistent lesions and associated ethical concerns (229,230).

LZ can provide cross-protection against VL and there is further evidence for cross-species protection:

i) *L. major*-infected CL patients seem to be protected against *L. infantum* infection (235,236);
ii) longitudinal studies have shown that humans in *L. major* endemic areas in Sudan are protected against *L. donovani* VL (237);
iii) a CL-causing *L. donovani* clinical isolate from Sri Lanka protects mice against visceral disease (238);
iv) in experimental VL, C57BL/B6 mice infected with *L. major* are protected from *L. infantum* VL, associated with the recruitment of IFN-γ producing Ly6C+ CD4+ cells to skin and visceral organs (162).

Also, several parasite-related factors support the development of a vaccine against leishmaniases (239,240):

i) No significant antigenic variation is observed between *Leishmania* spp. life stage forms (detailed in chapter II);
ii) The preferential host target cell is the macrophage, and current tools allow for modulation of macrophage intracellular killing and functional enhancement;
iii) High level of antigenic conservation among pathogenic species (detailed in chapter II);
iv) New suitable Th1-inducing adjuvants are today under development (see Table I.3);
v) Parasite – and vector-associated antigens can be used to induce protection.

The three available canine vaccines (CaniLeish®, LeishTec® and Letifend®) provide additional evidence that it is possible to induce protection against leishmaniasis through immunisation in the dog model.

Thus far, no major breakthroughs in the development of a second-generation vaccine were observed. There are many gaps in the knowledge about immune responses against *Leishmania*, correlates of protection and even knowledge on the number of natural challenges after the first inoculation. From 11 published studies testing first generation vaccines (fractions of the parasite or whole killed *Leishmania* with or without adjuvants), only 4 showed a decrease in human infection, measured by LST or Montenegro skin test (MST) seroconversion, highlighting the need for better methods to assess of T-cell mediated responses and correlates of protection (241).

Several inactivated/killed and subunit vaccines have all shown to be quite ineffective in human hosts, despite promising results in animal studies (242). Also, vaccines tested in endemic regions have not performed as well as when tested in healthy volunteers (226,242). There seem to be several underlying issues at play: the use of mouse models to validate antigen candidates (golden hamsters as promising alternative as an experimental VL model); intravenous challenges instead of natural challenge used in the experimental designs; the ability of the immune system to effectively recognize live, but not killed, parasites (242); and the methods to test and measure protective immune responses (226).

Vaccine feasibility:

- Natural protective immunity after recovery from infection;
- Leishmanization;
- Cost-effectiveness studies;
- Parasite-related features.

Main challenges for leishmaniasis vaccine development:

- No cellular immunity-inducing vaccine available;
- Lack of approved Th1-inducing adjuvants for human vaccine development;
- Lack of reliable correlates with immune protection;
- Poor translation from animal models;
- Undefined optimal administration route.

3.6. Immune correlates of protection against leishmaniasis

Specific immune correlates of protection, defined as measurable immune responses that are responsible for and statistically interrelated with protection (243), are largely unknown and non-standardised in vaccine development against *Leishmania*. The mechanisms behind protection, and their evaluation, must be understood so that an effective vaccine is developed.

As detailed above, the resolution of human CL depends on Th1 cells that secrete IFN-γ and activate macrophages for intracellular parasite killing. Similarly for VL, if Th1 cells are absent, or if Th2 cells are predominant, patients develop chronic progressive VL (160,244).

The major role attributed to TCD8+ cells during anti-*Leishmania* immune responses is *via* IFN-γ secretion that, in turn, will induce nitric oxide production and resistance against reinfection, as well as cytotoxic activity towards infected macrophages (244–246). TCD8+ cell-produced IFN-γ has been implicated in both resistance and immunopathology (247).

The known role of T_{CM} and T_{EM} populations in long-lasting protective responses should be explored to assess the quality of induced responses after immunisation and evaluate protection throughout time.

Antibody production, used as a correlate of protection for currently available antibody-based vaccines, does not provide quality evidence of protective responses against *Leishmania*. .
VL and DCL are associated with high parasite burdens and, as infection progresses, DTH responses become poor and total antibody titres rise (248). The anti-*Leishmania* antibodies detected by the diagnostic tests rK39 (India) or rK28 (Africa) persist many years after cure, which can be due to repeated exposure or parasite persistence, and cannot distinguish between asymptomatic, previous or current infection (248). The presence of IgG2 antibodies is detected in healed patients, however, they are not actively functional in infection control (60). On the contrary, there are reports showing a deleterious effect by IgG antibodies in animal models, through the production of IL-10 by macrophages after binding to the Fcγ receptor (249).

Da-Cruz et al performed studies with healed patients infected by *L. braziliensis*, and suggested two markers for protective immunity: i) cure is correlated with high, but decreasing, numbers of *Leishmania*-specific TCD4+ cells, in the absence of IL-4; and ii) decreasing levels IL-5 are associated with long-term cure (175). These data are confirmed by studies with healed patients after *L. major* infection (250). In MCL long-term healing, 1 to 17 years after treatment, there is a slight increase in TCD8+ subpopulation for a final ratio TCD4+/TCD8+ of around 1 (175). After treatment and parasite elimination, it is expected that the population of effector TCD8+ specific cells will retract. The maintenance of specific TCD8+ cells in circulation, and that can re-expand and differentiate upon restimulation, is probably mediated by the TCD4+ cell population. Interestingly, although MCL healed patients showed increasing IFN-γ levels, these were variable, with high and low responders. Probably, persistent Th1 CD4+ cells, together with decreasing proportions TCD4+ and TCD8+ specific populations in the absence of IL-4 production, are important for infection control and prevention of relapse (175).

Polyfunctional T cells produce several effector cytokines simultaneously and often in higher amounts, particularly pro-inflammatory cytokines IFN-γ, TNF-α, and IL-2. Polyfunctionality broadens the range of effector functions, and has been implicated in the development of vaccines against intracellular pathogens, i.e. targeting cellular-mediated immunity (251–253). Polyfunctionality correlates with T cell efficacy and can be used during vaccine development, for the design of candidates that target such responses and to analyse vaccine-induced responses (254).

The magnitude of response, as in the number of antigen-specific effector cells, is used for humoral responses as a correlate of protection (253). However, the quality of a given immune response, determined by the effector functions (proliferation, cytotoxicity, cytokine and chemokine production) and associated with polyfunctionality, is a better determinant for protection than the magnitude of response (253). Considering this, a high-quality response is one that specifically balances different effector functions to produce an effective and long-lasting protective immune response. Also, the diversity of antigens targeted by specific T cells (breadth of response) can be used as a correlate of protection – in studies with HIV patients, broader antigenic diversity correlates with protective immune responses (255).

Leishmania-specific Th1 cells that produce IL-2 and/or TNF-α are mostly CCR7+ T_{CM} cells; (256). It is possible that polyfunctional cells differentiate into IFN-γ-producing (only) T cells due to chronic antigen exposure (257), so vaccines must induce adequate Th1 responses, neither too low nor too high, to assure both effector functions and memory maintenance. These responses depend of antigen dose, stimulation duration, antigen-presenting cells targeted, and the early cytokine environment induced by the vaccine formulation (253,258).

Second and third generation vaccines typically do not induce antigen persistence, which is probably required to maintain anti-*Leishmania* immunity (157,218), as it occurs in natural infection or whole-organism vaccines. In highly endemic areas, natural boosts may occur, nevertheless, this is an important aspect to consider and include during vaccine design and preclinical development, and, for example, adapt vaccination schedules to include boost immunizations.

Finally, it has been demonstrated that memory responses in vaccinees correlate with the number of precursor cells in the naïve repertoire (259–261). This crucial feature should be exploited during the preclinical development of vaccine candidates through the comparison of immunogenicity profiles induced by the antigens in different immune backgrounds – naïve, healed VL, healed CL, etc. Antigens that correlate both with the presence of memory responses and the existence of naïve precursor cells are strong candidates for a successful vaccine formulation, able to induce memory in

naive individuals (prophylactic vaccine) or boost existing responses to help control infection (therapeutic vaccine).

In summary, the most important markers of immunity against leishmaniasis are: a predominant Th1 response and with the presence of polyfunctional TCD4+ cells (IFN-γ, IL-2, TNF-α), and with low IL-10 (80). This immune environment will ultimately favour the production of effector anti-leishmanial molecules such as reactive nitrogen and oxygen intermediates for parasite control (63,68,141). Also, a relevant aspect for the induction of immunity against *Leishmania* seems to be the need for a balanced ratio between IFN-γ/IL-12 and IL-10/IL-4 (159), which should be preserved to prevent immunopathology and assure memory generation and maintenance.

The lack of well-established correlates of protection hinders vaccine development against leishmaniasis:

- Mouse models are of limited extrapolation to human disease and should not be used to screen potential candidates designed for human vaccine formulations;
- Protective responses against *Leishmania* are characterised by mixed Th1/Th2 responses with predominant IFN-γ and low IL-10;
- Memory T cells (CD45RO+) present in healed individuals are expected to proliferate and acquire effector functions after vaccine antigen stimulation;
- Vaccine candidates should induce Th1 cellular responses, and respective associated cytokines, maintaining the balance high IFN-γ/IL-12 and low IL-10/IL-4;
- Cell surface phenotype should be performed to identify the cell populations involved in the induced responses;
- It is advisable to include a complete panel of induced cytokines (Th1, Th2, Th17 and Treg) to assess polyfunctionality and to better elucidate the immune mechanisms induced by the vaccine candidates.

4. Vaccines against human leishmaniasis

Since the end of leishmanization, several attempts were made to develop other first generation vaccines (killed or live attenuated *Leishmania* parasites), as well as second generation (recombinant proteins or peptide-based vaccines) or third generation vaccines (DNA-based vaccines) (106,262,263).

Several first generation vaccines using crude antigens from *Leishmania* promastigotes, prepared with or without BCG as the adjuvant, were tested in human clinical trials. These candidates were found to be safe but generally ineffective, as is the case of autoclaved *L. major* (ALM) vaccine which showed results comparable to BCG injection alone (169). Highly variable efficacy levels were reported against CL (0-75%) and very low efficacy against VL (169). Three other first generation vaccine candidates have been tested: 'Mayrink vaccine'; 'Convit vaccine' and 'Razi Institute vaccine' with inconclusive or negative results for prophylaxis potential, but somewhat encouraging for therapeutic indications (264). These studies, although unsuccessful in launching a vaccine against leishmaniasis, further established the vaccine feasibility. Moreover, they paved the way for the development of second and third generation vaccines. Second and third generation vaccines need to be formulated with an adjuvant to increase immunogenicity and response specificity.

4.1. Current vaccine pipeline

In the last decades, several antigens have been explored as candidate antigens for both prophylactic and therapeutic second-generation vaccines against leishmaniasis. These antigens will be further characterised in chapter II.

The most advanced vaccine candidates are described below.

The vaccine candidate Leish-111f (also named LEISH-F1 or MML) was the first subunit vaccine to progress to human phase I and II clinical trials, safety and immunogenicity testing in healthy subjects (265). Leish-111f consists of a chimeric recombinant protein containing three antigens in tandem (*Leishmania* elongation initiation factor, LeIF; *L. major* stress-inducible protein, *Lm*STI1; and thiol-specific antioxidant, TSA), combined with monophosphoryl lipid A-stable emulsion

(MPL-SE), and induces a potent Th1-type immune response (265). This formulation was tested in five clinical trials:

i) Phase I trials in 2005, "Study to Evaluate the Leish-111f + MPL-SE Vaccine in Healthy Adults Not Previously Exposed to *Leishmania* Parasite" and "Safety Study to Evaluate the Leish-111f + MPL-SE Vaccine in the Prevention of Cutaneous Leishmaniasis in Healthy Subjects Previously Exposed to the *Leishmania* Parasite" in Colombia (ClinicalTrials.gov Identifiers: NCT00121862 and NCT00121849, respectively);
ii) Phase I trials in 2005-2007, "Study to Evaluate the Leish-111F + MPL-SE Vaccine in the Treatment of Mucosal Leishmaniasis" in Peru, and "Study to Evaluate the Leish-111F + MPL-SE Vaccine in the Treatment of Cutaneous Leishmaniasis" in Brazil (ClinicalTrials.gov Identifiers: NCT00111514 and NCT00111553, respectively);
iii) Phase I trial in 2005-2006, "Safety Study to Evaluate the Leish-111f + MPL-SE Vaccine in the Prevention of Cutaneous Leishmaniasis in Healthy Subjects Previously Exposed to the *Leishmania* Parasite" (ClinicalTrials.gov Identifier: NCT00121849);
iv) Phase I trial in 2007-2015, "Open-Label Safety Study of Three-Antigen *Leishmania* Polyprotein with Adjuvant MPL-SE in Healthy Adults in India" (ClinicalTrials.gov Identifier NCT00486382).

The LEISH-F1+MPL-SE vaccine was partially protective in experimental VL, but ineffective against canine leishmaniasis in phase III field trials (266). In human trials, this formulation was found to be safe and well tolerated, and immunogenic against VL and CL (230). Furthermore, LEISH-F1 showed therapeutic efficacy against MCL, in terms of accelerated time to cure, when used in combination with chemotherapy (267).

Researchers at the Infectious Disease Research Institute (IDRI) redesigned Leish-111f into a new construct Leish-110f (also named LEISH-F2), to better comply with regulatory and manufacture concerns – the 6-Histidine tag near the amino terminus was removed, and the residue Lys274 was replaced by glutamine, so to eliminate an apparent proteolytic site (230,268). The safety, immunogenicity and efficacy of the LEISH-F2+MPL-SE candidate were tested in a phase I clinical trial with promising results, and in a phase II trial to assess its immunotherapeutic potential against CL in comparison with standard chemotherapy with SSG (showing very limited success):

i) Interventional Phase I trial in 2009-2011, "Safety and Immunogenicity of the LEISH-F2 + MPL-SE Vaccine with sodium stibogluconate (SSG) for Patients with PKDL" in Sudan (ClinicalTrials.gov Identifier NCT00982774);

ii) Interventional Phase II trials in 2009-2013, "A Study of the Efficacy and Safety of the LEISH-F2 + MPL-SE Vaccine for Treatment of Cutaneous Leishmaniasis" in Peru (ClinicalTrials.gov Identifier NCT01011309).

A third vaccine candidate was developed by IDRI, LEISH-F3. From previously identified 43 antigen candidates recognised by VL patients' serum from Sudan, researchers applied rational criteria (sequence conservation among species, low homology with human proteins, and proved efficacy in animal models) and selected two proteins - *L. infantum/donovani* nonspecific nucleoside hydrolase (NH) and sterol 24-c-methyltransferase (SMT). These proteins were combined to form the recombinant protein LEISH-F3 and adjuvanted with either MPL-SE, glucopyranosyl lipid A formulated as stable emulsion (GLA-SE), or second generation lipid adjuvant stable emulsion (SLA-SE) (269).

The LEISH-F3 vaccine candidate was tested in three clinical trials to assess safety and immunogenicity of the different formulations (270):

i) Interventional Phase I trial in 2011-2013, "Phase 1 LEISH-F3 Vaccine Trial in Healthy Adult Volunteers" in USA (ClinicalTrials.gov Identifier NCT01484548);
ii) Interventional Phase I trial in 2012-2016, "LEISH-F3 + GLA-SE and the LEISH-F3 + MPL-SE Vaccine" in USA (ClinicalTrials.gov Identifier NCT01751048);
iii) Interventional Phase I trial in 2014-2016, "Phase 1 LEISH-F3 + SLA-SE Vaccine Trial in Healthy Adult Volunteers" in USA (ClinicalTrials.gov Identifier NCT02071758).

Later, the LEISH-F3 candidate was slightly modified with the addition of the cysteine protease B (CPB) to increase antigen recognition and extend the geographical affectation of the vaccine (271). This new candidate was formulated with GLA-SE and is called LEISH-F3+. This vaccine has evaluated in experimental infection models (hamsters) and has shown to provide robust immunity, similar to LEISH-F3 (271). The LEISH-F3+GLA-SE vaccine candidate was also proven to be safe, and induced strong and specific immune responses, measured by cytokine production and immunoglobulin subclass information. This candidate should, therefore, be tested in additional clinical trials in endemic countries, in populations susceptible to *Leishmania* infection (269).

Another promising candidate in the pipeline is the ChAd63-KH vaccine candidate, also named Leish2a, developed by researchers at York University. It consists of a chimpanzee adenovirus-based

vaccine (ChAd63) carrying a synthetic HASPB gene linked to a KMP11 gene with a viral 2A sequence (272). This candidate progressed to a Phase I clinical trial where the safety profile was confirmed, and strong TCD8+ cellular responses were observed (273). The Leish2a vaccine was also assessed in a Phase II clinical trial in Sudan to test safety and induced immune responses in individuals with persistent PKDL infection (ClinicalTrials.gov Identifier: NCT02894008, "A Study of a New *Leishmania* Vaccine Candidate ChAd63-KH (Leish2a)").

In earlier stages of development, there is also a new candidate developed by the European FP7 network MuLeVaClin, which is based on defined *Leishmania* antigens KMP11 and LeishF3+, and sand fly saliva antigen LJL143, formulated into Influenza virus-like particles (VLP) and adjuvanted with GLA-SE, a TLR-4 agonist (274). Pre-clinical studies in the mouse model show this candidate is safe and immunogenic, particularly against the vector-derived antigen (274).

Furthermore, other polypeptidic/chimeric vaccines against leishmaniasis are in pre-clinical stages of development, and were validated in human samples.

The protein Nucleoside Hydrolase NH36 is the main component of the Leishmune® vaccine, once commercialised in Brazil. The immunogenicity of three NH36 polypeptides (N-terminal, central, C-terminal domains) was tested pre-clinically with samples from cured CL and VL patients, from areas endemic to *L. infantum* in Spain (275). Recombinant NH36 and the N-terminal domain (F1) induced lymphoproliferation in samples from cured and asymptomatic individuals, accompanied by increased IFN-γ, TNF-α and granzymeB production. Also, F1 induced IL-17 production cured CL patients and asymptomatic patients. Additional studies were performed with samples from healed and asymptomatic individuals, from areas endemic to *L. braziliensis* in Brazil (276). Here, the F1 domain induced Th1 and Th17 responses in cured/exposed patients infected with *L. (L.) infantum* (*chagasi).*

The polyprotein KSAC includes fragments of the proteins KMP11, SMT (24-c-methyltransferase), A2 (A2 amastigote-specific protein), and CPB (cysteine protease B) (277). KSAC formulated with monophosphoryl lipid A (MPL-SE) induced antigen-specific multifunctional Th1 cells and conferred protection against *L. infantum* and *L. major* challenge in mice. Post-challenge responses in the KSAC/MPL-SE-vaccinated mice show higher IFN-γ/IL-4 ratios and decreased IgG1 responses to SLA (268,277).

A recombinant chimeric protein composed by T cell epitopes specific to human and mice MHC alleles and derived from 4 *Leishmania infantum* proteins (LiHyp1, LiHyp6, LiHyV and HRF proteins)

was formulated with saponin and tested in BALB/c mice. Vaccinated mice and challenged with *L. amazonensis* were protected against infection. The candidate induced a specific IFN-γ, IL-12 and GM-CSF production, lower IL-4 and IL-10 levels, as well as high IgG2a isotype antibody levels. In addition, antileishmanial nitrite production by splenocytes was detected. Interestingly, the recombinant protein administered alone induced a partial protection against challenge (278).

4.2. *Leishmania* Excreted-Secreted Antigens as promising vaccine candidates and the successful canine vaccine CaniLeish®

Leishmania excreted-secreted proteins (ESP or ESA) are important virulence factors, and are implicated in early host-parasite interactions, establishment of infection and in the regulation of host immune responses (279–281). Their role in the modulation and evasion of the host immune responses is reviewed in Holzmuller et al 2018 (2).

In the last decades, particularly after the completion of genome and proteome annotation of several pathogenic *Leishmania* species, these excreted-secreted proteins were further characterised. Additionally, ESP have been shown to be processed and presented by antigen-presenting cells (127), as well as, to trigger protective cellular immune responses (282–285). Gour et al described the fractions F1 (11, 13 and 16 KDa) and F3 (26, 29 and 33 KDa) from *Leishmania* ESA, which contained Th1-inducing proteins (283). Interestingly, several of the antigens used in the most advanced vaccine candidates are found in the *Leishmania* secretome, such as KMP11, nucleoside hydrolase, thiol-specific antioxidant (TSA) or translation initiation factors (detailed in chapter III – Vaccine Antigen Selection).

The Lemesre lab (UMR177, IRD, France) developed a method for the production of naturally excreted-secreted antigens from culture supernatants of *L. infantum* promastigotes, and evaluated their potential use in a vaccine formulation. Notably, *Li*ESAp were shown to provide significant protection in dogs by inducing a strong and long-lasting Th1 response, leading to the development of the european canine vaccine Canileish® (284,286). The adaptation and scaling-up of this vaccine for human use is impossible due to elevated costs, and because the active principle remains undefined, although likely to contain a plethora of different protein antigens, which may induce non-specific responses and/or response variability. Also, this vaccine is formulated with QA-21 (purified extract of *Quillaja saponaria*) as adjuvant, which is currently not approved for human use. However, QA-21 is derived from the QS-21 adjuvant which has been approved for human use by the FDA.

The canine vaccine is well-tolerated and safety studies did not show any significant adverse effects (286,287). However, a study in Spain has revealed a high rate of adverse yet mild reactions in vaccine users included in the survey (82%), the most frequent being local reactions, apathy, fever and gastroenteritis (288). Still, one often overlooked aspect is that each vaccine dose contains a fixed amount of QA-21 adjuvant, which is not adapted to the dog's weight. This highlights the need for optimal selection and dosage of the adjuvant used human vaccine formulations. An epitope-based approach and new Th1-inducing adjuvants will greatly mitigate these concerns.

Some attempts were made to identify the protective antigens contained in total ESP preparations. From *Leishmania* excreted/secreted proteins, research developed at the Lemesre Lab showed that soluble Promastigote Surface Antigens (PSA) is an immunodominant component of *L. amazonensis* and *L. infantum* secretome, and highly conserved among *Leishmania* species (289). Moreover, the vaccinated dogs with recombinant PSA (rPSA) of *L. amazonensis* or its carboxy-terminal part, both combined with QA-21 as adjuvant, were protected at 78.8% and 80%, respectively (290). This cross-protection was associated with hallmarks of a dominant Th1-type immune response. In *L. infantum* and *L. major*-protected humans (healed individuals), it was also clearly demonstrated that the rPSA induced a dominant Th1 response associated to cytotoxicity *in vitro*. During the FP7-funded RAPSODI project (2009-2012), the Lemesre Lab and collaborators developed a vaccine candidate containing the Promastigote Surface Antigen (PSA) from *L. amazonensis*, whose immunogenicity was validated on human cells (282). However, the very low productivity of recombinant *La*PSA would result in an expensive vaccine, which is unfeasible for human mass vaccination. As peptide production is far less expensive, the adopted approach was to select only peptides derived from proteins of interest for vaccine development.

4.3. Peptide-based vaccines

The final formulation is of extreme importance for the development of an effective vaccine to be deployed in low-resource settings. The ideal target product profile (TPP) for a vaccine against leishmaniasis englobes:

i) good safety profile;
ii) a minimum number of immunizations and boosts needed;
iii) the absence of animal products and impurities;
iv) cost-effectiveness;
v) a prophylactic and therapeutic effectiveness;
vi) optimal delivery;
vii) preferably, the lack of cold-chain supply (169,270).

A peptide-based vaccine formulation would answer many of these requirements (291). Peptide-based vaccines explore the minimal immunogenic units necessary to induce protective immune responses (epitopes) and have become an increasingly important tendency in vaccine development (291,292). Peptide-based approaches allow for a highly specific stimulation of host immune responses, while decreasing the risk of unwanted cross-reactions and other adverse reactions (292).

Several epitope discovery strategies have been exploited in peptide-based vaccine design (293,294), particularly:

i) *in silico* epitope mapping through HLA-binding epitope prediction, selection of conserved antigens by phylogenetic analysis, protein localization prediction or through reverse vaccinology approaches;
ii) phage display assays;
iii) immunodominance assays and peptide competition assays.

An epitope-based approach, in contrast to subunit recombinant protein or inactivated live vaccines, offers several advantages, namely: i) easiness of production, ii) superior stability and no need for cold-chain supply; iii) higher purity and absence of potentially harmful substances; iv) lower antigen complexity; v) low scaling-up costs; vi) high response specificity; and vii) ability to combine epitopes to design multi-epitopic and/or multi-specific vaccines (292,295).

From the CaniLeish® and recombinant PSA development studies, our lab has identified the main immunogenic portion from carboxy-terminal region of rPSA, representing 19.4% of the rPSA sequence. Vaccinated dogs with three peptides derived from this region, combined with QA-21 as adjuvant, were protected at 60% (Petitdidier E. et al. submitted). This protection level is significant given the small size of the immunogenic peptides used in this vaccine formulation, in comparison to total rPSA (78.8% of protection) or PSA carboxy-terminal region (80% of protection). This study clearly demonstrates peptide-based vaccines are a promising approach for the design of vaccines against leishmaniases.

In human leishmaniases, we need to consider the diversity of *Leishmania* species and HLA diversity due to population variability from numerous endemic areas.

Altogether, peptide-based vaccines are a promising approach for the design of vaccines against leishmaniases, through the ability to include multiple immunogenic and conserved epitopes from various protein antigens, increasing the chances of immunogenicity and both parasite strain and target population coverage.

However, short peptides are known poor immunogens when administered alone. The main associated challenges are the avoidance of inactivation or degradation by the host immune system, and the enhancement of peptide immunogenicity. This low immunogenicity can be overcome with the addition of adjuvants in the final vaccine formulation that direct and boost the induced responses. There is the need to improve adjuvants for the final vaccine formulations, adjuvants that are more efficacious, less toxic, and approved for human use (226). Some new promising adjuvants are being developed, particularly Th1-inducing adjuvants and TLR agonists, which can be used in anti-*Leishmania* vaccine formulations (Table I.3.)

Table I. 3 Toll-like receptor agonists and Th1-inducing adjuvants for use in vaccine formulations, adapted from (239)***.***

Adjuvant	Class	Mechanism of action	Type of immune response
Lipid A analogues (e.g., MPL, GLA)	Immunomodulatory molecule	Toll-like receptor 4	Antibody, Th1
Imidazoquinolines (e.g.,Imiquimod, R848)	Immunomodulatory molecule	Toll-like receptor 7 / 8	Antibody, Th1
CpG ODN	Immunomodulatory molecule	Toll-like receptor 9	Antibody, Th1, Th2, TCD8+ cells
Saponins (e.g., QS21)	Immunomodulatory molecule	Unknown	Antibody and cell-mediated immune responses
Virosomes	Particulate formulation	Antigen delivery	Antibody, Th1, Th2
GLA-SE	Combination	Toll-like receptor 4	Antibody, Th1

In summary, the most important bottlenecks in peptide-based vaccine design are the epitope choice, the adjuvant selection and availability; and vaccine administration. Despite vaccines targeting cellular immunity are not yet available, recent advancements human Th1 inducing adjuvants make this objective are a closer reality.

Moreover, synthetic peptides present great versatility for adaptation to innovative delivery systems (292,295). The exploitation of innovative approaches should also be considered for anti-*Leishmania* vaccine development, such novel devices for optimal intradermal delivery (i.e. dermal patches, nano-delivery systems, etc) that ensure proper immune response priming throughout time (296). Also, chimeric proteins combining peptides, polyproteins and delivery vectors are promising formulations.

4.4. *Leishmania*-specific peptides tested in human cells

Potential peptides are tested for immunogenicity in mice or human PBMC. Most immunogenicity testing of candidate proteins and peptides in leishmaniasis vaccine development have been performed in mouse models. Although these models provided important insights into *Leishmania* pathogenesis and host immunity, they are inconclusive in terms of vaccine antigen design. Many vaccine candidates under development fail in early stages of clinical development probably due to weak antigenicity and lack of predictive animal models. Promising candidates against leishmaniasis developed with animal studies often were unsuccessful in human trials (297,298). The failure of protective vaccines in human trials may be because of differences in peptide processing (TAP binding) and presentation between murine H-2 and human HLA systems. In which case, humanized animal models may provide a useful tool for vaccine preclinical development.

Still, some peptide vaccine candidates against *Leishmania* were tested in human PBMC. These studies used mostly samples from LST+ or healed patients, and lymphoproliferation and cytokine production by ELISA as readout. Briefly, studies testing *Leishmania*-specific peptide immunogenicity in human immune cells are:

i) Russo 1993 (299): peptides from *L.major* GP63 (13 peptides with 14 a.a.) (300), three of which increased PBMC proliferation, and IFN-γ production. Epitope prediction was performed with an algorithm developed by Rothbard and Taylor (301) based on sequence motifs based on known human, mouse and guinea pig epitopes.

ii) Delgado G 2003 (302): peptides from *L. panamensis* KMP11 (6 peptides with 20 a.a. spanning the whole protein sequence) induced lymphoproliferation and IFN-γ production by T-cells. Authors assess peptide binding to HLA molecules with binding assays performed with four purified HLA-DR molecules (DRB1* 0101, 0401, 0701 and 1101). Donors were HLA-typed (PCR to amplify the exon 2 of the DRB gene) and the correlation between *in vitro* immune response and peptide binding assay predictions was ascertained.

iii) Basu R 2007 (303): peptides from *L. donovani* KMP11 (84 octamers spanning the whole protein sequence), some of which induced IFN-γ, IL-10 and IL-17 production by $TCD8^+$ cells and increased splenocyte proliferation. Authors propose potential HLA restriction according to the experimental data and donor HLA typing (HLA-A, -B and –C alleles).

iv) Seyed N 2011 (304): peptide pools from *L. major* antigens (18 nonamers from CPB, CPC, TSA, LeIF, LmSTI, LPG) increased IFN-γ production by $TCD8^+$ cells. The six protein antigens were examined for epitopes restricted to the HLA-A*0201 allele using *in silico* HLA-binding prediction algorithms (see Chapter IV). Authors used a 2-step epitope selection pipeline using 2 prediction algorithms (SYFPEITHI and BIMAS) and 5 five other algorithms (EpiJen, Rankpep, nHLApred, NetCTL and Multipred) which further filtered the peptide list and, finally, analysed with NetMHCpan1.1 to check for the possibility of binding to different alleles of HLA-A2 supertype (HLA- A*0202-A*0206 and A*0209). BLAST was then used to reject peptides that were 100% identical with mice and human proteins.

v) Elfaki ME 2012 (305): peptides from *L. donovani* GP63 (4 peptides with 15-21 a.a.), two of which increased IL-10 production, while another peptide and peptide pools decreased IL-10 production, with no significant IL-4 production. Epitope prediction was performed with EpiMatrix (www.immunome.org/iVAX/) for alleles DRB1*1101 and DRB1*0804 and a panel of 8 HLA-DR "supertype" alleles (> 90% of human populations worldwide). Peptide hydrophobicity analysis and BLAST against the human genome were performed.

vi) Naouar I 2016 (306): peptides from *L. major* antigens (78 nonamers from 33 proteins), six of which induced cytotoxicity, assessed by granzyme B production, but low IFN-γ and IL-10 levels were found in PBMC culture supernatants. Epitope prediction was performed for the HLA-A*0201 allele and proteasomal cleavage prediction (RANKPEP algorithm, imed.med.ucm.es/Tools/rankpep.HTML). Affinity to HLA-0201 alleles was assessed with stabilisation assays.

vii) Freitas e Silva R 2016 (307): ten peptides (15-mer) conserved in *L. braziliensis*, *L. major* and *L. infantum* discovered with a reverse vaccinology approach using proteomic datasets, five of which induced PBMC proliferation. Linear epitope prediction was performed for 9 HLA-class I supertypes (HLA-A1, -A2, -A3, -A24, -A26, -B7, -B27, -B44 and -B58) and for 4 HLA-class II supertypes (HLA-DPA, HLA-DPB, HLA-DQA, HLA-DRB). The algorithms NETMHC and NETCTL were used to predict HLA-class I-binding epitopes, and NETMHC Class II was used for a MHC II prediction. Additionally, the linear peptides were structurally modelled to MHC molecules to find the most stable *in silico* epitope + allele complexes. Structures from 33 different alleles of MHC I (21) and MHC II (12) were used.

viii) Mahantesh V 2017 (308): peptides from *L. donovani* 3'-nucleotidase (5 nonamers) induced IFN-γ and IL-2 production, increased T-cell proliferation in PBMC culture and CTL activity. Epitope prediction was performed for HLA-A*02 and HLA- B40 alleles with six HLA-binding prediction algorithms (SYFPEITHI, BIMAS, RankPep, NetCTL1.2, IEDB tools.iedb.org, and ProPredI). In the last epitope selection step, the selected peptides were analysed by NetMHCpan3.4 to check the binding affinity to other HLA class I alleles, and promiscuous epitopes that bind to at least 10 HLA-A*02 or 5 HLA-B40 were selected.

These studies are further detailed in Table I.4.

Table I. 4* Leishmania-*specific peptide vaccine candidates validated using human samples. 'Ref', bibliographical references.

Species	Peptides	Peptide design/Epitope prediction	Human samples	Assays performed	Peptide dose	Results	Ref
L. major	**13 peptides with 14 a.a.**	- Peptides used in this publication were described in a previous publication (Jardim A et al 1990); - Epitope prediction with an algorithm (Rothbard and Taylor 1988) which finds sequence motifs based on known human, mouse and guinea pig epitopes	- Whole blood samples from: patients with active or recently treated CL (*L. amazonensis*), active MCL (*L. braziliensis*) or cured VL patients (*L. chagasi*) at 3-6 months after standard chemotherapy. - Inclusion criteria: parasitologic and serologic evaluation; - Active CL patients presented active lesions, positive DTH, and *in vitro* proliferative responses to promastigote lysate.	- PBMC stimulation assays (5 days) - Lymphoproliferation assay (tritiated thymidine incorporation for the final 18 hours); - Cytokine ELISA to detect IFN-γ in culture supernatants; - Peptide-specific T cell line generation from PBMC from healthy donors and patients; - Cell surface phenotyping (flow cytometry).	- Peptides were tested at concentrations ranging from 10 to 100 µg/ml; - Optimal proliferative responses were obtained with 50 µg/ml: PT7,50 µg/ml; or PT4 50 µg/ml - native GP63 5 µg/ml.	- 3 peptides increased PBMC proliferation and IFN-γ production by T cells from leishmaniasis patients; - Peptide PT7 stimulated cells from all individuals tested (n=7); - Anti-PT7 T cell lines responded by proliferation and IFN-γ production to *in vitro* stimulation with *Leishmania* promastigote lysate; - Peptide PT1 stimulated PBMC from an *L. chagasi* patient although the corresponding *L. major*-derived peptide did not; - Both *L. major* PT7 and *L. chagasi* PT1 induced specific T cell lines from from healthy donors' PBMC.	Russo DM 1993
L. panamensis - L(V)p	**6 peptides with 20 a.a.**	- Six peptides spanning the whole KMP11 sequence; - Peptide binding assays performed with four purified HLA-DR molecules (DRB1* 0101, 0401, 0701 and 1101); - Donors were HLA-typed (PCR to amplify the DRB gene's exon 2) - Correlation between *in vitro* immune response and peptide binding assay predictions was evaluated	- Whole blood samples from: - individuals with ACTIVE disease (A1-4), with active CL lesion, untreated; - CURED individuals (P1-9), with healed skin lesions by L(V)p, clinical history of CL and without relapse 3 months after treatment; - Inclusion criteria: (i) individuals living in high L(V)p endemic areas; (ii) volunteers with positive LST. - CONTROLS (C1-5) from non-endemic L(V)p areas.	- Detection of rKMP11- and LSA-specific antibodies in the plasma (ELISA). - T-cell stimulation assays (T-cells co-cultured with pulsed mature DCs for 48 h); - Cytokine ELISA to detect IL-4 and IFN-γ in the culture supernatants; - Lymphoproliferation assay (tritiated thymidine added to the cultures for the last 16 hours); - Peptide competition assays (competitor peptide inhibiting biotinylated indicator peptide binding by 50% or more was indicative of efficient binding by a peptide to that particular MHC-class II allele).	- T-cell stimulation assays: immature DCs were treated with 2 µg/mL Poly I:C (DC maturation) and pulsed with 10 µg/mL of each antigen (synthetic peptides, L(V)p or rKMP-11).	- DCs efficiently stimulate proliferation of short-term, rKMP-11 specific T-cell lines; - DCs used as APC favoured identifying T-cell epitopes from KMP-11 in ACTIVE and CURED volunteers; - 2 peptides exhibited high binding capacity with HLA DRB1*0401 allele; - ALL peptides induced lymphoproliferation and IFN-γ production, in at least 2 (up to 6) CURED individuals (n=9); - IFN-γ production was more frequent in T-cells from CURED individuals; - IL-4 production was characteristic of T-cells from ACTIVE disease patients.	Delgado G 2003
L. donovani	**84 nonamer peptides**	- 9-mer peptide library spanning the whole KMP11 sequence - Donors were HLA typed (HLA-A, -B and –C alleles); - Inhibition assays with anti-HLA-class I antibody - Authors propose potential peptide HLA restriction according to the experimental data and based on the sequence motifs in the SYFPEITHI database.	- Total PBMC cells harvested from whole blood samples from 10 healthy donors.	- T cell priming cultures (CD8+ T cells isolated from the PBMCs by magnetic sorting. Remaining cells were irradiated and used as APCs. T CD8+ cells were tested in ELISpot assays for the frequencies of peptide-specific T cells); - Assays with infected macrophages (T cell lines specific to 4 peptides were added to iMφ, and the induction of IFN-γ was measured in ELISpot). - Responses to iMφ were analyzed without and with anti-HLA class I MAbs to test the natural presentation and the HLA class I restriction of the responses.	- Pools of every tenth peptide along the protein sequence were prepared at final concentration of 400 µg/mL total peptide or of 44 µg/mL for each individual peptide. - ELISpot assays and the monospecific T cell lines: peptides prepared individually in stock solutions of 20 µg/mL.	- 30 peptides induced T cell responses (IFN-γ production); - These 30 peptides are promiscuous (no clear correlation with HLA specificity); - KMP-11 is processed and presented by MHC-I molecules of infected cells; - 4 (in 5) peptide-specific T CD8+ cell lines responded to the infected macrophages with IFN-γ secretion	Basu R 2007

Table I.4 (continued)

Species	Peptides	Peptide design/Epitope prediction	Human samples	Assays performed	Peptide dose	Results	Ref
L. major	**18 nonamer peptides from CPB, CPC, TSA, LeIF, LmSTI, LPG**	- The six protein antigens were examined for epitopes restricted to the HLA-A*0201 allele using in silico HLA binding prediction algorithms (see Chapter IV). - Epitope selection pipeline using 2 prediction algorithms (SYFPEITHI and BIMAS) >> 5 five other algorithms (EpiJen, Rankpep, nHLApred, NetCTL and Multipred) >> NetMHCpan1.1 to check for binding to the HLA-A2 supertype (HLA- A*0202-A*0206 and A*0209). - BLAST was used to reject peptides 100% identical with mice and human proteins	- Whole blood samples from: - HLA-A2+ individuals recovered from L. major infection; - Controls: HLA-A2- individuals recovered from L. major and HLA-A2+ healthy donors.	- PBMC stimulation assays (10 days); - Intracellular cytokine assay (after 10 day culture cells were re-stimulated with peptide pools and anti-CD49d/anti-CD28 antibodies, then stained and analyzed by flow cytometry); - Cytokine ELISA to detect IFN-γ in the culture supernatants (before re-stimulation at day 10); - IFN-γ/IL-4 ELISpot assay after 10-day stimulation with peptide pools (positive responses= spots>mean+2SD of control)	Four peptide pools: - Pool I (3 peptides from CPB, 2 from CPC); - Pool II (4 peptides from LmSTI); - Pool III (2 peptides from TSA, 3 from LeIF); - Pool IV (4 peptides from LPG). - cells were stimulated with peptide pools at 10 µg/ml/peptide and also freeze/thawed antigens of L. major (10 mg/ml) as indicator of previous disease	- Specific response to pools II and IV (LmsTI-1 and LPG-3-related peptides) is specifically presented in HLA-A*0201 context; - Specific memory CD8+ T cells are detectable in PBMC of CL-recovered individuals by flow cytometry and ELISpot; - Secreted IFN-γ is produced by PBMCs from HLA-A2+ recovered individuals stimulated with peptide pools II and IV; - HLA-A2+ CL recovered individuals developed specific response against peptides from LmsTI-1 and LPG-3; - Peptides are promiscuous as positive responses were detected in HLA-A2-recovered individuals against other two protein groups (CPB, CPC, LeIF and TSA).	SeyedN 2011
L. donovani	**4 peptides with 15-21 a.a.**	- 9-mer peptide library spanning the whole GP63 sequence >> 4 peptides selected - Epitope prediction was performed with EpiMatrix (www.immunome.org/iVAX) - Predicted binding to alleles DRB1*1101 and DRB1*0804, and a panel of 8 HLA-DR "supertype" alleles (>90% of human populations) worldwide. - Hydrophobicity analysis - BLAST to confirm no significant homology to the human genome.	- Whole blood samples from healthy LST+ individuals (with or without a history of VL, and leishmanin skin induration of ≥ 8 mm)	- Whole blood stimulation assays (16-24h) - Cytokine ELISA to detect IFN-γ IL-4 and IL-10 in the culture supernatants	- 100 µg from each of the 4 predicted peptides, or - 100 µg of a pool composed of equal proportions of each peptide	- No clear correlation with Th1 responses and different responses observed after peptide or pool stimulation; - Pooled peptides produced a moderate IFN-γ increase in some volunteers; - Peptides P1 and P3 have a strong IL-10 upregulatory role; - Peptide P4 and peptide pools decreased IL-10 production; - No significant IL-4 production either with peptides or pool stimulation; - Mean IL-10 levels were significantly reduced for all individuals compared with controls after pool stimulation.	Elfaki ME 2012
L. major	**78 nonamer peptides**	- 33 *L. major* genes previously identified as potentially ESP were analysed; - Epitope prediction for the HLA-A*0201 allele and proteasomal cleavage prediction with RANKPEP; - Affinity to HLA-A*0201 alleles was assessed with stabilisation assays.	- Whole blood samples from: - 6 HLA-A*0201+ and 6 HLA-A*0201- donors recovered from ZCL, living in a *L. major* high transmission area (Tunisia); - Inclusion criteria: (i) presence of ZCL scars, (ii) positive LST, and/or (iii) positive lymphoproliferative response to SLA; - Controls: HLA-A*0201+ healthy individuals living outside endemic areas without any lymphoproliferative response to SLA.	- Peptide binding activity by peptide-HLA stabilization assay with TAP-deficient HLA-A*02+ T2 cells); - PBMC stimulation assays (5 days): induction of GrB and IFN-γ by stimulated PBMCs from healed ZCL individuals; - Cytokine ELISA to detect Granzyme B, IFN-γ, and IL-10 in the culture supernatants.	- 20 peptide pools were prepared (1 to 7 peptides each); - Final peptide concentration 1µg/mL; - In some experiments, peptides were added separetely at a concentration of 20µg/mL.	- 6 peptides (E2, E6, F6, G2, G3, and G4) among the 24 tested induced the highest levels of granzyme B in PBMCs from immune individuals; - low IFN-γ levels (MAX: 45 pg/mL) and low IL-10 (20-120 pg/mL) were detected in culture supernatants of PBMC stimulated with peptide pools, while SLA and PHA induced high IFN-γ responses; - Among these peptides, 3 showed highest affinity to HLA-A*0201; - The 6 immunogenic peptides derived from the sequence of the Pr37, Pr38, Pr78 (ucharacterised), and Pr74 (IF-1alpha) proteins.	Naouar I 2016

Table I.4 (continued)

Species	Peptides	Peptide design/Epitope prediction	Human samples	Assays performed	Peptide dose	Results	Ref
L. braziliensis + *L. major* + *L. infantum*	**10 peptides with 15 a.a.**	- The full proteomes (*L. braziliensis*, *L. major*, *L. infantum*) were analysed >> 10 peptides selected - Linear epitope prediction was performed for 9 HLA-class I supertypes with NetMHC and NetCTL (HLA-A1, -A2, -A3, -A24, -A26,-B7, -B27, -B44 and -B58), and for 4 HLA-class II supertypes with NetMHC Class II (HLA-DPA, HLA-DPB, HLA-DQA, HLA-DRB); - Linear peptides were structurally modelled to MHC molecules (21 MHC-class I and 12 MHC-class II alleles)	- PBMC derived from patients cured after treatment (n=10). - Inclusion criteria: PBMC proliferation when stimulated with total antigen from *L. braziliensis*, and non-stimulated cells had minimal levels of proliferation.	- The following bioinformatics analysis was performed: BLASTp (proteins with more than 60% conservancy) >> epitope prediction (NetMHC and NetCTL with processing; NetMHC-II) >> BLASTp (similarity ≥ 40% with human or mice proteins were excluded) >> subcellular localization predictions (TMHMM, WoLF PSORT) >> clustering (MySQL database) >> molecular docking (PDB files for 33 alleles +Pymol +GriDoMol +Rosetta's FixBB, Relax and FlexPepDock protocols) >> top 10 ranked 15-mer peptides were synthesized - Validation of peptide epitopes by PBMC proliferation assay (96 hours)	- CFSE-stained PBMC stimulated with 20 µg/mL of each peptide	- 5 peptides induced higher PBMC proliferation in samples from treated patients (peptides 1, 3, 5, 6, and 7); - These five peptides are derived from 3 out of 4 identified *L. braziliensis* proteins (hypothetical proteins), meaning positive responses are not associated with one specific protein.	Freitas e Silva R 2016
L. donovani	**5 nonamer peptides**	- The full sequence of L. donovani 3'-nucleotidase was analysed >> 5 peptides selected - Epitope prediction was performed for HLA-A*02 and HLA- B40 alleles with six HLA-binding prediction algorithms (SYFPEITHI, BIMAS, RankPep, NetCTL1.2, IEDB tools.iedb.org, and ProPredI); - Selected peptides were analysed by NetMHCpan3.4 to check the binding affinity to other HLA class I alleles >> promiscuous epitopes selected (predicted binding to minimum 10 HLA-A*02 or 5 HLA-B40)	- Whole bood samples from: - Cured VL patients: 25 successfully treated VL participants (16 HLA-A*02+ and 9 HLA- A*02-) without hepathosplenomegaly and fever; - Controls: 15 healthy donors with no apparent history of VL and tested negative in the rK39 serological test.	- Peptide binding activity by peptide-HLA stabilization assay with TAP-deficient HLA-A*02+ T2 cells). - PBMC stimulation assays (48 hours); - Cytokine ELISA to detect secretory IFN-γ, IL-10, and IL-12 (p70), and IL-2; - Intracellular IFN-γ Staining (6-8 hours) - T Cell Proliferation Assay (CFSE-stained lymphocytes co-cultured with adherent macrophages and stimulated with 10 µg/ml synthetic peptide for 96 hrs). - Cytotoxic activity (autologous T lymphocytes co-cultured with peptide-pulsed CFSE-stained macrophages for more 4-6 hours).	- PBMC stimulation: 10µg/ml peptide and pool - Intracellular cytokine production: whole blood stimulation with 10 µg/ml peptide and pool; - T cell proliferation assay: 10 µg/ml synthetic peptide; - Cytotoxicity assays: CFSE-stained macrophages incubated with 10 µg/ml synthetic peptide	- All peptides were found to significantly stimulate the proliferation of T lymphocytes from cured VL patients; - T lymphocytes from all cured VL subjects exhibited CTL activity against the peptide-pulsed macrophages; - Peptide stimulation induces a significantly higher IFN-γ, IL-12 and IL-2 production in PBMCs from cured VL subjects; - No differences in IL-10 production in response to peptide stimulation were observed; - Peptide stimulation significantly elevated the intracellular IFN-γ level in cured VL subjects (particularly, epitopes P1 and P3); - Results imply the presence of circulating epitope-specific memory T lymphocytes in cured VL subjects.	Mahantesh V 2017

5. Project objectives and approach

The present doctoral project was part of the Marie Sklodowska-Curie Innovative Training Network EUROLEISH.net (project 2.1 - Development of a multiepitope peptide-based vaccine against human leishmaniasis). Project 2.1 is integrated within the 15 EUROLEISH.net projects as a basic sciences project focused on leishmaniasis prevention.

The project's main objective is the development of a human-compatible second-generation peptide-based vaccine against leishmaniasis, with global coverage, and based on immunogenic epitopes conserved among several pathogenic *Leishmania* species.

The major aims are:

- To identify the antigenic proteins present in the ***Leishmania* secretome;**
- To design and elaborate species-conserved peptides with multiple and appropriate **immunogenic epitopes** to be used as vaccine antigens;
- To perform preclinical phase studies in human cells to evaluate the predictive peptide's **immunogenic** and **immunoprotective** properties;
- to evaluate the **immunoprotective profiles of exposed individuals** who have developed immunity to *Leishmania* infection.

Considering the limitations of previous candidates and restrictive budgets associated with NTD research, the proposed strategy aims combine several data and methods to increase cost-efficiency and shorten development time to maximise late-stage success of a vaccine against human leishmaniases (Figure I.19).

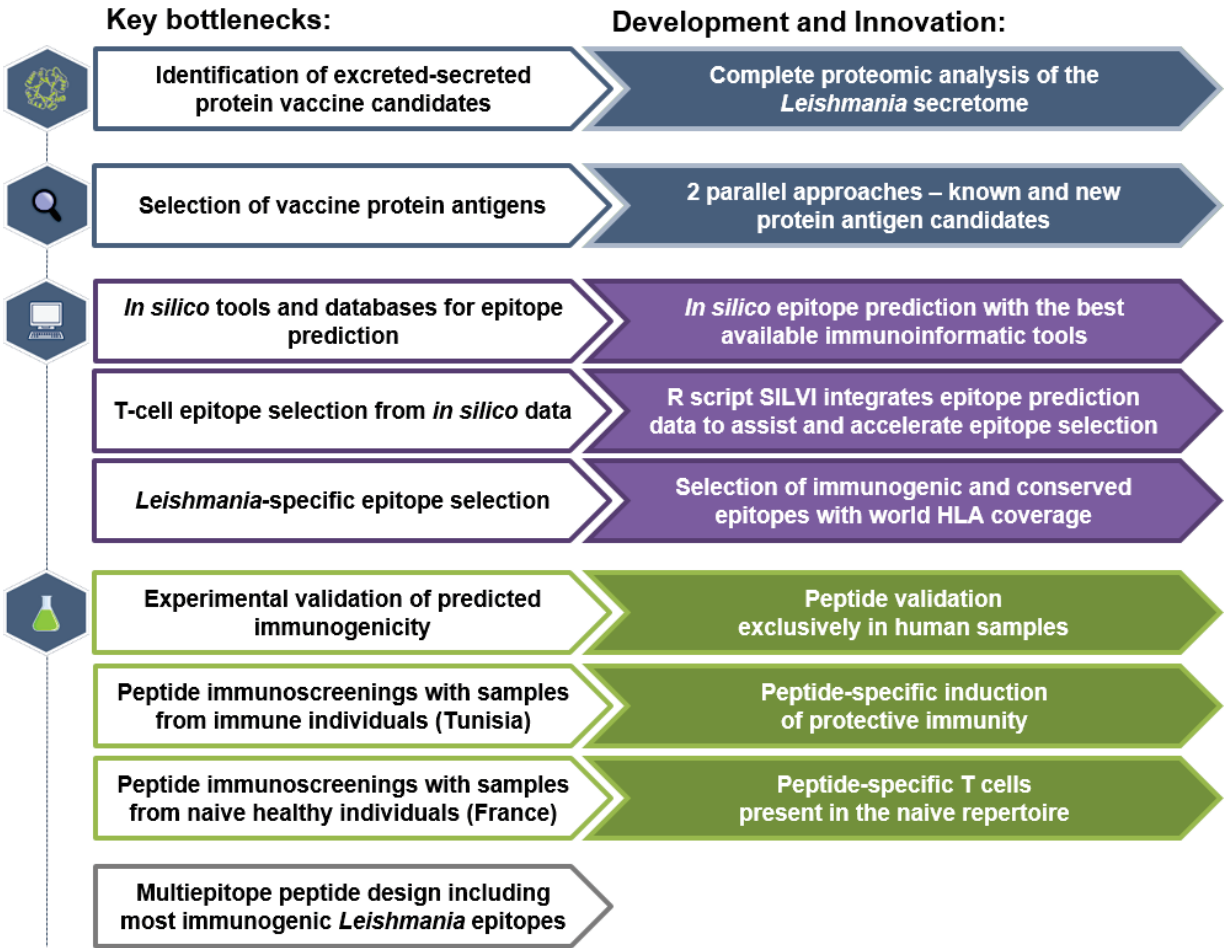


Figure I. 19 Overview of the project's work plan, the key challenges or bottlenecks and how these will be addressed.

6. Bibliography

1. Geiger A, Bossard G, Sereno D, Pissarra J, Lemesre J-L, Vincendeau P, et al. Escaping Deleterious Immune Response in Their Hosts: Lessons from Trypanosomatids. Front Immunol. 2016;7(May):1–21.
2. Holzmuller P, Geiger A, Nzoumbou-Boko R, Pissarra J, Hamrouni S, Rodrigues V, et al. Trypanosomatid Infections: How Do Parasites and Their Excreted–Secreted Factors Modulate the Inducible Metabolism of l-Arginine in Macrophages? Front Immunol. Frontiers; 2018 Apr 20;9:778.
3. Oral Presentation Sessions. Trop Med Int Heal. 2017 Oct;22:2–114.
4. Hailu T. Challenges in visceral leishmaniasis control and elimination in the developing countries: A review. 2016;
5. Okwor I, Uzonna J. Social and Economic Burden of Human Leishmaniasis. Am J Trop Med Hyg. The American Society of Tropical Medicine and Hygiene; 2016 Mar;94(3):489–93.
6. Savoia D. Recent updates and perspectives on leishmaniasis. J Infect Dev Ctries. 2015 Jul 4;9(06):588–96.
7. Al-Salem WS, Pigott DM, Subramaniam K, Haines LR, Kelly-Hope L, Molyneux DH. Cutaneous leishmaniasis and conflict in Syria. Emerg Infect Dis. 2016;22.
8. Du R, Hotez PJ, Al-Salem WS, Acosta-Serrano A. Old World cutaneous leishmaniasis and refugee crisis in the Middle East and North Africa. PLoS Negl Trop Dis. 2016;10.
9. Al-Salem W, Herricks JR, Hotez PJ. A review of visceral leishmaniasis during the conflict in South Sudan and the consequences for East African countries. Parasites and Vectors. 2016;9(460).
10. Hayani K, Dandashli A, Weisshaar E. Cutaneous Leishmaniasis in Syria: Clinical Features, Current Status and the Effects of War. Acta Derm Venereol. 2015;95(1):62–6.
11. WHO. WHO: Weekly epidemiological record: Global leishmaniasis update, 2006-2015, a turning point in leishmaniasis surveillance. World Heal Organ. 2017;92(38):557–72.
12. Burza S, Croft SL, Boelaert M. Leishmaniasis. Vol. 392, The Lancet. Elsevier; 2018. p. 951–70.
13. Savioli L, Daumerie D, Crompton DW. Accelerating Work to Overcome the Global Impact of Neglected Tropical Diseases: A Roadmap for Implementation. World Health Organization. 2012.
14. The Sixtieth World Health Assembly. Resolution WHA60.13 Control of leishmaniasis. 2007.
15. Lacerda MM. The Brazilian Leishmaniasis Control Program. Mem Inst Oswaldo Cruz. 1994

Sep;89(3):489–95.

16. Gavgani ASM, Hodjati M, Mohite H, Davies C. Effect of insecticide-impregnated dog collars on incidence of zoonotic visceral leishmaniasis in Iranian children: a matchedcluster randomised trial. Lancet. Elsevier; 2002 Aug 3;360(9330):374–9.
17. Dye C. The logic of visceral leishmaniasis control. Am J Trop Med Hyg. 1996 Aug;55(2):125–30.
18. Ready PD. Epidemiology of visceral leishmaniasis. Clin Epidemiol. 2014;6(1):147–54.
19. Picado A, Ostyn B, Rijal S, Sundar S, Singh SP, Chappuis F, et al. Long-lasting Insecticidal Nets to Prevent Visceral Leishmaniasis in the Indian Subcontinent; Methodological Lessons Learned from a Cluster Randomised Controlled Trial. PLoS Negl Trop Dis. 2015 Apr;9(4).
20. Olliaro PL, Shamsuzzaman TAKM, Marasini B, Dhariwal AC, Be-Nazir A, Mondal D, et al. Investments in Research and Surveillance Are Needed to Go Beyond Elimination and Stop Transmission of Leishmania in the Indian Subcontinent. Lockwood DNJ, editor. PLoS Negl Trop Dis. Public Library of Science; 2017 Jan 26;11(1):e0005190.
21. Chappuis F, Sundar S, Hailu A, Ghalib H, Rijal S, Peeling RW, et al. Visceral leishmaniasis: what are the needs for diagnosis, treatment and control? Nat Rev Microbiol. 2007 Nov;5(11):873–82.
22. Biswas S, Subramanian A, Elmojtaba IM, Chattopadhyay J, Sarkar RR. Optimal combinations of control strategies and cost-effective analysis for visceral leishmaniasis disease transmission. 2017;
23. Kamhawi S. The yin and yang of leishmaniasis control. Aksoy S, editor. PLoS Negl Trop Dis. Public Library of Science; 2017 Apr 20;11(4):e0005529.
24. Sunyoto T, Potet J, Boelaert M. Visceral leishmaniasis in Somalia: A review of epidemiology and access to care. Satoskar AR, editor. PLoS Negl Trop Dis. Public Library of Science; 2017 Mar 9;11(3):e0005231.
25. Sunyoto T, Potet J, Boelaert M. Why miltefosine—a life-saving drug for leishmaniasis—is unavailable to people who need it the most. BMJ Glob Heal. BMJ Specialist Journals; 2018 May 3;3(3):e000709.
26. Ghorbani M, Farhoudi R. Leishmaniasis in humans: drug or vaccine therapy? Drug Des Devel Ther. Dove Press; 2017 Dec 22;Volume 12:25–40.
27. Fox-Rushby JA, Kaddar M, Levine R, Brenzel L. The economics of vaccination in low- and middle-income countries. Bull World Health Organ. 2004;82(9):640.
28. Hotez PJ, Bottazzi ME, Strych U. New Vaccines for the World's Poorest People. Annu Rev Med.

2016;67(1):405–17.

29. Lee BY, Bacon KM, Shah M, Kitchen SB, Connor DL, Slayton RB. The economic value of a visceral leishmaniasis vaccine in Bihar State, India. Am J Trop Med Hyg. 2012;86(3):417–25.
30. Bacon KM, Hotez PJ, Kruchten SD, Kamhawi S, Bottazzi ME, Valenzuela JG, et al. The potential economic value of a cutaneous leishmaniasis vaccine in seven endemic countries in the Americas. Vaccine. Elsevier; 2013 Jan 7;31(3):480–6.
31. Magill AJ. Chapter 277 - Leishmania Species: Visceral (Kala-Azar), Cutaneous, and Mucosal Leishmaniasis. In: Bennett JE, Dolin R, Blaser MJ, editors. Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases. 8th ed. Elsevier Saunders; 2015. p. 3493–525.
32. Bates PA. Transmission of Leishmania metacyclic promastigotes by phlebotomine sand flies. Int J Parasitol. 2007 Aug;37(10):1097–106.
33. Akhoundi M, Kuhls K, Cannet A, Votýpka J, Marty P, Delaunay P, et al. A Historical Overview of the Classification, Evolution, and Dispersion of Leishmania Parasites and Sandflies. Bañuls A-L, editor. PLoS Negl Trop Dis. Public Library of Science; 2016 Mar 3;10(3):e0004349.
34. Dey A, Singh S. Transfusion transmitted leishmaniasis: a case report and review of literature. Indian J Med Microbiol. Medknow Publications; 2006 Jul;24(3):165–70.
35. Meinecke CK, Schottelius J, Oskam L, Fleischer B. Congenital transmission of visceral leishmaniasis (Kala Azar) from an asymptomatic mother to her child. Pediatrics. American Academy of Pediatrics; 1999 Nov 1;104(5):e65.
36. St. C. Symmers W. LEISHMANIASIS ACQUIRED BY CONTAGION: A CASE OF MARITAL INFECTION IN BRITAIN. Lancet. Elsevier; 1960 Jan 16;275(7116):127–32.
37. Harhay MO, Olliaro PL, Lamounier Costa D, Nery Costa CH. Urban parasitology: visceral leishmaniasis in Brazil. Trends Parasitol. 2011;27:403–9.
38. Andrade BB, de Oliveira CI, Brodskyn CI, Barral A, Barral-Netto M. Role of Sand Fly Saliva in Human and Experimental Leishmaniasis: Current Insights. Scand J Immunol. John Wiley & Sons, Ltd; 2007 Aug 1;66(2–3):122–7.
39. Lestinova T, Rohousova I, Sima M, de Oliveira CI, Volf P. Insights into the sand fly saliva: Blood-feeding and immune interactions between sand flies, hosts, and Leishmania. Milon G, editor. PLoS Negl Trop Dis. Public Library of Science; 2017 Jul 13;11(7):e0005600.
40. Murphy K, Weaver C. Chapter 3 - The Induced Responses of Innate Immunity. In: Janeway's Immunobiology. 9th ed. Garland Science; 2017. p. 77–137.
41. Oryan A, Akbari M. Worldwide risk factors in leishmaniasis. Asian Pac J Trop Med. 2016;9(10):925–32.

42. Aronson N, Herwaldt BL, Libman M, Pearson R, Lopez-Velez R, Weina P, et al. Diagnosis and Treatment of Leishmaniasis: Clinical Practice Guidelines by the Infectious Diseases Society of America (IDSA) and the American Society of Tropical Medicine and Hygiene (ASTMH). Am J Trop Med Hyg. 2017;96(1):24–45.
43. Torres-Guerrero E, Quintanilla-Cedillo MR, Ruiz-Esmenjaud J, Arenas R. Leishmaniasis: a review. F1000Research. 2017 May 26;6:750.
44. Singh OP, Hasker E, Sacks D, Boelaert M, Sundar S. Asymptomatic Leishmania infection: a new challenge for Leishmania control. Clin Infect Dis. 2014 May 15;58(10):1424–9.
45. Zijlstra EE, el-Hassan AM, Ismael A, Ghalib HW. Endemic kala-azar in eastern Sudan: a longitudinal study on the incidence of clinical and subclinical infection and post-kala-azar dermal leishmaniasis. Am J Trop Med Hyg. 1994 Dec;51(6):826–36.
46. Schaefer KU, Kurtzhals JA, Gachihi GS, Muller AS, Kager PA. A prospective sero-epidemiological study of visceral leishmaniasis in Baringo District, Rift Valley Province, Kenya. Trans R Soc Trop Med Hyg. 89(5):471–5.
47. Ali A, Ashford RW. Visceral leishmaniasis in Ethiopia. IV. Prevalence, incidence and relation of infection to disease in an endemic area. Ann Trop Med Parasitol. 1994 Jun;88(3):289–93.
48. Evans TG, Teixeira MJ, McAuliffe IT, Vasconcelos I, Vasconcelos AW, Sousa A de A, et al. Epidemiology of visceral leishmaniasis in northeast Brazil. J Infect Dis. 1992 Nov;166(5):1124–32.
49. Moral L, Rubio EM, Moya M. A leishmanin skin test survey in the human population of l'Alacantí region (Spain): implications for the epidemiology of Leishmania infantum infection in southern Europe. Trans R Soc Trop Med Hyg. 96(2):129–32.
50. Bern C, Haque R, Chowdhury R, Ali M, Kurkjian KM, Vaz L, et al. The epidemiology of visceral leishmaniasis and asymptomatic leishmanial infection in a highly endemic Bangladeshi village. Am J Trop Med Hyg. 2007 May;76(5):909–14.
51. Ostyn B, Gidwani K, Khanal B, Picado A, Chappuis F, Singh SP, et al. Incidence of Symptomatic and Asymptomatic Leishmania donovani Infections in High-Endemic Foci in India and Nepal: A Prospective Study. Milon G, editor. PLoS Negl Trop Dis. Public Library of Science; 2011 Oct 4;5(10):e1284.
52. Topno RK, Das VNR, Ranjan A, Pandey K, Singh D, Kumar N, et al. Asymptomatic infection with visceral leishmaniasis in a disease-endemic area in bihar, India. Am J Trop Med Hyg. The American Society of Tropical Medicine and Hygiene; 2010 Sep;83(3):502–6.
53. Brake DA. Parasites and Immune Responses: Memory Illusion? DNA Cell Biol.

2003;22(6):405–19.

54. Solbach W, Laskay T. The Host Response to Leishmania Infection. In: Advances in Immunology. Academic Press; 1999. p. 275–317.
55. Murphy K, Weaver C. Glossary. In: Janeway's Immunobiology. 9th ed. Garland Science; 2017. p. 818–54.
56. SASSI A, LOUZIR H, BEN SALAH A BEN, MOKNI M, BEN OSMAN A BEN, DELLAGI K. Leishmanin skin test lymphoproliferative responses and cytokine production after symptomatic or asymptomatic Leishmania major infection in Tunisia. Clin Exp Immunol. Wiley/Blackwell (10.1111); 1999 Apr 1;116(1):127–32.
57. Rodrigues-Neto JF, Monteiro GR, Keesen TSL, Lacerda HG, Carvalho EM, Jeronimo SMB. CD45RO+ T Cells and T Cell Activation in the Long-Lasting Immunity after Leishmania infantum Infection. Am J Trop Med Hyg. 2018;98(3):875–82.
58. Gradoni L, Gramiccia M. Chapter 2.1.11 Leishmaniosis. In: Manual of Diagnostic Tests and Vaccines for Terrestrial Animals. 8th ed. World Organisation for Animal Health (OIE); 2018.
59. Michel G, Pomares C, Ferrua B, Marty P. Importance of worldwide asymptomatic carriers of Leishmania infantum (L. chagasi) in human. Acta Trop. Elsevier; 2011 Aug 1;119(2–3):69–75.
60. Singh OP, Gidwani K, Kumar R, Nylén S, Jones SL, Boelaert M, et al. Reassessment of immune correlates in human visceral leishmaniasis as defined by cytokine release in whole blood. Clin Vaccine Immunol. American Society for Microbiology; 2012 Jun 1;19(6):961–6.
61. Mougneau E, Bihl F, Glaichenhaus N. Cell biology and immunology of Leishmania. Immunol Rev. 2011 Mar;240(1):286–96.
62. Walker DM, Oghumu S, Gupta G, McGwire BS, Drew ME, Satoskar AR. Mechanisms of cellular invasion by intracellular parasites. Cell Mol Life Sci. NIH Public Access; 2014 Apr;71(7):1245–63.
63. de Morais CGV, Castro Lima AK, Terra R, dos Santos RF, Da-Silva SAG, Dutra PML. The Dialogue of the Host-Parasite Relationship: *Leishmania* spp. and *Trypanosoma cruzi* Infection. Biomed Res Int. 2015;2015:1–19.
64. Davies DH. Immune System. In: Encyclopedia of Life Sciences. John Wiley & Sons, Ltd; 2013.
65. Murphy K, Weaver C. Chapter 2 - Innate immunity: First Lines of Defense. In: Janeway's Immunobiology. 9th ed. Garland Science; 2017. p. 37–76.
66. Gurung P, Kanneganti TD. Innate immunity against Leishmania infections. Cell Microbiol. 2015 Sep 1;17(9):1286–94.
67. Domínguez M, Moreno I, López-Trascasa M, Toraño A. Complement interaction with

trypanosomatid promastigotes in normal human serum. J Exp Med. Rockefeller University Press; 2002 Feb 18;195(4):451–9.

68. Cecí-lio P, Pérez-Cabezas B, Santarém N, Maciel J, Rodrigues V, Cordeiro da Silva A. Deception and Manipulation: The Arms of Leishmania, a Successful Parasite. Front Immunol. 2014;5(October):1–16.
69. Rossi M, Fasel N. How to master the host immune system? Leishmania parasites have the solutions!
70. Kumar R, Nylén S. Immunobiology of visceral leishmaniasis. Front Immunol. Frontiers; 2012;3:251.
71. Puentes SM, Da Silva RP, Sacks DL, Hammer CH, Joiner KA. Serum resistance of metacyclic stage Leishmania major promastigotes is due to release of C5b-9. J Immunol. American Association of Immunologists; 1990 Dec 15;145(12):4311–6.
72. Hermoso T, Fishelson Z, Becker SI, Hirschberg K, Jaffe CL. Leishmanial protein kinases phosphorylate components of the complement system. EMBO J. 1991;10(13):4061–7.
73. Ueno N, Wilson ME. Receptor-mediated phagocytosis of Leishmania: Implications for intracellular survival. Trends Parasitol. Elsevier Ltd; 2012;28(8):335–44.
74. Brittingham A, Morrison CJ, McMaster WR, McGwire BS, Chang K-P, Mosser DM. Role of the Leishmania Surface Protease gp63 in Complement Fixation, Cell Adhesion, and Resistance to Complement-Mediated Lysis. J Immunol . 1995;155:3102–11.
75. Carter CR, Whitcomb JP, Campbell JA, Mukbel RM, McDowell MA. Complement receptor 3 deficiency influences lesion progression during Leishmania major infection in BALB/c mice. Infect Immun. American Society for Microbiology Journals; 2009 Dec 1;77(12):5668–75.
76. Marth T, Kelsall BL. Regulation of interleukin-12 by complement receptor 3 signaling. J Exp Med. Rockefeller University Press; 1997 Jun 2;185(11):1987–95.
77. Culley FJ, Harris RA, Kaye PM, McAdam KP, Raynes JG. C-reactive protein binds to a novel ligand on Leishmania donovani and increases uptake into human macrophages. J Immunol. American Association of Immunologists; 1996 Jun 15;156(12):4691–6.
78. Woelbing F, Kostka SL, Moelle K, Belkaid Y, Sunderkoetter C, Verbeek S, et al. Uptake of Leishmania major by dendritic cells is mediated by Fcgamma receptors and facilitates acquisition of protective immunity. J Exp Med. Rockefeller University Press; 2006 Jan 23;203(1):177–88.
79. Charmoy M, Auderset F, Allenbach C, Tacchini-Cottier F. The prominent role of neutrophils during the initial phase of infection by leishmania parasites. J Biomed Biotechnol.

2010;2010:2–9.

80. Kaye P, Scott P. Leishmaniasis: complexity at the host–pathogen interface. Nat Rev Microbiol. Nature Publishing Group; 2011;9(8):604–15.
81. Denkers E, Abi Abdallah D. Neutrophils cast extracellular traps in response to protozoan parasites. Front Immunol. 2012;3:382.
82. Sousa LMA, Carneiro MBH, Resende ME, Martins LS, Dos Santos LM, Vaz LG, et al. Neutrophils have a protective role during early stages of Leishmania amazonensis infection in BALB/c mice. Parasite Immunol. 2014 Jan;36(1):13–31.
83. Chang KP. Leishmanicidal mechanisms of human polymorphonuclear phagocytes. Am J Trop Med Hyg. 1981;30(2):322–33.
84. Pearson RD, Steigbigel RT, Zandbergen G van, Laufs H, Hansen B, Müller K, et al. Phagocytosis and killing of the protozoan Leishmania donovani by human polymorphonuclear leukocytes. J Immunol. American Association of Immunologists; 1981 Oct 15;127(4):1438–43.
85. Guimarães-Costa AB, Nascimento MTC, Froment GS, Soares RPP, Morgado FN, Conceicao-Silva F, et al. Leishmania amazonensis promastigotes induce and are killed by neutrophil extracellular traps. Proc Natl Acad Sci. 2009;106(16):6748–53.
86. Morgado FN, Nascimento MTC, Saraiva EM, De Oliveira-Ribeiro C, De Fátima Madeira M, Da Costa-Santos M, et al. Are neutrophil extracellular traps playing a role in the parasite control in active American tegumentary leishmaniasis lesions? PLoS One. 2015;10(7).
87. de Souza Carmo EV, Katz S, Barbiéri CL. Neutrophils reduce the parasite burden in Leishmania (Leishmania) amazonensis-infected macrophages. PLoS One. Public Library of Science; 2010 Nov 3;5(11):e13815.
88. Carlsen ED, Liang Y, Shelite TR, Walker DH, Melby PC, Soong L. Permissive and protective roles for neutrophils in leishmaniasis. Clin Exp Immunol. 2015 Nov;182(2):109–18.
89. Aga E, Katschinski DM, van Zandbergen G, Laufs H, Hansen B, Müller K, et al. Inhibition of the spontaneous apoptosis of neutrophil granulocytes by the intracellular parasite Leishmania major. J Immunol. American Association of Immunologists; 2002 Jul 15;169(2):898–905.
90. Filardy AA, Pires DR, Nunes MP, Takiya CM, Freire-de-Lima CG, Ribeiro-Gomes FL, et al. Proinflammatory clearance of apoptotic neutrophils induces an IL-12(low)IL-10(high) regulatory phenotype in macrophages. J Immunol. American Association of Immunologists; 2010 Aug 15;185(4):2044–50.
91. John B, Hunter CAH. Neutrophil Soldiers or Trojan Horses? Science (80-). 2008;321:9–10.
92. Ribeiro-Gomes FL, Peters NC, Debrabant A, Sacks DL. Efficient Capture of Infected

Neutrophils by Dendritic Cells in the Skin Inhibits the Early Anti-Leishmania Response. Müller I, editor. PLoS Pathog. Public Library of Science; 2012 Feb 16;8(2):e1002536.

93. Ritter U, Frischknecht F, van Zandbergen G. Are neutrophils important host cells for Leishmania parasites? Trends Parasitol. 2009 Nov;25(11):505–10.
94. Wenzel UA, Bank E, Florian C, Forster S, Zimara N, Steinacker J, et al. Leishmania major parasite stage-dependent host cell invasion and immune evasion. FASEB J. 2012;26(1):29–39.
95. van Zandbergen G, Bollinger A, Wenzel A, Kamhawi S, Voll R, Klinger M, et al. Leishmania disease development depends on the presence of apoptotic promastigotes in the virulent inoculum. Proc Natl Acad Sci. 2006 Sep 12;103(37):13837–42.
96. Peters NC, Egen JG, Secundino N, Debrabant A, Kimblin N, Kamhawi S, et al. In vivo imaging reveals an essential role for neutrophils in leishmaniasis transmitted by sand flies. Science. American Association for the Advancement of Science; 2008 Aug 15;321(5891):970–4.
97. Michael A. J. Ferguson. The structure, biosynthesis and functions of glycosylphosphatidylinositol anchors, and the contributions of trypanosome research. J Cell Sci. 1999;112(Pt17):2799–809.
98. Sundar S, Singh B. Identifying vaccine targets for anti-leishmanial vaccine development. Expert Rev Vaccines. 2014;13(4):489–505.
99. Soehnlein O, Steffens S, Hidalgo A, Weber C. Neutrophils as protagonists and targets in chronic inflammation. Nat Rev Immunol. Nature Publishing Group; 2017;advance on.
100. Silva MT, Correia-Neves M. Neutrophils and macrophages: the main partners of phagocyte cell systems. Front Immunol. 2012 Jan;3:174.
101. Hurrell BP, Schuster S, Grün E, Coutaz M, Williams RA, Held W, et al. Rapid Sequestration of Leishmania mexicana by Neutrophils Contributes to the Development of Chronic Lesion. PLoS Pathog. Public Library of Science; 2015 May 28;11(5):e1004929.
102. Ribeiro-Gomes FL, Sacks D. The influence of early neutrophil-Leishmania interactions on the host immune response to infection. Front Cell Infect Microbiol. Frontiers Media SA; 2012 Jan;2:59.
103. Shio MT, Hassani K, Isnard A, Ralph B, Contreras I, Gomez MA, et al. Host cell signalling and leishmania mechanisms of evasion. Vol. 2012, Journal of Tropical Medicine. Hindawi; 2012. p. 819512.
104. Ritter U, Moll H, Laskay T, Brocker E-B, Velazco O, Becker I, et al. Differential Expression of Chemokines in Patients with Localized and Diffuse Cutaneous American Leishmaniasis. Vol.

173, The Journal of Infectious Diseases. 1995.

105. Ribeiro-Gomes FL, Moniz-de-Souza MCA, Alexandre-Moreira MS, Dias WB, Lopes MF, Nunes MP, et al. Neutrophils activate macrophages for intracellular killing of Leishmania major through recruitment of TLR4 by neutrophil elastase. J Immunol. American Association of Immunologists; 2007 Sep 15;179(6):3988–94.
106. Jain K, Jain NK. Vaccines for visceral leishmaniasis: A review. J Immunol Methods. 2015;422:1–12.
107. Flynn B, Wang V, Sacks DL, Seder RA, Verthelyi D. Prevention and treatment of cutaneous leishmaniasis in primates by using synthetic type D/A oligodeoxynucleotides expressing CpG motifs. Infect Immun. American Society for Microbiology Journals; 2005 Aug 1;73(8):4948–54.
108. Chauhan P, Shukla D, Chattopadhyay D, Saha B. Redundant and regulatory roles for Toll-like receptors in Leishmania infection. Clin Exp Immunol. Wiley-Blackwell; 2017;190(2):167–86.
109. Paul J, Naskar K, Chowdhury S, Alam N, Chakraborti T, De T. TLR4-mediated activation of MyD88 signaling induces protective immune response and IL-10 down-regulation in Leishmania donovani infection. Vol. 51, Indian Journal of Biochemistry & Biophysics. 2014.
110. Eberl G, Colonna M, Santo JPD, McKenzie ANJ. Innate lymphoid cells: A new paradigm in immunology. Science (80-). 2015;348(6237).
111. Mckenzie ANJ, Spits H, Eberl G. Innate Lymphoid Cells in Inflammation and Immunity. 2014;
112. Liese J, Schleicher U, Bogdan C. The innate immune response against Leishmania parasites. Immunobiology. Urban & Fischer; 2008 May 14;213(3–4):377–87.
113. Becker I, Salaiza N, Aguirre M, Delgado J, Carrillo-Carrasco N, Kobeh LG, et al. Leishmania lipophosphoglycan (LPG) activates NK cells through toll-like receptor-2. Mol Biochem Parasitol. Elsevier; 2003 Aug 31;130(2):65–74.
114. Maroof A, Beattie L, Zubairi S, Svensson M, Stager S, Kaye PM. Posttranscriptional Regulation of Il10 Gene Expression Allows Natural Killer Cells to Express Immunoregulatory Function. Immunity. Cell Press; 2008 Aug 15;29(2):295–305.
115. Kaye PM, Svensson M, Ato M, Maroof A, Polley R, Stager S, et al. The immunopathology of experimental visceral leishmaniasis. Immunol Rev. 2004;201:239–53.
116. Belkaid Y, Mendez S, Lira R, Kadambi N, Milon G, Sacks D. A Natural Model of Leishmania major Infection Reveals a Prolonged "Silent" Phase of Parasite Amplification in the Skin Before the Onset of Lesion Formation and Immunity. J Immunol. American Association of Immunologists; 2000 Jul 15;165(2):969–77.

117. von Stebut E, Tenzer S. Cutaneous leishmaniasis: Distinct functions of dendritic cells and macrophages in the interaction of the host immune system with Leishmania major. Int J Med Microbiol. Urban & Fischer; 2017 Jan 1;308(1):206–14.
118. Iwasaki A, Medzhitov R. Control of adaptive immunity by the innate immune system. Nat Immunol. NIH Public Access; 2015 Apr;16(4):343–53.
119. Pedersen AE. Immunity to Infection. In: Encyclopedia of Life Sciences. John Wiley & Sons, Ltd; 2001.
120. Kedzierski L, Evans KJ. Immune responses during cutaneous and visceral leishmaniasis. Parasitology. 2014;1–19.
121. Murphy K, Weaver C. Chapter 9 - T-cell-Mediated Immunity. In: Janeway's Immunobiology. 9th ed. Garland Science; 2017. p. 345–98.
122. Parham P. The Immune System. 2nd ed. Garland Science; 2005.
123. Murphy K, Weaver C. Chapter 6 - Antigen Presentation to T Lymphocytes. In: Janeway's Immunobiology. 9th ed. Garland Science; 2017. p. 213–56.
124. Zhang L, Udaka K, Mamitsuka H, Zhu S. Toward more accurate pan-specific MHC-peptide binding prediction: A review of current methods and tools. Brief Bioinform. 2012 May 1;13(3):350–64.
125. Zhang W, Moldovan I, Targoni OS, Subbramanian RA, Lehmann P V. How much of virus-specific CD8 T cell reactivity is detected with a peptide pool when compared to individual peptides? Viruses. Multidisciplinary Digital Publishing Institute; 2012 Oct 29;4(11):2636–49.
126. Kreer C, Rauen J, Zehner M, Burgdorf S. Cross-presentation: How to get there - or how to get the ER. Front Immunol. 2012;2(January):1–10.
127. Goldszmid RS, Sher A. Processing and presentation of antigens derived from intracellular protozoan parasites. Curr Opin Immunol. 2010;22(1):118–23.
128. Bertholet S, Sacks D, Desjardins M, Debrabant A, Goldszmid R, Morrot A, et al. Leishmania Antigens Are Presented to CD8+ T Cells by a Transporter Associated with Antigen Processing-Independent Pathway In Vitro and In Vivo. J Immunol. American Association of Immunologists; 2014 Sep 15;177(6):3525–33.
129. Marsh SGE, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich H a., et al. Nomenclature for factors of the HLA system, 2010. Tissue Antigens. 2010;75(4):291–455.
130. Marsh SGE, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA, et al. Nomenclature for factors of the HLA system, 2010. Tissue Antigens. John Wiley & Sons, Ltd (10.1111); 2010 Apr 1;75(4):291–455.

131. Wang P, Sidney J, Dow C, Mothé B, Sette A, Peters B. A Systematic Assessment of MHC Class II Peptide Binding Predictions and Evaluation of a Consensus Approach. PLoS Comput Biol. 2008;4(4):e1000048.
132. Wang M, Claesson MH. Chapter 17 - Classification of human leukocyte antigen (HLA) supertypes. In: Tomar N, De RK, editors. Immunoinformatics Methods in Molecular Biology (Methods and Protocols). 2nd ed. Humana Press; 2014. p. 309–17.
133. Sidney J, Peters B, Frahm N, Brander C, Sette A. HLA class I supertypes: a revised and updated classification. BMC Immunol. 2008 Jan;9:1.
134. Sette A, Sidney J. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. Immunogenetics. Springer-Verlag; 1999 Nov 22;50(3–4):201–12.
135. Lund O, Nielsen M, Kesmir C, Petersen AG, Lundegaard C, Worning P, et al. Definition of supertypes for HLA molecules using clustering of specificity matrices. Immunogenetics. Springer-Verlag; 2004 Mar 1;55(12):797–810.
136. Panina-Bordignon P, Tan A, Termijtelen A, Demotz S, Corradin G, Lanzavecchia A. Universally immunogenic T cell epitopes: promiscuous binding to human MHC class II and promiscuous recognition by T cells. Eur J Immunol. John Wiley & Sons, Ltd; 1989 Dec 1;19(12):2237–42.
137. Fakiola M, Langford C, Pontes NN, Edkins S, Mathew CG, Giannoulatou E, et al. Common variants in the HLA-DRB1–HLA-DQA1 HLA class II region are associated with susceptibility to visceral leishmaniasis. Nat Genet. Nature Publishing Group; 2013 Feb 6;45(2):208–13.
138. Piñero J, Martínez E, Pacheco R, Carmelo E, Zurita A, De Armas F, et al. Lack of association between HLA class II antigens polymorphism and leishmaniosis in a perivian endemic region. Rev Ibérica Parasitol. 2002;62(1–2):15–9.
139. Faghiri Z, Tabei SZ, Taheri F. Study of the Association of HLA Class I Antigens with Kala-Azar. Hum Hered. Karger Publishers; 1995;45(5):258–61.
140. Roy S, Mukhopadhyay D, Mukherjee S, Ghosh S, Kumar S, Sarkar K, et al. A Defective Oxidative Burst and Impaired Antigen Presentation are Hallmarks of Human Visceral Leishmaniasis. J Clin Immunol. 2015;35(1):56–67.
141. Rodrigues V, Cordeiro-da-Silva A, Laforge M, Ouaissi A, Akharid K, Silvestre R, et al. Impairment of T cell function in parasitic infections. PLoS Negl Trop Dis. Public Library of Science; 2014 Feb;8(2):e2567.
142. Klaus GG. Lymphocyte development. In: Encyclopedia of Life Sciences. John Wiley & Sons, Ltd; 2001.

143. Ademokun AA, Dunn-Walters D. Immune Responses: Primary and Secondary. In: Encyclopedia of Life Sciences. John Wiley & Sons, Ltd; 2010.
144. Lever M, Maini PK, van der Merwe PA, Dushek O. Phenotypic models of T cell activation. Nat Rev Immunol. Nature Publishing Group; 2014;14(9):619–29.
145. Udono H, Wang JO, Watanabe T. Antigen Presentation to Lymphocytes. In: Encyclopedia of Life Sciences. John Wiley & Sons, Ltd; 2007.
146. Stone JD, Chervin AS, Kranz DM. T-cell receptor binding affinities and kinetics: impact on T-cell activity and specificity. Immunology. 2009;126(2):165–76.
147. Leggatt G. Peptide Dose and/or Structure in Vaccines as a Determinant of T Cell Responses. Vaccines. Multidisciplinary Digital Publishing Institute; 2014 Jul 2;2(3):537–48.
148. Kersh EN, Kaech SM, Onami TM, Moran M, Wherry EJ, Miceli MC, et al. TCR Signal Transduction in Antigen-Specific Memory CD8 T Cells. J Immunol. American Association of Immunologists; 2003 Jun 1;170(11):5455–63.
149. Huber M, Timms E, Mak TW, Röllinghoff M, Lohoff M. Effective and Long-Lasting Immunity against the Parasite Leishmania major in CD8-Deficient Mice. Infect Immun. American Society for Microbiology Journals; 1998 Aug 1;66(8):3968–70.
150. Belkaid Y, Von Stebut E, Mendez S, Lira R, Caler E, Bertholet S, et al. CD8+ T Cells Are Required for Primary Immunity in C57BL/6 Mice Following Low-Dose, Intradermal Challenge with Leishmania major. J Immunol. 2002;168:3992–4000.
151. Kaufmann SH, Kabelitz D. The Immune Response to Infectious Agents. In: Kaufmann SH, Kabelitz D, editors. Immunology of Infection. 3rd ed. Elsevier; 2010.
152. McMahon-Pratt D, Alexander J. Does the Leishmania major paradigm of pathogenesis and protection hold for New World cutaneous leishmaniases or the visceral disease? Immunol Rev. Wiley/Blackwell (10.1111); 2004 Oct 1;201(1):206–24.
153. Mosmann TR, Coffman RL. TH1 and TH2 Cells: Different Patterns of Lymphokine Secretion Lead to Different Functional Properties. Annu Rev Immunol. 2003 Apr;7(1):145–73.
154. Bradley DJ. Genetics of susceptibility and resistance in the vertebrate host. In: Peters W, Killick-Kendrick R, editors. The Leishmaniases in Biology and Medicine 551. 1987.
155. Feng G-J, Goodridge HS, Harnett MM, Wei X-Q, Nikolaev A V., Higson AP, et al. Extracellular Signal-Related Kinase (ERK) and p38 Mitogen-Activated Protein (MAP) Kinases Differentially Regulate the Lipopolysaccharide-Mediated Induction of Inducible Nitric Oxide Synthase and IL-12 in Macrophages: Leishmania Phosphoglycans Subvert Macr. J Immunol. American Association of Immunologists; 1999;163:6403–3412.

156. Okwor I, Uzonna JE. Pathways leading to interleukin-12 production and protective immunity in cutaneous leishmaniasis. Cell Immunol. 2016;309:32–6.
157. Okwor I, Uzonna J. Persistent parasites and immunologic memory in cutaneous leishmaniasis: implications for vaccine designs and vaccination strategies. Immunol Res. 2008 Jun 4;41(2):123–36.
158. Murphy K, Weaver C. Chapter 11 - Integrated Dynamics of Innate and Adaptive Immunity. In: Janeway's Immunobiology. 9th ed. Garland Science; 2017. p. 445–91.
159. Stober CB, Lange UG, Roberts MTM, Alcami A, Blackwell JM. IL-10 from regulatory T cells determines vaccine efficacy in murine Leishmania major infection. J Immunol. American Association of Immunologists; 2005 Aug 15;175(4):2517–24.
160. Reiner SL, Locksley RM. The Regulation of Immunity to Leishmania Major. Annu Rev Immunol. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA ; 1995 Apr 28;13(1):151–77.
161. Peters NC, Pagán AJ, Lawyer PG, Hand TW, Henrique Roma E, Stamper LW, et al. Chronic Parasitic Infection Maintains High Frequencies of Short-Lived Ly6C+CD4+ Effector T Cells That Are Required for Protection against Re-infection. Gazzinelli RT, editor. PLoS Pathog. Public Library of Science; 2014 Dec 4;10(12):e1004538.
162. Romano A, Doria NA, Mendez J, Sacks DL, Peters NC. Cutaneous Infection with Leishmania major Mediates Heterologous Protection against Visceral Infection with Leishmania infantum. J Immunol. American Association of Immunologists; 2015 Oct 15;195(8):3816–27.
163. Aebischer T, Matuschewski K, Hartmann S. Editorial: Parasite Infections: From Experimental Models to Natural Systems. Front Cell Infect Microbiol. Frontiers; 2018 Feb 2;8:12.
164. Pirmez C. Cytokine Patterns in the Pathogenesis of Human Leishmaniasis. 1993;
165. Kharazmi A. T-cell response in human leishmaniasis. 1999;
166. Loeuillet C, Bañuls A-L, Hide M. Study of Leishmania pathogenesis in mice: experimental considerations. Parasit Vectors. BioMed Central; 2016 Dec 11;9(1):144.
167. Loria-Cervera EN, Andrade-Narvaez FJ. Animal Models for the Study of Leishmaniasis Immunology. Rev Inst Med Trop Sao Paulo. 2014;56(1):1–11.
168. Hosein S, Blake DP, Solano-Gallego L. Insights on adaptive and innate immunity in canine leishmaniosis. 2019;
169. Coler RN, Reed SG. Second-generation vaccines against leishmaniasis. Trends Parasitol. Elsevier Current Trends; 2005 May 1;21(5):244–9.
170. Ghalib HW, Piuvezam MR, Skeiky YA, Siddig M, Hashim FA, el-Hassan AM, et al. Interleukin 10

production correlates with pathology in human Leishmania donovani infections. J Clin Invest. American Society for Clinical Investigation; 1993 Jul 1;92(1):324–9.

171. Cummings HE, Tuladhar R, Satoskar AR. Cytokines and their STATs in cutaneous and visceral leishmaniasis. J Biomed Biotechnol. Hindawi; 2010 Mar 15;2010:294389.
172. Kumar R, Goto Y, Gidwani K, Cowgill KD, Sundar S, Reed SG. Evaluation of ex vivo human immune response against candidate antigens for a visceral leishmaniasis vaccine. Am J Trop Med Hyg. The American Society of Tropical Medicine and Hygiene; 2010 May;82(5):808–13.
173. Nylén S, Sacks D. Interleukin-10 and the pathogenesis of human visceral leishmaniasis.pdf. Trends Immunol. 2007;28(9):378–84.
174. Kumar R, Chauhan SB, Ng SS, Sundar S, Engwerda CR. Immune checkpoint targets for host-directed therapy to prevent and treat leishmaniasis. Front Immunol. Frontiers; 2017 Nov 8;8(NOV):1492.
175. Da-Cruz AM, Bittar R, Mattos M, Oliveira-Neto MP, Nogueira R, Pinho-Ribeiro V, et al. T-cell-mediated immune responses in patients with cutaneous or mucosal leishmaniasis: long-term evaluation after therapy. Clin Diagn Lab Immunol. American Society for Microbiology; 2002 Mar 1;9(2):251–6.
176. Katara GK, Raj A, Kumar R, Avishek K, Kaushal H, Ansari NA, et al. Analysis of localized immune responses reveals presence of Th17 and Treg cells in cutaneous leishmaniasis due to Leishmania tropica. BMC Immunol. 2013;14(Il):52.
177. Belkaid Y. Regulatory T cells and infection: a dangerous necessity. Nat Rev Immunol. Nature Publishing Group; 2007 Nov 1;7(11):875–88.
178. Costa DL, Guimarães LH, Cardoso TM, Queiroz A, Lago E, Roselino AM, et al. Characterization of regulatory T cell (Treg) function in patients infected with Leishmania braziliensis. Hum Immunol. 2013;72(12).
179. Maurya R, Kumar R, Prajapati VK, Manandhar KD, Sacks D, Sundar S, et al. Human visceral leishmaniasis is not associated with expansion or accumulation of Foxp3+ CD4 cells in blood or spleen. Parasite Immunol. John Wiley & Sons, Ltd (10.1111); 2010 Mar 19;32(7):479–83.
180. Stäger S, Maroof A, Zubairi S, Sanos SL, Kopf M, Kaye PM. Distinct roles for IL-6 and IL-12p40 in mediating protection againstLeishmania donovani and the expansion of IL-10+ CD4+ T cells. Eur J Immunol. John Wiley & Sons, Ltd; 2006 Jul 1;36(7):1764–71.
181. Zijlstra EE. The immunology of post-kala-azar dermal leishmaniasis (PKDL). Parasit Vectors. BioMed Central; 2016;9(1):464.
182. Katara GK, Ansari NA, Singh A, Ramesh V, Salotra P. Evidence for Involvement of Th17 Type

Responses in Post Kala Azar Dermal Leishmaniasis (PKDL). Kamhawi S, editor. PLoS Negl Trop Dis. Public Library of Science; 2012 Jun 19;6(6):e1703.

183. Katara GK, Ansari NA, Verma S, Ramesh V, Salotra P. Foxp3 and IL-10 Expression Correlates with Parasite Burden in Lesional Tissues of Post Kala Azar Dermal Leishmaniasis (PKDL) Patients. Engwerda CR, editor. PLoS Negl Trop Dis. Public Library of Science; 2011 May 31;5(5):e1171.
184. Salih MA, Ibrahim ME, Blackwell JM, Miller EN, Khalil EAG, ElHassan AM, et al. IFNG and IFNGR1 gene polymorphisms and susceptibility to post-kala-azar dermal leishmaniasis in Sudan. Genes Immun. Nature Publishing Group; 2007 Jan 30;8(1):75–8.
185. Bacellar O, Lessa H, Schriefer A, Machado P, Ribeiro de Jesus A, Dutra WO, et al. Up-regulation of Th1-type responses in mucosal leishmaniasis patients. Infect Immun. American Society for Microbiology Journals; 2002 Dec 1;70(12):6734–40.
186. Muniz AC, Bacellar O, Lima Lago E, Carvalho AM, Carneiro PP, Guimarães LH, et al. Immunologic markers of protection in leishmania (Viannia) braziliensis infection: A 5-year cohort study. J Infect Dis. Oxford University Press; 2016 Aug 15;214(4):570–6.
187. Soria-Guerra RE, Nieto-Gomez R, Govea-Alonso DO, Rosales-Mendoza S, Duarte A, Queiroz AT, et al. The baculovirus expression vector system: A commercial manufacturing platform for viral vaccines and gene therapy vectors. Vaccine. Elsevier Ltd; 2015;14(2):1–12.
188. Convit J, Kerdel-Vegas F, Gordon B. Disseminated Anergic Cutaneous Leishmaniasis. Br J Dermatol. John Wiley & Sons, Ltd (10.1111); 1962 Apr 1;74(4):132–5.
189. Hashiguchi Y, Gomez EL, Kato H, Martini LR, Velez LN, Uezato H. Diffuse and disseminated cutaneous leishmaniasis: clinical cases experienced in Ecuador and a brief review. Trop Med Health. BioMed Central; 2016 Dec 14;44(1):2.
190. Pinto L, Covas MJ, Victorino RM. Loss of CD45RA and gain of CD45RO after in vitro activation of lymphocytes from HIV-infected patients. Immunology. Wiley-Blackwell; 1991 Jun;73(2):147–50.
191. Farber DL, Yudanin NA, Restifo NP. Human memory T cells: generation, compartmentalization and homeostasis. Nat Rev Immunol. 2014 Jan;14(1):24–35.
192. Mahnke YD, Brodie TM, Sallusto F, Roederer M, Lugli E. The who's who of T-cell differentiation: Human memory T-cell subsets. Eur J Immunol. 2013;43(11):2797–809.
193. Golubovskaya V, Wu L. Different subsets of T cells, memory, effector functions, and CAR-T immunotherapy. Cancers (Basel). 2016;8(3).
194. Seder RA, Darrah PA, Roederer M. T-cell quality in memory and protection: implications for

vaccine design. Nat Rev Immunol. Nature Publishing Group; 2008 Apr 1;8(4):247–58.

195. Rodríguez-Cortés A, Carrillo E, Martorell S, Todolí F, Ojeda A, Martínez-Flórez A, et al. Compartmentalized Immune Response in Leishmaniasis: Changing Patterns throughout the Disease. Stäger S, editor. PLoS One. Public Library of Science; 2016 May 12;11(5):e0155224.
196. Jaimes MC, Rojas OL, González AM, Cajiao I, Charpilienne A, Pothier P, et al. Frequencies of Virus-Specific CD4+ and CD8+ T Lymphocytes Secreting Gamma Interferon after Acute Natural Rotavirus Infection in Children and Adults. J Virol. American Society for Microbiology Journals; 2002 May 15;76(10):4741–9.
197. González PA, Prado CE, Leiva ED, Carreño LJ, Bueno SM, Riedel CA, et al. Respiratory Syncytial Virus Impairs T Cell Activation by Preventing Synapse Assembly with Dendritic Cells. Proc Natl Acad Sci U S A. 2008;105(39):14999–1500.
198. Glennie ND, Scott P. Memory T cells in cutaneous leishmaniasis. Cell Immunol. 2016;
199. Schenkel JM, Fraser KA, Beura LK, Pauken KE, Vezys V, Masopust D, et al. T cell memory. Resident memory CD8 T cells trigger protective innate and adaptive immune responses. Science. American Association for the Advancement of Science; 2014 Oct 3;346(6205):98–101.
200. Khamesipour A, Nateghi Rostami M, Tasbihi M, Miramin Mohammadi A, Shahrestani T, Sarrafnejad A, et al. Phenotyping of circulating CD8 + T cell subsets in human cutaneous leishmaniasis. Microbes Infect. Elsevier Masson; 2012 Aug 1;14(9):702–11.
201. Boelaert M, Sundar S. Leishmaniasis. In: Farrar J, Hotez PJ, Junghanss T, Kang G, Lalloo D, White NJ, editors. Manson's Tropical Infectious Diseases. Elsevier; 2014. p. 631–651e4.
202. Keshavarz Valian H, Nateghi Rostami M, Tasbihi M, Miramin Mohammadi A, Eskandari SE, Sarrafnejad A, et al. CCR7+ central and CCR7 effector memory CD4+ T cells in human cutaneous leishmaniasis. J Clin Immunol. 2013;33(1):220–34.
203. Mary C, Auriault V, Faugère B, Dessein AJ. Control of Leishmania infantum infection is associated with CD8(+) and gamma interferon- and interleukin-5-producing CD4(+) antigen-specific T cells. Infect Immun. American Society for Microbiology Journals; 1999 Nov 1;67(11):5559–66.
204. Carvalho AM, Magalhães A, Carvalho LP, Bacellar O, Scott P, Carvalho EM. Immunologic response and memory T cells in subjects cured of tegumentary leishmaniasis. BMC Infect Dis. BioMed Central; 2013 Dec 9;13(1):529.
205. Bittar RC, Nogueira RS, Vieira-Gonçalves R, Pinho-Ribeiro V, Mattos MS, Oliveira-Neto MP, et al. T-cell responses associated with resistance to Leishmania infection in individuals from

endemic areas for Leishmania (Viannia) braziliensis. Mem Inst Oswaldo Cruz. 2007;102(5):625–30.

206. Pereira-Carvalho R, Mendes-Aguiar CO, Oliveira-Neto MP, Covas CJF, Bertho ÁL, Da-Cruz AM, et al. Leishmania braziliensis-Reactive T Cells Are Down-Regulated in Long-Term Cured Cutaneous Leishmaniasis, but the Renewal Capacity of T Effector Memory Compartments Is Preserved. Rodrigues MM, editor. PLoS One. Public Library of Science; 2013 Nov 26;8(11):e81529.
207. de Oliveira Mendes-Aguiar C, Vieira-Gonçalves R, Guimarães LH, de Oliveira-Neto MP, Carvalho EM, Da-Cruz AM. Effector memory CD4+ T cells differentially express activation associated molecules depending on the duration of American cutaneous leishmaniasis lesions. Clin Exp Immunol. John Wiley & Sons, Ltd (10.1111); 2016 Aug 1;185(2):202–9.
208. Mendonça VR de, Barral-Netto M. Immunoregulation in human malaria: the challenge of understanding asymptomatic infection. Memórias do Inst Oswaldo Cruz. Instituto Oswaldo Cruz; 2015 Dec;110(8):945–55.
209. De Luca PM, Macedo ABB. Cutaneous Leishmaniasis Vaccination: A Matter of Quality. Front Immunol. Frontiers; 2016 Jan 21;7:151.
210. Hammarlund E, Lewis MW, Hansen SG, Strelow LI, Nelson JA, Sexton GJ, et al. Duration of antiviral immunity after smallpox vaccination. Nat Med. Nature Publishing Group; 2003 Sep 17;9(9):1131–7.
211. Wilkinson TM, Li CKF, Chui CSC, Huang AKY, Perkins M, Liebner JC, et al. Preexisting influenza-specific CD4+ T cells correlate with disease protection against influenza challenge in humans. Nat Med. Nature Publishing Group; 2012 Feb 29;18(2):274–80.
212. Henrickson SE, Perro M, Loughhead SM, Senman B, Stutte S, Quigley M, et al. Antigen availability determines CD8+T cell-dendritic cell interaction kinetics and memory fate decisions. Immunity. NIH Public Access; 2013 Sep 19;39(3):496–507.
213. Wherry EJ, Teichgräber V, Becker TC, Masopust D, Kaech SM, Antia R, et al. Lineage relationship and protective immunity of memory CD8 T cell subsets. 2003;
214. Okhrimenko A, Grün JR, Westendorf K, Fang Z, Reinke S, von Roth P, et al. Human memory T cells from the bone marrow are resting and maintain long-lasting systemic memory. Proc Natl Acad Sci U S A. 2014 Jun 24;111(25):9229–34.
215. Mandell MA, Beverley SM, Lin P, Flynn J, Taylor M, Hoerauf A, et al. Concomitant Immunity Induced by Persistent Leishmania major Does Not Preclude Secondary Re-Infection: Implications for Genetic Exchange, Diversity and Vaccination. Debrabant A, editor. PLoS Negl

Trop Dis. Public Library of Science; 2016 Jun 28;10(6):e0004811.

216. Aebischer T, Moody SF, Handman E. Persistence of virulent Leishmania major in murine cutaneous leishmaniasis: a possible hazard for the host. Infect Immun. American Society for Microbiology Journals; 1993 Jan 1;61(1):220–6.
217. Kimblin N, Peters N, Debrabant A, Secundino N, Egen J, Lawyer P, et al. Quantification of the infectious dose of Leishmania major transmitted to the skin by single sand flies. Proc Natl Acad Sci U S A. 2008 Jul 22;105(29):10125–30.
218. Belkaid Y, Piccirillo CA, Mendez S, Shevach EM, Sacks DL. CD4+ CD25+ regulatory T cells control Leishmania major persistence and immunity. Nature. 2002;420.
219. Diro E, van Griensven J, Mohammed R, Colebunders R, Asefa M, Hailu A, et al. Atypical manifestations of visceral leishmaniasis in patients with HIV in north Ethiopia: a gap in guidelines for the management of opportunistic infections in resource poor settings. Lancet Infect Dis. Elsevier; 2015 Jan 1;15(1):122–9.
220. Távora LGF, Nogueira MB, Gomes ST. Visceral Leishmaniasis/HIV co-infection in northeast Brazil: evaluation of outcome. Vol. 19, The Brazilian Journal of Infectious Diseases. 2015.
221. Desjeux P, Alvar J. *Leishmania* /HIV co-infections: epidemiology in Europe. Ann Trop Med Parasitol. 2003 Oct 2;97(sup1):3–15.
222. Stenger S, Donhauser N, Thüring H, Röllinghoff M, Bogdan C. Reactivation of latent leishmaniasis by inhibition of inducible nitric oxide synthase. J Exp Med. Rockefeller University Press; 1996 Apr 1;183(4):1501–14.
223. Zaph C, Uzonna J, Beverley SM, Scott P. Central memory T cells mediate long-term immunity to Leishmania major in the absence of persistent parasites. Nat Med. Nature Publishing Group; 2004 Oct 26;10(10):1104–10.
224. Mendez S, Reckling SK, Piccirillo CA, Sacks D, Belkaid Y. Role for CD4(+) CD25(+) regulatory T cells in reactivation of persistent leishmaniasis and control of concomitant immunity. J Exp Med. Rockefeller University Press; 2004 Jul 19;200(2):201–10.
225. Martínez-López M, Soto M, Iborra S, Sancho D. Leishmania Hijacks myeloid cells for immune escape. Front Microbiol. Frontiers; 2018 May 7;9(MAY):883.
226. Srivastava S, Shankar P, Mishra J, Singh S, Alvar J, Velez I, et al. Possibilities and challenges for developing a successful vaccine for leishmaniasis. Parasit Vectors. BioMed Central; 2016 Dec 12;9(1):277.
227. Mandell MA, Beverley SM. Continual renewal and replication of persistent Leishmania major parasites in concomitantly immune hosts. Proc Natl Acad Sci U S A. National Academy of

Sciences; 2017 Jan 31;114(5):E801–10.

228. Glennie ND, Volk SW, Scott P. Skin-resident CD4+T cells protect against Leishmania major by recruiting and activating inflammatory monocytes. PLoS Pathog. 2017;13(4).
229. Product Development for Vaccines Advisory Committee W. Status of Vaccine Research and Development of Vaccines for Leishmaniasis. 2011;1–4.
230. Gillespie PM, Beaumier CM, Strych U, Hayward T, Hotez PJ, Bottazzi ME. Status of vaccine research and development of vaccines for leishmaniasis. Vaccine. 2016;34(26):2992–5.
231. Khamesipour A. Leishmanization: Use of an old method for evaluation of candidate vaccines against leishmaniasis. 2005;
232. Noazin S, Modabber F, Khamesipour A, Smith PG, Moulton LH, Nasseri K, et al. First generation leishmaniasis vaccines: A review of field efficacy trials. Vaccine. Elsevier; 2008 Dec 9;26(52):6759–67.
233. Handman E. Leishmaniasis: current status of vaccine development. Clin Microbiol Rev. American Society for Microbiology Journals; 2001 Apr 1;14(2):229–43.
234. Dunning N. Leishmania vaccines: from leishmanization to the era of DNA technology. Biosci Horizons. Oxford University Press; 2009 Feb 17;2(1):73–82.
235. Streit JA, Recker TJ, Filho FG, Beverley SM, Wilson ME. Protective Immunity Against the Protozoan Leishmania chagasi Is Induced by Subclinical Cutaneous Infection with Virulent But Not Avirulent Organisms. J Immunol. American Association of Immunologists; 2001 Feb 1;166(3):1921–9.
236. Cecílio P, Oliveira F, da Silva AC. Vaccines for Human Leishmaniasis: Where Do We Stand and What Is Still Missing? In: Leishmaniases as Re-emerging Diseases. InTech; 2018.
237. Khalil EAG, Zijlstra EE, Kager PA, El Hassan AM. Epidemiology and clinical manifestations of Leishmania donovani infection in two villages in an endemic area in eastern Sudan. Trop Med Int Heal. John Wiley & Sons, Ltd (10.1111); 2002 Jan 1;7(1):35–44.
238. McCall LI, Zhang WW, Dejgaard K, Atayde VD, Mazur A, Ranasinghe S, et al. Adaptation of leishmania donovani to cutaneous and visceral environments: In vivo selection and proteomic analysis. J Proteome Res. 2015;14(2):1033–59.
239. Reed S, Coler R, Mondal D, Kamhawi S, Valenzuela J. Leishmania vaccine development: exploiting the host–vector–parasite interface. Expert Rev Vaccines. 2015;0584(November 2015):1–10.
240. Duthie MS, Van Hoeven N, MacMillen Z, Picone A, Mohamath R, Erasmus J, et al. Heterologous Immunization With Defined RNA and Subunit Vaccines Enhances T Cell Responses That

Protect Against Leishmania donovani. Front Immunol. Frontiers; 2018 Oct 17;9:2420.

241. Stockdale L, Newton R. A Review of Preventative Methods against Human Leishmaniasis Infection. Büscher P, editor. PLoS Negl Trop Dis. Public Library of Science; 2013 Jun 20;7(6):e2278.
242. Mendonça SCF. Differences in immune responses against Leishmania induced by infection and by immunization with killed parasite antigen: Implications for vaccine discovery. Parasites and Vectors. 2016 Dec 6;9(1):492.
243. Plotkin SA. Correlates of Protection Induced by Vaccination. Clin Vaccine Immunol. American Society for Microbiology (ASM); 2010;17(7):1055.
244. Röllinghoff M, Bogdan C, Gessner A, Lohoff M. Immunity to protozoa. In: Encyclopedia of Life Sciences. 2001.
245. Müller I, Kropf P, Louis JA, Milon G. Expansion of gamma interferon-producing CD8+ T cells following secondary infection of mice immune to Leishmania major. Infect Immun. American Society for Microbiology Journals; 1994 Jun 1;62(6):2575–81.
246. Brodskyn CI, Barral A, Boaventura V, Carvalho E, Barral-Netto M. Parasite-driven in vitro human lymphocyte cytotoxicity against autologous infected macrophages from mucosal leishmaniasis. J Immunol. American Association of Immunologists; 1997 Nov 1;159(9):4467–73.
247. Novais FO, Carvalho LP, Graff JW, Beiting DP, Ruthel G, Roos DS, et al. Cytotoxic T Cells Mediate Pathology and Metastasis in Cutaneous Leishmaniasis. PLoS Pathog. 2013;9(7).
248. Gidwani K, Picado A, Ostyn B, Singh SP, Kumar R, Khanal B, et al. Persistence of Leishmania donovani antibodies in past visceral leishmaniasis cases in India. Clin Vaccine Immunol. American Society for Microbiology (ASM); 2011 Feb;18(2):346–8.
249. Buxbaum LU. A detrimental role for IgG and FcgammaR in Leishmania mexicana infection. Immunol Res. Humana Press Inc; 2008 Oct 11;42(1–3):197–209.
250. Ajdary S, Alimohammadian MH, Eslami MB, Kemp K, Kharazmi A. Comparison of the immune profile of nonhealing cutaneous leishmaniasis patients with those with active lesions and those who have recovered from infection. Infect Immun. 2000;68(4):1760–4.
251. Griffiths KL, Khader SA. Novel vaccine approaches for protection against intracellular pathogens. Curr Opin Immunol. Elsevier Ltd; 2014;28(1):58–63.
252. Lewinsohn DA, Lewinsohn DM, Scriba TJ. Polyfunctional CD4+T cells as targets for tuberculosis vaccination. Vol. 8, Frontiers in Immunology. Frontiers; 2017. p. 1262.
253. Morrow WJW, Sheikh NA, Schmidt CS, Davies DH. Vaccinology: Principles and Practice.

Vaccinology: Principles and Practice. Wiley-Blackwell; 2012.

254. Boyd A, Almeida JR, Darrah PA, Sauce D, Seder RA, Appay V, et al. Pathogen-specific T cell polyfunctionality is a correlate of T cell efficacy and immune protection. PLoS One. 2015;10(6).

255. Goulder PJR, Watkins DI. Impact of MHC class I diversity on immune control of immunodeficiency virus replication. Nat Rev Immunol. Nature Publishing Group; 2008 Aug 1;8(8):619–30.

256. Darrah PA, Patel DT, De Luca PM, Lindsay RWB, Davey DF, Flynn BJ, et al. Multifunctional TH1 cells define a correlate of vaccine-mediated protection against Leishmania major. Nat Med. Nature Publishing Group; 2007 Jul 10;13(7):843–50.

257. Iezzi G, Karjalainen K, Lanzavecchia A. The duration of antigenic stimulation determines the fate of naive and effector T cells. Immunity. Cell Press; 1998 Jan 1;8(1):89–95.

258. O'Garra A. Cytokines Induce the Development of Functionally Heterogeneous T Helper Cell Subsets. Immunity. Cell Press; 1998 Mar 1;8(3):275–83.

259. Kwok WW, Tan V, Gillette L, Littell CT, Soltis MA, LaFond RB, et al. Frequency of epitope-specific naive CD4(+) T cells correlates with immunodominance in the human memory repertoire. J Immunol. American Association of Immunologists; 2012 Mar 15;188(6):2537–44.

260. Moon JJ, Chu HH, Pepper M, McSorley SJ, Jameson SC, Kedl RM, et al. Naive CD4+ T Cell Frequency Varies for Different Epitopes and Predicts Repertoire Diversity and Response Magnitude. Immunity. Cell Press; 2007 Aug 24;27(2):203–13.

261. Besse MA, Levy Y, Schwartz O, Casartelli BN, Grygar C, Schneider A, et al. Hierarchy of CD4 T Cell Epitopes of the ANRS Lipo5 Synthetic Vaccine Relies on the Frequencies of Pre-Existing Peptide-Specific T Cells in Healthy Donors. J Immunol. 2013;190:5757–63.

262. Taslimi Y, Zahedifard F, Rafati S. Leishmaniasis and various immunotherapeutic approaches. 2016;

263. Singh B, Sundar S. Leishmaniasis: Vaccine candidates and perspectives. Vaccine. Elsevier Ltd; 2012;30(26):3834–42.

264. WHO. Control of the leishmaniases. World Health Organization technical report series. 2010.

265. Coler RN, Goto Y, Bogatzki L, Raman V, Reed SG. Leish-111f, a recombinant polyprotein vaccine that protects against visceral leishmaniasis by elicitation of CD4+T cells. Infect Immun. 2007 Sep 1;75(9):4648–54.

266. Gradoni L, Foglia Manzillo V, Pagano A, Piantedosi D, De Luna R, Gramiccia M, et al. Failure of

a multi-subunit recombinant leishmanial vaccine (MML) to protect dogs from Leishmania infantum infection and to prevent disease progression in infected animals. Vaccine. Elsevier; 2005 Nov 1;23(45):5245–51.

267. Nascimento E, Fernandes DF, Vieira EP, Campos-Neto A, Ashman JA, Alves FP, et al. A clinical trial to evaluate the safety and immunogenicity of the LEISH-F1 + MPL-SE vaccine when used in combination with meglumine antimoniate for the treatment of cutaneous leishmaniasis. Vaccine. Elsevier; 2010 Sep 14;28(40):6581–7.
268. Bertholet S, Goto Y, Carter L, Bhatia A, Howard RF, Carter D, et al. Optimized subunit vaccine protects against experimental leishmaniasis. Vaccine. Elsevier; 2009 Nov 23;27(50):7036–45.
269. Coler RN, Duthie MS, Hofmeyer KA, Guderian J, Jayashankar L, Vergara J, et al. From mouse to man: safety, immunogenicity and efficacy of a candidate leishmaniasis vaccine LEISH-F3+GLA-SE. Clin Transl Immunol. Wiley-Blackwell; 2015 Apr;4(4):e35.
270. Didwania N, Shadab M, Sabur A, Ali N. Alternative to Chemotherapy—The Unmet Demand against Leishmaniasis. Front Immunol. Frontiers; 2017 Dec 21;8:1779.
271. Duthie MS, Pereira L, Favila M, Hofmeyer KA, Reed SJ, Metangmo S, et al. A defined subunit vaccine that protects against vector-borne visceral leishmaniasis. npj Vaccines. 2017;2(1):23.
272. Kreutzer S. Nursing Body and Soul in the Parish: Lutheran Deaconess Motherhouses in Germany and the United States. Nurs Hist Rev. 2010;18(1):134–50.
273. Osman M, Mistry A, Keding A, Gabe R, Cook E, Forrester S, et al. A third generation vaccine for human visceral leishmaniasis and post kala azar dermal leishmaniasis: First-in-human trial of ChAd63-KH. McDowell MA, editor. PLoS Negl Trop Dis. Public Library of Science; 2017 May 12;11(5):e0005527.
274. Cecílio P, Pérez-Cabezas B, Fernández L, Moreno J, Carrillo E, Requena JM, et al. Pre-clinical antigenicity studies of an innovative multivalent vaccine for human visceral leishmaniasis. Louzir H, editor. PLoS Negl Trop Dis. Public Library of Science; 2017 Nov 27;11(11):e0005951.
275. Carrillo E, Fernandez L, Ibarra-Meneses AV, Santos MLB, Nico D, de Luca PM, et al. F1 Domain of the Leishmania (Leishmania) donovani Nucleoside Hydrolase Promotes a Th1 Response in Leishmania (Leishmania) infantum Cured Patients and in Asymptomatic Individuals Living in an Endemic Area of Leishmaniasis. Front Immunol. 2017;8:750.
276. Santos MLB, Nico D, de Oliveira FA, Barreto AS, Palatnik-de-Sousa I, Carrillo E, et al. Leishmania donovani nucleoside Hydrolase (NH36) domains induce T-cell cytokine responses in human visceral leishmaniasis. Front Immunol. Frontiers; 2017 Mar

7;8(MAR):227.

277. Goto Y, Bhatia A, Raman VS, Liang H, Mohamath R, Picone AF, et al. KSAC, the first defined polyprotein vaccine candidate for visceral leishmaniasis. Clin Vaccine Immunol. American Society for Microbiology; 2011 Jul 1;18(7):1118–24.
278. Martins VT, Lage DP, Duarte MC, Carvalho AMRS, Costa LE, Mendes TAO, et al. A recombinant fusion protein displaying murine and human MHC class I- and II-specific epitopes protects against Leishmania amazonensis infection. Cell Immunol. 2016;
279. Lambertz U, Silverman JM, Nandan D, McMaster WR, Clos J, Foster LJ, et al. Secreted virulence factors and immune evasion in visceral leishmaniasis. J Leukoc Biol. 2012;91(6):887–99.
280. Hassani K, Antoniak E, Jardim A, Olivier M. Temperature-Induced Protein Secretion by Leishmania mexicana Modulates Macrophage Signalling and Function. PLoS One. 2011;6(5):e18724.
281. Rosa R, Roos Rodrigues O, Marques C, Santos-Gomes GM. Leishmania infantum: Soluble proteins released by the parasite exert differential effects on host immune response. Exp Parasitol. 2005;109(2):106–14.
282. Chamakh-Ayari R, Bras-Gonçalves R, Bahi-Jaber N, Petitdidier E, Markikou-Ouni W, Aoun K, et al. In vitro evaluation of a soluble Leishmania promastigote surface antigen as a potential vaccine candidate against human leishmaniasis. PLoS One. 2014;9(5):1–12.
283. Gour JK, Kumar V, Singh N, Bajpai S, Pandey HP, Singh RK. Identification of Th1-responsive leishmanial excretory–secretory antigens (LESAs). Exp Parasitol. Elsevier Inc.; 2012;132(3):355–61.
284. Lemesre J-L, Holzmuller P, Cavaleyra M, Gonçalves RB, Hottin G, Papierok G. Protection against experimental visceral leishmaniasis infection in dogs immunized with purified excreted secreted antigens of Leishmania infantum promastigotes. Vaccine. 2005;23(22):2825–40.
285. Tonui WK, Mejia JS, Hochberg L, Mbow ML, Ryan JR, Chan AST, et al. Immunization with Leishmania major exogenous antigens protects susceptible BALB/c mice against challenge infection with L. major. Infect Immun. 2004;72(10):5654–61.
286. Lemesre JL, Holzmuller P, Gonçalves RB, Bourdoiseau G, Hugnet C, Cavaleyra M, et al. Long-lasting protection against canine visceral leishmaniasis using the LiESAp-MDP vaccine in endemic areas of France: Double-blind randomised efficacy field trial. Vaccine. 2007;25(21):4223–34.
287. Oliva G, Nieto J, Foglia Manzillo V, Cappiello S, Fiorentino E, Di Muccio T, et al. A randomised,

double-blind, controlled efficacy trial of the LiESP/QA-21 vaccine in naïve dogs exposed to two leishmania infantum transmission seasons. PLoS Negl Trop Dis. Public Library of Science; 2014 Oct;8(10):e3213.

288. Lladró S, Picado A, Ballart C, Portús M, Gállego M. Management, prevention and treatment of canine leishmaniosis in north-eastern Spain: An online questionnaire-based survey in the province of Girona with special emphasis on new preventive methods (CaniLeish vaccine and domperidone). Vet Rec. 2017 Jan 14;180(2):47.
289. Bras-Gonçalves R, Petitdidier E, Pagniez J, Veyrier R, Cibrelus P, Cavaleyra M, et al. Identification and characterization of new Leishmania promastigote surface antigens, LaPSA-38S and LiPSA-50S, as major immunodominant excreted/secreted components of L. amazonensis and L. infantum. Infect Genet Evol. Elsevier; 2014 Jun 1;24:1–14.
290. Petitdidier E, Pagniez J, Papierok G, Vincendeau P, Lemesre J-L, Bras-Gonçalves R. Recombinant Forms of Leishmania amazonensis Excreted/Secreted Promastigote Surface Antigen (PSA) Induce Protective Immune Responses in Dogs. Louzir H, editor. PLoS Negl Trop Dis. World Intellectual Property Organization; 2016 May 25;10(5):e0004614.
291. Series W technical R. Synthetic Peptide Vaccines.Pdf. 1999.
292. Skwarczynski M, Toth I. Peptide-based synthetic vaccines. Chem Sci. Royal Society of Chemistry; 2016 Jan 26;7(2):842–54.
293. De Brito RCF, Cardoso JM d. O, Reis LES, Vieira JF, Mathias FAS, Roatt BM, et al. Peptide vaccines for leishmaniasis. Vol. 9, Frontiers in Immunology. Frontiers; 2018. p. 1043.
294. Pietersz GA, Pouniotis DS, Apostolopoulos V. Design of peptide-based vaccines for cancer. Curr Med Chem. 2006;13(14):1591–607.
295. De Brito RCF, Cardoso JMDO, Reis LES, Vieira JF, Mathias FAS, Roatt BM, et al. Peptide Vaccines for Leishmaniasis. Front Immunol. Frontiers; 2018 May 11;9:1043.
296. Hickling J, Jones K, Friede M, Zehrung D, Chen D, Kristensen D. Intradermal delivery of vaccines: potential benefits and current challenges. Bull World Health Organ. World Health Organization; 2011 Mar 1;89(3):221–6.
297. Kumar R, Engwerda C. Vaccines to prevent leishmaniasis. Clin Transl Immunol. Nature Publishing Group; 2014;3(3):e13.
298. Costa CHN, Peters NC, Maruyama SR, de Brito EC, de Miranda Santos IKF, Santos IKF de M. Vaccines for the Leishmaniases: Proposals for a Research Agenda. Louzir H, editor. PLoS Negl Trop Dis. Public Library of Science; 2011 Mar 29;5(3):e943.
299. Russo DM, Jardim A, Carvalho EM, Sleath PR, Armitage RJ, Olafson RW, et al. Mapping human

T cell epitopes in leishmania gp63 - Identification of cross-reactive and species-specific epitopes. J Immunol. 1993;150:932–9.

300. Jardim A, Alexander J, Teh HS, Ou D, Olafson RW. Immunoprotective Leishmania major synthetic T cell epitopes. J Exp Med. Rockefeller University Press; 1990 Aug 1;172(2):645–8.
301. Rothbard JB, Taylor WR. A sequence pattern common to T cell epitopes. EMBO J. John Wiley & Sons, Ltd; 1988 Jan 1;7(1):93–100.
302. Delgado G, Vargas LE, Torres A, Alonso C, Spinel C, Patarroyo ME. Characterizing cellular immune response to kinetoplastid membrane protein-11 (KMP-11) during. Parasite Immunol. 2003;(26):199–209.
303. Basu R, Roy S, Walden P. HLA Class I–Restricted T Cell Epitopes of the Kinetoplastid Membrane Protein–11 Presented by *Leishmania donovani* –Infected Human Macrophages. J Infect Dis. Oxford University Press; 2007 May 1;195(9):1373–80.
304. Seyed N, Zahedifard F, Safaiyan S, Gholami E, Doustdari F, Azadmanesh K, et al. In silico analysis of six known leishmania major antigens and in vitro evaluation of specific epitopes eliciting HLA-A2 restricted CD8 T cell response. Louzir H, editor. PLoS Negl Trop Dis. 2011;5(9):e1295.
305. Elfaki MEE, Khalil EAG, De Groot AS, Musa AM, Gutierrez A, Younis BM, et al. Immunogenicity and immune modulatory effects of in silico predicted L. donovani candidate peptide vaccines. Hum Vaccines Immunother. 2012;8(12):1769–74.
306. Naouar I, Boussoffara T, Chenik M, Gritli S, Ben Ahmed M, Belhadj Hmida N, et al. Prediction of T Cell Epitopes from Leishmania major Potentially Excreted/Secreted Proteins Inducing Granzyme B Production. PLoS One. Public Library of Science; Jan;11(1):e0147076.
307. de Freitas e Silva R, Ferreira LFGR, Hernandes MZ, de Brito MEF, de Oliveira BC, da Silva AA, et al. Combination of in silico methods in the search for potential CD4+ and CD8+ T cell epitopes in the proteome of Leishmania braziliensis. Front Immunol. 2016;7(327).
308. Mahantesh V, Amit A, Dikhit MR, Singh AK, Venkateshwaran T, Das VNR, et al. Immuno-informatics based approaches to identify CD8+ T cell epitopes within the Leishmania donovani 3-ectonucleotidase in cured visceral leishmaniasis subjects. Microbes Infect. 2017;19(6):358–69.

CHAPTER II

THE LEISHMANIA SECRETOME

1. Introduction

In summary, the principal *Leishmania* species causing cutaneous leishmaniasis (CL) in the Mediterranean Basin, Africa, Middle East, and India are *L. (L.) major* (zoonotic CL, ZCL), *L. (L.) tropica* (anthroponotic CL); whereas in Central and South America, the main CL-causing species are *L. (L.) mexicana*, *L. (L.) amazonensis*, and *L. (V.) braziliensis* (1). VL is caused by *L. (L.) donovani* parasites in the Indian and African regions, and *L. (L.) infantum* parasites in the Mediterranean Basin, Central and South America and China (1–3). Other *Leishmania* species infect non-human vertebrate hosts, as is the case of *L. tarentolae* whose natural host is the lizard (4). For pathogenesis and epidemiological aspects see Chapter I – General Introduction.

The term "co-evolution" was actually used for the first time to describe the relationship between *Leishmania* parasites and host sand fly species (5). Environmental cues lead to the differentiation from procyclic promastigote to metacyclic promastigote inside the sand fly, and to intracellular amastigote in the vertebrate host (temperature, pH, exposure to reactive oxygen and reactive nitrogen species, extracellular proteolytic activity, or nutritional requirements) (6,7). *Leishmania* parasites have co-evolved with vertebrate hosts since the Jurassic period, and have adapted to diverse epidemiological scenarios divergently for 20-100 million years, and yet still share very high levels of genomic homology and synteny (6).

The first full genome sequence of a *Leishmania* parasite was published in 2005 (8) marking the beginning of the post-genomic era for these parasites. *L. major* parasites have a 32,8 megabase haploid genome, containing over 8200 protein-coding genes, and over 94% of all mRNA molecules are constitutively expressed between promastigotes and amastigotes (9,10). Since then, three other *Leishmania* species genomes have been sequenced and annotated – *L. (L.) infantum* (11), *L. (V.) braziliensis* (11), and *L. (L.) mexicana* (12). Comparative analysis showed a high level of genetic conservation among species (7392 common genes), but quite remarkable large-scale genetic differences in terms of ploidy and gene copy number (12).

The high level of conservancy observed among *Leishmania* genomes, with minor changes in mRNA levels between promastigotes and amastigotes, suggests that differentiation, virulence and pathogenesis depend on post-transcriptional and post-translational expression control mechanisms (6,9,13–18). Predicted genes are expressed as polycistronic units present on the same DNA strand, much alike bacterial operons. However, genes grouped in a common orientation are not regulated by common mechanisms nor are they functionally related (19,20). Regulation of single gene expression does not rely in RNA polymerase (RNApol) promoter control – control of gene expression is performed at multiple levels (*trans*-splicing, mRNA polyadenylation, mRNA stability, transcript

elongation, and protein stability) (19–21). Accordingly, very few RNApol promoters were identified, none for RNApol II, except for the splice leader RNA promoter and RNApol I promoter sites in the rDNA locus (10).

Post-translational modifications (PTM) are proposed to be a major mechanism in *Leishmania* control of gene expression (6,22,23). Over 200 different types of PTM have been generally described, and *Leishmania* proteins can be cleaved eliminating signal sequences, pro-peptides and initiator methionines, or cleaved by specific proteases to generate final functional protein forms (10). *Leishmania* protein modifications include methylation, acetylation, phosphorylation, hexosylation, fucosylation; the addition of complex molecules such as lipids, sugars or glycolipids; ubiquination, nytrosylation and hyperoxidation of sulfhydryl group (6,10,22).

Proteomic studies are therefore optimal tools to study *Leishmania* pathogenesis since gene expression is for the most part regulated post-transcriptionally (6). The first proteomic studies compared promastigote and amastigote forms and were performed for *L. (L.) infantum* (24), *L. (L.) donovani* (25), *L. (L.) mexicana* (26), and *L. (V.) panamensis* (27).

The main pathways upregulated in amastigotes are fatty acid oxidation and gluconeogenesis, mitochondrial respiration, proteins with basic pH, and some specific proteins such as tryparedoxin peroxidase, methylthioadenosine phosphatase, required for survival under oxidative conditions in the host cell (6,25,26,28). By contrast, *Leishmania* promastigotes rely mostly on glycolysis and amino-acid metabolism for energy generation and are also capable of hydrolysing dissacharides. In promastigotes, the glycolytic pathway is upregulated, and in metacyclic promastigotes the proteins PFR1D, alpha- and beta-tubulin, cysteine protease B, trypanoredoxin, GP63, and GP46 show higher expression levels (6,25,26). Other differentially expressed proteins are associated with diverse metabolic functions – cytoskeleton components, stress responses, amino acid and carbohydrate metabolism, detoxification and proteolysis (6,25,26). Likewise, some modest differences were found between cutaneous- and visceral-adapted *L. donovani* strains, the latter having increased translation, biosynthetic processes, antioxidant protection and signalling functions (29).

The specific activities of some glycolytic enzymes are slightly decreased (hexokinase, phosphofructokinase and glucose-6-phophate dehydrogenase) or greatly reduced (pyruvate kinase) in amastigotes. In the amastigote form, parasites use mostly fatty acids and amino-acid metabolism for energy generation, as access to glucose and other sugars is limited (6,25,30). The capacity for sugar uptake is also reduced in amastigotes in comparison to promastigote forms, so glycolytic and pentose-phosphate pathways are present and functional but with lower activity levels. Also, glucose degradation results in succinate production (instead of pyruvate) through the action of phosphoenolpyruvate carboxykinase/malate dehydrogenase enzymes (6,25,30).

Sugar residues required by *Leishmania* amastigotes are generated *de novo* through gluconeogenesis (6,30). Both metacyclic promastigotes and amastigotes accumulate β-mannan, a complex polyssacharide, as intracellular energy storage. However, contrary to promastigotes, this molecule is essential for amastigote survival and replication and is dependent on hexose uptake (6,30).

The term 'secretome' was firstly defined by Tjalsma et al 2000 as a subset of the proteome comprising both all secreted proteins and the pathways and machinery required for protein transport/secretion (31). There is increasing interest in the secretome and excreted-secreted proteins as they play a pivotal role in virulence and interaction with the host, in this case, for *Leishmania* establishment of infection and immunomodulation (32–40). These early secreted proteins allow *Leishmania* parasites to survive in an otherwise lethal environment, inside the phagolysosome, as well as to modulate host response by suppressing host cell signalling and macrophage activation, while deviating the host's immune response towards a permissive anti-inflammatory phenotype (32–40).

Proteomic analysis of the secretome of *L. (L.) donovani* promastigotes (41), *L. (L.) mexicana* promastigotes (42), *L. (V.) braziliensis* promastigotes (43), and *L. (L.) major* promastigotes inside the sand fly midgut (44) were performed, contributing to an increased knowledge about the function of these proteins. These studies show the main secretion pathway in *Leishmania* parasites is through exosome/micro-vesicle production, which are crucial for parasite virulence (37,40,45). Only a small proportion of proteins possesses a N-terminal secretion signal peptide or is non-classically (ER/Golgi-independently) secreted (46). Interestingly, amastigotes continue to secrete exosomes to the cytoplasm of infected and neighbouring cells (46). Exosomes are increasingly implicated in host-pathogen interactions and are determinant for the promotion or inhibition of host immunity (36,47).

A plethora of candidate virulence factors has been identified in the *Leishmania* secretome, meaning functional characterization is extremely difficult. It is generally accepted that the *Leishmania* pathogenesis depends on parasite-mediated early immune priming and continuous immunomodulation, as well as maintenance of the parasitophorous vacuole conditions (30,32,48,49). Molecules found in the secretome can be associated to any of these essential functions.

Furthermore, other than virulence factors necessary for parasite survival, many interesting molecules have been identified in the secretome, namely drug and vaccine candidates. As referred in Chapter I – General Introduction, *Leishmania* secreted antigens have been shown to trigger Th1 responses in several disease models (39,50–52) and several antigen candidates in the human vaccine

development pipeline are known to be secreted. However, the total content of the naturally secreted *Li*ESAp, the active principle of the CaniLeish® canine vaccine, remains largely undefined. The development of a second-generation vaccine based on antigens present in the *Leishmania* secretome requires the comprehensive characterisation of the antigenic proteins responsible for the observed immune responses. The present analysis aims to fill this knowledge gap, to identify all the proteins present the *L. infantum* secretome produced in aseric conditions. Furthermore, we provide the first simultaneous analysis of five other major pathogenic *Leishmania*.

The present chapter describes the identification of the total secretome of six *Leishmania* species, exclusive human pathogens, as well as *L. tarentolae*, non-pathogenic to humans. The secretome was prepared according to the method used to produce *Li*ESAp from promastigote cultures growing in aseric medium, used in the CaniLeish® vaccine.

2. Material and Methods

2.1. *Leishmania* parasite cultures in aseric medium, generation and purification of total excreted-secreted proteins

The *in vitro* culture of *Leishmania* promastigotes (Table II.1) was performed in a completely chemically defined medium (CDM/LP) free of serum, macromolecules, proteins and peptides as previously described (53,54). When parasite concentration reached 2–3×10^7 promastigotes per ml in a 6-days period, culture was centrifuged (2000×g, 20 min, 4 ◦C) to remove parasites. The supernatant was collected, filtered (0.2-µm-pore-size filter, Millipore) to eliminate removal promastigotes, concentrated approximately 100-fold and dialysed by ultrafiltration with a 3-kDa-cutoff filter unit (Pall). Protein concentration was determined according to Bradford method (Bio-Rad Laboratories). The purified naturally excreted secreted antigens from *Leishmania* promastigotes are designated as ESAp (LiESAps are used as the CaniLeish® vaccine antigen). *Leishmania* ESAps were stored at -80°C until use.

Table II. 1 Leishmania spp. strains cultured for the generation of ESAp.

***Leishmania* spp. strains cultured for the generation of ESAp**
Leishmania infantum MHOM/MA/67/ITMAP-263
Leishmania major MHOM/SY/91/LEM-2420
Leishmania tropica MHOM/SY/90/LEM-2067
Leishmania donovani MHOM/IN/80/DD8
Leishmania braziliensis MHOM/BR/75/M-2904
Leishmania amazonensis MHOM/BR/76/LTB-012
Leishmania tarentolae WT commercial strain

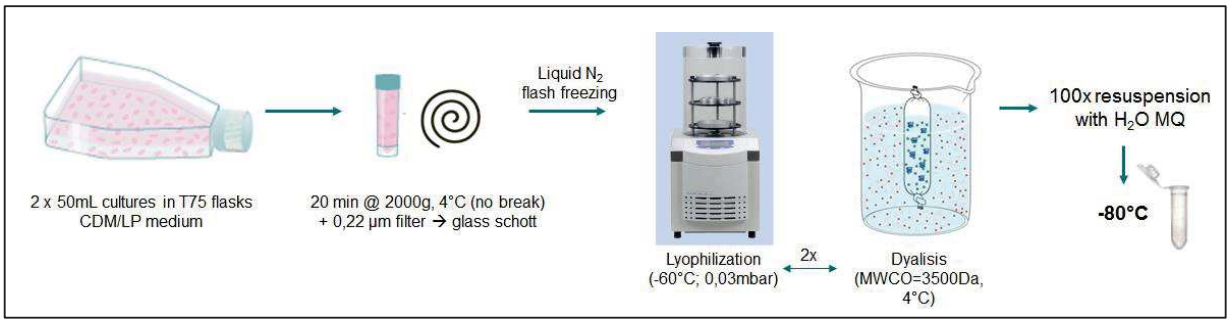


Figure II. 1 General workflow for the preparation of the Leishmania *promastigote secretome.*

2.2. Protein separation

Proteins were separated on 1D SDS-PAGE gels (12 % polyacrylamide, Mini-PROTEAN® TGX™ Precast Gels, Bio-Rad, Hercules USA). Gels were stained with PageBlue Protein Staining Solution (Fermentas) and scanned using a computer-assisted densitometer (Epson Perfection V750 PRO). Gel lanes were cut into 12 gel pieces (11 bands plus the well fraction) and destained with three washes in 50% acetonitrile and 50 mM triethylammonium bicarbonate buffer (TEABC). After protein reduction (with 10 mM dithiothreitol in 50mM TEABC at 56 °C for 45 min) and alkylation (55 mM iodoacetamide TEABC at room temperature for 30 min) proteins were digested in-gel using trypsin (800 ng/band, Gold, Promega, Madison USA) as previously described (1). Desalting and pre-concentration of samples were performed with a ZipTip C18. Digest products were dehydrated in a vacuum centrifuge and reduced to 3 µL.

2.3. High-performance liquid chromatography and MS measurements

Peptide samples were dehydrated in a vacuum centrifuge, solubilized in 3 µl of 0.1% formic acid-2% acetonitrile. Three µL were analyzed online by nano-flow HPLC-nanoelectrospray ionization using a LTQ Orbitrap XL mass spectrometer (LTQ Orbitrap XL, Thermo Fisher Scientific, San Jose, CA) coupled with an Ultimate 3000 HPLC (Dionex). Desalting and pre-concentration of samples were performed on-line on a Pepmap® precolumn (0.3 mm x 10 mm). A gradient consisting of 0-40% A in 30 min, 80% B in 15 min (A = 0.1% formic acid, 2% acetonitrile in water; B = 0.1 % formic acid in acetonitrile) at 300 nl/min was used to elute peptides from the capillary (0.075 mm x 150 mm) reverse-phase column (Pepmap®, Dionex). LC-MS/MS experiments comprised cycles of 5 events; an MS1 scan with orbitrap mass analysis at 60000 resolution followed by CID of the five most abundant precursors. Fragment ions generated by CID were detected at the linear trap. Normalized collision energy of 35 eV and activation time of 30 ms were used for CID. Spectra were acquired with the instrument operating in the information-dependent acquisition mode throughout the HPLC gradient. The mass scanning range was m/z 400-2000 and standard mass spectrometric conditions for experiments were: spray voltage, 1.9 to 2.4 kV; no sheath and auxiliary gas flow; heated capillary temperature, 200°C; capillary voltage, 40 V and tube lens, 120 V. For all full scan measurements with the Orbitrap detector, a lock-mass ion from ambient air (m/z 445.120024) was used as an internal calibrant as described (55).

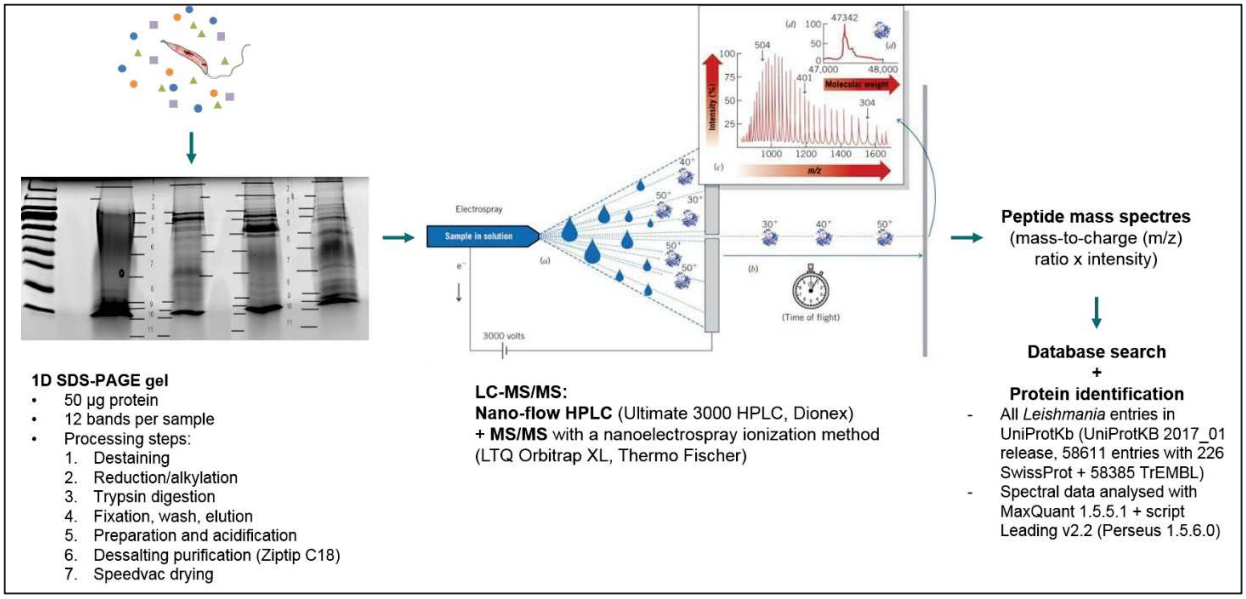


Figure II. 2 General workflow for the mass spectrometry analysis of the* Leishmania *promastigote secretome.

2.4. Bioinformatics analysis

Raw MS spectra were processed using the MaxQuant environment (56) and Andromeda for database search with iBAQ algorithm enabled (57). The MS/MS spectra were matched against the *Leishmania* entries of UniProtKB SwissProt and TrEMBL sections (release 2017_01; http://www.uniprot.org) and 250 frequently observed contaminants (MaxQuant contaminants database) as well as reversed sequences of all entries. The following settings were applied for database interrogation: mass tolerance of 7 ppm (MS) and 0.5 Th (MS/MS), trypsin/P enzyme specificity, up to two missed cleavages allowed for protease digestion, minimal peptide length at 7, cysteine carbamidomethylation as fixed modification and oxidation of methionine as variable modification. FDR was set at 0.01 for peptides and proteins.

A representative protein ID in each protein group was automatically selected using in-house bioinformatics tools (leading v2.2 and multi-species script), developed by Oana Vigy (*Plateforme de Proteomique Fonctionnelle de Montpellier*), and the Perseus software (version 1.5.3.0). First, proteins with the most numerous identified peptides are isolated in a "match group" (proteins from the "Protein IDs" column with the maximum number of "peptides counts (all)"). For the match groups where more than one protein ID is present after filtering (no specific tryptic peptides), the "leading" protein is firstly chosen as the best annotated protein in UniProtKB (reviewed SwissProt entries rather than automatic TrEMBL entries) corresponding to the 'Leading_AUTO' identification status. In case of ambiguous peptide identifications, proteins are identified by a given species and group-

species order preference, corresponding to a 'Leading_SPECIES' identification status (Table II.2). If ambiguous identifications cannot be distinguished with specific peptides or taxonomy, these remain identified as a match group ('MultiSpecies' identification status).

Table II. 2 Ordered species and group-species list established to define a "leading" protein based on the taxonomy criteria (only for match groups cases where more than one protein ID are present after filtering).

Taxonomy	Group Species Order
Leishmania infantum	**1**
Leishmania chagasi	2
Leishmania donovani (strain BPK282A1)	4
Leishmania donovani	**3**
Leishmania donovani donovani	4
Leishmania donovani archibaldi	4
Leishmania donovani complex sp. CR-2013	4
Leishmania major	**5**
Leishmania gerbilli	6
Leishmania turanica	6
Leishmania arabica	6
Leishmania tropica	**7**
Leishmania tropica complex sp. CR-2013	8
Leishmania aethiopica	8
Leishmania killicki	8
Leishmania mexicana (strain MHOM/GT/2001/U1103)	10
Leishmania mexicana	10
Leishmania amazonensis	**9**
Leishmania mexicana mexicana	10
Leishmania mexicana venezuelensis	10
Leishmania pifanoi	10
Leishmania braziliensis	**11**
Leishmania braziliensis complex EV-2015	12
Leishmania guyanensis	12
Leishmania peruviana	12
Leishmania panamensis	12
Leishmania shawi	12
Leishmania lainsoni	13
Leishmania naiffi	13
Leishmania utingensis	13
Leishmania lindenbergi	13
Leishmania sp. MHOM/BR/2002/NMT-RBO004	14
Leishmania sp. CR-2014	14
Leishmania sp.	14

2.5. Secretion pathway analysis (SecretomeP)

SecretomeP is a sequence-based method for the prediction of eukaryotic secreted proteins targeted to a non-classical secretory pathway, i.e. proteins without an N-terminal signal peptide (58). SecretomeP queries other feature prediction servers to obtain information on various post-translational and localisational aspects of the protein, which are integrated into the final secretion prediction. This method is also capable of predicting (signal peptide-containing) secretory proteins where only the mature part of the protein has been annotated, or cases where the signal peptide remains uncleaved.

The SecretomeP 1.0 standalone version was downloaded (http://www.cbs.dtu.dk/cgi-bin/nph-sw_request?secretomep) and fasta files with protein sequences corresponding to total identifications per sample were retrieved from UniProtKB. Each sample list was analysed with SecretomeP 1.0 (Signal-P and non-classical eukaryotic secretion predictions were performed). Proteins with NNscore > 0.5 were considered non-classically secreted.

2.6. Exosome marker analysis

The secretome datasets were screened for exosome cargo protein markers, from two different sources: i) ExoCarta, an exosome database, which includes the contents that were identified in exosome cargo in multiple organisms (59); and ii) Silverman et al 2008 (41) describes 26 *Leishmania*-secreted proteins associated with exosome-like and glycosomal vesicles.

To find the *Leishmania* genes ortholog to the most common exosome markers described by the ExoCarta database, a search strategy in OrthoMCL database was performed (Figure II.3). Briefly, 100 exosome markers (GenBank gene accessions) were downloaded from the Exocarta database (59). Firstly, a text search ('Identify Sequences based on text terms') was performed at OrthoMCL database web server with the 100 Gene Symbols. This search retrieved 3631 sequences (step 1), 83 of which correspond to human-specific proteins (step 2). These sequences were transformed to groups (step 3) to compare with four *Leishmania* phyletic ortholog groups in the database (step 4). These *Leishmania* phyletic groups were then retransformed into sequences (step 5), and 72 sequences from *L. infantum*, *L. mexicana*, *L. major* and *L. braziliensis* were retrieved (step 6) (Figure II.3 panel A). In parallel, 2140 accessions identified in the secretome proteomic datasets were converted to GeneDB IDs in the UNIProtKB database web server. 1860 out of 2140 identifiers from UniProtKB accessions identified in the secretome were successfully mapped to 1988 GeneDB Ids. This list was then

compared (list intersection) with the 72 *Leishmania* ortholog exosome marker genes in TriTrypDB (Figure II.3 panel B).

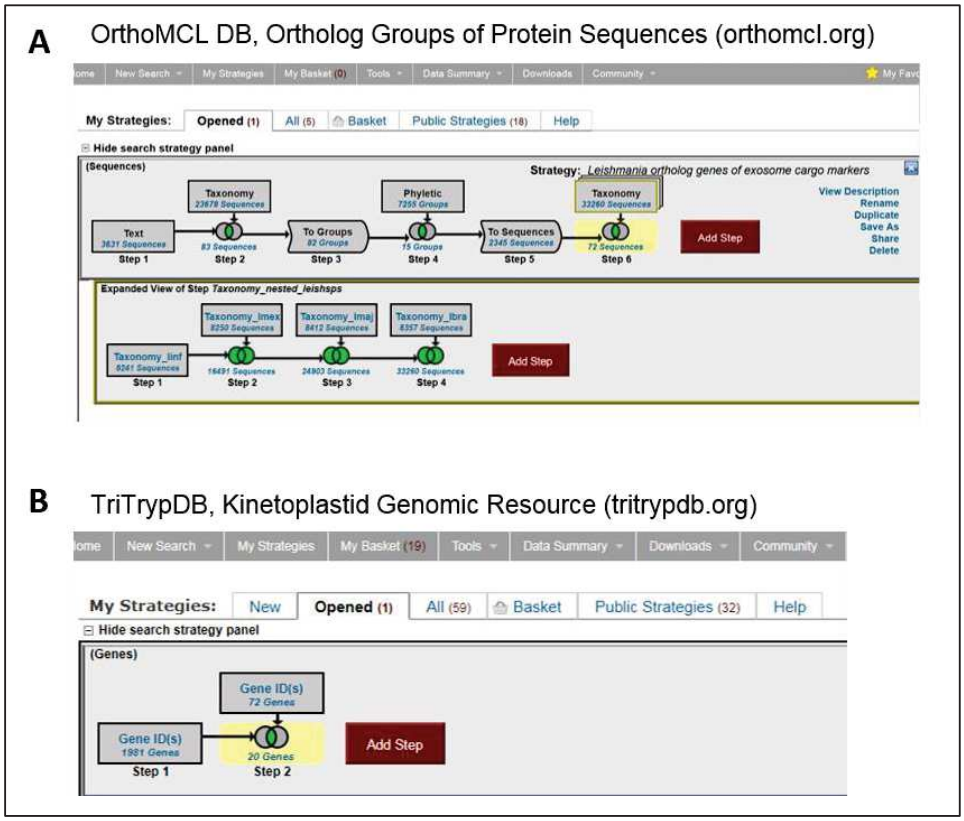


Figure II. 3 Exosome marker analysis of 100 exosome protein markers from ExoCarta. *A) OrthoMCL search strategy for* Leishmania *ortholog genes of exosome cargo markers from ExoCarta. B) Intersection between 72* Leishmania *ortholog genes and 1981 GeneDB IDs identified in the secretome proteomic analysis.*

3. Results

3.1. Proteins found in the secretome correspond to 12-17% of total *Leishmania* proteins

The method used for the preparation of the secretome is analogous to the large-scale production of *Li*ESAps for the CaniLeish® vaccine, which allows the production and analysis of naturally excreted-secreted proteins. *Leishmania* promastigotes are cultured in a completely defined aseric medium until the growth stationary phase (day 6).

To identify the proteins present in the *Leishmania* secretome, 11 bands (plus the well fraction) from the 1D gel were carefully isolated and analysed per sample, covering all proteins per lane (Figure II.4 panel A). A total of 2140 unique proteins were identified (UniProtKB database release 2017_01), between 909 and 1357 total accessions per sample (Figure II.4 Panel B), considering all identification categories defined by the bioinformatics identification script. Amongst the identified proteins, only a portion correspond to characterised proteins with SwissProt annotation (2.8%). Remaining proteins correspond to transcripts with an *in silico* characterisation in the TrEMBL database.

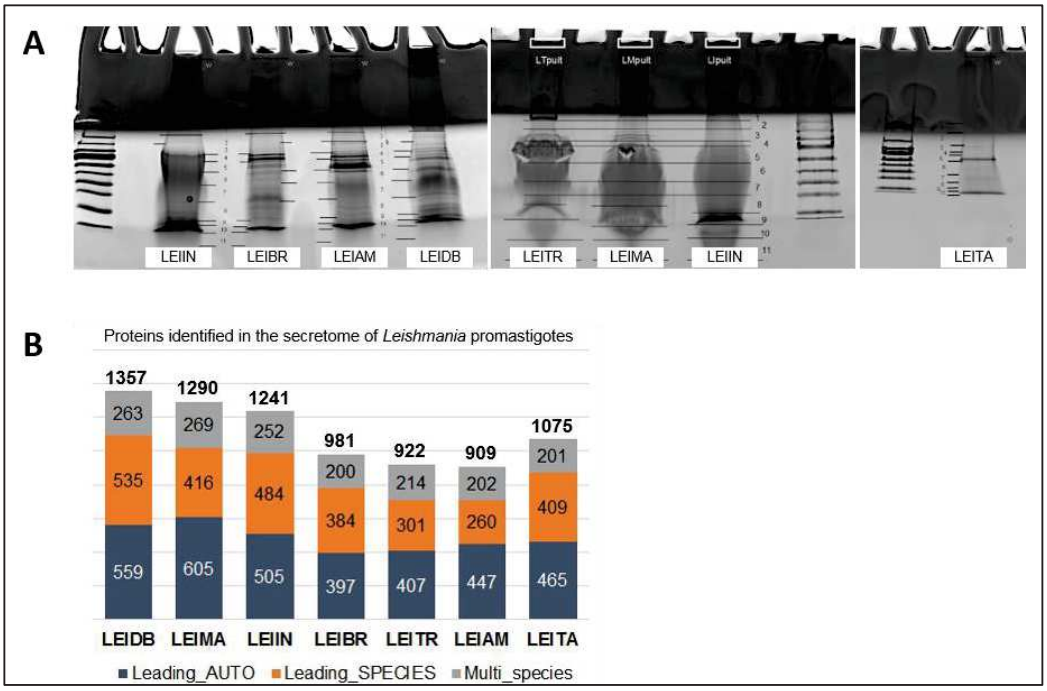


Figure II. 4* Leishmania *secretome protein identifications. *A) Synthetic images of one-dimensional gel electrophoresis of total secretome of different* Leishmania *species. B) Total number of protein accessions identified. Total number of accessions identified per sample, according to the bioinformatic script: 'Leading_AUTO' (accessions unique to the tested species), 'Leading_SPECIES' (ambiguous identifications assigned according to tested species), and 'Multi_species' (ambiguous identifications, impossible to distinguish between species).*

All six human pathogenic species analysed share a high number of common accessions (326), corresponding to 24% up to 35.9% of total identifications per sample (Figure II.5 panel A). The number of commonly identified proteins decreases to 306 accessions when considering all seven *Leishmania* species, meaning 20 proteins were not identified in the *L. tarentolae* sample.

The 20 accessions identified only in human pathogenic species are: beta-fructofuranosidase-like protein (EC 3.2.1.26); chaperonin HSP60, mitochondrial; cystathionine beta-synthase (EC 4.2.1.22); putative cytochrome c; GDP-mannose pyrophosphorylase (EC 2.7.7.22); GDPMP protein (EC 2.7.7.13); mannose-1-phosphate guanyltransferase (EC 2.7.7.13); paraflagellar_rod_component_-_putative; putative 2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60); putative ribosomal protein S6; putative small GTP-binding protein Rab1; S-methyl-5'-thioadenosine phosphorylase (EC 2.4.2.28) (MTA phosphorylase); superoxide dismutase (EC 1.15.1.1); three uncharacterised proteins (A4HIL9; E9AZC3; E9B159) and one hypothetical conserved protein (A4I5W4).

As expected, the *L. braziliensis* and *L. major* samples, two species with reference proteomes in UniProtKb (LEIBR and LEIMA), have the highest number of specific proteins (Figure II.2 panel B). Interestingly, LEIAM and LEITA, although underrepresented in the databases, show a high number of specific proteins, respectively, 66 and 63 accessions. Also, shared proteins between LEIIN and LEIDO samples are more numerous than LEIIN-specific proteins (Figure II.2 panel B).

Considering the reference proteomes available in UniprotKB, an average total of 8034 proteins are encoded in the *Leishmania* genome. The proteins identified in the secretome correspond to 12% up to 17% of total proteins (Table II.3).

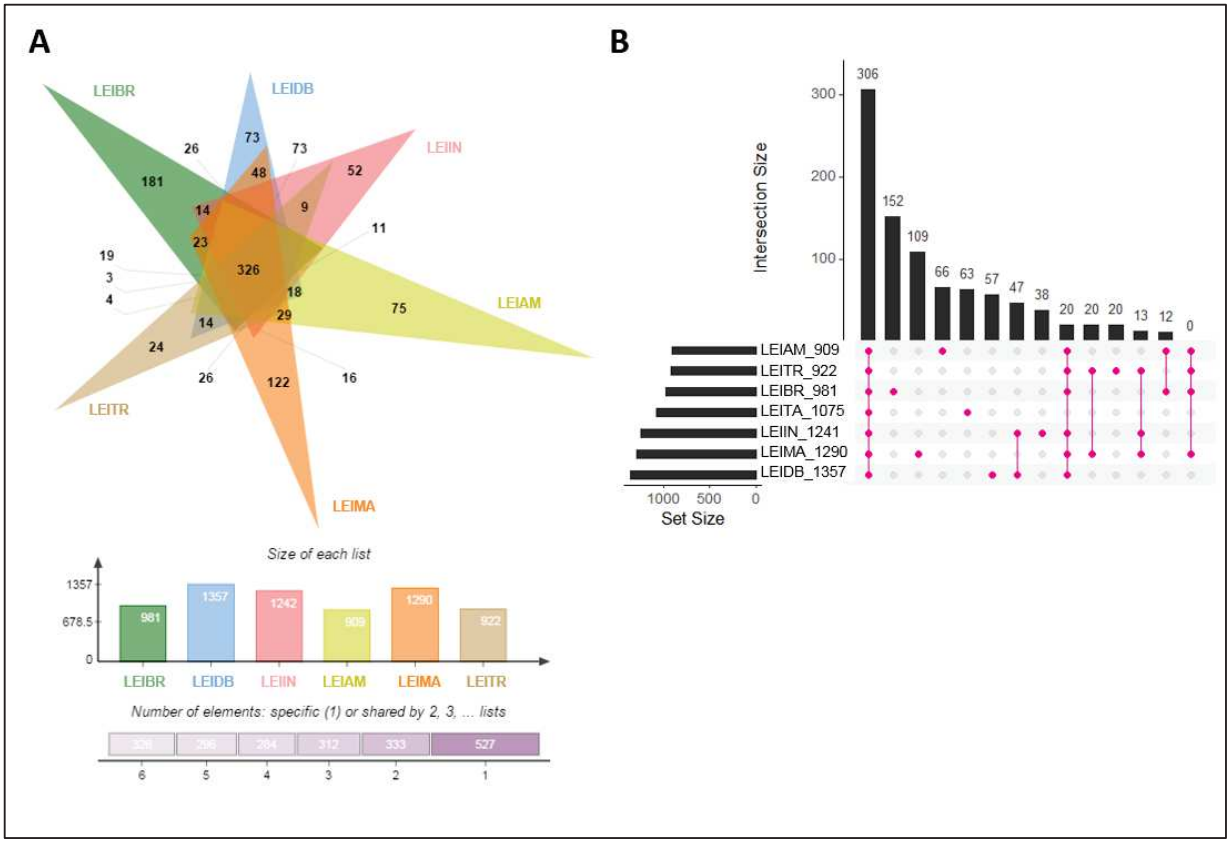


Figure II. 5 Comparison of protein identification lists. *A) Venn diagram comparing proteins identified in six human pathogenic species. B) Intersection lists between all* Leishmania *species, individual unique lists, and relevant groups: human pathogens (all except* L. tarentolae*), VL-causing species (LEIIN+LEIDB), Old World CL-causing species (LEIMA+LEITR), Mediterranean species (LEITR+LEIIN+LEIMA), New World CL-causing species (LEIAM+LEIBR), CL-causing species (LEIAM+LEITR+LEIBR+LEIMA).*

Table II. 3* Leishmania *reference proteomes, respective protein counts, and proteins found in the secretome.

Proteome ID	Organism	Organism ID	Protein count	Secretome protein counts	% of total proteins
UP000000542	*Leishmania major* (Strain: MHOM/IL/81/Friedlin)	5664	8038	1290	16,1
UP000008153	*Leishmania infantum* (Strain: JPCM5)	5671	8045	1241	15,4
UP000007258	*Leishmania braziliensis* (Strain: MHOM/BR/75/M2904)	5660	8084	981	12,1
UP000008980	*Leishmania donovani* (strain BPK282A1)	981087	7960	1357	17,1

3.2. The majority of *Leishmania* excreted-secreted proteins are non-classically secreted

A bioinformatics analysis of the secretome protein lists was performed with SecretomeP to identify proteins containing secretion signal peptides and proteins which are non-classically secreted (also known as leaderless secretion, unconventional or non-conventional secretory pathway) (58).

The consensus regarding protein secretion in eukaryotes is that membrane-associated or secreted proteins reach their final localisation via the classical secretory pathway. Proteins contain a signal peptide and/or transmembrane domain which targets them to the ER, from where they exit in vesicles, reaching the Golgi and eventually the membrane. Yet, there are extracellularly active proteins which do not enter the ER/Golgi pathway (60).

Results show that only a few proteins identified in the *Leishmania* secretome contain a classical secretion signal peptide (from 4.8% up to 12.9%). These results agree with previous proteomic studies which have shown as well that classical secretion is not the main secretion pathway in *Leishmania* parasites (33,46).

Over 42% of proteins are predicted to be secreted through non-classical pathways for all *Leishmania* species analysed. 'Leaderless' proteins are extracellularly active despite not having a signal peptide or a transmembrane domain. So far, four non-conventional mechanisms have been described in eukaryotes: i) pore-mediated translocation (type I); ii) ABC transporter-based secretion (type II or membrane flip-flop) of acylated proteins; iii) organelle-based translocation or autophagosome/endosome-based secretion (type III) by diverting endosomal and other membrane-bound compartments from their normal function to become secretory (60). The fourth unconventional protein secretion mechanism occurs when proteins containing signal peptides or transmembrane domains bypass the Golgi and are still secreted (type IV) (60). Interestingly, the type II membrane flip-flop mechanism (membrane translocation through N-terminus dual acylation) although not yet well-studied, has been demonstrated for *Leishmania* protein HASPB (61).

Importantly, the proteins secreted via non-classical secretion are predicted to be 3 to 9 times more frequent than proteins containing signal peptides for classical secretion.

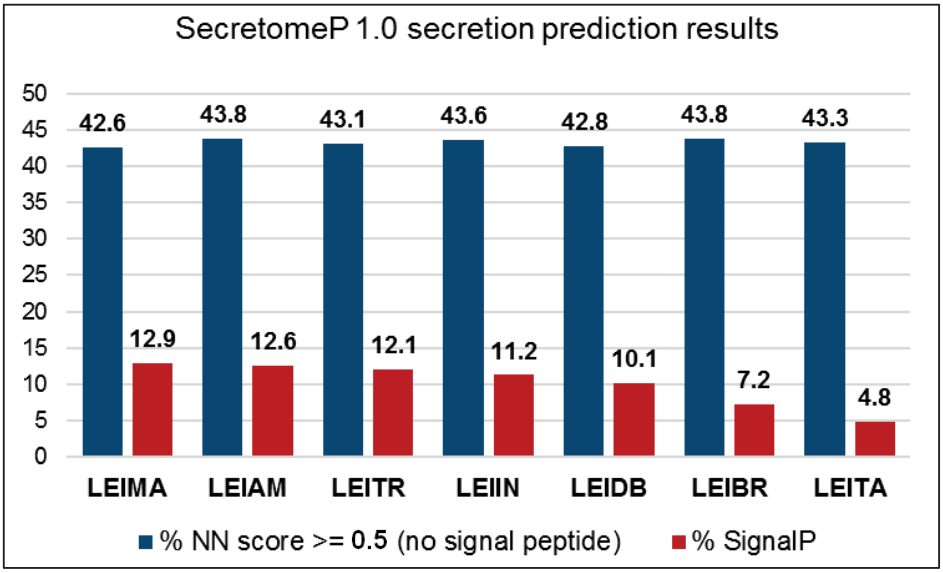


Figure II. 6 Secretion pathway analysis. *Total accessions identified in each sample were analysed with SecretomeP 1.0 (mammalian proteins). Percentage of proteins with NN score > 0,5 and no predicted signal peptide are predicted to be non-classically secreted (blue bars). SignalP analysis predicts the presence of signal peptides which target proteins for classical secretion (percentage of total proteins, red bars).*

As indicated by the SecretomeP results, and as previous secretome studies have shown, *Leishmania*-secreted proteins are associated with exosome-like vesicles as well as glycosomal vesicles. In the latter case, either *Leishmania* parasites secrete whole glycosomes or glycosomal cargo into the extracellular moiety (41). Exosomes are 30-100 nm membrane vesicles of endocytic origin secreted by most cell types *in vitro* (47).

Interestingly, in the present study we find 8 proteins associated with vesicle-based secretion, from the 26 proteins described by Silverman *et al* 2008 (Table II.4). Furthermore, several *Leishmania* proteins ortholog to exosome protein markers from other eukaryotic cell types were identified in the secretome datasets (Table II.5). The ExoCarta database describes the 100 most common proteins often associated with exosome cargo from different tissues and cell types, and across 10 different species, including humans (59). Remarkably, a high number of ortholog genes were found in *Leishmania*, 18 orthologs per species for a total of 72 genes identified. These correspond to 16 different proteins, 9 of which are found in the secretome datasets (Table II.5).

Table II. 4 List of* Leishmania*-secreted proteins associated with exosome-like and glycosomal vesicles found in the secretome datasets (Silverman et al 2008). *AP, adipocyte exosome (adiposome); GLY,* Leishmania *glycosome; BC, B-cell lymphocyte exosome; DC, dendritic cell exosome*

GeneDB accession number	Protein name	Microvesicle association
LmjF28.2860	Cytosolic malate dehydrogenase, putative	AP
LmjF24.2060	Transketolase, putative	GLY
LmjF28.2770	Heat-shock protein hsp70, putative	BC, DC, AP
LmjF14.1160	Enolase	BC, DC, AP
LmjF05.0350	Trypanothione reductase	GLY
LmjF16.0540	Aspartate carbamoyltransferase, putative	GLY
LmjF31.1070	Biotin/lipoate protein ligase-like protein	AP
LmjF36.3210	14-3-3 Protein-like protein	DC, AP

Table II. 5 List of exosome markers and genes found in the secretome. *From the 100 most abundant proteins associated with exosome cargo (Exocarta database)* (59)*, 72* Leishmania *ortholog genes were found (18 per species), and 20 of these were found in this dataset through OrthoMCL and TriTrypDB searches, corresponding to 9 different proteins. Yellow = proteins with signal peptide prediction (SignalP).*

GeneIDs	Found in Secretome	[Gene Product]	[PFam Domains]	Domain name
LbrM.15.1040; LbrM.33.0760; LinJ.15.1060; LinJ.33.0770; LmjF.15.1000; LmjF.33.0720; LmxM.15.1000; LmxM.32.0720	**LinJ.15.1060; LinJ.33.0770; LmjF.15.1000; LmjF.33.0720**	60S ribosomal protein L6, putative	PF01159	Ribosomal protein L6e
LbrM.30.0540; LinJ.30.0470; LmjF.30.0460; LmxM.29.0460;	**LinJ.30.0470; LmxM.29.0460**	aspartyl-tRNA synthetase, putative	PF00152, PF01336	tRNA synthetases class II (D, K and N), OB-fold nucleic acid binding domain
LbrM.32.0470; LinJ.32.0410; LmjF.32.0400; LmxM.31.0400; LinJ.15.1060	**LinJ.32.0410; LmxM.31.0400**	ATP-dependent RNA helicase, putative	PF00270, PF00271	DEAD/DEAH box helicase, Helicase conserved C-terminal domain
LbrM.34.3140; LinJ.35.3280; LmjF.35.3230; LmxM.34.3230; LmxM.15.1000	**LbrM.34.3140; LinJ.35.3280**	cystathione gamma lyase, putative	PF01053	Cys/Met metabolism PLP-dependent enzyme
LbrM.29.2180; LinJ.29.2310; LmjF.29.2200; LmxM.08_29.2200; LmjF.35.3230	**LbrM.29.2180; LinJ.29.2310; LmjF.29.2200; LmxM.08_29.2200**	GTP-binding protein, putative	PF00350, PF02212, PF01031	Biotin-requiring enzyme
LbrM.32.4110; LmxM.31.3870; LinJ.32.4020; LmjF.32.3870; LmjF.36.2660	**LinJ.32.4020; LmxM.31.3870**	myosin XXI	PF00063	Myosin head (motor domain)
LbrM.11.0920; LinJ.11.1140; LmjF.11.1160; LmxM.11.1160; LinJ.32.4020	**LmjF.11.1160**	protein transport protein Sec31, putative	PF00400	WD domain, G-beta repeat
LbrM.19.0760; LinJ.19.0440; LmjF.19.0440; LmxM.19.0440; LmjF.36.3660	**LbrM.19.0760**	nucleosome assembly protein, putative	PF00956	Nucleosome assembly protein (NAP)
LbrM.31.1980; LinJ.31.1770; LmjF.31.1750; LmxM.30.1750; LmjF.11.1160	**LinJ.31.1770; LmjF.11.1160**	nucleosome assembly protein-like protein	PF00956	Nucleosome assembly protein (NAP)
LbrM.14.0470; LinJ.14.0470; LmjF.14.0460; LmxM.14.0460; LmjF.30.0460	None found	cystathionine beta-lyase-like protein	PF01053	Cys/Met metabolism PLP-dependent enzyme
LbrM.24.1520; LinJ.24.1400; LmjF.24.1360; LmxM.24.1360; LmjF.14.0460	None found	hypothetical protein, conserved	null	
LbrM.20.2180; LinJ.34.2450; LmjF.34.2620; LmxM.33.2620; LmjF.29.2200	None found	RNA helicase, putative,mitochondrial, putative	PF00271	Helicase conserved C-terminal domain
LbrM.35.3890; LinJ.36.3840; LmjF.36.3660; LmxM.36.3660; LmjF.24.1360	None found	nudix hydrolase-like protein	PF00293	NUDIX domain
LbrM.29.2720; LinJ.13.0260; LmjF.13.0260; LmxM.13.0260; LmjF.34.2620	None found	hypothetical protein, conserved	PF00583	Acetyltransferase (GNAT) family
LbrM.31.2920; LinJ.31.2650; LmjF.31.2580; LmxM.30.2580; LmjF.13.0260	None found	ubiquinol-cytochrome-c reductase-like protein	PF02320	Ubiquinol-cytochrome C reductase hinge protein
LbrM.20.0631; LinJ.34.0740; LmjF.34.0705; LmxM.33.0705; LmjF.19.0440	None found	hypothetical protein, conserved	PF09799	Predicted membrane protein

3.3. Functionality of excreted-secreted proteins is conserved among *Leishmania* species

To better elucidate the specialised functions performed by the *Leishmania*-secreted proteins, a Gene Ontology (GO) annotation analysis was performed. Gene Ontology provides a representation or categorisation of the properties assigned to different gene products, and covers three main domains which represent how individual genes contribute to organism biology: Molecular function, the molecular activities of individual gene products; Cellular component, where the gene products are active; and Biological process, the pathways to which the gene product's activities contribute (62).

Results show secreted proteins by all seven *Leishmania* species analysed have conserved functions, across all GO domains (Figure II.7).

In all analysed *Leishmania* species, the vast majority of secreted *Leishmania* proteins possess, regarding molecular function pathways, either **binding** (GO:0005488, refers to the selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule) or **catalytic activity functions** (GO:0003824, refers to the catalysis of a biochemical reaction at physiological temperatures; gene products possess specific binding sites for substrates, and are usually composed wholly or largely of protein). The remaining molecular functions identified are **structural molecule activity** (GO:0005198, refers to the action of a molecule that contributes to the structural integrity of a complex or its assembly within or outside a cell) or **antioxidant activity** (GO:0016209, refers to the inhibition of the reactions brought about by dioxygen (O_2) or peroxides; usually the antioxidant is effective because it can itself be more easily oxidized than the substance protected) (Figure II.7 panel A). These results agree entirely with previous proteomic analysis of *L. donovani* and *L. braziliensis* secretomes (41,43).

Regarding the biological process domain, a high number of proteins are involved in **cellular** (GO:0009987, refers to any process that is carried out at the cellular level, but not necessarily restricted to a single cell, for example, cell communication) **or metabolic processes** (GO:0008152, refers to the chemical reactions and pathways, including anabolism and catabolism, by which living organisms transform chemical substances)**, regulation** (GO:0065007, refers to any process that modulates a measurable attribute of any biological process, quality or function.) and **response to stimulus** (GO:0050896, refers to Any process that results in a change in state or activity of a cell or an organism as a result of a stimulus in terms of movement, secretion, enzyme production, gene expression, etc.) (Figure II.7 panel B).

Cellular compartment GO annotations are more disperse, but still conserved among *Leishmania* species (Figure II.7 panel C). Most gene products are described to be active in the **cytoplasm** (GO:0005737, refers to all of the contents of a cell excluding the plasma membrane

and nucleus, but including other subcellular structures) and **nucleus** (GO:0005634, refers to membrane-bounded organelle of eukaryotic cells in which chromosomes are housed and replicated). The main gene products which possess these functions are: tubulin, elongation factors, proteasome-associated proteins and histones.

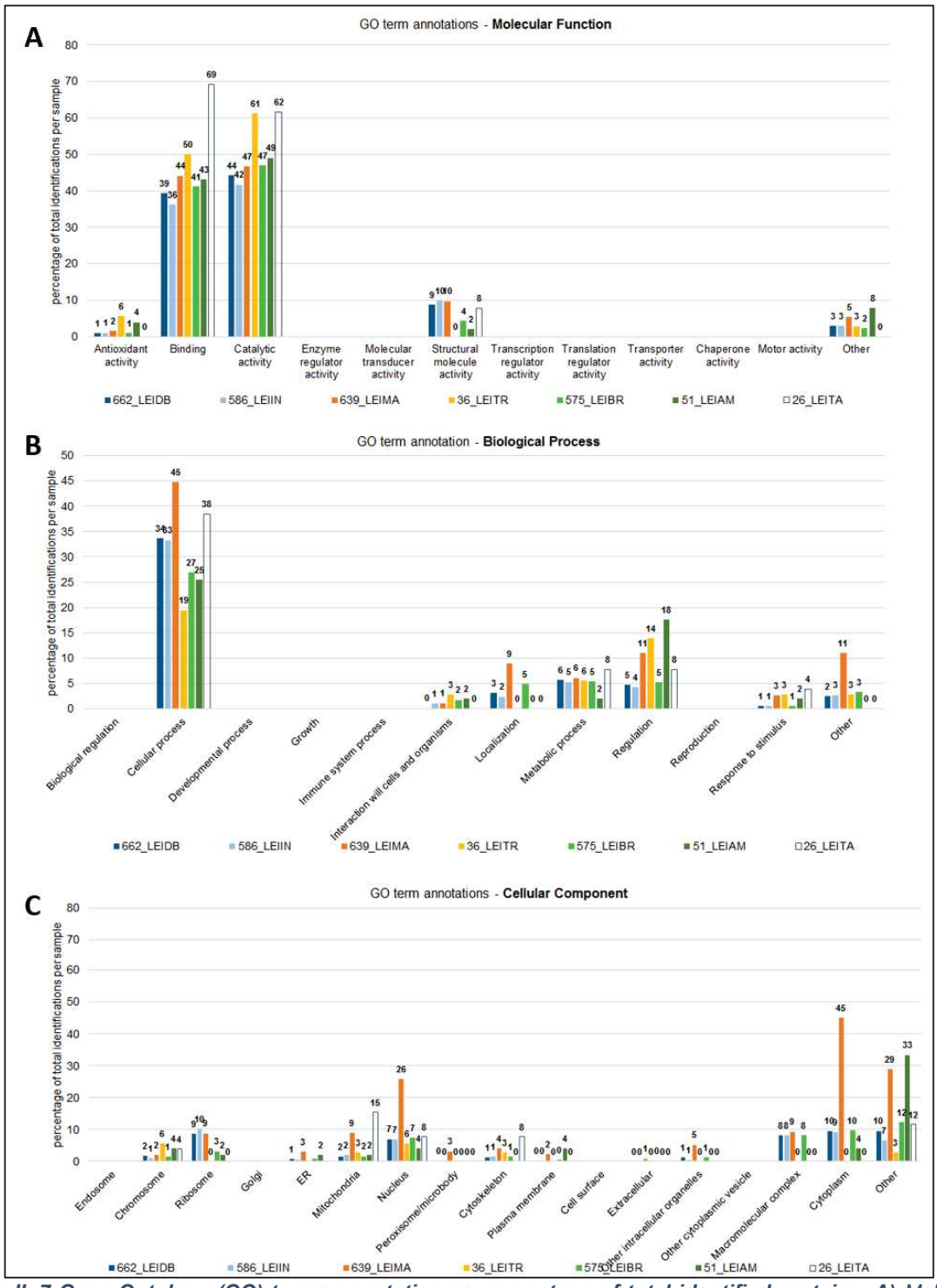


Figure II. 7 Gene Ontology (GO) terms annotation as percentage of total identified proteins. A) *Molecular function GO terms. B) Cellular component GO terms. C) Biological process GO terms.*

Additionally, multifunctional proteins have been described, particularly, moonlighting proteins that perform multiple independent, and often unrelated, functions, in the absence of gene fusion, splicing variants or different catalytic domains (63). Several moonlighting proteins have been described in both eukaryotes and prokaryotes which were incorporated in the MoonProt Database, a Database for Moonlighting Proteins (64). This database includes four *Leishmania*-specific proteins with described moonlighting functions (Table II.6). These four proteins were found in secretome datasets from all analysed *Leishmania* species.

Table II. 6 **Leishmania** ***proteins with moonlighting functions in the MoonProt database*** (64)**.**

UniProt ID	Protein Name	Function 1 (primary function)	Function 2 (moonlighting function)	Species Name
Q3HL75	**Enolase**	enolase, enzyme reversible conversion of 2-phosphoglycerate to phosphoenolpyruvate	binds plasminogen	*Leishmania mexicana*
E9BTJ1	**fructose-1,6-bisphosphate aldolase**	aldolase, cleavage of fructose-1,6-bisphosphate to glyceraldehyde 3-phosphate and dihydroxyacetone phosphate, in glycolysis	Activation of mouse (host) macrophage protein tyrosine phosphatase-1 (SHP-1), causes macrophage disfunction	*Leishmania donovani*
Q95VF2	**elongation factor 1-alpha**	translation elongation factor, the rate and fidelity of protein translation	binds to and activates Src homology 2 domain containing tyrosine phosphatase-1 (SHP-1) in host macrophages, inhibits activity of infected macrophages	*Leishmania donovani*
Q95U89	**mitochondrial 2-cysteine peroxiredoxin**	peroxidase activity, detoxification of reactive oxygen species (ROS), removal of peroxide, use redox active cysteine residue (peroxidatic Cys) to reduce substrates like H2O2	chaperone and activators of signal transduction cascades, prevents thermal aggregation of citrate synthase in vitro, lack of expression makes promastigoes more sensitive to temperature in the mammalian host (37°C)	*Leishmania infantum*

3.4. Several important virulent factors are found among the most abundant proteins (iBAQ analysis)

The analysis of the proteomic data (UniProtKB database release 2017_01) resulted in a total of 2140 accessions identified, from all identification status (Leading_AUTO, Leading_SPECIES, Multispecies = LEADING_CHECK). A normalised iBAQ value of 1 corresponds to the overall iBAQ average (absolute iBAQ average $1{,}81 \times 10^7$; Log_{10} iBAQ average 6,195). 582 accessions with normalised iBAQ values over 1,1 were considered as abundant (see Appendix II.1).

The most abundant protein overall is **nucleoside diphosphate kinase**, the only identified accession with a normalised iBAQ value over 1,5 (Figure II.8). This enzyme is required for the synthesis of nucleoside triphosphates (other than ATP) and catalyses the reaction transforming ATP and nucleoside diphosphate in ADP and nucleoside triphosphate. The generated nucleoside triphosphates are then used for nucleic acid, lipid or polyssacharide synthesis, protein elongation, signal transduction and microtubule polymerisation. It is likely this protein possesses additional

moonlighting functions such as protein binding, similarly to other kinases described in the MoonProt database.

Several important well-documented virulence factors were identified in the *Leishmania* promastigote secretome. *Leishmania* parasites secrete virulence factors to the host cell cytoplasm where they actively modulate host signalling molecules and immune responses (33,65).

There are fourteen highly abundant proteins in the secretome, with normalised iBAQ values over 1,4, including well-known virulence factors (Figure II.8, dark orange). By descending order: soluble **promastigote surface antigen PSA-38S**, **tubulin** beta chain (2 accessions), tubulin alpha chain, putative **small myristoylated protein-3**, **ubiquitin-60S ribosomal protein L40**, an **uncharacterized protein**, **infective insect stage-specific protein**, **peroxidoxin 2**, soluble **promastigote surface antigen PSA-34S** (Fragment), **adenosylhomocysteinase** (2 accessions), putative **heat-shock protein hsp70**, and **elongation factor 1-alpha** (Appendix II.1).

Among the most abundant proteins, we wish to highlight **HSP70**, a protein chaperone with key function for parasite's adaptation to higher temperatures in vertebrate hosts; and **ubiquitin-60S ribosomal protein L40** involved in gene regulation. Also, the **promastigote surface antigen** is a major surface glycoprotein, related to glycoprotein 46 (gp46), whose functions are not fully understood, but is associated to the evasion of complement lysis (66) and parasite virulence, as evidenced by its upregulation in PKDL-causing *Leishmania* parasites (67)**.** Furthermore, the **elongation factor 1-alpha** contributes to parasite persistence by inhibiting macrophage function through SHP-1-mediated disruption of JAK1, JAK2, STAT1 phosphorylation (68). Additionally, EF1alpha interacts with **ribosomal subunits**, for example, the 60S subunit also identified here, associated with protein translation (65). In turn, ribosomal subunits also interact with **S-adenosylhomocysteine hydrolase**, also very abundant in the secretome (65). The infective insect stage-specific protein expressed by gene META1 is specific to infective metacyclic promastigotes (69).

Several of these proteins have also been described as important antigens, namely, the protein **PSA-38S** (51) and **elongation factor 1-alpha** (70). Interestingly, the enzyme **adenosylhomocysteinase**, involved in amino-acid biosynthesis (synthesis of L-homocysteine from S-adenosyl-L-homocysteine) was found to be highly abundant, which is expected for *Leishmania* promastigotes, and it has also been shown to be immunogenic in hamsters and human PBMC stimulation assays (71). The **uncharacterised protein** (A4H3U0) identified among the most abundant proteins is actually the most abundant in the LEIBR sample. This protein is predicted to have a metalloendopeptidase activity with cell adhesion functions and could be analysed further as a potential antigen candidate.

Moreover, heat-shock proteins, other cytoplasmic and cytoskeleton proteins, such as tubulin and actin, as well as membrane-associated proteins are ubiquitous proteins identified as

exosome cargo. Indeed, among the most abundant proteins identified (normalised iBAQ value over 1.4) the proteins tubulin (alpha and beta chains), heat-shock protein 70 and surface antigens are found. The high abundance of these proteins further supports the importance of exosome-based secretion in *Leishmania* parasites.

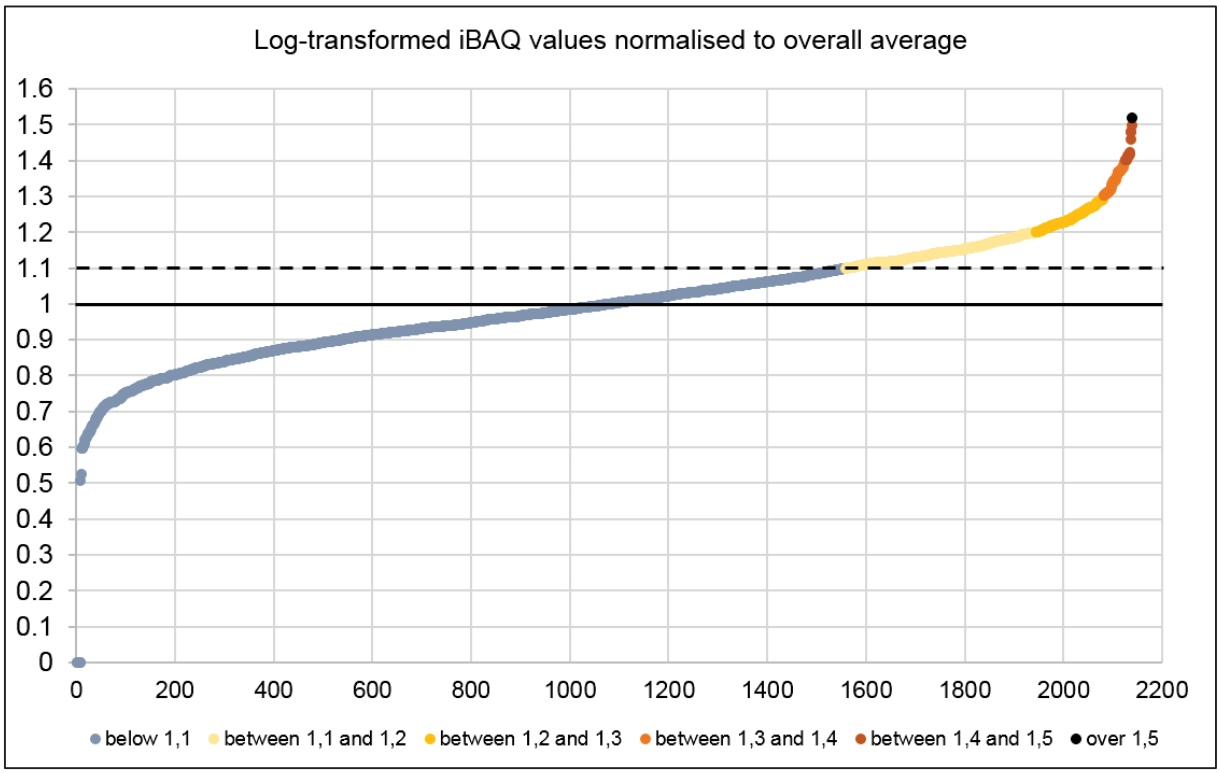


Figure II. 8 Log-transformed iBAQ values normalised to overall average. *The iBAQ values from 2140 accessions identified in the* Leishmania *secretome were analysed. A normalised iBAQ value superior to 1,1 is associated to abundant proteins.*

Among the second most abundant accessions with normalised iBAQ values above 1,3, the main virulence factors identified were **promastigote surface antigens** and soluble forms of the two major *Leishmania* virulence factors, **glycoprotein 63 (GP63) or leishmanolysin**. GP63 is a zinc-dependent surface metalloprotease with a major role in promastigote evasion of the immune system, as well as adhesion and invasion of host macrophages (72). Other important virulence factors include several **cysteine proteases,** enzymes involved in the *Leishmania* **antioxidant system** (superoxide dismutase, trypanothione reductase, tryparedoxin peroxidase, and thiol specific antioxidant); and other enzymes such as **peptidyl-prolyl cis-trans isomerase** (cyclophilin), and **enolase** with probable key roles in tissue invasion and virulence (73,74). The protein **aldolase** (putative 2,4-dihydroxyhept-2-ene-1,7-dioic acid aldolase) binds host tyrosine phosphatase SHP-1, inducing macrophage dysfunction and promoting parasite persistence (75). The putative **histidine secretory acid phosphatase** is found on the parasite's surface and is

continuously secreted by all species analysed to date and, although its role in pathogenesis is for the most part unknown, it is implicated in nutrient scavenging as well as modulation of host environment (72).

The identification of **histone** proteins in stationary phase samples can be associated with the release of apoptotic vesicles. Indeed, several histone proteins are found in all datasets - Histone H4 and Histone H2B are the most abundant (normalised iBAQ values between 1.37 and 1.38). Furthermore, we also found Histone H3 (1.29), Histone H2A (1.22), and Histone H2A.1 (1.18). The histones are core components of the nucleosome, the DNA packaging unit in eukaryotes consisting of histones bound to the DNA strands (chromatin). In the MoonProt database, the only histone described to have a moonlighting function is mouse histone H1 which acts as a thyroglobulin receptor.

4. Discussion

To our knowledge, this is the first proteomic analysis of the secretome comparing multiple *Leishmania* species, pathogenic and non-pathogenic to humans (publication in preparation). Furthermore, the method used to produce naturally secreted proteins using aseric *in vitro* growth conditions is recommended to analyse such proteins. Seldom the secretome analysed in proteomic studies is prepared from cultured parasites with medium containing serum, which negatively affects protein secretion. The lack of duplicates for each sample was due to cost constraints and it can pose some limitations to the present study. However, previous experiments from the development of the canine vaccine show limited variability between secretome batches (vaccine quality control, unpublished data).

The present study focuses on the *Leishmania* promastigote secretome, a critical player in early stage host-pathogen interactions. *Leishmania* parasites do undergo differentiation to amastigotes inside the vertebrate host, however, previous studies show low levels of differential gene expression, with a big role for post-translational regulation of the stage-specific gene expression and survival in the intracellular environment. As gene and protein annotation improve, so will our understanding of all players involved in host-parasite interactions. The key to potentially discover targets or markers and develop the much-needed improved treatments and diagnostics.

As expected, classical secretion mediated by N-terminal signal peptides is the least important secretion mechanism in *Leishmania* parasites and most proteins are predicted to be secreted through non-conventional pathways. Still, a large portion of the proteins are not predicted to be secreted despite being identified in the secretome. Despite its usefulness in deciphering the protein secretion pathways, it has been demonstrated that the SecretomeP prediction algorithm presents some limitations in predicting non-classical secretion for plant proteins and other eukaryotic organisms (76,77). These limitations may also apply to *Leishmania* proteins because of the algorithm training data. Possibly, *Leishmania* proteins possess unexpected features or share features with proteins from both eukaryotes and prokaryotes which are not considered by the algorithm. To evaluate this possibility, an additional bioinformatics analysis with SecretomeP predictions for gram-negative/gram-positive bacteria non-classical secretion will be performed and included in the publication.

Nevertheless, the results presented here are completely aligned with previous studies on *Leishmania* secreted proteins and the importance of vesicle-based secretion in *Leishmania* parasites (33,37,40,45). Exosome cargo is increasingly studied and implicated in host-host and host-pathogen interactions (35,47,78), and in this database we find proteins that are identified often in exosomes. We found many exosome marker proteins in the *Leishmania* secretome (9

different proteins out of 16 *Leishmania* orthologs to genes described in the Exocarta). This suggests a highly conserved cell communication mechanism across eukaryotic organisms, to which the parasites perfectly adapted to interrupt and manipulate so to enable their replication and continued transmission.

Several important and well-described virulence factors are among the most abundant proteins (PSA, leishmanolysin, HSP70, and others). Some of these proteins have not only been described to be key players in early establishment of infection but also as potential vaccine antigen candidates (79–81).

The biological reason for the presence of certain proteins or protein families remains elusive. We expect to find a mix of virulence factors, immunomodulators, and excretion products. Interestingly, *Leishmania* proteins with moonlighting functions have been described, all of which are found in the secretome. Considering the existence of these moonlighting functions, the remaining proteins and virulence factors identified in the secretome are probably performing additional functions which remain unidentified.

Importantly, the functions attributed to the secreted proteins are highly conserved among *Leishmania* species. *L. braziliensis* is the most divergent species among the human infective species analysed and *L. tarentolae* which is non-infective to human. Still, they share gene product functions across all Gene Ontology domains. The secretome data generated deserves further in-depth analysis to evaluate if certain GO terms are enriched in this biological compartment. This analysis will be performed and discussed in the manuscript in preparation, aiming at unravelling the essentiality of excreted-secreted proteins in host-pathogen interactions. Nevertheless, we provide evidence of the high degree of *Leishmania* species conservation on the proteomic level (active end products of genome expression) which further supports the discovery of conserved virulence factors and/or antigens.

The datasets generated with the proteomic analysis constitute an ancillary expected outcome of the present project. Hopefully, the detailed characterization of the secretome will contribute to novel discoveries on potential connections between geographical distribution, clinical presentation, immunomodulation and pathogenesis. Also, by including a non-pathogenic *Leishmania* species (*L. tarentolae*), we hope to contribute to increase our knowledge about interspecies variability and possibly reveal novel important virulence factors that can contribute to leishmaniasis diagnosis or drug development.

The high conservation of the proteins identified in the *Leishmania* secretome datasets further corroborates the proposed strategy of using promastigotes secreted proteins as a source for antigen discovery and pan-*Leishmania* vaccine development. The following chapter will explore these datasets to highlight relevant vaccine antigen candidates.

5. References

1. Burza S, Croft SL, Boelaert M. Leishmaniasis. Vol. 392, The Lancet. Elsevier; 2018. p. 951–70.
2. Magill AJ. Chapter 277 - Leishmania Species: Visceral (Kala-Azar), Cutaneous, and Mucosal Leishmaniasis. In: Bennett JE, Dolin R, Blaser MJ, editors. Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases. 8th ed. Elsevier Saunders; 2015. p. 3493–525.
3. Zijlstra EE. The immunology of post-kala-azar dermal leishmaniasis (PKDL). Parasit Vectors. BioMed Central; 2016;9(1):464.
4. Akhoundi M, Kuhls K, Cannet A, Votýpka J, Marty P, Delaunay P, et al. A Historical Overview of the Classification, Evolution, and Dispersion of Leishmania Parasites and Sandflies. Bañuls A-L, editor. PLoS Negl Trop Dis. Public Library of Science; 2016 Mar 3;10(3):e0004349.
5. Killick-Kendrick R. Some epidemiological consequences of the evolutionary fit between Leishmaniae and their phlebotomine vectors. Bull Soc Pathol Exot Filiales. 1985;78(5 Pt 2):747–55.
6. Batista de Jesus J, Mesquita-Rodrigues C, Cuervo P. Chapter 14 - Proteomics Advances in the Study of Leishmania Parasites and Leishmaniasis. In: Santos AL, Branquinha MH, D'Avila-Levy CM, Kneipp LF, Sodré CL, editors. Proteins and Proteomics of Leishmania and Trypanosoma. 1st ed. Springer; 2014. p. 323–49.
7. Besteiro S, Williams RAM, Coombs GH, Mottram JC. Protein turnover and differentiation in Leishmania. Int J Parasitol. 2007;37(10):1063–75.
8. Ivens AC, Peacock CS, Worthey E a, Murphy L, Aggarwal G, Berriman M, et al. The genome of the kinetoplastid parasite, Leishmania major. Science. 2005;309(5733):436–42.
9. Leifso K, Cohen-Freue G, Dogra N, Murray A, McMaster WR. Genomic and proteomic expression analysis of Leishmania promastigote and amastigote life stages: The Leishmania genome is constitutively expressed. Mol Biochem Parasitol. 2007;152(1):35–46.
10. Fasel N, Acestor N, El Fadili-Kundig A, Gonzalez I, Masina S. The Leishmania proteome. In: Leishmania. 2008.
11. Peacock CS, Seeger K, Harris D, Murphy L, Ruiz JC, Quail MA, et al. Comparative genomic analysis of three Leishmania species that cause diverse human disease. Nat Genet. Nature Publishing Group; 2007 Jul 17;39(7):839–47.
12. Rogers MB, Hilley JD, Dickens NJ, Wilkes J, Bates PA, Depledge DP, et al. Chromosome and gene copy number variation allow major structural change between species and strains of

Leishmania. Genome Res. 2011 Dec;21(12):2129–42.

13. Teixeira SM, de Paiva RMC, Kangussu-Marcolino MM, Darocha WD. Trypanosomatid comparative genomics: Contributions to the study of parasite biology and different parasitic diseases. Genet Mol Biol. Sociedade Brasileira de Genética; 2012 Jan;35(1):1–17.
14. Rosenzweig D, Smith D, Opperdoes F, Stern S, Olafson RW, Zilberstein D. Retooling Leishmania metabolism: from sand fly gut to human macrophage. FASEB J. 2007 Sep 20;22(2):590–602.
15. Alcolea PJ, Alonso A, Gómez MJ, Postigo M, Molina R, Jiménez M, et al. Stage-specific differential gene expression in Leishmania infantum: from the foregut of Phlebotomus perniciosus to the human phagocyte. BMC Genomics. BioMed Central; 2014 Jan 3;15(1):849.
16. Cohen-Freue G, Holzer TR, Forney JD, McMaster WR. Global gene expression in Leishmania. Int J Parasitol. 2007;37(10):1077–86.
17. Rochette A, Raymond F, Corbeil J, Ouellette M, Papadopoulou B. Whole-genome comparative RNA expression profiling of axenic and intracellular amastigote forms of Leishmania infantum. Mol Biochem Parasitol. 2009;165(1):32–47.
18. Wenzel UA, Bank E, Florian C, Forster S, Zimara N, Steinacker J, et al. Leishmania major parasite stage-dependent host cell invasion and immune evasion. FASEB J. 2012;26(1):29–39.
19. Clayton CE. Life without transcriptional control? From fly to man and back again. EMBO J. EMBO Press; 2002 Apr 15;21(8):1881–8.
20. Campbell DA, Thomas S, Sturm NR. Transcription in kinetoplastid protozoa: why be normal? Microbes Infect. Elsevier Masson; 2003 Nov 1;5(13):1231–40.
21. Palenchar JB, Bellofatto V. Gene transcription in trypanosomes. Mol Biochem Parasitol. Elsevier; 2006 Apr 1;146(2):135–41.
22. Rosenzweig D, Smith D, Myler PJ, Olafson RW, Zilberstein D. Post-translational modification of cellular proteins duringLeishmania donovani differentiation. Proteomics. WILEY-VCH Verlag; 2008 May;8(9):1843–50.
23. Oliveira I a, Freire-de-lima L, Penha LL, Dias WB, Todeschini AR. Proteins and Proteomics of Leishmania and Trypanosoma. Vol. 74, Subcellular Biochemistry. 2014. 181-201 p.
24. El Fakhry Y, Ouellette M, Papadopoulou B. A proteomic approach to identify developmentally regulated proteins in Leishmania infantum. Proteomics. Wiley-Blackwell; 2002 Aug 1;2(8):1007.
25. Bentel M, Harder S, Wiesgigl M, Heukeshoven J, Gelhaus C, Krause E, et al. Developmentally induced changes of the proteome in the protozoan parasite Leishmania donovani. In: Proteomics. Wiley-Blackwell; 2003. p. 1811–29.

26. Nugent PG, Karsani SA, Wait R, Tempero J, Smith DF. Proteomic analysis of Leishmania mexicana differentiation. Mol Biochem Parasitol. Elsevier; 2004 Jul 1;136(1):51–62.
27. Walker J, Vasquez J-J, Gomez MA, Drummelsmith J, Burchmore R, Girard I, et al. Identification of developmentally-regulated proteins in Leishmania panamensis by proteome profiling of promastigotes and axenic amastigotes. Mol Biochem Parasitol. Elsevier; 2006 May 1;147(1):64–73.
28. Pescher P, Blisnick T, Bastin P, Späth GF. Quantitative proteome profiling informs on phenotypic traits that adapt Leishmania donovani for axenic and intracellular proliferation. Cell Microbiol. John Wiley & Sons, Ltd; 2011 Jul 1;13(7):978–91.
29. McCall LI, Zhang WW, Dejgaard K, Atayde VD, Mazur A, Ranasinghe S, et al. Adaptation of leishmania donovani to cutaneous and visceral environments: In vivo selection and proteomic analysis. J Proteome Res. 2015;14(2):1033–59.
30. Sampaio P, Veras T, Perrone J, De Menezes B. Using Proteomics to Understand How Leishmania Parasites Survive inside the Host and Establish Infection. 2016;
31. Tjalsma H, Bolhuis A, Jongbloed JDH, Bron S, van Dijl JM. Signal Peptide-Dependent Protein Transport in Bacillus subtilis: a Genome-Based Survey of the Secretome. Microbiol Mol Biol Rev. 2000 Sep 1;64(3):515–47.
32. Lambertz U, Silverman JM, Nandan D, McMaster WR, Clos J, Foster LJ, et al. Secreted virulence factors and immune evasion in visceral leishmaniasis. J Leukoc Biol. 2012;91(6):887–99.
33. Corrales RM, Sereno D, Mathieu-Daudé F. Deciphering the Leishmania exoproteome: what we know and what we can learn. FEMS Immunol Med Microbiol. 2010;58(1):27–38.
34. Rosa R, Roos Rodrigues O, Marques C, Santos-Gomes GM. Leishmania infantum: Soluble proteins released by the parasite exert differential effects on host immune response. Exp Parasitol. 2005;109(2):106–14.
35. Silverman JM, Reiner NE. Exosomes and other microvesicles in infection biology: organelles with unanticipated phenotypes. Cell Microbiol. Blackwell Publishing Ltd; 2011 Jan;13(1):1–9.
36. Twu O, Johnson PJ, Hargett L, Bauer N, Harding C, Heuser J, et al. Parasite Extracellular Vesicles: Mediators of Intercellular Communication. Knoll LJ, editor. PLoS Pathog. Public Library of Science; 2014 Aug 28;10(8):e1004289.
37. Silverman JM, Reiner NE. Leishmania Exosomes Deliver Preemptive Strikes to Create an Environment Permissive for Early Infection. Front Cell Infect Microbiol. 2012;1(January):1–8.
38. Markikou-Ouni W, Drini S, Bahi-Jaber N, Chenik M, Meddeb-Garnaoui A, Alvar J, et al. Immunomodulatory Effects of Four Leishmania infantum Potentially Excreted/Secreted

Proteins on Human Dendritic Cells Differentiation and Maturation. Haziot A, editor. PLoS One. Public Library of Science; 2015 Nov 18;10(11):e0143063.

39. Gour JK, Kumar V, Singh N, Bajpai S, Pandey HP, Singh RK. Identification of Th1-responsive leishmanial excretory–secretory antigens (LESAs). Exp Parasitol. Elsevier Inc.; 2012;132(3):355–61.
40. Silverman JM, Clos J, Horakova E, Wang AY, Wiesgigl M, Kelly I, et al. Leishmania Exosomes Modulate Innate and Adaptive Immune Responses through Effects on Monocytes and Dendritic Cells. J Immunol. 2010 Nov 1;185(9):5011–22.
41. Silverman JM, Chan SK, Robinson DP, Dwyer DM, Nandan D, Foster LJ, et al. Proteomic analysis of the secretome of Leishmania donovani. Genome Biol. BioMed Central; 2008;9(2):R35.
42. Hassani K, Antoniak E, Jardim A, Olivier M. Temperature-Induced Protein Secretion by Leishmania mexicana Modulates Macrophage Signalling and Function. PLoS One. 2011;6(5):e18724.
43. Cuervo P, De Jesus JB, Saboia-Vahia L, Mendonça-Lima L, Domont GB, Cupolillo E. Proteomic characterization of the released/secreted proteins of Leishmania (Viannia) braziliensis promastigotes. J Proteomics. Elsevier; 2009 Nov 2;73(1):79–92.
44. Atayde VD, Aslan H, Townsend S, Hassani K, Kamhawi S, Olivier M. Exosome Secretion by the Parasitic Protozoan Leishmania within the Sand Fly Midgut. Cell Rep. 2015 Oct;13(5):957–67.
45. Atayde VD, Hassani K, da Silva Lira Filho A, Borges AR, Adhikari A, Martel C, et al. Leishmania exosomes and other virulence factors: Impact on innate immune response and macrophage functions. Cell Immunol. 2016;
46. Silverman JM, Clos J, De'Oliveira CC, Shirvani O, Fang Y, Wang C, et al. An exosome-based secretion pathway is responsible for protein export from Leishmania and communication with macrophages. J Cell Sci. 2010;123(6):842–52.
47. Schorey JS, Cheng Y, Singh PP, Smith VL. Exosomes and other extracellular vesicles in host-pathogen interactions. EMBO Rep. EMBO Press; 2015 Jan 5;16(1):24–43.
48. Hassani K, Olivier M. Immunomodulatory impact of leishmania-induced macrophage exosomes: a comparative proteomic and functional analysis. PLoS Negl Trop Dis. Public Library of Science; 2013 Jan 2;7(5):e2185.
49. Tsigankov P, Gherardini PF, Helmer-Citterich M, Zilberstein D. What has proteomics taught us about Leishmania development? Parasitology. 2012;139(09):1146–57.
50. Tonui WK, Mejia JS, Hochberg L, Mbow ML, Ryan JR, Chan AST, et al. Immunization with Leishmania major exogenous antigens protects susceptible BALB/c mice against challenge infection with L. major. Infect Immun. 2004;72(10):5654–61.

51. Chamakh-Ayari R, Bras-Gonçalves R, Bahi-Jaber N, Petitdidier E, Markikou-Ouni W, Aoun K, et al. In vitro evaluation of a soluble Leishmania promastigote surface antigen as a potential vaccine candidate against human leishmaniasis. PLoS One. 2014;9(5):1–12.
52. Lemesre J-L, Holzmuller P, Cavaleyra M, Gonçalves RB, Hottin G, Papierok G. Protection against experimental visceral leishmaniasis infection in dogs immunized with purified excreted secreted antigens of Leishmania infantum promastigotes. Vaccine. 2005;23(22):2825–40.
53. Merlen T, Sereno D, Brajon N, Rostand F, Lemesre J-LL. Leishmania spp.: Completely defined medium without serum and macromolecules (CDM/LP) for the continuous in vitro cultivation of infective promastigote forms. Am J Trop Med Hyg. 1999;60(1):41–50.
54. Lemesre J-L. Method for the culture in vitro of different stages of tissue parasites. Organisation Mondiale de la Propriete Intellectuelle; WO 94/26899, 1994.
55. Olsen J V, de Godoy LMF, Li G, Macek B, Mortensen P, Pesch R, et al. Parts per Million Mass Accuracy on an Orbitrap Mass Spectrometer via Lock Mass Injection into a C-trap. Mol Cell Proteomics. American Society for Biochemistry and Molecular Biology; 2005 Dec 1;4(12):2010–21.
56. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. Nature Publishing Group; 2008 Dec 30;26(12):1367–72.
57. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen J V., Mann M. Andromeda: A peptide search engine integrated into the MaxQuant environment. J Proteome Res. American Chemical Society; 2011 Apr;10(4):1794–805.
58. Bendtsen JD, Jensen LJ, Blom N, Von Heijne G, Brunak S. Feature-based prediction of non-classical and leaderless protein secretion. Protein Eng Des Sel. Narnia; 2004 May 4;17(4):349–56.
59. Keerthikumar S, Chisanga D, Ariyaratne D, Al Saffar H, Anand S, Zhao K, et al. ExoCarta: A Web-Based Compendium of Exosomal Cargo. J Mol Biol. Academic Press; 2016 Feb 22;428(4):688–92.
60. Rabouille C. Pathways of Unconventional Protein Secretion. Trends Cell Biol. 2017;27(3):230–40.
61. Denny PW, Gokool S, Russell DG, Field MC, Smith DF. Acylation-dependent protein export in Leishmania. J Biol Chem. American Society for Biochemistry and Molecular Biology; 2000 Apr 14;275(15):11017–25.
62. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: Tool for the unification of biology. Nat Genet. 2000 May;25(1):25–9.
63. Huberts DHEW, van der Klei IJ. Moonlighting proteins: An intriguing mode of

multitasking. Biochim Biophys Acta - Mol Cell Res. Elsevier; 2010 Apr 1;1803(4):520–5.

64. Chen C, Zabad S, Liu H, Wang W, Jeffery C. MoonProt 2.0: an expansion and update of the moonlighting proteins database. Nucleic Acids Res. Narnia; 2018 Jan 4;46(D1):D640–4.
65. Da Fonseca Pires S, Fialho LC, Silva SO, Melo MN, De Souza CC, Tafuri WL, et al. Identification of virulence factors in leishmania infantum strains by a proteomic approach. J Proteome Res. 2014;13(4):1860–72.
66. Lincoln LM, Ozaki M, Donelson JE, Beetham JK. Genetic complementation of Leishmania deficient in PSA (GP46) restores their resistance to lysis by complement. Mol Biochem Parasitol. Elsevier; 2004 Sep 1;137(1):185–9.
67. Salotra P, Duncan RC, Singh R, Subba Raju BV, Sreenivas G, Nakhasi HL. Upregulation of surface proteins in Leishmania donovani isolated from patients of post kala-azar dermal leishmaniasis. Microbes Infect. 2006 Mar;8(3):637–44.
68. Nandan D, Reiner NE. Attenuation of gamma interferon-induced tyrosine phosphorylation in mononuclear phagocytes infected with Leishmania donovani: selective inhibition of signaling through Janus kinases and Stat1. Infect Immun. American Society for Microbiology Journals; 1995 Nov 1;63(11):4495–500.
69. Nourbakhsh F, Uliana SRB, Smith DF. Characterisation and expression of a stage-regulated gene of Leishmania major. Mol Biochem Parasitol. Elsevier; 1996 Feb 1;76(1–2):201–13.
70. Naouar I, Boussoffara T, Chenik M, Gritli S, Ben Ahmed M, Belhadj Hmida N, et al. Prediction of T Cell Epitopes from Leishmania major Potentially Excreted/Secreted Proteins Inducing Granzyme B Production. PLoS One. Public Library of Science; Jan;11(1):e0147076.
71. Khare P, Jaiswal AK, Tripathi CDP, Sundar S, Dube A. Immunoprotective responses of T helper type 1 stimulatory protein-S-adenosyl-L-homocysteine hydrolase against experimental visceral leishmaniasis. Clin Exp Immunol. Wiley-Blackwell; 2016;185(2):165–79.
72. Olivier M, Atayde VD, Isnard A, Hassani K, Shio MT. Leishmania virulence factors: Focus on the metalloprotease GP63. Microbes Infect. 2012;14(15):1377–89.
73. Avilán L, Gualdrón-López M, Quiñones W, González-González L, Hannaert V, Michels PAM, et al. Enolase: a key player in the metabolism and a probable virulence factor of trypanosomatid parasites-perspectives for its use as a therapeutic target. Enzyme Res. Hindawi Limited; 2011;2011:932549.
74. Yurchenko V, Xue Z, Sherry B, Bukrinsky M. Functional analysis of Leishmania major cyclophilin. Int J Parasitol. NIH Public Access; 2008 May;38(6):633–9.
75. Nandan D, Tran T, Trinh E, Silverman JM, Lopez M. Identification of leishmania fructose-1,6-bisphosphate aldolase as a novel activator of host macrophage Src homology 2

domain containing protein tyrosine phosphatase SHP-1. Biochem Biophys Res Commun. 2007 Dec 21;364(3):601–7.

76. Lonsdale A, Davis MJ, Doblin MS, Bacic A. Better Than Nothing? Limitations of the Prediction Tool SecretomeP in the Search for Leaderless Secretory Proteins (LSPs) in Plants. Front Plant Sci. Frontiers; 2016 Sep 27;7:1451.
77. Nielsen H, Petsalaki EI, Zhao L, Stühler K. Predicting eukaryotic protein secretion without signals. Biochim Biophys Acta - Proteins Proteomics. Elsevier; 2018 Dec 4;
78. van Niel G, D'Angelo G, Raposo G. Shedding light on the cell biology of extracellular vesicles. Nat Rev Mol Cell Biol. Nature Publishing Group; 2018 Jan 17;19(4):213–28.
79. Jain K, Jain NK. Vaccines for visceral leishmaniasis: A review. J Immunol Methods. 2015;422:1–12.
80. Singh B, Sundar S. Leishmaniasis: Vaccine candidates and perspectives. Vaccine. Elsevier Ltd; 2012;30(26):3834–42.
81. Paape D, Aebischer T. Contribution of proteomics of Leishmania spp. to the understanding of differentiation, drug resistance mechanisms, vaccine and drug development. J Proteomics. 2011;74(9):1614–24.

CHAPTER III

VACCINE ANTIGEN SELECTION

Objective: to search for *Leishmania*-specific vaccine antigens that will advance for epitope prediction analysis for the development of a human vaccine formulation against human leishmaniases.

Aims:

- To screen a maximum number of antigens, so to fully explore the potential of the secretome proteomic datasets and increase the chances of finding strong binding epitopes;
- To list and search previously described antigens in *Leishmania* secretome, which will allow the identification of known vaccine candidates, according to the literature;
- To perform a reverse vaccinology (RV) approach which will allow the identification of novel protein antigen candidates based on (low) homology with human host proteins.

1. Introduction

An antigen is a molecule that can bind specifically to an antibody or generate peptide fragments that are recognised by a T cell receptor, both produced by adaptive immune responses (1). Immunogens are molecules that, on their own, can elicit an adaptive immune response on injection into a person or animal (1). All antigens have the potential of binding specific antibodies or TCR, however, not all antigens are immunogenic, as they can be recognised by specific antibodies or TCR without inducing an immune response. In this sense, all immunogens are antigens, but not all antigens are immunogens, and vaccine candidates must be immunogenic.

As mentioned in the previous chapter, there is a high level of antigenic conservation among pathogenic species, as all species share high genome homology and synteny. Also, there is not a significant antigenic variation among *Leishmania* parasite forms (promastigote and amastigote).

Over the last decades, *Leishmania* vaccine antigens have been increasingly detailed, from non-defined crude antigen mixtures to the generation of recombinant proteins. Several factors have greatly contributed to these advances: i) the availability of pathogen- and host-specific genome and proteome information; ii) increased knowledge on mechanisms behind antigen recognition, presentation and subsequent immune activation cascades; and iii) the improvement of computational tools for epitope mapping and prediction.

2. Methods

2.1. Bioinformatic analysis of proteomic datasets

Raw MS spectra (AESLMTS_110414, AESLMTS_151026, AESLMTS_160115) were processed using the MaxQuant environment (version 1.5.5.0), as in Chapter II. The MS/MS spectra were matched against the *Leishmania* entries of UniProtKB SwissProt and TrEMBL sections (release 2016_05: 50997 entries with 229 SwissProt + 50768 TrEMBL). Again, a representative protein ID in each protein group was automatically selected using in-house bioinformatics tools (leading v2.2 and multi-species scripts) developed by Oana Vigy (*Plateforme de Proteomique Fonctionnelle de Montpellier*), and the Perseus software (version 1.5.3.0).

All *Leishmania* entries available in UniProtKB (UniProtKB release 2016_05) were used for tryptic peptide identification: four reference proteomes from *Leishmania infantum* strain JPCM5 (RefProteome_LEIIN-all_2016_01.fasta), *Leishmania major* strain MHOM/IL/81/Friedlin (RefProteome_LEIMA-all_2016_01.fasta), *Leishmania braziliensis* strain MHOM/BR/75/M2904 (RefProteome_LEIBR-all_2016_01.fasta), *Leishmania mexicana* strain MHOM/GT/2001/U1103, *Leishmania donovani* strain BPK282A1 (Proteome_LEIDB-all_2016_01.fasta). Also, all *L. amazonensis* and *L. tropica* proteins in the UniProtKB were included (Uniprot_LEIAM-all_2016_01.fasta; Uniprot_LEITR-all_2016_01.fasta).

2.2. Protein set A – previously described vaccine antigen candidates

The complete list of previously known antigen candidates (Table III.1) was retrieved from the following publications:

i) Singh B, Sundar S. 2012. Leishmaniasis: Vaccine candidates and perspectives. Vaccine 30:3834– 3842 (2);
ii) Kumar R, Engwerda C. 2014. Vaccines to prevent leishmaniasis. Clinical & Translational Immunology 3, e13 (3);
iii) Lakshmi BS, Wang R, Madhubala R. 2014. *Leishmania* genome analysis and high-throughput immunological screening identifies tuzin as a novel vaccine candidate against visceral leishmaniasis. Vaccine 32:3816–3822 (4);
iv) Sundar S, Singh B. 2014. Identifying vaccine targets for anti-leishmanial vaccine development. Expert Rev Vaccines. 13(4): 489–505 (5);
v) Alvar J, Croft S, Kaye P, Khamesipour A, Sundar S, Reed SG. 2013. Case study for a vaccine against leishmaniasis. Vaccine 31S B244–B249 (6);

vi) Mou Z et al. 2015 Identification of broadly conserved cross-species protective *Leishmania* antigen and its responding $CD4^+$ T cells. Sci. Transl. Med. 7, 310ra167 (7).

Table III. 1 Initial list of vaccine antigen candidates found in the literature.

Reference	Leishmania vaccine antigen candidates
(2)	Heat-shock proteins | PSA-2 and GP46 (surface glycoproteins) | A2 & HASPB | Cysteine proteases | LACK | CRK (creatine kinase) & MAPK (mitogen-activated protein kinase) | Histones | Glycolytic enzymes & SMT (sterol 24-c-methyltransferase) | Recombinant antigens (Leishmune/fucose mannose ligand, MML-MPL) | PFR | LPG | Initiation factors | GSH (glutathione) & TSH (trypanothione reductase) | LmSTI1 | TSA Topoisomerase and RIC (RNA import complex) | Lcr | Ldp23 KMP11 | Purine salvage system & nucleoside hydrolase (NH)
(3)	fucose mannose ligand (leishmune) | KMP11 | SMT | A2 | cysteine protease B | Lb EIF | HASPB | LACK | PSA | NH | GP63 | Leish-111f (TSA::STI1::EIF)
(4)	Tuzin | FGP | phospholipase A1-like protein (PLA1) | potassium voltage-gated channel protein (K VOLT)
(5)	Leish-111f |Leishmune (FML vaccine) | Leish-Tec (adenovirus with *L. donovani* A2 protein) | H1 | CPa+CPb | LACK | P0 | A2 | HASPB | LPG | LiESA/QA-21
(6)	LACK | Leishmanolysin (GP63) | TSA | STI1 | EIF | HASPB | SMT | KMP11 | A2 | Cysteine protease B | NH | Methionine aminopeptidase 45 | Protein disulfide isomerase | Elongation Factor-2
(7)	Glycosomal phosphoenolpyruvate carboxykinase (gPEPCK)

After listing all vaccine antigen candidates cited in the selected publications, a manual search was performed in the secretome datasets, either by protein accession or by protein name, as follows:

i) A word search by protein name was performed on the 3 already available secretome proteomic datasets from experiment AESLMTS_110414 (*L. infantum, L. major* and *L. tropica* datasets analysed with the UniProtKB database release 2011_04 (SwissProt_TrEMBL_2011_04: 15,082,690 sequences; 4,872,194,356 residues / *Leishmania* 42,035 sequences). All associated identifiers (protein entry numbers) were retrieved from all match groups and respective identified tryptic peptides, so that each protein contains at least 3 species-specific sequences (protein x = ID1 + ID2 + ID3 + …).

ii) The identified tryptic peptides were aligned with the UniProtKB protein sequences (through ClustalOmega alignments) to check if all peptides are found and to check for unique residues or peptides.

iii) For *L. tropica*, with less annotated proteins in UniProtKB, proteins are often identified as one of the species with complete or reference proteome in UniProtKB. To include the maximal number of species-specific sequences and include *L. tropica*-specific sequences, the coding genes were searched in the TriTrypDB. The corresponding translated proteins were retrieved and aligned with identified tryptic peptides – if the sequences contained the tryptic peptides, they were included as *L. tropica* species-specific sequence.

iv) The ortholog sequences for the remaining 3 species datasets (*L. donovani*, *L. braziliensis*, *L. amazonensis*) were retrieved with a UniProtKB BLAST search. Briefly, the *L. infantum* accession for each protein of interest was 'blasted' to find other *Leishmania*-specific sequences through sequence similarity. All '*Leishmania*' accessions were retrieved, including protein accession numbers, organism, and sequence identity (%).

v) All accession numbers per protein of interest were searched in the secretome proteomic datasets generated by the analysis with the UniProtKB release 2016_05, including the additional species *L. infantum*, *L. donovani*, *L. braziliensis*, *L. amazonensis* and *L. tarentolae* (experiments AESLMTS_110414, AESLMTS_151026, AESLMTS_160115). All identified tryptic peptide counts per sequence were retrieved, as well as respective identification status assigned by the script leading v2.2 (LEADING_AUTO / LEADING_SPECIES / Multispecies). Species-specific sequences with higher homology levels and containing the highest number of identified peptides were included.

vi) The final list for protein Set A contains 4 to 6 species-specific sequences per protein antigen, meaning only proteins present in all studied species were selected;

2.3. Protein set B - reverse vaccinology approach for antigen selection

A total of 618 accessions were identified in all 6 proteomic datasets from pathogenic *Leishmania* species (UniProtKB release 2016_05) (Figure III.1 panel A).

All 618 protein sequences were submitted to BLASTp analysis against the human proteome (*Homo sapiens* taxid:9606 RefSeq protein database from ftp.ncbi.nlm.nih.gov/genomes/H_sapiens/protein/). This analysis was performed using the standalone BLAST tool for Windows™, with imported search parameters from the online version, to keep standard alignment parameters (Figure III.1 panel B). Briefly, to import the default BLASTp search strategy, one random accession was run on the BLASTp online tool, against Human RefSeq proteins, with standard algorithm parameters. The search strategy was saved to a local directory as an ".asn" file and used in the standalone BLAST analysis (Figure III.1 panel B).

The alignment results were retrieved using the tabular output format is a ".csv" file (*outfmt*=6), containing the columns (Figure III.1 panel B): query sequence name (*qseqid*), length (*qlen*), sequence (*qseq*); subject sequence name (*sseqid*), alignment length (*length*), start of alignment in query (*qstart*), end of alignment in query (*qend*), expect value (*evalue*), bit score (*bitscore*), percentage of identical matches (*pident*), number of identical matches (*nident*), percentage of positive-scoring matches (*ppos*), number of positive-scoring matches (*positive*).

All accessions with "no hits found" were included in the protein set B.

Sequences were retrieved from UniProtKB and aligned to identify protein groups. In order to complete the protein groups with species-specific sequences, accessions per protein group were 'blasted' in UniProtKB. Species-specific sequences with higher homology levels and containing the highest number of identified peptides were included, resulting in 4 to 6 species-specific sequences per protein.

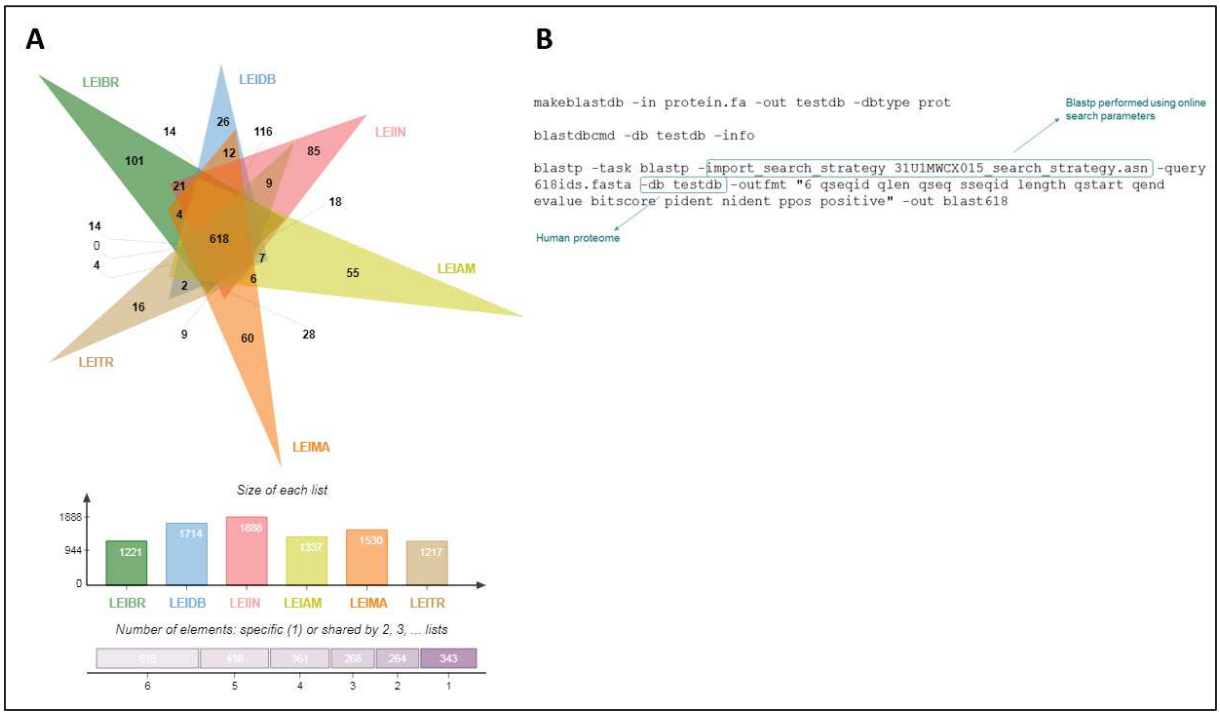


Figure III. 1 A reverse vaccinology approach to find novel excreted-secreted* Leishmania *antigen candidates. *A) Venn diagram comparing total identified proteins from 6* Leishmania *species* – L. infantum *(LEIIN),* L. major *(LEIMA),* L. donovani *(LEIDB),* L. braziliensis *(LEIBR),* L. tropica *(LEITR) and* L. amazonensis *(LEIAM) showing 618 accessions were identified in all samples. B) Standalone blastp code (MS-DOS) used to blast 618 common accessions against the human proteome (RefSeq proteome) using default parameters from the online BLASTp server.*

2.4. Evaluation of protein relative abundance with iBAQ

To rank the relative abundance of different proteins, an intensity-based absolute quantification (iBAQ) algorithm was used. Raw data files were analysed with the MaxQuant software package (version 1.5.5.0). As detailed in the previous chapter, the iBAQ value is obtained by dividing peptide intensities by the number of theoretically observable peptides of the protein (all fully tryptic peptides with 6 to 30 amino acids calculated by *in silico* protein digestion).

The iBAQ values corresponding to 2333 LEADING_CHECK accessions were retrieved. In order to estimate relative antigen abundance of protein sets A and B, absolute iBAQ values (2333 total LEADING_CHECK accessions, 88 set A accessions and 68 set B accessions) were log-

transformed (log10) and normalised to overall iBAQ average (average iBAQ value $2,4x10^7$; log10 average iBAQ value = 6,3302).

3. Results

3.1. Around 40% of previously described *Leishmania* vaccine antigens are found in the secretome (protein set A)

All previously described vaccine antigens were manually searched in the secretome datasets, either by protein accession or protein name (Table III.1). Only proteins present in all studied species were included in protein set A.

An initial list of 72 protein antigens was retrieved from the literature (Table III.2). These candidates include kinases, translation initiation factors, several proteases and other virulence factors.

The 72 proteins listed from the database (Table III.2) were manually searched in the proteomic datasets. From these, 36 were found in the secretome, and 28 are present in all tested species which corresponds roughly to 40% of the initial list (Table III.3).

Table III. 2 A total of 72 protein antigens listed from the selected publications was searched in the secretome proteomic datasets (2–7).

Antigen category	Individual proteins to search in proteomic datasets	Description
Heat-shock proteins	HSP70	Heat-shock proteins (HSP) were one of the first proposed antigen candidates. These proteins are encoded by a single copy gene, and the temperature shift the parasites undergo increases HSP expression, to adapt to survival in the vertebrate host. HSP90 is the most abundant form, with 17 tandem gene copies. HSP70 has been identified has the immunodominant antigen in antibody responses against *Leishmania* and has shown to stimulate DC maturation and cytokine production. HSP90 and HSP70 are B-cell mitogens.
	HSP71	
	HSP60 mitochondrial	
	HSP90	
Enzymes involved in the trypanothione antioxidant system	Trypanothione reductase	A recombinant *L. donovani* trypanothione reductase (LdTPR) induces lymphoproliferation and NO production *in vitro*, and protects hamsters against challenge, but also, in human samples from active and cured VL patients. LdTPR induces PBMC lymphoproliferation, even if slightly milder. In cured VL samples, LdTPR also induced IL-12 and IFN-γ production, accompanied by low IL-10 levels, contrary to soluble *L. donovani* antigens, which induce a mixed Th1/Th2 response with higher IL-10 levels and low IFN-γ.
	Peroxidoxin 1 and 2	Pxn1 and Pxn2 tested in mice exhibit differences in the elicited response. Both are immunogenic, however Pxn1 induced a predominant Th2 response and high IgG1 whereas Pxn2 induced a mixed Th1/Th2 response with high IgG2a antibody levels.
	Tryparedoxin peroxidase (also known as TSA, thiol-specific antioxidant antigen) and mitochondrial tryparedoxin peroxidase	3 forms in *L. infantum* with different immunological properties: cytosolic tryparedoxin peroxidase (LicTXNPx), mitochondrial tryparedoxin peroxidase (LimTXNPx), and tryparedoxin (LiTXN1). LiTXNPx elicits strong humoral response and has no influence on cytokine production. LmTXNPx, non-secreted, decreases IL-4 secretion both *in vitro* and *in vivo*. LiTXN1 is poorly immunogenic and promotes induces IL-10 secretion both in vitro and in vivo favouring parasite internalization and survival. TSA elicits strong Th1 responses in infected BALB/c mice infected by L. major, and there is evidence for its immunogenicity in humans. TSA is abundantly present and homogeneously distributed on the promastigote surface. TSA-based DNA vaccine induced cytotoxicity and increased levels of IgG1 and IgG2a.
Surface glycoprotein	GP46	GP46, also known as M-2, is a promastigote surface membrane glycoprotein. When administered with adjuvant, it induces protective cellular immune responses.
	Promastigote surface antigen-2	Promastigote Surface Antigen (PSA) exists in soluble and membrane-anchored form, and contains a leucine rich repeat (LRR) region which is an immunogenic epitope after protein fragmentation, inducing IFN-γ production and Th1 responses. Sequence analysis between PSA molecules from 9 different *Leishmania* species revealed highly conserved segments in the N-terminal region, while the central LRR domain and C-terminal regions are more divergent
	Putative surface antigen protein 2	
	Putative surface antigen protein	
Flagellar and flagelle-associated proteins	PFR-2C (paraflagellar rod protein)	Flagelar proteins are upregulated in promastigote forms. PFR-2 is a highly conserved immunogen inducing a Th1 immune response in dogs and mice. FPG was identified in a high-throughput *in vitro* immunological screening (with mouse splenocytes) and induced a predominant Th1 response (IL-12 and IFN-γ). Lcr1 is present in *L. infantum* amastigotes, it stimulates Th1 cytokines (IFN-γ) and antibodies, partially protecting BALB/c mice.
	PFR-1D (paraflagellar rod protein)	
	Flagellar glycoprotein-like protein (FPG)	
	Lcr1 peptide (T-cell proliferation stimulating peptide)	

Table III.2 (continued)

Antigen category	Individual proteins to search in proteomic datasets	Description
Histones	H2A	Histones are evolutionary conserved DNA-binding nuclear proteins. In *Leishmania*, these present relatively low homology with human proteins and are stably expressed by both amastigotes and promastigotes. H2B has been shown to be a promising candidate, inducing PBMC proliferation, IFN-γ production, promoting a Th1 protective immune response. H1 from *L. infantum* and *L. braziliensis* induce humoral responses
	H2B	
	H3	
	H4	
Translation initiation factors	LeIF (Leishmania elongation initiation factor)	LeIF protein belongs to the DEAD box protein family, it is homologous to eIF4A. There is evidence LeIF functions as a Th1-type natural adjuvant, mimicking IL-12-mediated downregulation of IL-4 production in lymph node cultures.
	Putative eukaryotic translation initiation factor 2 subunit	
	Putative eukaryotic translation initiation factor 3 subunit 8	
	Putative eukaryotic translation initiation factor 3 subunit 7	
	Probable eukaryotic initiation factor 4A; Short=eIF-4A; EC=3.6.4.13	
	Eukaryotic initiation factor 5a	
Glycolytic enzymes	Hexokinase (EC=2.7.1.1)	*Leishmania* glycolytic enzymes are compartmentalized in glycosomes, unique trypanosomatid organelles, and are phylogenetically distant from mammalian host enzymes, making them useful vaccine antigen candidates.
	Glucose-6-phosphate isomerase	
	Phosphofructokinase	
	Fructose-biphosphate aldolase	
	Triosephosphate isomerase	
	Glyceraldehyde-3-phosphate dehydrogenase	
	Phosphoglycerate kinase and mutase	
	Phosphopyrvate hydratase (Enolase)	
	Pyruvate kinase (EC=2.7.1.40)	
Other kinases	mitogen activated protein (MAP) kinase (Putative MAP kinase)	
	CRK (creatine kinase)	
	Putative fucose kinase; EC=2.7.1.52;	
	Glycosomal phosphoenolpyruvate carboxykinase (EC=4.1.1.-)	
	Putative nima-related protein kinase; EC=2.7.11.1;	
Enzymes involved in the purine salvage pathway	Nucleoside diphosphate kinase; EC=2.7.4.6;	*Leishmania* parasites depend upon the purine salvage system to utilize purine bases from their mammalian hosts; for this purpose they utilize a variety of nucleoside transporters that can be the potent targets for vaccine development. During parasite-delayed macrophage apoptosis, *Leishmania* nucleoside diphosphate kinase can inhibit ATP binding to P2X receptors thereby preventing apoptosis. The nucleoside hydrolase (NH) is a glycoprotein, part of the *Leishmania* fucose-mannose ligand (FML) complex, and essential for early infection establishment. NH is a known vaccine antigen, known to induce protective immune responses in mice and dogs against experimental infection.
	Putative adenosine kinase; EC=2.7.1.20;	
	Nucleotide hydrolase	
Topoisomerases and other isomerases	DNA topoisomerase	DNA topoisomerases in the kinetoplastid parasites are mainly involved in kDNA replication, essential for the survival of parasite. These are, however, distinct from other eukaryotic counterparts, making them interesting antigen candidates.
	Protein disulfide isomerase (PDI)	PDI catalyses thiol-disulfide interchange to prevent cell toxicity associated with ER stress and protein misfolding.

Table III.2 (continued)

Antigen category	Individual proteins to search in proteomic datasets	Description
Cyclophilins (peptidyl-prolyl-isomerases)	Peptidyl-prolyl cis-trans isomerase; EC=5.2.1.8;	Cyclophilins possess peptidyl-prolyl cis-trans isomerase activity, can perform different functions and are known receptors for cyclosporine, an immunosuppressive drug. *Leishmania* presents different cyclophilin isoforms, not fully characterized. Immunization with the *L. infantum* recombinant cyclophilin protein-1 confers partial protection in mice and generates specific memory T cells.
	Peptidyl-prolyl cis-trans isomerase; EC=5.2.1.8; (Cyclophilin-40)	
Sterol 24-c methyltransferase and other transferases	SMT (sterol 24-c-methyltransferase)	SMT is involved in the final step of the sterol biosynthetic pathway in *Leishmania*, leading to the production of ergosterol.
	arginine methyltransferase	
	Putative carnitine/choline acetyltransferase; EC=2.3.1.-	
	Aspartate aminotransferase	
	serine hydroxy methyl transferase	
Proteases	Cysteine protease A and B	*Leishmania* cysteine protease is the major protease involved in survival and adaptation to the host cell. Cysteine proteases have been shown to stimulate Th2 responses (IL-4 and IL-1). However, *L. mexicana* cysteine protease is a T-cell immunogen causes the development of protective Th1 responses (IL-12 and IFN-γ). GP63 or major surface protease (MSP) is a zinc protease and a key virulence factor responsible for evasion of the immune system, and degradation of fibronectin enhancing movement within the connective tissue (see chapter I). The cpL and cpB cysteine proteases are responsible for the lack of MHC-II invariant chain, therefore, modulating the cytokine production (increased IL-4 and IL-1) to the parasite's benefit. Also, the human host's papain-like cysteine proteases cathepsin A-like and B-like proteases, are targets of *Leishmania* immunomodulation and induction of Th2 responses. Ldp45 is involved in protein maturation. It is recognized by human T-cells and protects hamsters when formulated with BCG.
	GP63 (surface metalloprotease) / MSP (major surface protease) / leishmanolysin	
	Cathepsin B-like (cpB) and cathepsin L-like (cpL) cysteine proteases	
	Ldp45 (Methionine aminopeptidase 45)	
	Putative aminopeptidase (Metallo-peptidase, clan mf, family m17; EC=3.4.11.)	
	Putative aminopeptidase; (Metallo-peptidase, clan ma(E), family m1; EC=3.4.-.-)	
	Putative aminopeptidase P (Metallo-peptidase, clan mg, family m24)	
	Putative carboxypeptidase (Metallo-peptidase, clan ma(E), family 32)	
	Putative thimet oligopeptidase (Metallo-peptidase, clan ma(E), family m3; EC=3.4.24.15)	
	Putative dipeptidyl-peptidase III (Metallo-peptidase, clan m-, family m49; EC=3.4.14.4)	
	Putative peptidyl dipeptidase (Metallo-peptidase, Clan MA(E), Family M3,putative, partial)	

Table III.2 (continued)

Antigen category	Individual proteins to search in proteomic datasets	Description
Other antigens	KMP11 (kinetoplastid membrane protein-11)	KMP11 is a widely known amastigote and promastigote antigen, non-covalently bound to LPG, and conserved in trypanosomatids. B and T cell immunogen. There is evidence it is processed by DCs and presented by MHC-II molecules, activating specific T cells. The IFN-γ produced is enough to foster TNF-α and NO production resulting in parasite death inside macrophages. In susceptible individuals, KMP11 induces IL-4, IL-10 and IL-13 production perpetuating macrophage infection.
	LACK (Leishmania analogue of the kinase receptor C)	LACK is present in both amastigotes and promastigotes, and is highly conserved among *Leishmania* species. It is homologue of mammalian RACK1. LACK belongs to the WD repeat protein family. It is important for DNA replication, RNA synthesis, signal transduction and cell cycle, and also in the differentiation from metacyclic promastigote to amastigote, as there is a minimum threshold of this protein to establish infection. LACK is known to induce Th2 responses (IL-4 and IL-10 production).
	Elongation factor 2	EF-2 is recognised by human T cells and provides protection in hamsters.
	Ldp23 or 23 kDa cell surface protei (probable 60S ribosomal protein)	Ldp23 is present on the surface of *L.donovani* and *L. major* amastigotes and promastigotes, it accelerates IFN-y production and inhibits IL-4 production.
	A-2	A2 is the major antigen in the Leishmune® canine vaccine, it is only present in amastigotes. It contains multiple copies of a 10 amino-acid repeat, and is known to induce Th1 responses and provide partial protection against *Leishmania* species.
	HASPB and HASPA (hydrophylic acylated surface proteins A and B)	*Leishmania* HASPB is a lipoprotein that is exported via an unconventional secretory pathway. Recombinant HASPB1 confers protection against experimental challenge in mice.

Table III. 3 Protein description of the 28 selected antigen candidates (Set A). *In yellow, the antigens used in current vaccine formulation. Significant differential expression information from Alcolea P. et al 2010. Transcriptomics throughout the life cycle of Leishmania infantum: high down-regulation rate in the amastigote stage. Int J Parasitol. 40(13):1497-516* (8).

Protein name	REFS	differential expression in stationary phase promastigotes VS amastigotes?	Protein function (associated GO terms)
Flagellar glycoprotein-like protein (FPG)	1, 3		integral component of membrane [GO:0016021]; motile cilium [GO:0031514]
Trypanothione reductase	1	-25,7 +- 1,4	cell [GO:0005623]; disulfide oxidoreductase activity [GO:0015036]; flavin adenine dinucleotide binding [GO:0050660]; trypanothione-disulfide reductase activity [GO:0015042]; cell redox homeostasis [GO:0045454]
Cysteine protease_cathepsinB-like	1, 4	-	cysteine-type peptidase activity [GO:0008234]
LmSTI1 (Lm stress-inducible 1) homolog	1 - 5	-	
TSA (thiol-specific antigen) / Tryparedoxin peroxidase	1 - 5	-	peroxidase activity [GO:0004601]; peroxiredoxin activity [GO:0051920]
LACK (leishmania analogue of the kinase receptor C)	1 - 5	-	ribosome [GO:0005840]; kinase activity [GO:0016301]; regulation of cytokinesis [GO:0032465]
Protein disulfide isomerase-2	4, 5	-	endoplasmic reticulum [GO:0005783]; protein disulfide isomerase activity [GO:0003756]; cell redox homeostasis [GO:0045454]
KINASES			
Nucleoside diphosphate kinase; EC=2.7.4.6; (Purine salvage system)	1	14+-0	ATP binding [GO:0005524]; nucleoside diphosphate kinase activity [GO:0004550]; CTP biosynthetic process [GO:0006241]; GTP biosynthetic process [GO:0006183]; UTP biosynthetic process [GO:0006228]
Putative adenosine kinase; EC=2.7.1.20; (Purine metabolism)	1	-	adenosine kinase activity [GO:0004001]; purine ribonucleoside salvage [GO:0006166]
Pyruvate kinase; EC=2.7.1.40; (glycolytic enzyme)	1	-	kinase activity [GO:0016301]; magnesium ion binding [GO:0000287]; potassium ion binding [GO:0030955]; pyruvate kinase activity [GO:0004743]
Putative glycosomal phosphoenolpyruvate carboxykinase; EC=4.1.1.-;	6	N.D.	ATP binding [GO:0005524]; kinase activity [GO:0016301]; phosphoenolpyruvate carboxykinase (ATP) activity [GO:0004612]; gluconeogenesis [GO:0006094]
Putative hexokinase; EC=2.7.1.1; (glycolytic enzyme)	1		cell [GO:0005623]; ATP binding [GO:0005524]; glucose binding [GO:0005536]; hexokinase activity [GO:0004396]; cellular glucose homeostasis [GO:0001678]; glycolytic process [GO:0006096]
INITIATION FACTORS			
Putative eukaryotic translation initiation factor 2 subunit	1, 2	N.D.	GTP binding [GO:0005525]; GTPase activity [GO:0003924]; translation initiation factor activity [GO:0003743]

Table III.3 (continued)

Protein name	REFS	differential expression in stat.promast vs amastigotes? [Alcolea PJ 2010]	Protein function (associated GO terms)
Putative eukaryotic translation initiation factor 5a	1, 2	N.D.	ribosome binding [GO:0043022]; translation elongation factor activity [GO:0003746]; translation initiation factor activity [GO:0003743]; positive regulation of translational elongation [GO:0045901]; positive regulation of translational termination [GO:0045905]; translational frameshifting [GO:0006452]
eIF4A	1, 2, 5	-	ATP binding [GO:0005524]; helicase activity [GO:0004386]; translation initiation factor activity [GO:0003743]
HISTONES			
Histone2A	1, 4	-	nucleosome [GO:0000786]; nucleus [GO:0005634]; DNA binding [GO:0003677]
Histone H2B	1, 4	2,6 +- 0,2	nucleosome [GO:0000786]; nucleus [GO:0005634]; DNA binding [GO:0003677]
Histone H4	1, 4	-	nucleosome [GO:0000786]; nucleus [GO:0005634]; DNA binding [GO:0003677]
OTHER			
Triosephosphate isomerase	1	-	triose-phosphate isomerase activity [GO:0004807]; gluconeogenesis [GO:0006094]; glycolytic process [GO:0006096]; pentose-phosphate shunt [GO:0006098]
Nucleoside hydrolase	1, 2, 5	-	hydrolase activity [GO:0016787]
PFR-2C	1	-	motile cilium [GO:0031514]
HSP60		-	cytoplasm [GO:0005737]; ATP binding [GO:0005524]; protein refolding [GO:0042026]
GP63	5	-	membrane [GO:0016020]; metalloendopeptidase activity [GO:0004222]; cell adhesion [GO:0007155]
Metallo-peptidase, clan mf, family m17); EC=3.4.11.-;		-	cytoplasm [GO:0005737]; aminopeptidase activity [GO:0004177]; manganese ion binding [GO:0030145]; metalloexopeptidase activity [GO:0008235]
Putative aminopeptidase (Metallo-peptidase, clan ma(E), family m1); EC=3.4.-.-;		-	aminopeptidase activity [GO:0004177]; metallopeptidase activity [GO:0008237]; zinc ion binding [GO:0008270]
Putative dipeptidyl-peptidase III (Metallo-peptidase, clan m-, family m49); EC=3.4.14.4;		-	cytoplasm [GO:0005737]; dipeptidyl-peptidase activity [GO:0008239]; metal ion binding [GO:0046872]
Putative thimet oligopeptidase (Metallo-peptidase, clan ma(E), family m3); EC=3.4.24.15;			metal ion binding [GO:0046872]; metalloendopeptidase activity [GO:0004222]
Elongation factor 2	1	-	GTP binding [GO:0005525]; GTPase activity [GO:0003924]; translation elongation factor activity [GO:0003746]
KMP11	1 - 5		
Calpain-like cysteine protease	1		

At times, because proteins were manually searched by protein name, redundant proteins and isoform sequences are found, as for the case of peroxidoxins. Through sequence alignment and gene IDs three groups for peroxidoxins were generated according to their function in the anti-oxidation pathway – tryparedoxin peroxidase, peroxidoxin and tryparedoxin (Table III.4).

Table III. 4 Deconvolution of peroxidoxin sequences annotated in the UniProt and Genbank databases. *Underlined sequences were selected.*

Species	Name	UniProt Acc Nr	UniProt GeneID	Genbank ID
L. major	thiol-specific antioxidant antigen	**A9LJZ6_LEIMA**	TSA	ABX11567.1.
	thiol-specific antioxidant antigen	**Q4QF68_LEIMA**	TRYP7	AAC31146.1
	tryparedoxin peroxidase	**Q4QF74_LEIMA**	TRYP4	CAJ03332.1
	tryparedoxin peroxidase	**Q4QF80_LEIMA**	TRYP1	CAJ03330.1
	tryparedoxin peroxidase	**Q4QF76_LEIMA**	TRYP3	CAJ03334.1
	tryparedoxin peroxidase	**A9LNR9_LEIMA**	TRYP6	ABX26130.1
	peroxidoxin	**Q9TZS4_LEIMA**	N/A	AAC79432.1
	peroxidoxin	**Q4QBH2_LEIMA**	LMJF_23_0040	CAJ03825.1
	Tryparedoxin	**E9ADX3_LEIMA**	TXN2 / LMJF_29_1150	CBZ12452.1
L. infantum	thiol-specific antioxidant antigen	**Q8MU50_LEIIN**	N/A	AAK58478.1
	tryparedoxin peroxidase	**A4HWK3_LEIIN**	TRYP / LINJ_15_1120	CAM66832.1
	tryparedoxin peroxidase	**A4HWK2_LEIIN**	TRYP / LINJ_15_1100	CAM66830.1
	cytosolic peroxiredoxin	**Q95NF5_LEIIN**	TRYP / LINJ_15_1140	AAL25847.1
	putative mitochondrial peroxiredoxin	**Q95U89_LEIIN**	mTXNPx / LINJ_23_0050	AAL25846.1
	Tryparedoxin	**Q6RYT3_LEIIN**	TXN1 / LINJ_29_1250	AAS48350.1./ CAM69684.1.
L. tropica	**peroxidoxin 1**	Q07DU6_LEITR	Pxn1	AAZ23600.1
	peroxidoxin 2	Q07DU5_LEITR	Pxn2	AAZ23601.1

Each protein antigen in set A includes 4 to 6 species-specific sequences (Table III.5). The proteins with 5 and/or 6 species-specific sequences include proteins specific to *L. tropica* and/or *L. amazonensis* samples, which are automatically identified as one of the reference proteome species by the bioinformatics analysis. The gene search in TriTrypDB and respective translated proteins allowed the inclusion of species-specific sequences.

Finally, the total 28 protein antigens from protein set A include 3 protein antigens with 4 species-specific sequences, 14 protein antigens with 5 species-specific sequences, and 11 protein antigens with 6 species-specific sequences (Table III.5).

In summary, the protein set A includes i) potential vaccine antigens; ii) antigens present in the secretome; and iii) antigens common to six pathogenic *Leishmania* species.

Table III. 5 Protein Set A includes 28 antigenic proteins (A1 to A32) described in the literature as potential vaccine antigens against* Leishmania *spp. *Each protein antigen includes 4 to 6 species-specific sequences. The peptide counts allowing the identification of each sequence in the several samples are shown, as well as the protein length, percentage of identity among sequences and the identification (ID) status according the bioinformatic script used (see chapter II).* L. braziliensis *(LEIBR),* L. donovani *(LEIDB),* L. infantum *(LEIIN),* L. amazonensis *(LEIAM),* L. major *(LEIMA) and* L. tropica *(LEITR).*

protein	length	%identity	LEIBR	LEIDB	LEIIN	LEIAM	LEIMA	LEITR	ID status
A1	593	100.0%	1	2	4	0	0	2	LEADING_SPECIES
	597	99.0%	1	2	4	0	0	2	Match group (LEADING_SPECIES)
	573	85.8%	0	0	0	0	5	0	LEADING_AUTO
	602	82.8%	0	0	0	2	0	0	LEADING_AUTO
A2	491	100.0%	4	19	19	1	7	4	Match group (LEADING_AUTO)
	491	99.6%	4	19	19	1	7	4	LEADING_AUTO
	491	95.7%	3	7	6	2	15	4	LEADING_AUTO
	476	90.8%	1	0	0	12	2	0	Match group (LEADING_AUTO)
	491	84.1%	11	3	3	1	2	0	LEADING_AUTO
A4	340	100.0%	0	11	5	2	3	2	Match group (LEADING_AUTO)
	340	100.0%	0	11	5	2	3	2	LEADING_SPECIES
	340	99.7%	0	10	4	2	4	2	LEADING_AUTO
	340	91.5%	1	4	2	1	7	1	LEADING_AUTO
	340	85.3%	0	2	1	4	1	1	MULTISPECIES
	340	68.6%	3	0	0	0	0	0	LEADING_SPECIES
A5	255	100.0%	2	6	3	3	5	4	LEADING_SPECIES
	255	100.0%	2	6	3	3	5	4	Match group (LEADING_SPECIES)
	255	97.6%	2	6	3	3	5	4	Match group (LEADING_SPECIES)
	255	94.5%	2	3	1	5	3	2	LEADING_AUTO
	255	84.6%	5	1	1	1	2	1	LEADING_AUTO
A6	199	100.0%	1	9	9	4	4	6	Match group (LEADING_AUTO)
	199	99.5%	1	9	9	4	4	6	LEADING_AUTO
	199	94.9%	2	6	6	5	4	10	LEADING_AUTO
	199	91.4%	1	3	5	5	12	5	MULTISPECIES
	199	89.4%*	1	5	5	9	5	6	Match group (LEADING_AUTO)
	199	83.8%*	7	4	2	6	1	2	LEADING_AUTO
A7	312	100.0%	3	10	8	3	8	6	Match group (LEADING_AUTO)
	312	100.0%	3	10	8	3	8	6	LEADING_AUTO
	312	99.7%	4	9	8	2	9	7	Match group (LEADING_AUTO)
	312	99.7%	3	10	8	3	8	6	Match group (LEADING_AUTO)
	312	99.4%	4	9	8	2	9	7	Match group (LEADING_AUTO)
	312	97.1%	5	6	5	1	6	4	LEADING_AUTO
A8	477	100.0%	1	14	6	3	5	0	Match group (LEADING_AUTO)
	477	99.4%	1	14	6	3	5	0	Match group (LEADING_AUTO)
	477	93.7%	1	7	5	4	11	1	Match group (MULTISPECIES)
	473	87.4%	1	3	3	4	5	2	Match group (LEADING_AUTO)
	473	86.6%	1	3	3	4	5	2	LEADING_AUTO
	470	77.8%	10	1	0	0	0	0	LEADING_AUTO
A9	151	100.0%	2	13	12	7	11	13	Match group (LEADING_SPECIES)
	151	100.0%	2	13	12	7	11	13	LEADING_SPECIES
	151	99.3%	2	12	11	6	12	12	LEADING_AUTO
	151	94.0%	1	8	6	8	5	6	LEADING_AUTO
	151	90.7%	9	4	2	3	3	4	MULTISPECIES
A10	345	100.0%	0	3	6	1	3	6	LEADING_SPECIES
	345	97.4%	0	2	3	1	5	3	LEADING_AUTO
	345	97.4%	0	3	6	1	3	6	Match group (LEADING_SPECIES)
	345	94.5%	0	3	6	1	3	6	Match group (LEADING_SPECIES)
	345	84.6%	2	1	0	0	1	0	LEADING_AUTO
A11	507	100.0%	6	19	12	8	10	10	LEADING_SPECIES
	499	99.8%	6	19	12	8	10	10	Match group (LEADING_SPECIES)
	499	98.6%	7	15	11	8	13	11	LEADING_AUTO
	498	92.3%	6	13	8	10	9	8	LEADING_AUTO
	454	92.7%	16	5	2	2	3	2	LEADING_AUTO
	524	98.1%	7	15	11	8	13	11	Match group (LEADING_AUTO)
A12	525	100.0%	3	11	8	1	3	3	LEADING_AUTO
	525	99.2%	3	11	8	1	3	3	Match group (LEADING_AUTO)
	525	97.9%	5	7	5	2	5	3	MULTISPECIES
	524	92.0%	7	3	3	2	5	3	LEADING_AUTO
	524	91.2%	8	2	2	2	4	3	LEADING_AUTO
A14	471	100.0%	3	12	11	2	8	3	Match group (LEADING_AUTO)
	471	99.8%	3	12	11	2	8	3	Match group (LEADING_AUTO)
	471	98.7%	3	12	11	2	8	3	Match group (LEADING_AUTO)
	471	98.5%	3	12	11	2	8	3	Match group (LEADING_AUTO)
	471	90.0%	6	3	3	0	1	1	LEADING_AUTO
	434	99.5%	3	12	11	2	8	3	Match group (LEADING_AUTO)
A16	257	100.0%	2	2	4	2	2	2	Match group (LEADING_AUTO)
	166	100.0%	2	2	4	2	2	2	Match group (LEADING_AUTO)
	166	98.2%	2	2	4	2	2	2	Match group (LEADING_AUTO)
	166	97.6%	2	2	4	2	2	2	Match group (LEADING_AUTO)
	166	97.0%	2	2	4	2	2	2	Match group (LEADING_AUTO)

Table III.5 (continued)

protein	length	%identity	LEIBR	LEIDB	LEIIN	LEIAM	LEIMA	LEITR	ID status
A18	135	100.0%	0	4	3	2	3	3	MULTISPECIES
	107	100.0%	0	4	3	2	3	3	Match group (MULTISPECIES)
	111	97.2%	0	3	3	2	5	5	MULTISPECIES
	111	94.4%	0	3	3	2	5	5	MULTISPECIES
	107	96.3%	0	4	3	2	3	3	Match group (MULTISPECIES)
	182	75.8%	4	0	0	0	1	0	Match group (LEADING_AUTO)
A19	100	100	2	9	6	3	7	4	LEADING_SPECIES
	100	98	2	9	6	3	7	4	Match group (LEADING_SPECIES)
	100	94	7	3	2	5	3	3	LEADING_AUTO
	100	92	7	3	2	5	3	3	Match group (LEADING_AUTO)
	100	98	2	7	5	3	7	4	LEADING_AUTO
A20	251	100.0%	3	10	10	5	5	5	LEADING_SPECIES
	252	99.2%	3	10	10	5	5	5	Match group (LEADING_SPECIES)
	251	98.0%	3	7	7	4	8	7	LEADING_AUTO
	251	91.2%	7	2	3	4	2	2	LEADING_SPECIES
	251	94.4%	3	5	6	9	4	4	LEADING_AUTO
A23	403	100.0%	10	11	15	7	11	7	Match group (LEADING_SPECIES)
	456	100.0%	10	11	15	7	11	7	Match group (LEADING_SPECIES)
	403	100.0%	10	11	15	7	11	7	Match group (LEADING_SPECIES)
	403	100.0%	10	11	15	7	11	7	LEADING_SPECIES
	403	98.3%	12	10	13	7	10	6	LEADING_AUTO
A24	352	100.0%	1	7	8	4	7	3	MULTISPECIES
	352	99.7%	1	7	8	4	7	3	Match group (MULTISPECIES)
	352	96.9%	1	6	6	4	9	3	LEADING_AUTO
	320	98.4%	1	6	6	4	9	3	Match group (LEADING_AUTO)
	352	87.2%	3	1	1	1	1	0	LEADING_AUTO
A25	599	100.0%	1	6	6	1	4	3	LEADING_SPECIES
	599	100.0%	1	6	6	1	4	3	Match group (LEADING_SPECIES)
	599	99.8%	1	6	6	1	4	3	Match group (LEADING_SPECIES)
	599	98.8%	0	0	0	0	1	0	peptide manual alignment
A26	562	100.0%	6	6	11	4	3	3	LEADING_AUTO
	562	98.6%	6	6	11	4	3	3	Match group (LEADING_AUTO)
	565	98.4%	6	6	11	4	3	3	Match group (LEADING_AUTO)
	562	94.7%	6	6	11	4	3	3	Match group (LEADING_AUTO)
	566	87.2%	3	7	14	3	3	3	LEADING_SPECIES
	566	87.2%	3	7	14	3	3	3	Match group (LEADING_SPECIES)
A27	599	100.0%	2	19	12	3	6	5	Match group (LEADING_AUTO)
	599	99.5%	2	19	12	3	6	5	Match group (LEADING_AUTO)
	602	86.0%	1	2	2	2	7	2	Match group (MULTISPECIES)
	602	81.7%	1	3	3	11	1	3	LEADING_AUTO
	657	79.6%	1	3	4	3	2	7	LEADING_AUTO
	589	71.2%	12	0	1	1	1	2	LEADING_AUTO
A28	535	100.0%	3	11	7	2	8	4	LEADING_AUTO
	538	96.3%	3	7	4	2	11	6	LEADING_AUTO
	535	89.9%	2	2	1	3	2	1	LEADING_AUTO
	538	85.0%	13	2	1	2	3	1	LEADING_AUTO
A29	868	100.0%	2	10	4	3	3	3	LEADING_AUTO
	868	99.5%	2	12	3	3	2	2	LEADING_AUTO
	868	95.6%	1	7	3	2	4	3	LEADING_AUTO
	868	91.8%	1	6	1	3	1	1	LEADING_AUTO
	868	82.8%	14	1	0	0	0	0	LEADING_AUTO
A30	679	100.0%	6	24	30	8	13	8	LEADING_AUTO
	679	99.9%	6	25	29	8	13	8	LEADING_AUTO
	679	96.9%	5	17	23	6	16	10	LEADING_AUTO
	679	94.7%	6	15	19	11	11	7	LEADING_AUTO
	679	89.0%	20	5	5	2	4	2	LEADING_SPECIES
A31	685	100.0%	1	13	7	7	5	4	LEADING_AUTO
	685	99.7%	3	15	6	7	5	3	LEADING_AUTO
	685	97.5%	1	8	5	8	8	4	LEADING_AUTO
	685	95.0%	2	7	4	12	4	3	LEADING_AUTO
	607	90.9%	4	4	2	4	1	2	LEADING_SPECIES
A32	845	100.0%	20	32	30	17	20	14	Match group (LEADING_SPECIES)
	845	100.0%	20	32	30	17	20	14	LEADING_SPECIES
	845	99.9%	20	31	29	17	21	15	LEADING_AUTO
	845	99.2%	19	27	24	20	15	10	LEADING_AUTO
	845	98.1%	27	23	22	13	15	13	MULTISPECIES
A34	92	100.0%	1	4	3	4	2	2	LEADING_SPECIES
	92	98.9%	1	4	3	4	2	2	Match group (LEADING_SPECIES)
	92	98.9%	1	4	3	4	2	2	Match group (LEADING_SPECIES)
	92	97.8%	1	4	3	4	2	2	Match group (LEADING_SPECIES)
	92	96.7%	1	4	3	4	2	2	Match group (LEADING_SPECIES)
	92	94.6%	3	1	1	1	1	1	MULTISPECIES
A35	115	100.0%	1	9	9	5	7	4	Match group (LEADING_SPECIES)
	115	100.0%	1	9	9	5	7	4	LEADING_SPECIES
	115	99.1%	1	9	9	5	7	4	Match group (LEADING_SPECIES)
	115	98.3%	1	7	7	5	9	6	LEADING_AUTO
	115	91.3%	6	2	2	1	2	2	MULTISPECIES

3.2. A reverse vaccinology approach allowed the identification of 24 novel antigen candidates, including 3 antigens previously described in the literature

The BLASTp analysis of the common 618 accessions generated a total of 15720 alignment results for 542 unique protein sequences. Proteins with no significant homology with human proteins were selected – a total of 76 accessions with "no hits found". The 76 accessions include 43 annotated sequences, corresponding to 14 proteins, and 33 accessions corresponding to uncharacterised proteins (UP). These 33 UP sequences were aligned with the ClustalOmega alignment tool, and 10 clusters were identified (Table III.6).

Considering 14 proteins and 10 UP, protein set B contains 24 protein antigen candidates. The protein set B includes 7 protein antigens with 4 species-specific sequences, 16 proteins antigens with 5, and one protein antigen with 6 species-specific sequences (Table III.7).

Among the proteins selected through this approach, 4 accessions were already included in the protein set A. These accessions correspond to 3 different proteins: i) *L. infantum* nucleoside hydrolase-like protein and *L. major* putative inosine-guanine nucleoside hydrolase; ii) *L. infantum* paraflagellar rod protein 2C; and iii) *L. infantum* kinetoplastid membrane protein 11C.

Table III. 6 ClustalOmega alignment results (percent identity matrix) for 33 uncharacterized proteins identified with the reverse vaccinology approach

cluster1	**100,00**	**94,03**	11,97	16,11	19,88	17,33	16,96	17,83	13,92	14,56	12,50	16,13	16,13	16,13	14,29	10,00	10,00	13,30	16,67	12,62	11,50	7,14	7,14	10,00	9,09	4,83	8,54	8,54	10,13	9,60	11,69	0,00	33,33
	94,03	**100,00**	10,56	15,44	19,30	17,33	16,96	17,83	13,92	14,56	9,38	16,13	16,13	16,13	14,29	20,00	20,00	12,88	10,00	11,65	10,50	7,14	7,14	9,09	8,18	4,83	8,54	8,54	9,49	8,59	10,82	0,00	33,33
UP	11,97	10,56	**100,00**	10,74	16,03	26,67	18,80	19,62	15,24	14,63	4,55	9,09	9,09	9,09	9,52	3,33	3,33	8,22	13,33	9,85	9,96	10,20	6,12	7,03	6,25	9,26	9,91	9,44	8,85	12,32	10,71	3,23	4,84
UP	16,11	15,44	10,74	**100,00**	14,91	23,70	13,17	12,46	17,33	17,33	17,17	15,22	17,39	17,39	8,33	0,00	0,00	9,93	7,69	9,41	9,95	6,06	6,06	6,25	6,25	9,47	9,72	9,72	13,02	8,64	10,15	0,00	10,00
UP	19,88	19,30	16,03	14,91	**100,00**	13,87	17,97	18,30	17,62	17,62	11,70	11,63	11,63	11,63	0,00	14,29	14,29	12,59	10,64	11,81	14,40	5,68	6,82	10,77	9,23	6,57	6,72	6,72	10,60	12,62	9,88	0,00	0,00
UP	17,33	17,33	26,67	23,70	13,87	**100,00**	20,55	19,18	17,89	14,74	13,56	16,67	12,50	14,58	33,33	-nan	-nan	11,18	4,76	8,89	8,96	6,38	6,38	5,00	5,00	10,26	7,69	7,69	10,88	11,11	12,40	-nan	-nan
cluster2	16,96	16,96	18,80	13,17	17,97	20,55	**100,00**	**88,79**	20,62	20,06	15,00	14,68	14,68	14,68	12,00	10,69	12,21	10,70	1,92	13,75	13,54	10,00	9,00	11,62	11,62	10,36	12,69	12,91	9,19	14,09	12,12	7,78	6,67
	17,83	17,83	19,62	12,46	18,30	19,18	**88,79**	**100,00**	22,47	21,35	15,83	13,76	13,76	13,76	14,67	12,98	14,50	11,33	3,85	13,46	14,10	10,00	9,00	12,06	12,06	10,66	12,99	12,61	9,51	13,82	11,69	8,89	7,78
cluster3	13,92	13,92	15,24	17,33	17,62	17,89	20,62	22,47	**100,00**	**91,54**	17,80	21,50	22,43	20,56	0,00	10,00	10,00	10,91	6,25	8,78	8,23	8,42	8,42	10,28	10,28	11,52	10,45	10,45	5,31	13,39	9,59	0,00	0,00
	14,56	14,56	14,63	17,33	17,62	14,74	20,06	21,35	**91,54**	**100,00**	19,49	21,50	23,36	21,50	0,00	10,00	10,00	9,73	6,25	9,54	9,52	8,42	8,42	8,41	8,41	12,12	9,70	9,70	4,90	12,13	9,25	0,00	0,00
UP	12,50	9,38	4,55	17,17	11,70	13,56	15,00	15,83	17,80	19,49	**100,00**	32,46	35,09	35,96	-nan	0,00	0,00	9,26	8,33	10,00	11,43	16,13	16,13	7,69	7,69	7,69	22,22	22,22	5,95	9,62	5,19	-nan	-nan
cluster4	16,13	16,13	9,09	15,22	11,63	16,67	14,68	13,76	21,50	21,50	32,46	**100,00**	**88,70**	**85,22**	-nan	14,29	14,29	13,27	5,56	7,94	9,52	19,05	14,29	0,00	0,00	13,33	28,57	28,57	11,84	10,87	5,63	-nan	-nan
	16,13	16,13	9,09	17,39	11,63	12,50	14,68	13,76	22,43	23,36	35,09	**88,70**	**100,00**	**94,78**	-nan	14,29	14,29	13,27	5,56	7,94	9,52	19,05	14,29	0,00	0,00	13,33	42,86	42,86	10,53	13,04	5,63	-nan	-nan
	16,13	16,13	9,09	17,39	11,63	14,58	14,68	13,76	20,56	21,50	35,96	**85,22**	**94,78**	**100,00**	-nan	14,29	14,29	13,27	5,56	7,94	9,52	14,29	9,52	0,00	0,00	13,33	42,86	42,86	10,53	10,87	4,23	-nan	-nan
UP	14,29	14,29	9,52	8,33	0,00	33,33	12,00	14,67	0,00	0,00	-nan	-nan	-nan	-nan	**100,00**	13,75	12,50	17,58	19,35	19,88	19,77	0,00	5,00	10,94	10,94	10,34	11,35	12,06	8,96	14,71	14,77	15,00	13,33
cluster5	10,00	20,00	3,33	0,00	14,29	-nan	10,69	12,98	10,00	10,00	0,00	14,29	14,29	14,29	13,75	**100,00**	**95,11**	16,77	20,00	21,43	21,43	20,00	20,00	7,14	7,14	20,00	7,07	7,92	13,79	9,52	12,50	6,67	6,67
	10,00	20,00	3,33	0,00	14,29	-nan	12,21	14,50	10,00	10,00	0,00	14,29	14,29	14,29	12,50	**95,11**	**100,00**	17,39	20,00	22,08	22,08	20,00	20,00	7,14	7,14	20,00	6,06	6,93	13,79	9,52	11,36	6,67	6,67
UP	13,30	12,88	8,22	9,93	12,59	11,18	10,70	11,33	10,91	9,73	9,26	13,27	13,27	13,27	17,58	16,77	17,39	**100,00**	13,28	19,57	20,66	17,12	15,65	12,50	11,93	15,38	11,95	13,53	16,55	13,87	15,80	6,90	3,45
UP	16,67	10,00	13,33	7,69	10,64	4,76	1,92	3,85	6,25	6,25	8,33	5,56	5,56	5,56	19,35	20,00	20,00	13,28	**100,00**	25,78	25,58	2,86	2,86	17,39	13,04	18,75	9,68	12,90	15,71	6,76	11,88	0,00	14,29
cluster6	12,62	11,65	9,85	9,41	11,81	8,89	13,75	13,46	8,78	9,54	10,00	7,94	7,94	7,94	19,88	21,43	22,08	19,57	25,78	**100,00**	**84,25**	10,69	10,62	10,67	10,33	17,91	13,05	13,54	12,90	11,95	14,88	5,88	7,35
	11,50	10,50	9,96	9,95	14,40	8,96	13,54	14,10	8,23	9,52	11,43	9,52	9,52	9,52	19,77	21,43	22,08	20,66	25,58	**84,25**	**100,00**	12,50	12,42	9,88	9,88	18,03	13,50	14,29	13,08	11,97	14,29	5,88	5,88
cluster7	7,14	7,14	10,20	6,06	5,68	6,38	10,00	10,00	8,42	8,42	16,13	19,05	19,05	14,29	0,00	20,00	20,00	17,12	2,86	10,69	12,50	**100,00**	**92,67**	12,35	12,35	10,99	13,39	13,39	16,02	16,56	20,42	13,33	13,33
	7,14	7,14	6,12	6,06	6,82	6,38	9,00	9,00	8,42	8,42	16,13	14,29	14,29	9,52	5,00	20,00	20,00	15,65	2,86	10,62	12,42	**92,67**	**100,00**	12,35	12,35	12,09	12,50	12,50	16,48	16,56	20,83	13,33	13,33
cluster8	10,00	9,09	7,03	6,25	10,77	5,00	11,62	12,06	10,28	8,41	7,69	0,00	0,00	0,00	10,94	7,14	7,14	12,50	17,39	10,67	9,88	12,35	12,35	**100,00**	**97,60**	16,00	13,47	14,95	17,08	16,17	17,99	21,74	19,57
	9,09	8,18	6,25	6,25	9,23	5,00	11,62	12,06	10,28	8,41	7,69	0,00	0,00	0,00	10,94	7,14	7,14	11,93	13,04	10,33	9,88	12,35	12,35	**97,60**	**100,00**	14,86	13,99	13,92	17,08	16,17	18,52	21,74	19,57
UP	4,83	4,83	9,26	9,47	6,57	10,26	10,36	10,66	11,52	12,12	7,69	13,33	13,33	13,33	10,34	20,00	20,00	15,38	18,75	17,91	18,03	10,99	12,09	16,00	14,86	**100,00**	21,37	22,90	14,81	22,14	18,39	9,09	0,00
cluster9	8,54	8,54	9,91	9,72	6,72	7,69	12,69	12,99	10,45	9,70	22,22	28,57	42,86	42,86	11,35	7,07	6,06	11,95	9,68	13,05	13,50	13,39	12,50	13,47	13,99	21,37	**100,00**	**85,15**	13,04	16,63	18,38	10,48	11,43
	8,54	8,54	9,44	9,72	6,72	7,69	12,91	12,61	10,45	9,70	22,22	28,57	42,86	42,86	12,06	7,92	6,93	13,53	12,90	13,54	14,29	13,39	12,50	14,95	13,92	22,90	**85,15**	**100,00**	14,13	18,05	19,57	10,38	9,43
UP	10,13	9,49	8,85	13,02	10,60	10,88	9,19	9,51	5,31	4,90	5,95	11,84	10,53	10,53	8,96	13,79	13,79	16,55	15,71	12,90	13,08	16,02	16,48	17,08	17,08	14,81	13,04	14,13	**100,00**	17,14	19,49	11,11	11,11
UP	9,60	8,59	12,32	8,64	12,62	11,11	14,09	13,82	13,39	12,13	9,62	10,87	13,04	10,87	14,71	9,52	9,52	13,87	6,76	11,95	11,97	16,56	16,56	16,17	16,17	22,14	16,63	18,05	17,14	**100,00**	21,83	13,04	9,78
UP	11,69	10,82	10,71	10,15	9,88	12,40	12,12	11,69	9,59	9,25	5,19	5,63	5,63	4,23	14,77	12,50	11,36	15,80	11,88	14,88	14,29	20,42	20,83	17,99	18,52	18,39	18,38	19,57	19,49	21,83	**100,00**	24,79	23,14
cluster10	0,00	0,00	3,23	0,00	0,00	-nan	7,78	8,89	0,00	0,00	-nan	-nan	-nan	-nan	15,00	6,67	6,67	6,90	0,00	5,88	5,88	13,33	13,33	21,74	21,74	9,09	10,48	10,38	11,11	13,04	24,79	**100,00**	**87,60**
	33,33	33,33	4,84	10,00	0,00	-nan	6,67	7,78	0,00	0,00	-nan	-nan	-nan	-nan	13,33	6,67	6,67	3,45	14,29	7,35	5,88	13,33	13,33	19,57	19,57	0,00	11,43	9,43	11,11	9,78	23,14	**87,60**	**100,00**

Table III. 7 Protein set B includes 14 annotated proteins and 10 uncharacterised proteins (B1 to B24). *Each protein antigen includes 4 to 6 species-specific sequences. The peptide counts allowing the identification of each sequence in the several samples are shown, as well as the protein length, percentage of identity among sequences and the identification (ID) status according the bioinformatic script used (see chapter II).* L. braziliensis *(LEIBR),* L. donovani *(LEIDB),* L. infantum *(LEIIN),* L. amazonensis *(LEIAM),* L. major *(LEIMA) and* L. tropica *(LEITR).*

protein	length	%identity	LEIBR	LEIDB	LEIIN	LEIAM	LEIMA	LEITR	ID status
B1	597	99.5%	5	15	17	8	11	12	LEADING_AUTO
	641	83.5%	26	5	6	4	5	6	LEADING_AUTO
	643	100.0%	5	15	19	9	12	12	LEADING_AUTO
	636	91.2%	4	7	9	17	8	9	MULTISPECIES
	640	94.1%	5	10	14	8	16	17	LEADING_AUTO
	636	91.2%	4	7	9	17	8	9	MULTISPECIES
B2	499	99.8%	2	20	17	9	12	7	LEADING_AUTO
	498	85.3%	12	2	1	1	2	1	LEADING_AUTO
	499	100.0%	2	19	17	9	12	7	LEADING_AUTO
	499	96.4%	3	13	11	14	10	7	LEADING_AUTO
B3	503	99.6%	1	22	18	10	9	4	LEADING_AUTO
	503	100.0%	1	20	20	9	8	3	LEADING_AUTO
	503	95.6%	2	14	12	12	5	1	LEADING_AUTO
	503	95.4%	1	12	11	6	13	5	LEADING_AUTO
	503	80.5%	7	2	2	3	1	1	LEADING_AUTO
B4	1020	89.8%	4	5	6	3	5	3	LEADING_AUTO
	1020	100.0%	3	23	17	7	17	9	LEADING_SPECIES
	1020	97.3%	3	17	12	9	25	11	LEADING_AUTO
	1020	99.7%	3	23	17	7	17	9	Match group (LEADING_SPECIES)
B5	279	95.0%	4	3	4	3	4	3	LEADING_SPECIES
	279	100.0%	3	4	7	4	6	3	LEADING_AUTO
	279	98.2%	3	4	7	4	7	3	LEADING_AUTO
	279	99.6%	3	4	7	4	7	3	Match group (LEADING_AUTO)
	279	98.6%	3	4	7	4	7	3	Match group (LEADING_AUTO)
B6	710	100.0%	5	6	4	2	3	1	LEADING_AUTO
	710	99.4%	5	6	4	2	3	1	Match group (LEADING_AUTO)
	712	95.2%	5	6	4	2	3	1	Match group (LEADING_AUTO)
	710	94.9%	5	6	4	2	3	1	Match group (LEADING_AUTO)
	710	87.0%	5	6	4	2	3	1	Match group (LEADING_AUTO)
B7	131	100.0%	4	6	5	4	6	5	LEADING_AUTO
	131	99.2%	4	6	4	4	5	4	LEADING_AUTO
	131	96.9%	4	5	5	4	6	5	LEADING_AUTO
	131	96.2%	5	4	4	4	4	4	LEADING_SPECIES
B8	731	100.0%	5	6	6	1	5	4	LEADING_SPECIES
	731	92.6%	3	10	11	3	9	7	Match group (LEADING_AUTO)
	731	92.6%	3	10	11	3	9	7	Match group (LEADING_AUTO)
	730	92.2%	3	10	11	3	9	7	Match group (LEADING_AUTO)
	731	92.3%	3	10	11	3	9	7	LEADING_AUTO
B9	568	100.0%	1	2	2	2	3	2	LEADING_SPECIES
	568	92.2%	0	10	10	4	8	4	Match group (MULTISPECIES)
	568	92.2%	0	10	10	4	8	4	MULTISPECIES
	568	92.1%	0	8	8	5	9	4	Match group (MULTISPECIES)
	568	91.5%	0	8	8	5	9	4	Match group (MULTISPECIES)
B10	365	100.0%	2	4	8	2	1	1	LEADING_AUTO
	384	98.6%	2	4	8	2	1	1	Match group (LEADING_AUTO)
	385	93.4%	2	0	1	0	6	0	LEADING_AUTO
	384	88.8%	2	0	2	3	0	0	LEADING_AUTO
	382	83.8%	10	0	2	1	1	0	LEADING_SPECIES
B11	195	100.0%	1	2	1	1	2	1	LEADING_AUTO
	204	97.9%	2	3	2	0	1	0	LEADING_SPECIES
	204	97.9%	2	3	2	0	1	0	Match group (LEADING_SPECIES)
	191	93.2%	4	2	1	0	0	0	LEADING_SPECIES
B12	412	100.0%	2	10	7	1	2	1	LEADING_SPECIES
	412	99.8%	2	10	7	1	2	1	Match group (LEADING_SPECIES)
	412	98.8%	2	10	7	1	2	1	Match group (LEADING_SPECIES)
	412	93.2%	2	10	7	1	2	1	Match group (LEADING_SPECIES)
B13	235	100.0%	1	6	7	2	2	2	LEADING_SPECIES
	235	100.0%	1	6	7	2	2	2	Match group (LEADING_SPECIES)
	235	96.2%	0	3	3	1	3	4	LEADING_AUTO
	235	93.6%	1	3	3	2	0	0	LEADING_AUTO

Table III.7 (continued)

protein	length	%identity	LEIBR	LEIDB	LEIIN	LEIAM	LEIMA	LEITR	ID status
B14	615	100.0%	4	12	6	3	5	3	LEADING_SPECIES
	615	100.0%	4	12	6	3	5	3	Match group (LEADING_SPECIES)
	615	95.4%	4	12	6	3	5	3	Match group (LEADING_SPECIES)
	616	94.8%	4	12	6	3	5	3	Match group (LEADING_SPECIES)
	614	85.3%	5	4	3	1	3	3	LEADING_SPECIES
B15	501	100.0%	3	11	9	1	6	5	LEADING_SPECIES
	501	99.8%	3	11	9	1	6	5	Match group (LEADING_SPECIES)
	501	98.6%	3	11	9	1	6	5	Match group (LEADING_SPECIES)
	501	97.6%	3	7	6	2	4	5	LEADING_AUTO
	501	93.4%	4	4	4	0	3	5	LEADING_AUTO
B16	233	100.0%	2	5	7	3	5	4	LEADING_SPECIES
	233	100.0%	2	5	7	3	5	4	Match group (LEADING_SPECIES)
	233	98.7%	2	5	7	3	5	4	Match group (LEADING_SPECIES)
	233	97.4%	2	5	7	3	5	4	Match group (LEADING_SPECIES)
	232	92.7%	4	2	4	2	2	3	LEADING_SPECIES
B17	803	100.0%	1	20	6	2	4	2	LEADING_SPECIES
	803	99.9%	1	20	6	2	4	2	Match group (LEADING_SPECIES)
	807	88.5%	1	5	3	1	11	3	LEADING_AUTO
	801	82.2%	1	0	0	10	1	0	LEADING_AUTO
	878	66.7%	7	1	0	0	0	0	LEADING_AUTO
B18	115	100.0%	1	5	4	3	4	3	LEADING_SPECIES
	115	100.0%	1	5	4	3	4	3	Match group (LEADING_SPECIES)
	115	95.7%	1	5	4	3	4	3	Match group (LEADING_SPECIES)
	115	94.8%	1	3	3	4	3	2	LEADING_AUTO
	115	88.7%	5	1	1	1	1	1	LEADING_AUTO
B19	335	100.0%	1	6	4	1	1	1	LEADING_SPECIES
	335	100.0%	1	6	4	1	1	1	Match group (LEADING_SPECIES)
	335	99.4%	1	6	4	1	1	1	Match group (LEADING_SPECIES)
	335	98.8%	1	6	4	1	1	1	Match group (LEADING_SPECIES)
	335	94.0%	2	3	2	1	1	1	LEADING_SPECIES
B20	449	100.0%	1	3	2	1	2	1	LEADING_SPECIES
	449	100.0%	1	3	2	1	2	1	Match group (LEADING_SPECIES)
	449	91.5%	1	1	1	1	3	1	LEADING_AUTO
	468	74.2%	5	2	1	1	2	0	LEADING_SPECIES
B21	121	100.0%	2	4	3	1	1	1	LEADING_SPECIES
	121	99.2%	2	4	3	1	1	1	Match group (LEADING_SPECIES)
	121	95.0%	2	4	3	1	1	1	Match group (LEADING_SPECIES)
	121	93.4%	2	4	3	1	1	1	Match group (LEADING_SPECIES)
	121	87.6%	4	2	1	1	1	1	LEADING_AUTO
B22	184	100.0%	1	2	4	1	1	1	LEADING_SPECIES
	184	100.0%	1	2	4	1	1	1	Match group (LEADING_SPECIES)
	184	95.1%	1	1	1	1	2	1	LEADING_AUTO
	184	92.9%	1	1	1	1	2	1	Match group (LEADING_AUTO)
	184	84.2%	1	1	1	1	2	1	Match group (LEADING_AUTO)
B23	1006	100.0%	1	3	2	1	1	1	LEADING_SPECIES
	966	76.4%	1	1	1	1	5	1	Match group (LEADING_AUTO)
	930	78.3%	1	1	1	1	5	1	LEADING_AUTO
	730	97.2%	1	3	2	1	1	1	Match group (LEADING_SPECIES)
	918	60.8%	5	0	0	0	0	0	LEADING_SPECIES
B24	595	100.0%	5	8	8	2	5	5	MULTISPECIES
	595	100.0%	5	8	8	2	5	5	Match group (MULTISPECIES)
	595	99.5%	5	8	8	2	5	5	Match group (MULTISPECIES)
	595	98.3%	5	8	8	2	5	5	Match group (MULTISPECIES)
	595	97.0%	5	8	8	2	5	5	Match group (MULTISPECIES)

3.3. Most protein antigens selected are well represented in the secretome

Normalisation of iBAQ values allowed the determination of the relative abundance of a given protein, providing a rough estimation of overall antigen abundance. The absolute iBAQ values for the 2333 LEADING accessions identified with the proteomic analysis range from 2524,8 (10^3) to 3526300000 (10^9). Nine protein accessions have iBAQ values of 0, including 4 uncharacterized proteins, Calpain-like cysteine peptidase (Fragment), Putative ubiquitin conjugation factor E4 B; Tubulin binding cofactor A-like protein; Serine/threonine-protein phosphatase and DNA-directed RNA polymerase subunit beta. The null iBAQ values imply the detection of specific tryptic peptides but with very low intensity, still, these values were included in the analysis. After normalisation, iBAQ values range from 0,537 to 1,508, and proteins with normalised iBAQ values over 1 can be considered abundant in the secretome.

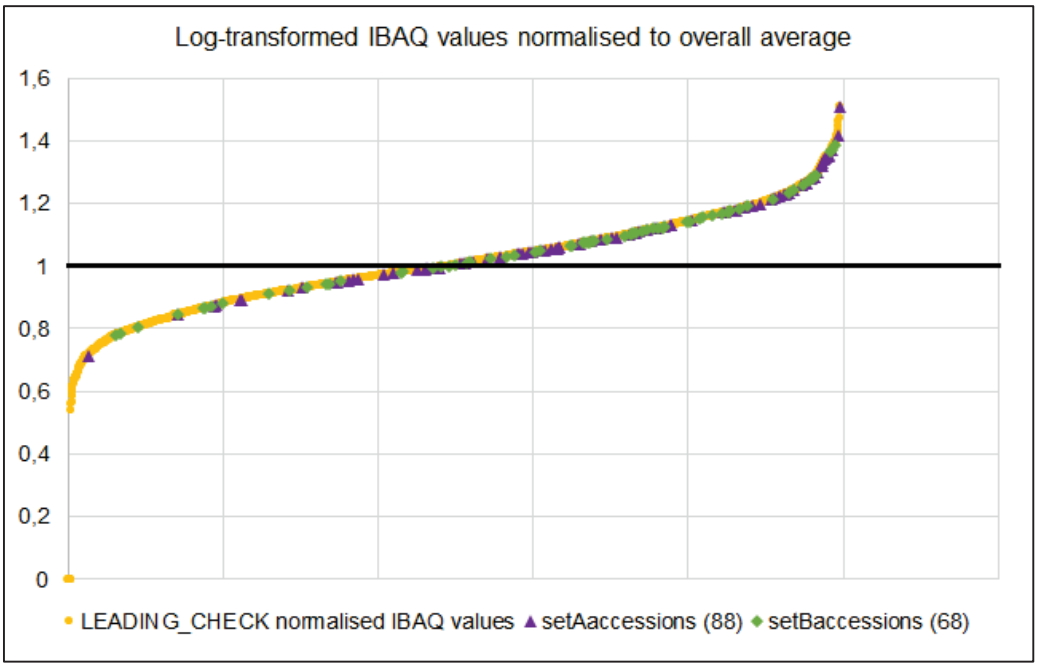


Figure III. 2 Normalised iBAQ values show most of the selected protein antigens are more abundant than average. *Absolute iBAQ values were log-transformed (log10) and normalized to overall log10-iBAQ average value. Total identified proteins (yellow), 88 set A protein antigen sequences (purple), and 68 set B protein antigen sequences (green).*

From 147 unique protein accessions in the protein set A, 88 iBAQ values are available for the LEADING_CHECK accessions, with absolute values ranging from 32613 (10^4) to 3526300000 (10^9). The normalised iBAQ values for protein set A range from 0,713 to 1,508, including one highly abundant protein with a normalised iBAQ value above 1,5. The protein set A contains 67 accessions with normalised iBAQ values above 1, which corresponds to 76% of total accessions,

indicating most set A proteins are well represented in the secretome (Figure III.3). The protein set A normalised iBAQ average is 1,117 ($1{,}06x10^8$).

The most abundant protein in set A is also the most abundant protein in the overall proteomic analysis. The nucleoside diphosphate kinase (Q9GP00; LEADING_SPECIES) is the only protein with normalised iBAQ value superior to 1,5. There are 10 other very abundant proteins in protein set A with normalised iBAQ values superior to 1,319 (trypanothione reductase; triosephosphate isomerase; elongation factor 2; probable eukaryotic initiation factor 4A; leishmanolysin; histone H2B; histone H2B; nucleoside diphosphate kinase; histone H4; putative small myristoylated protein-3). The least abundant protein in set A, with normalised IBAQ value under 0,8, is a putative aminopeptidase (E9ALJ9; LEADING_AUTO).

The protein set B includes 114 unique accessions, 68 of which correspond to LEADING_CHECK identifications with iBAQ values. The absolute iBAQ values range from 86138 (10^4) to 604490000 (10^8). The normalised iBAQ values for protein set B range from 0,78 to 1,39. 73,5% of all accessions in protein set B are abundant, with 50 accessions presenting normalised iBAQ values above 1 (Figure III.3). The protein set B normalised iBAQ average is 1,077 ($4{,}1x10^7$).

The most abundant proteins in set B have normalised iBAQ values above 1,35 (putative aldolase; putative beta-fructofuranosidase; putative small myristoylated protein-1). The least abundant proteins in the set B, with normalised iBAQ values under 0,8, are a putative eukaryotic translation initiation factor 3 and glutamate dehydrogenase.

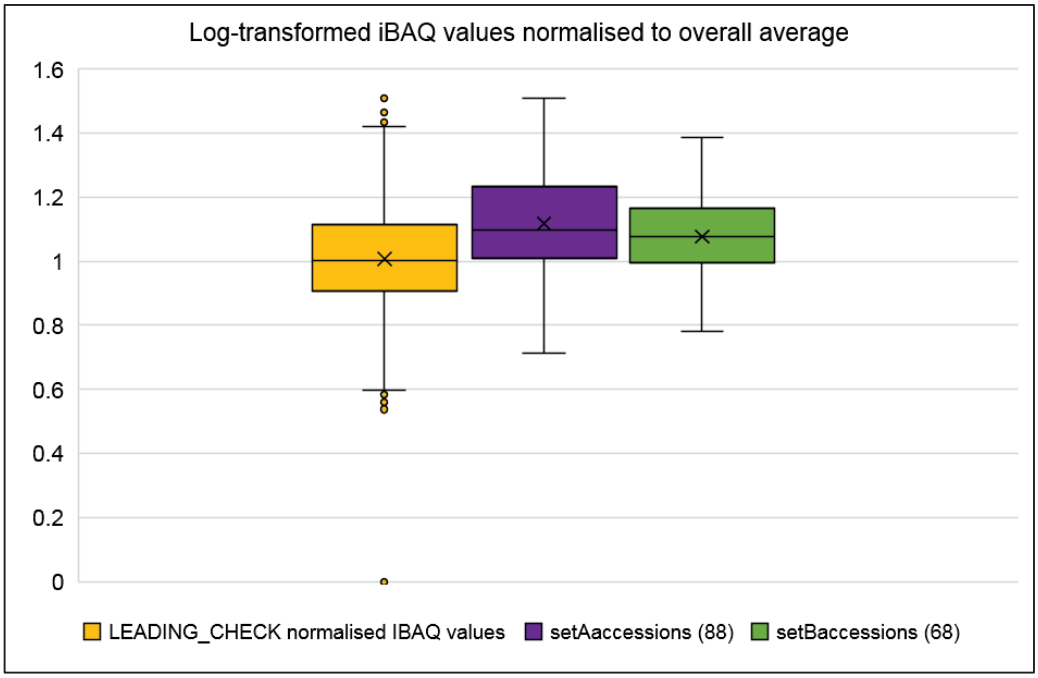


Figure III. 3 Estimation of selected protein antigens' abundance with normalised iBAQ values.

4. Discussion

A potential limitation of the overall analysis of the promastigote secretome is to overlook amastigote-specific proteins, potentially very important in a vaccine formulation. Some amastigote-specific proteins have been described and are discussed in the selected publications from which antigens were listed, notably the A2 protein. These proteins evidently were not found present in the promastigote secretome. However, there is evidence that antigens present in the promastigote secretome can provide immunity against intracellular forms and that many secreted antigens are continuously secreted during intracellular life stages.

The simultaneous high of conservation between *Leishmania* species, long-term co-evolution and selective pressure by the immune system suggests the parasite relies on early immune responses to establish infection. The vertebrate hosts are exposed to metacyclic promastigote forms in early infection, when the immune response cascade is activated. This provides a unique opportunity for antigen discovery since early immune responses against promastigotes will dictate the fate of the adaptive immune response. Moreover, gene expression remains stable after parasite differentiation into intracellular amastigote forms, providing the advantage of targeting the parasites in early stages of infection as well as intracellular replicative forms (9–12). Additionally, it is established from the dog infection model that the *Leishmania* promastigote secretome is a great antigenic source for *Leishmania* recognition and induction of protective immune responses (13) (see Chapter I).

In this light, it is not surprising that multiple known candidates are found in the secretome. Precisely, 36 proteins out of 72 potential antigen candidates are excreted-secreted proteins. These include several protein antigens currently in use by the most advanced vaccine candidates against leishmaniasis – initiation factors, cysteine proteases; TSA (thiol-specific antigen); LACK; protein disulfide isomerase-2; Putative glycosomal phosphoenolpyruvate carboxykinase; nucleoside hydrolase; GP63; and KMP11, among others.

The second-generation vaccine LEISH-F3, the most advanced human vaccine in clinical trial, is based on a fusion recombinant protein combining the Nucleoside hydrolase from *L. donovani*, Sterol-24-c-methyl-transferase from *L. infantum*; delta cysteine protease "B" from *L. infantum*. Interestingly, both nucleoside hydrolase and cysteine protease B antigens are found in the *Leishmania* secretome. Furthermore, the other antigen candidates included in vaccines LEISH-F1 and -F2, the *Leishmania* elongation initiation factor (LeIF) and thiol-specific antioxidant (TSA) are also present in the *Leishmania* secretome.

In eukaryotic organisms, glutathione (GSH) is an antioxidant that works as a redox buffer during the glutathione redox cycle, protecting the cells from reactive oxygen species. GSH reduces disulfide bonds of cytoplasmic proteins becoming oxidized (GSSG) and the ratio between GSH and GSSG is a measure of cell toxicity. Instead, *Leishmania* parasites have trypanothione (TSH), trypanothione reductase (TryR), tryparedoxin, tryparedoxin peroxidase also known as TSA (thiol-specific antigen), peroxidoxin, and mitochondrial peroxiredoxin. These proteins make-up their anti-oxidant system for protection against reactive oxygen and nitrogen species, and are therefore essential for parasite survival inside the phagolysosomal vesicles. Two of these were found in the secretome of all 6 species – trypanothione reductase and tryparedoxin peroxidase, also known as TSA (thiol-specific antigen). These proteins are important virulence factors and targets particularly for drug development, but little information exists on protein immunogenicity and sequence annotation is still incomplete for the many existing isoforms.

Proteases are ubiquitous enzymes that catalyse the hydrolysis of peptide bonds and are important in several biological activities. There are at least 6 classes of proteases classified according to the nucleophilic group responsible for the first step in the proteolysis: serine, cysteine, metallo, aspartate, glutamate, and threonine proteases. The two major types in eukaryotes are cysteine (papain-like) proteases and serine (trypsin-like) proteases. Cysteine proteases are categorized into 72 families, but not all are represented in protozoan parasites. The most abundant and well-characterized cysteine proteases are the clan CA papain-family enzymes.

Cysteine proteases are vital virulence factors that ensure parasite survival and establishment of infection. Moreover, some proteases were described as potential vaccine antigen candidates, namely GP63, methionine aminopeptidase p45, and cathepsin-like proteases. Although not specifically described as potential candidates, other *Leishmania* proteases were also searched (aminopeptidases and metallo-proteases) to expand the protease antigens included in the protein set A.

The antigens trypanothione reductase and glycosomal phosphoenolpyruvate carboxykinase (gPEPCK), although not used in vaccine formulation, are among the most studied vaccine antigens against *Leishmania*, showing very promising pre-clinical results (7). gPEPCK is upregulated in stationary phase promastigotes and *L. donovani* amastigotes probably due to gluconeogenesis activation (14).

The recombinant fusion protein Q is the main component of the latest canine vaccine approved in Europe (Letifend®), shown to be effective in pre-clinical trials in dogs. This fusion

protein contains portions of *L. infantum* histone H2A and ribosomal proteins (p2a, p2b, and p0). Again, these proteins are found in the *Leishmania* secretome.

To expand the antigen port-folio to proceed for *in silico* epitope prediction, we adopted a reverse vaccinology approach, using low human host homology as selection criterium. This approach allows the comprehensive exploration of the secretome proteomic analysis data thus allowing the screening of numerous protein antigen candidates.

Results show there were 618 accessions common to six *Leishmania* species (UniProtKB release 2016_05). These 618 common protein accessions correspond to about a third of total identifications, indicative of the high conservation among pathogenic species, and in agreement with previous studies studying excreted-secreted proteins (see chapter II). Common identified accessions between the tested samples also imply that peptides that lead to each protein identification are found in all samples (peptide counts) and are from conserved regions. The total number of species-specific sequences per protein (4 to 6) takes into account sequence variability for all given antigens.

The BLASTp alignments are a powerful tool to find interesting protein candidates which share little homology with host proteins. Several cut-off values have been suggested for selection, according to Evalue, Bit score, minimum alignment length, and/or minimum sequence identity (15,16). In the present analysis we selected only the proteins with no significant homology found, so no cutoff values were applied. The generated proteomic datasets could be further explored by including proteins with a related host protein but still with significant partial sequence divergence.

Some candidates found with the RV approach were already described as *Leishmania* antigen candidates – KMP11, nucleoside hydrolase, paraflagellar rod proteins and calpain-like cysteine peptidase. We believe the identification of these proteins through the RV approach further supports the validity of the proposed approach.

Some uncharacterised and hypothetical proteins have been proposed as *Leishmania* antigen candidates, such as the hypothetical protein LiHyR (XP_001568689.1), which corresponds to the protein A4HNR3 in the UniProtKB database (17). LiHyR induced Th1 responses and reduced parasite burden in mice. More importantly, it is also immunogenic in human PBMC from healthy donors and cured VL patients. This protein not found in the secretome, nor any of its related proteins (A4ICT2, E9ASH2, E9BTB7, A0A088S1U1, Q4Q223). The high number of hypothetical proteins in the databases is a result of the lack of protein annotation and functional characterisation of *Leishmania* proteins. However, this should not deter the study of

their immunogenic properties. Accordingly, the 10 uncharacterised proteins identified with the reverse vaccinology approach were included in the protein set B.

Antigen abundance is a pre-requisite for the induction of immune responses. the secretome as an antigen source contains over 1000 proteins, around 15% of total *Leishmania* proteins. The non-quantitative approach used does not allow to ascertain in absolute terms the amount of each identified protein. However, relative protein abundance across all tested species can be used as an indicator of overall abundance. We find most protein antigens are well represented in the secretome, 76% for set A and 74% for set B proteins. This analysis is merely descriptive, as antigen abundance was not used as an inclusion or exclusion criterion. However, this information may be useful in later stages, to favour or eliminate epitopes from more or less abundant proteins, respectively. Also, these proteomic datasets can be further explored to find new and highly abundant antigens (e.g. protein set C).

The total number of protein antigens to proceed for *in silico* prediction analysis is limited by the lack of high-throughput epitope selection analysis tools. Still, a total of 52 protein antigen candidates from the *Leishmania* secretome were selected and will proceed for epitope prediction analysis (Figure III.4).

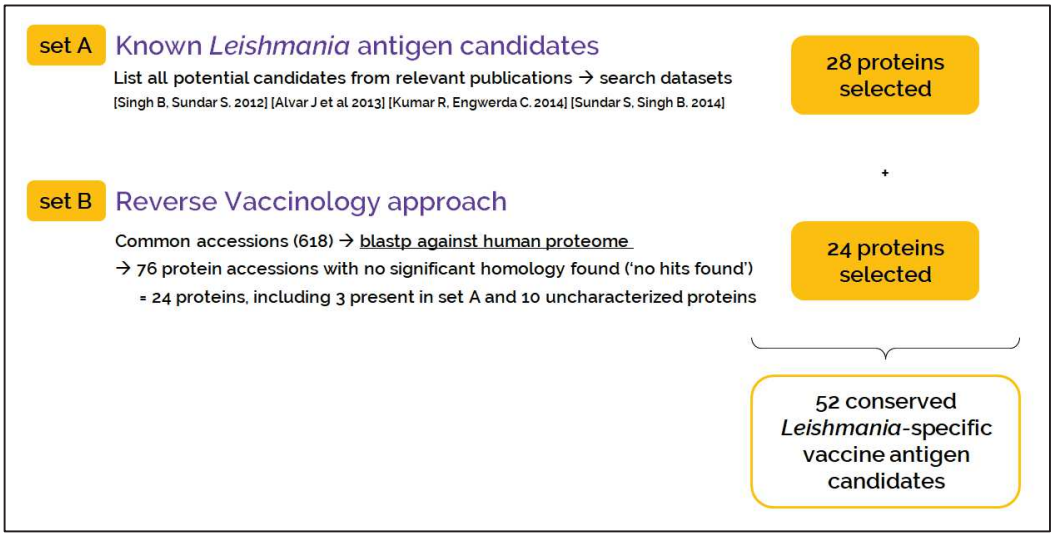


Figure III. 4 Overall protein antigen selection results from the secretome proteomic data.

5. Bibliography

1. Murphy K, Weaver C. Glossary. In: Janeway's Immunobiology. 9th ed. Garland Science; 2017. p. 818–54.
2. Singh B, Sundar S. Leishmaniasis: Vaccine candidates and perspectives. Vaccine. Elsevier Ltd; 2012;30(26):3834–42.
3. Kumar R, Engwerda C. Vaccines to prevent leishmaniasis. Clin Transl Immunol. Nature Publishing Group; 2014;3(3):e13.
4. Lakshmi BS, Wang R, Madhubala R. Leishmania genome analysis and high-throughput immunological screening identifies tuzin as a novel vaccine candidate against visceral leishmaniasis. Vaccine. Elsevier Ltd; 2014;32(30):3816–22.
5. Sundar S, Singh B. Identifying vaccine targets for anti-leishmanial vaccine development. Expert Rev Vaccines. 2014;13(4):489–505.
6. Alvar J, Croft SL, Kaye P, Khamesipour A, Sundar S, Reed SG. Case study for a vaccine against leishmaniasis. Vaccine. 2013 Apr 18;31 Suppl 2:B244-9.
7. Mou Z, Li J, Boussoffara T, Kishi H, Hamana H, Ezzati P, et al. Identification of broadly conserved cross-species protective Leishmania antigen and its responding CD4 + T cells. Sci Transl Med. 2015;7(310):1–13.
8. Alcolea PJ, Alonso A, Gómez MJ, Moreno I, Domínguez M, Parro V, et al. Transcriptomics throughout the life cycle of Leishmania infantum: High down-regulation rate in the amastigote stage. Int J Parasitol. 2010;40(13):1497–516.
9. Rosa R, Marques C, Rodrigues OR, Santos-Gomes GM. Leishmania infantum released proteins specifically regulate cytokine expression and production patterns by CD4+ and CD8+ T cells. Acta Trop. 2006;97(3):309–17.
10. Lambertz U, Silverman JM, Nandan D, McMaster WR, Clos J, Foster LJ, et al. Secreted virulence factors and immune evasion in visceral leishmaniasis. J Leukoc Biol. 2012;91(6):887–99.
11. Rosa R, Roos Rodrigues O, Marques C, Santos-Gomes GM. Leishmania infantum: Soluble proteins released by the parasite exert differential effects on host immune response. Exp Parasitol. 2005;109(2):106–14.
12. Gour JK, Kumar V, Singh N, Bajpai S, Pandey HP, Singh RK. Identification of Th1-responsive leishmanial excretory–secretory antigens (LESAs). Exp Parasitol. Elsevier Inc.; 2012;132(3):355–61.
13. Hosein S, Blake DP, Solano-Gallego L. Insights on adaptive and innate immunity in canine leishmaniosis. 2019;
14. Alcolea PJ, Alonso A, Gómez MJ, Postigo M, Molina R, Jiménez M, et al. Stage-specific

differential gene expression in Leishmania infantum: from the foregut of Phlebotomus perniciosus to the human phagocyte. BMC Genomics. BioMed Central; 2014 Jan 3;15(1):849.

15. Pearson WR. An Introduction to Sequence Similarity (“ Homology ”) Searching. Curr Protoc Bioinforma. 2013;1–9.
16. Carmona SJ, Sartor PA, Leguizamón MS, Campetella OE, Agüero F. Diagnostic Peptide Discovery: Prioritization of Pathogen Diagnostic Markers Using Multiple Features. PLoS One. 2012;7(12):e50748.
17. Ribeiro PAF, Dias DS, Lage DP, Costa LE, Martins VT, Tavares GSV, et al. Evaluation of a Leishmania hypothetical protein administered as DNA vaccine or recombinant protein against Leishmania infantum infection and its immunogenicity in humans. Cell Immunol. Academic Press; 2018 Sep 1;331:67–77.

CHAPTER IV

IN SILICO EPITOPE DISCOVERY AND SELECTION

Objective: to discover T-cell epitopes present in antigens from the *Leishmania* secretome capable of inducing cellular immune responses mediated by $TCD8^+$ and $TCD4^+$ cells.

Aims:

- To select the best performing immunoinformatic tools;
- To perform epitope prediction targeting world HLA coverage;
- To select immunogenic epitopes according to *in silico* prediction data.

1. Introduction

An epitope, also called an antigenic determinant, is the site on an antigen recognised by an antibody or an antigen receptor – T-cell epitopes are short peptides bound to MHC molecules, and B-cell epitopes are typically structural motifs on the antigen surface (1). Epitopes are amino-acid sequences characterised by a low level of similarity to the proteome of the immunoreacting host.

The epitopes to include in epitope-based vaccine development should be (2):

i) Conserved among pathogen species;
ii) High affinity binders to HLA molecules;
iii) Presentable in target host population (HLA restriction);
iv) Immunogenic (able to induce T cell activation);
v) Able to induce long-term protection (memory responses);
vi) Not cross-reactive with self-proteins (autoimmunity).

Immunodominant epitopes present in an antigen are preferentially recognised by T cells, such that T cells specific for those epitopes come to dominate the immune response (1). Immunodominance depends both on antigen-related factors and on T cell-related factors (Table IV.1). Discovering immunodominant epitopes is desirable in vaccine development, as these will elicit the strongest immune responses. However, in the case of mutation, a vaccine based only on immunodominant epitopes would be rendered ineffective. Moreover, while the observed responses against dominant epitopes will be the strongest, they are more likely to present high response variability among individuals and, therefore, potentially generate an uneven vaccine performance.

Table IV. 1 Determinants of immunodominance. *Adapted from* (3).

Antigen-related factors	T cell-related factors
Kinetics of transcription and translation; Antigen processing and transport; Antigen abundance; Affinity to MHC molecules; Stability and turnover of pMHC complexes; Type of APCs	Timing of CTL clonal expansion; T cell precursor frequency; TCR repertoire; Signal strength; TCR affinity and avidity; Proliferative capacity; Intrinsic ability to respond; Pre-conditioning or priming; Dwell time on APCs; Elimination of APCs; Competition for resources (cytokines, antigens or physical niche).

On the other hand, **immunoprevalent** T-cell epitopes are frequently immunogenic in the context of multiple MHC alleles. These epitopes can induce specific IFN-γ responses with high responding T-cell frequency within the repertoire and are common across individuals with different HLA types (4). In summary, while immunodominant epitopes will be recognised more vigorously, immunoprevalent epitopes will be recognised more frequently. Epitope-based vaccines should privilege immunoprevalent epitopes, particularly vaccines targeting large populations or against pathogens with multiple antigenic sources.

Immunodominance and immunoprevalence are relative terms rather than absolute, as they always depend on the total composition of antigenic molecules and on competitive high-affinity binding to MHC molecules. Both should be considered when designing peptide-based vaccines.

1.1. T cell epitope prediction and immunoinformatics

The knowledge of immunogenicity determinants and immunogenic epitopes increased greatly in recent years, demonstrated by available databases such as IEDB (Immune Epitope DataBase) (5–8).

Table IV. 2 Available T-cell epitope databases. *General epitope databases, and pathogen- and tumor-specific databases of T-cell epitopes* (9).

General Database	Description	Availability
MHCPEP	Database of MHC-binding peptides	ftp://ftp.webi.edu.au/pub/biology/mhcpep
SYFPEITHI	Database of MHC ligand and peptide motifs	http://www.syfpeithi.de
AntiJen (JenPep)	Quantitative immunology database	http://www.ddg-pharmfac.net/antijen
MHCBN	Database of MHC/TAP-binding peptides and T-cell epitopes	http://www.imtech.res.in/raghava/mhcbn
EPIMHC	Database for customised computational vaccinology	http://bio.dfci.harvard.edu/epimhc/
IEDB	Immune epitope database	http://www.iedb.org
IMGT/HLA	IMGT/HLA database	http://www.ebi.ac.uk/ipd/imgt/hla/
Pathogen- and tumor-specific databases		
AntigenDB	Database of pathogen antigens	http://www.imtech.res.in/raghava/antigendb
Protegen	Database of protective antigens	http://violinet.org/protegen
HIV Molecular immunology database	HIV database	http://hiv.lanl.gov/content/immunology
HCV immunology database	HCV database	http://cancerimmunity.org/peptide/
TANTIGEN	Database of tumor T-cell antigens	http://cvc.dfci.harvard.edu/tadb

The IEDB is the largest and most complete epitope database, capturing epitope information from 99% of all publications describing immune epitopes, all except HIV- and cancer-specific epitopes (10,11). The IEDB contains both epitope and assay information regarding epitopes from infectious diseases, autoimmune and allergic diseases, and alloantigens for humans, primates, mice and other host species (12).

The IPD-IMGT/HLA (or IPD-IMGT/MH-DB) database is one of the seven databases from the the Immuno Polymorphism Database (IPD) (13). The IPD-IMGT/HLA database is part of the International Immunogenetics Information System® (www.imgt.org) and is the main specialised repository for the sequence data of polymorphic gene sequences (14). It contains allelic sequences of HLA genes and official sequences named by the WHO Nomenclature Committee for Factors of the HLA System (as of April 2019, release 3.36.0 contains 16,200 HLA Class I Alleles, 6,162 HLA Class II Alleles and 186 other non-HLA Alleles).

The MHCPEP database contains both naturally processed and synthetic peptides (over 4000 binder peptides) (15), whereas SYFPEITHI is a manually curated database from the published literature and exclusively contains naturally processed peptides (over 7000 MHC ligands, motifs and epitopes), but both these databases are no longer updated (16). The MHCBN is a curated database

containing binder and non-binder epitopes (over 25800 peptide entries) with information about TAP interactions and MHC-linked autoimmune diseases, last updated in 2006 (17).

Parallel to databases, analysis resources allow the prediction of T cell epitopes. Several sophisticated immunoinformatics tools, of variable performance, are available to evaluate some of the characteristics associated with immunogenic epitopes (7,8,18,19). A determinant aspect for algorithm performance is the amount and quality of the training datasets. As epitope data increases and improves, so will the performance of epitope prediction algorithms.

T-cell epitopes prediction can be performed through **direct** (predicting T-cell receptor, TCR recognition) or **indirect methods** (predicting epitope binding to MHC/HLA molecules), the latter extensively more accurate than the former (20).

HLA-binding affinity has become the first criterion when trying to predict if a given peptide sequence constitutes an epitope, since it is the first requirement for T-cell activation and it correlates with peptide linear sequences (7,20). The first algorithms developed used basic motif listings to prediction T-cell epitopes. Subsequently, HLA-binding predictions based on machine-learning algorithms were developed, based on Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs), which display better predictive performance (4,21–23) (Table IV.3). HLA-binding predictions are affected by two main issues – prediction accuracy and sensitivity, and the MHC-class I and -class II alleles for which predictions are available (10).

The performance of the several predictors was performed through benchmark studies, using curated datasets different from the algorithm's training datasets (24–27). The evaluated predictors show in general somewhat similar performance, but ANN-based and consensus predictions have the best overall performance (10,27). The IEDB analysis resource centre has the highest number of predictors, including consensus predictions, the most effective prediction algorithm (23,25,28).

Table IV. 3 The most commonly used sequence-based algorithms for T cell epitope prediction (29)**.**

Server name	Link	HLA-class I alleles	HLA-class II alleles	Predictive method	Database	Comments
MHCPred	http://www.ddg-pharmfac.net/mhcpred/MHCPred/	11	3	Additive method to predict the binding affinity of MHC-class I and II molecules and TAP processing . Allele-specific Quantitative Structure Activity Relationship (QSAR) models.	MHCPred models generated from IC50 values obtained from radioligand competition assays characterising pMHC affinity. Values collated from the literature and stored in the JenPep database.	Maximum of 1000 residues plain text (does not accept fasta sequences). Limited number of available alleles. HTML output with IC50 predictions.
SYFPEITHI	http://www.syfpeithi.de/bin/MHCServer.dll/EpitopePrediction.htm	33	6	Published motifs (two-dimensional data array to score individual residues).	SYFPEITHI database, based on previous publications on T-cell epitopes and MHC ligands (Rammensee et al 1995 Immunogenetics; Rammensee et al 1997 Landes Bioscience)	HTML output. Peptide score (no IC50 prediction). SYFPEITHI achieves 85% reliability within the top 10%.
EpiVax	http://www.epivax.com/	6	8	Epimatrix algorithm, motif matrices for MHC-class I molecules		Restricted-access software (mandatory training course by EpiVax)
RANKPEP	http://bio.dfci.harvard.edu/RANKPEP/	77	50	Position-Specific Scoring Matrices (PSSM)	Matrices generated using ungapped block alignments from alignments of MHC-ligands collected from the MHCPEP database (13423 high and moderate binding peptide entries distributed between 281 MHC alleles). Training sets for proteasomal cleavage prediction from a database contaning the C-terminus and flanking regions of 332 antigens restricted by HLA-class I molecules.	HTML output. Peptide score (no IC50 predictions). Peptides produced by proteasomal cleavage prediction highlighted in violet.
EpiJen	http://www.ddg-pharmfac.net/epijen/EpiJen/EpiJen.htm	18		Multi-step algorithm based on quantitative matrices (proteasome cleavage, TAP transport, MHC binding and epitope selection). EpiJen is a further development of MHCPred. Written in Perl, with an interface written in HTML.	development of the MHC-binding models with 1371 peptides extracted from AntiJen and SYFPEITHI databases; training set for the proteasome cleavage model with 489 naturally processed T-cell epitopes associated with HLA-A and HLA-B molecules; training set for the TAP processing model includes 163 poly-Alanine 9-mer peptides.	Performs better than other integrated methods (NetCTL, WAPP and SMM). Only accepts plain text sequences (does not accept fasta sequences). Limited number of available alleles. Percentile cutoff (up to 5%) and results in HTML output with IC50 predictions.
nHLAPred	http://www.imtech.res.in/raghava/nhlapred/	26		Neural network-based MHC Class-I Binding Peptide Prediction Server (ANN and QM). The server consists of two major parts: ANNPred (ANN-based epitope prediction for 30 alleles) and ComPred (quantitative matrix-based prediction for 37 alleles) with proteasome and immunoproteasome filters.	The dataset of MHC binders for the training data (both ANN and QM) obtained from MHCBN database.	HTML output. Peptide score (no IC50 prediction).

Table IV. 3 (continued)

Server name	Link	HLA-class I alleles	HLA-class II alleles	Predictive method	Database	Comments
ProPred I	http://www.imtech.res.in/raghava/propred1/	39		Promiscuous MHC Class-I Binding Peptide Prediction Server (QM). It allows matrix-based prediction of MHC binders for 47 MHC-I alleles and proteasome cleavage site, simultaneously.	Matrices obtained from BIMAS server and from the literature. The weight matrices for both standard proteasome and immunoproteasome were derived from Toes et al., 2001.	HTML/tabular output. Peptide score (no IC50 predictions).
MMBPred	http://www.imtech.res.in/raghava/mmbpred/	39		Prediction of mutated MHC binders (QM). Quantitative matrices for 47 MHC alleles	MHC binder (and non-binder) data of 9-mer peptides for each MHC alleles obtained from MHCBN database. Otherwise 9-mer peptides are randomly chosen from the SwissProt database.	HTML output. Peptide score (no IC50 predictions)
NetMHC	http://www.cbs.dtu.dk/services/NetMHC/	81		ANN-based method trained for 81 different Human MHC alleles including HLA-A, -B, -C and -E. For alleles not included in the list use NetMHCpan.	NetMHC was trained on MHC peptide binding data contained in the IEDB plus 100 random natural peptides (8 to 11-mer) as artificial negatives for each allele.	Accepts multiple fasta sequences (up to 5000). HTML or csv output format. Binding affinity IC50 predictions.
NetMHCpan	http://www.cbs.dtu.dk/services/NetMHCpan/	12 supertype representatives		ANN-based method predicts binding of peptides to any MHC molecule of known sequence	Trained on a combination of over 180000 quantitative binding data and MS-derived MHC eluted ligands. The binding affinity data covers 172 MHC molecules.	Accepts multiple fasta sequences (up to 5000). HTML or csv output format. Binding affinity IC50 predictions.
NetCTL	http://www.cbs.dtu.dk/services/NetCTL	12 supertypes		ANN-regression, it integrates prediction of peptide MHC class I binding , proteasomal C terminal cleavage and TAP transport efficiency. CTL epitopes restricted to 12 MHC class I supertypes.	Combines binding predictions by NetMHC, proteasomal cleavage prediction by NetCHOP, and TAP transport prediction with a weight matrix-based method described by Peters et al. 2003 J Immunol.	Accepts multiple fasta sequences (up to 5000). HTML or csv output format. Individual and combined prediction scores.
IEDB MHC-I binding	http://tools.iedb.org/mhci/	56		Methods included: ANN, SMM, SMM with a pMHC Binding Energy Covariance matrix (SMMPMBEC), Scoring Matrices derived from Combinatorial Peptide Libraries (Comblib_Sidney2008), NetMHCpan, NetMHCcons, PickPocket, NetMHCstabpan, and consensus predictions. IEDB recommended uses (by order of preference according to queried alleles): Consensus > ANN > SMM > NetMHCpan > CombLib.	Quantitative experimental data from IEDB	Accepts multiple fasta sequences. HTML or csv output format. Binding affinity IC50 predictions from multiple algorithms.
BIMAS	http://www-bimas.cit.nih.gov/molbio/hla_bind/	33		Bioinformatics Molecular Analysis Section - Published coefficient tables to rank potential peptides (8 to 10-mer) based on a predicted half-time of dissociation to HLA class I molecules.	Training dataset from quantitative experimental assays (stability of HLA-A2 complexes containing specific peptides assessment).	HTML output. Peptide score (estimate of Half Time of Disassociation of a Molecule Containing This Sub-sequence).

Table IV. 3 (continued)

Server name	Link	HLA-class I alleles	HLA-class II alleles	Predictive method	Database	Comments
MHC2Pred	http://www.imtech.res.in/raghava/mhc2pred/		38	SVM-based method to predict of promiscuous MHC class II binding peptides	Training data extracted from MHCBN and JenPep (AntiJen) databases	HTML/tabular output. Peptide score (no IC50 predictions).
IMTECH	http://www.imtech.res.in/raghava/mhc		3	Matrix Optimization Technique for Predicting MHC binding (QM), a matrix with 97 to 99% accuracy on HLA-DR1, HLA-DR2 and HLA-DR5 alleles.	Training dataset from MHCPEP database (binder/non-binder data for HLA-class II DR1, DR2 and DR5 isotypes)	HTML output or Graphical view of predicted binder regions (indicated "*****"). Very limited number of available alleles. No score and no IC50 predictions.
ProPred	http://www.imtech.res.in/raghava/propred/		51	MHC-class II binding peptide prediction server (QM). It employs amino acid / position coefficient tables deduced from literature by Sturniolo et al., 1999, in a linear prediction model	Matrices for 51 HLA-DR alleles fom a pocket profile database described in Sturniolo et al (1999). Threshold values from TEPITOPE (25 alleles) and PIR database (26 alleles).	HTML or graphical output (Predicted binders are marked with '*' or displayed in color). Peptide score (no IC50 predictions).
NetMHCII	http://www.cbs.dtu.dk/services/NetMHCII/		54	ANN-based binding prediction of peptides to HLA-DR, HLA-DQ, HLA-DP HLA-class II alleles (NN-align). Maximum alleles 15 per query.	5 quantitative HLA-DR, -DP, and -DQ restricted peptide-binding datasets from IEDB (2016 dataset: 134 281 data points, covering 36 HLA-DR, 27 HLA-DQ, 9 HLA-DP and 8 H-2 molecules)	Accepts multiple fasta sequences (up to 5000). HTML output format. Binding affinity IC50 predictions.
NetMHCIIpan	http://www.cbs.dtu.dk/services/NetMHCIIpan/		3 isotypes	ANN-based predictions for HLA-class II isotypes HLA-DR, HLA-DP and HLA-DQ.	5 quantitative HLA-DR, -DP, and -DQ restricted peptide-binding datasets from IEDB (2016 dataset: 134 281 data points, covering 36 HLA-DR, 27 HLA-DQ, 9 HLA-DP and 8 H-2 molecules)	Accepts multiple fasta sequences (up to 5000). HTML or csv output format. Binding affinity IC50 predictions.
IEDB MHC-II binding	http://tools.iedb.org/mhcii/		>700	Methods included: Combinatorial library, NN-align (netMHCII-2.2), SMM-align (netMHCII-1.1), Sturniolo, and NetMHCIIpan, and Consensus predictions. IEDB recommended uses (by order of preference according to queried alleles): Consensus combination of any three of the four methods (NN-align > SMM-align > CombLib > Sturniolo) >NetMHCIIpan.	Quantitative experimental data from IEDB	Accepts multiple fasta sequences. HTML or csv output format. Binding affinity IC50 predictions from multiple algorithms (Sturniolo is given as raw score).

1.2. *In silico* epitope predictions and implications for vaccine development

MHC-class II binding predictions must take into account the high peptide length variability due to the conformation of the binding groove, and are currently slightly less accurate than class I binding predictions (25,30). Due to the longer length of class II epitopes (around 15 to 25-mer), several binding registers or cores may be present in the same peptide (29,31).

Proteasomal cleavage analysis and TAP-transport prediction, although very informative in theory, do not improve MHC-binding predictions (10,32). These algorithms often display low prediction efficiency, and there are still significant knowledge gaps regarding protein intracellular processing.

Also, for the particular case of parasite-delivered antigens, the mechanisms responsible for cross-presentation remain poorly understood (32,33). Professional APC, namely DC are the main cells capable of cross-presenting antigens. However, due to the extensive modulation by *Leishmania* parasites and preference for macrophage infection these mechanisms remain largely unknown (33). In this case, the use of experimental data detailing peptide-specific immune responses is indispensable to ascertain which peptides are indeed associated with natural protection and can be used in a vaccine formulation.

Additional aspects related with immunogenicity assist in epitope prediction and selection. Analysis of subcellular localization, protein abundance and good expression dynamics are the filters with the highest selective power (32). Highly conserved epitopes are ideal for vaccine development because of pan-specific protection across multiple strains of a given pathogen (34). Furthermore, combinatorial approaches that use multiple predictors are beneficial since they increase the confidence level in the peptides' predicted binding affinity/HLA restriction (31,35).

Using homology to host proteins as rejection criterion is an unreliable filter. Self-recognition depends on the TCR-pMHC interaction which allows a reasonable amount of molecular mimicry, and therefore difficult to predict (36–38). However, potential interferences resulting in auto-immunity are correlated with epitope conservancy. BLASTp alignments can be used to compare and describe similarities pathogen- and host-derived peptide sequences.

The ability to analyse the vast amount of data generated by immunoinformatic algorithms is an additional challenge researchers must overcome (39–41).

Ultimately, *in silico* immunogenicity predictions on the epitope level should privilege:

i) antigen abundance, subcellular localization and expression dynamics, as abundant and early expressed pathogen-specific epitopes have increased chances of being processed and presented (32,42);
ii) peptide-MHC complex (pMHC) binding affinity and stability (2,43,44);
iii) homology, either as positive selection criterium of conserved sequences among pathogenic species, or as negative selection criterium of sequences homologous to host proteins (20);

Other biochemical properties can also be taken into account for optimal formulation and handling, such as solubility, since peptides should be soluble in aqueous solution (2,21).

We propose an epitope selection pipeline that starts with HLA-binding affinity prediction analysis, by at least two different algorithms, of a strong antigen pool of conserved, exposed and accessible proteins, expressed from early infection (antigens present in the *Leishmania* secretome). Additional adjustable filters are homology to host proteins, promiscuity, binding affinity, and solubility, with which we can rank epitopes.

2. Methods

2.1. Selection of HLA allele lists

The allele lists to include in the epitope prediction algorithms were defined from the IEDB reference sets with maximal population coverage.

a) HLA-class I alleles

The HLA-class I reference allele list includes 28 alleles from HLA-A and HLA-B genes corresponding to >97% world population coverage (file from IEDB: hla_ref_set.class_i.txt) (45,46).

To the 28 alleles included in the IEDB HLA-class I reference list, 8 more were included, for a total of 36 alleles included in the HLA-binding prediction analysis. The IEDB reference stipulates epitopes with variable length, with 8 up to 11 amino-acid long epitopes. However, in the present study only 9-mer epitopes were searched.

The HLA-class I alleles selected are: IEDB MHC-I binding predictions for 36 alleles (11 supertypes), SYFPEITHI predictions for 22 alleles (11 supertypes); and NetMHCpan predictions for 11 alleles (11 supertype representatives) (Table IV.4).

***Table IV. 4 HLA-class I alleles included in T-cell epitope prediction* (45).**

Supertype group	**Allele**	**Predictor** (I = IEDB/ N=NetMHCpan / S=SYFPEITHI)	Population frequency of allele	**Allele specific affinity cut-off** (IC50 nM)
A1	A*0101	I + N + S	16.2	884
	A*3002	I	5	674
	A*3201	I	5.7	131
A2	A*0201	I + N + S	25.2	255
	A*0203	I	3.3	92
	A*0206	I	4.9	60
	A*0211	I		500
A3	A*0301	I + N + S	15.4	602
	A*1101	I + S	12.9	382
	A*3001	I	5.1	109
	A*3101	I	4.7	329
	A*3301	I	3.2	606
	A*6601	I		500
	A*6801	I + S	4.6	197
A24	A*2301	I	6.4	740
	A*2402	I + N + S	16.8	849
A26	A*2601	I + N + S	4.7	815
B7	B*0702	I + N + S	13.3	687
	B*3501	I + S	6.5	348
	B*3503	I	1.2	888
	B*5101	I + S	5.5	939
	B*5301	I + S	5.4	538
B8	B*0801	I + N + S	11.5	663
B27	B*1402	I + S	2.8	700
	B*2705	I + N + S	2	584
	B*3801	I + S	2	944
	B*4801	I	1.8	887
B44	B*1801	I + S	4.4	732
	B*4001	I + N + S	10.3	639
	B*4402	I + S	9.2	904
	B*4403	I	7.6	780
	B*4501	I + S		500
B58	B*5801	I + N + S	3.6	446
	B*5701	I + S	3.2	716
	B*1517	I		500
B62	B*1501	I + N + S	5.2	528

b) HLA-class II alleles

The HLA-class II reference allele list includes 27 alleles from HLA-DRB, and -DQ and -DP (alpha and beta chain combinations) genes corresponding to >99% world population coverage (file from IEDB: hla_ref_set.class_ii.txt) (46).

The HLA-DQ alleles were excluded from this dataset due to their association with recognition of human epitopes and the induction of immune tolerance, to prevent a bias toward cross-reacting epitopes (47). The removal of these alleles does not greatly reduce population coverage, as only the eight common DR alleles (DRB1*0101, DRB1*0301, DRB1*0401, DRB1*0701, DRB1*0801, DRB1*1101, DRB1*1301, and DRB1*1501) cover around 97% of human populations worldwide (48).

The HLA-class II alleles selected: IEDB MHC-II binding predictions for 21 alleles, and NetMHCIIpan predictions for 21 alleles (same allele lists) (Table IV.4).

Table IV. 5 HLA-class II alleles included in T-cell epitope prediction (46).

HLA-class II alleles
HLA-DPA1*01:03/DPB1*02:01
HLA-DPA1*01/DPB1*04:01
HLA-DPA1*02:01/DPB1*01:01
HLA-DPA1*02:01/DPB1*05:01
HLA-DPA1*02:01/DPB1*14:01
HLA-DPA1*03:01/DPB1*04:02
HLA-DRB1*01:01
HLA-DRB1*03:01
HLA-DRB1*04:01
HLA-DRB1*04:05
HLA-DRB1*07:01
HLA-DRB1*08:02
HLA-DRB1*09:01
HLA-DRB1*11:01
HLA-DRB1*12:01
HLA-DRB1*13:02
HLA-DRB1*15:01
HLA-DRB3*01:01
HLA-DRB3*02:02
HLA-DRB4*01:01
HLA-DRB5*01:01

2.2. Selection and usage of *in silico* HLA-binding prediction algorithms

The following publications were reviewed for algorithm selection:

- Tung CW. Chapter 6 - Databases for T-Cell Epitopes. Immunoinformatics 2014, 2nd Ed (9);
- Desai SV, Kulkarni-Kale. Chapter 19 - T-Cell Epitope Prediction Methods: An Overview. Immunoinformatics 2014, 2nd Ed (20);
- Wang M, Mogens MH. Chapter 17 - Classification of Human Leukocyte Antigen (HLA) Supertypes;
- Algorithm benchmarking studies – Lin HH et al 2008 (24), Wang P et al 2008 (25), Peters B et al 2006 (26), Trolle T et al 2015 (27);
- Castelli M et al 2013 Clin Transl Immunology ID:521231 (29).

To improve prediction sensitivity, the main criteria used for the epitope prediction and selection were i) to include at least 2 different epitope databases; ii) to include at least 2 non-redundant algorithms; iii) to use ANN- and SVM-based algorithms (9,20).

Some analysis tools were explored but not selected (data not shown). EPIBOT does provide consensus predictions (NetMHC, SYFPEITHII, BIMAS, SVMHC and IEDB) but revealed no advantage over IEDB_consensus since it is based in a very restricted database of mouse-restricted epitopes (831 known epitopes from 397 proteins from IEDB). EPIMHC was excluded because it is based on a relatively limited database of 4867 distinct peptide sequences, that users must customise, and uses Position-Specific Scoring Matrices (PSSM) (RANKPEP). CTLPred is the only algorithm allowing predictions from the MHCBN database, but it was excluded because it does not preclude HLA restriction, and there are quite divergent results between the in-built ANN and SVM algorithms (data not shown). Although SYFPEITHI performs PSSM-based predictions, the associated epitope database is unique and valuable as it is manually curated and complementary to IEDB.

HLA-class I binding predictions were performed with: i) NetMHCpan 3.0 predictions for 11 supertype representative alleles (49), ii) IEDB MHC-I binding, prediction Method Version 2013-02-22, recommended predictions (consensus > ANN > SMM > NetMHCpan > CombLib) for 36 alleles (50), and iii) SYFPEITHI predictions (default predictions for 22 alleles) (16) (Table IV.6). Allele-specific binding affinity cut-off valued were applied (Table IV.4), and when not available, the general cut-off value of 500 nM was applied.

Table IV. 6 Selected databases and HLA-class I binding prediction algorithms.

Database	Prediction algorithm	Epitope selection
IEDB	MHC-I Binding IEDBrecommended	ANN_IC50 below allele-specific cut-off
IEDB + IMGT/HLA	NetMHCpan (ANN)	IC50 below allele-specific cut-off
SYFPEITHI	SYFPEITHI (PSSM)	No cutoff applied

HLA-class II binding predictions were performed with: i) NetMHCIIpan, pan-specific predictions for 21 alleles (51), and ii) IEDB MHC-II binding, recommended predictions (Consensus approach considers a combination of any three of the four methods, NN-align, SMM-align, CombLib and Sturniolo) for 21 alleles, and nn_align core and IC50 values (22,25) (Table IV.7). The top 10% of binding predictions was selected.

Table IV. 7 Selected databases and HLA-class II binding prediction algorithms.

Database	Prediction algorithm	Epitope selection
IEDB	MHC-II Binding IEDBrecommended	Top 10% predictions
IEDB + IMGT/HLA	NetMHCIIpan (ANN)	

The sequences of all 52 protein antigens, including 4 to 6 species-specific sequences, were retrieved from UniProtKB in fasta format. HLA-binding predictions were performed separately for each protein from Sets A and B.

HLA-binding predictions were performed through the online servers, where the fasta files were uploaded per protein. Results were downloaded as .csv or .html files. Result tables were further formatted to comply with the input data settings imposed by the epitope selection script developed in R (see below):

- HLA-class I predictions files contain 4 columns (allele|peptide|seq_num|score), and are named as 'proteincode_predictorcode.csv', e.g. "a1_iedb.csv".
- HLA-class II prediction files contain 6 columns (allele|seq_num|full_peptide|rank|core|ic50) and are named as 'proteincode_predictorcode.csv', e.g. "a1_iedbii.csv".

The data generated by each algorithm for each protein antigen were saved as .csv files and integrated in the epitope selection script developed in R.

2.3. Integration of HLA-binding prediction data

To merge and analyse protein-specific epitope prediction data, a selection script was developed in R (Figure IV.1 panel A). Briefly, the selection pipeline allows the integration of HLA-binding prediction data from multiple algorithms, and a first selection of conserved peptides (100% conserved among the six *Leishmania* species tested) and predicted by at least 2 non-redundant algorithms.

After this first selection step, short-BLASTp results are added (position-specific total mismatches and/or anchor position mismatches compared to host proteins). Briefly, in the BLASTp online server, the human RefSeq proteins were used as the host proteome (*Homo sapiens* Taxid:9606). The corresponding .txt result file was downloaded to the working directory and renamed "alignment_I.txt" or "alignment_II.txt" (Figure IV.1 panel A).

The result file generated by the selection script after these 2 selection steps was further analysed with the help of spreadsheet analysis software Excel® (Figure IV.1 panel B).

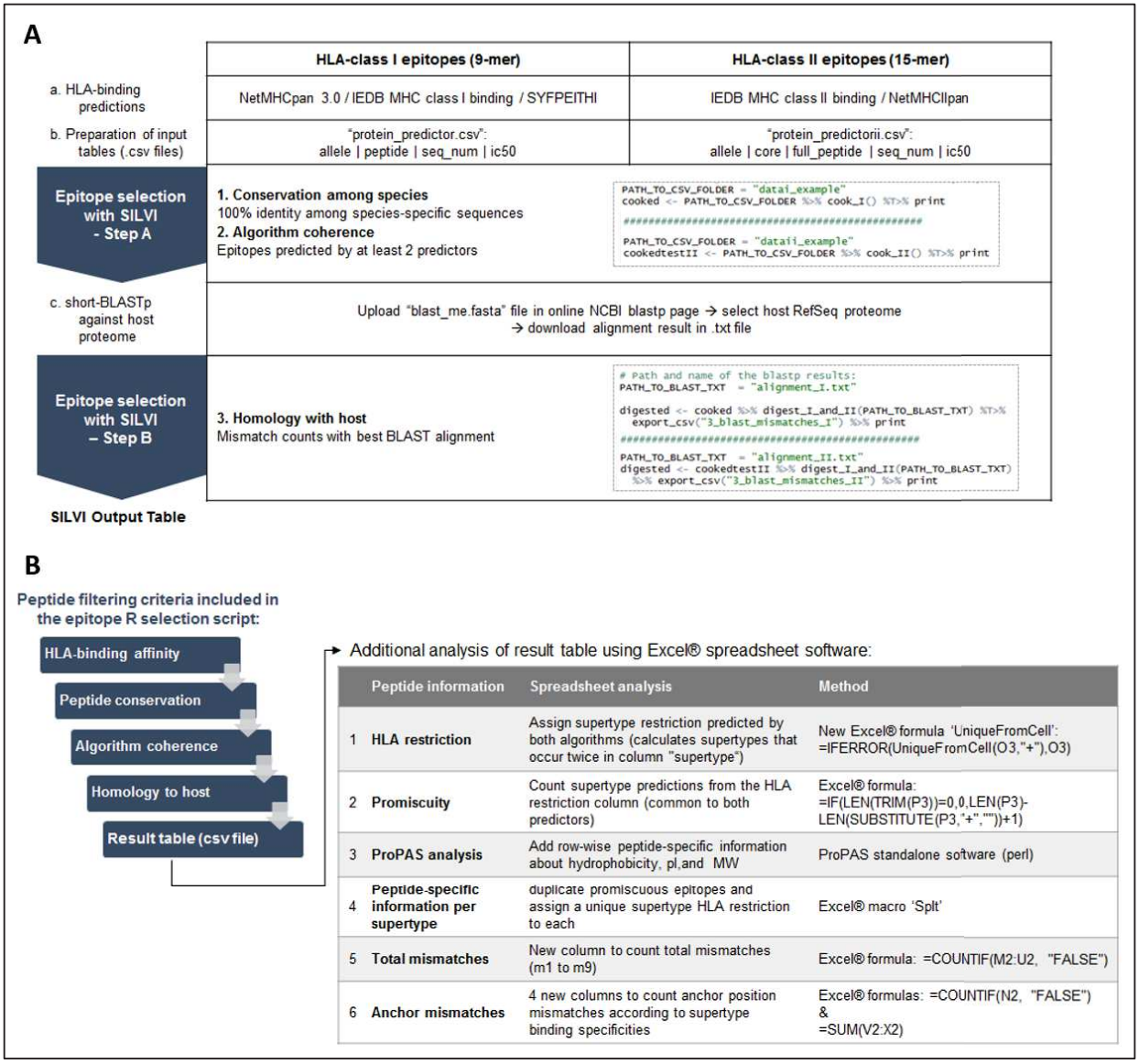


Figure IV. 1 T-cell epitope selection pipeline. *A) Epitopes were selecting through the application of relevant filters by a selection script in R. B) The result table containing conserved peptides was further analysed with spreadsheet software Excel®.*

The result table from the selection script contains the columns: peptide (epitope sequence), blastp results (middle; middle2; subject); position-specific mismatches (m_1 up to m_9), and protein supertype. Then, using Excel®, relevant physical-chemical properties were added (MW, hydrophobicity, pI) to the result table from the selection script after analysis of the peptide list with the ProPAS software (52). Moreover, epitope promiscuity and HLA restriction were calculated (Excel function 'UniqueFromCell' – Figure IV.2 panel A) and used to duplicate peptide information (Excel macro 'Splt' – Figure IV.2 panel B), so to generate a filterable table per supertype.

A

```
Function UniqueFromCell(rngCell, splitString)

Dim myCol
Set myCol = CreateObject("Scripting.Dictionary")
Dim duplicates As New Collection
Dim itmCol
Dim i As Long

Dim arrTMP As Variant
arrTMP = Split(rngCell, splitString)

For i = 0 To UBound(arrTMP)
    If myCol.Exists(CStr(arrTMP(i))) Then
        On Error Resume Next
        duplicates.Add arrTMP(i), CStr(arrTMP(i))
        On Error GoTo 0
    Else
        myCol.Add arrTMP(i), CStr(arrTMP(i))
    End If
Next i

If duplicates.Count <> 0 Then

    Dim result
    For Each itmCol In duplicates
        result = result & itmCol & splitString
    Next

    UniqueFromCell = Left(result, Len(result) - Len(splitString))
Else

    UniqueFromCell = ""

End If
End Function
```

B

```
Sub Splt()
Dim LR As Long, i As Long
Dim X As Variant
Application.ScreenUpdating = False
LR = Range("P" & Rows.Count).End(xlUp).Row
Columns("P").Insert
For i = LR To 1 Step -1
    With Range("Q" & i)
        If InStr(.Value, "+") = 0 Then
            .Offset(, -1).Value = .Value
        Else
            X = Split(.Value, "+")
            .Offset(1).Resize(UBound(X)).EntireRow.Insert
            .Offset(, -1).Resize(UBound(X) - LBound(X) + 1).Value =
Application.Transpose(X)
        End If
    End With
Next i
Columns("Q").Delete
LR = Range("P" & Rows.Count).End(xlUp).Row
With Range("A1:W" & LR)
    On Error Resume Next
    .SpecialCells(xlCellTypeBlanks).FormulaR1C1 = "=R[-1]C"
    On Error GoTo 0
    .Value = .Value
End With
Application.ScreenUpdating = True
End Sub
```

Figure IV. 2 Excel® software code (VBA code) for analysis of final peptide list, from the selection script in R. *A) VBA code for the creation of the 'UniqueFromCell' function, that identifies common HLA restriction between IEDB and NetMHCpan predictions. B) VBA code for the 'Splt' macro that duplicates peptide rows according to predicted (consensual) promiscuity.*

2.4. SILVI – an open-source pipeline for T-cell epitope selection

The final selection steps performed in the spreadsheet analysis software were integrated in the R selection script for publication purposes. The package Peptides was used for the addition of physical-chemical properties (MW, pI and hydrophobicity). Also, promiscuity and HLA restriction are now automatically calculated by the script. The final development version of the SILVI epitope selection pipeline will be published in Pissarra J et al 2019 (under review for publication in PLoS One).

2.5. Peptide toxicity assays

Cytotoxicity assays were performed with the PrestoBlue® viability dye. Briefly, $2x10^5$ human total PBMC were seeded per well in 96-well plates and incubated overnight at 37°C, 5% CO_2. Individual peptides were added at 3 different concentrations in triplicate wells (1 or 5, 10 and 25 μM). After 4 hours, 10% final PrestoBlue® (Life Technologies, Switzerland) was added to each well and the plates were returned to the incubator for 16 hours. The fluorescence intensity (bottom-read) was measured using a multiwell plate reader (EnVision 2105 Multimode Plate Reader, Perkin Elmer) excitation 560

nm, emission 595 nm. 560 nm absorbance values were normalized to the 595 nm values for the experimental wells.

3. Results

3.1. The T-cell epitope selection pipeline greatly reduced initial epitope lists

A T-cell epitope selection pipeline was developed *in-house* to optimise the process of epitope selection from vast amount of data produced by available algorithms, and to add extra relevant information, thus helping to refine the search of the most relevant epitopes through the application of filtering criteria. The R script reads epitope binding prediction data from different predictors, processes and compares data, assimilates BLASTp alignment results (53). The selection script feeds a final table with all relevant information to perform the desired selections. Additional epitope-specific information was added afterwards, in the spreadsheet analysis software Excel® (molecular weight, pI, hydrophobicity, NetMHCpan score).

The script successfully integrated epitope prediction information from all 52 proteins. Considering only one species-specific sequence per protein, these 52 proteins comprise 20712 9-mer overlapping peptides, and 20412 15-mer overlapping peptides (Figure IV.3). Peptides that are 100% identical in all species-specific sequences per protein were selected in the first selection step, so each peptide in the "blast_me.fasta" file are protein- and *Leishmania*-specific.

The result file "3_blast_mismatches_I.csv", after the second R script selection step, includes HLA-class I binding predictions for 1048 unique peptides from Set A proteins and 1069 unique peptides from Set B proteins. The total 2117 peptides correspond to 2277 predictions, after considering promiscuous peptides. These predictions are common between at least two algorithms (I+S, I+N, I+N+S).

The result file "3_blast_mismatches_II.csv", after the second R script selection step, includes HLA-class II binding predictions for 847 unique cores. These epitope core predictions are common among IEDB MHC-II binding and NetMHCIIpan (I+N).

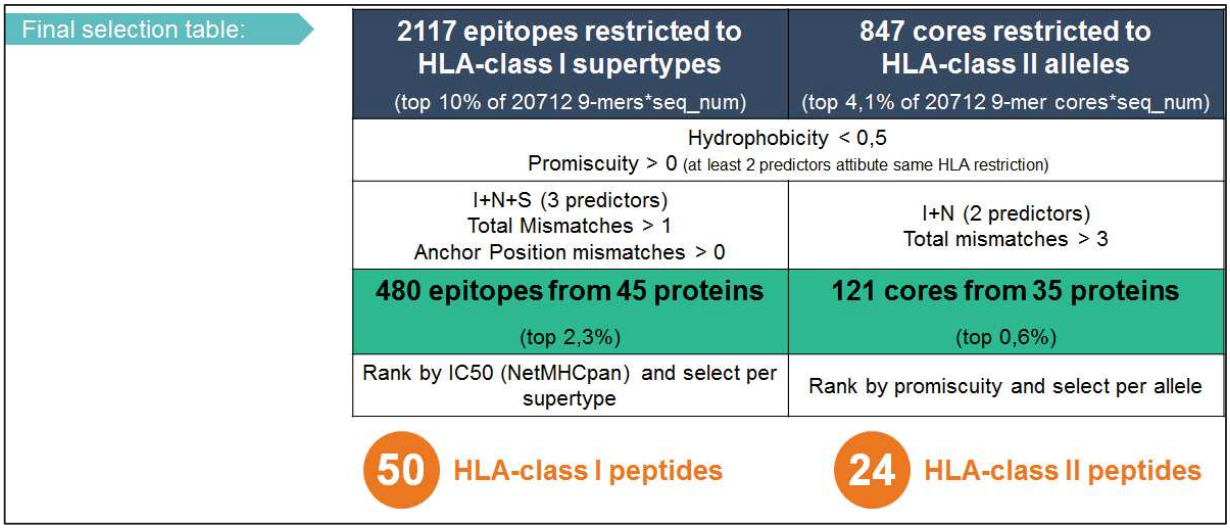


Figure IV. 3 Summary of selection filters and list reduction for HLA-class I and –class II binding predictions.

For HLA-class I predictions, the 20712 overlapping 9-mer peptides from 52 protein antigens were analysed. The allele-specific binding cut-off together with first script selection step reduced this list to 2117 unique peptides, conserved among species and predicted by three predictors (I+N+S) (Figure IV.3). The HLA-class I epitope list was reduced further with promiscuity and hydrophobicity filters (hydrophobity < 0,5; promiscuity > 0) after the first selection step. Moreover, we only selected peptides with at least 2 total mismatches and at least 1 anchor position mismatch. These filters reduced the list to 480 unique epitopes (Figure IV.3).

For HLA-class II binding predictions, the 20412 overlapping 15-mer full-length peptides from 52 protein antigens were analysed, each containing a 9-mer binding core. The top 10% predictions together with the first script selection step, conserved among species and predicted by two algorithms (I+N), reduced this list to 847 unique cores (Figure IV.3). The HLA-class II epitope list was further reduced to 121 unique cores with promiscuity, hydrophobicity and total mismatch filters (hydrophobity < 0,5; promiscuity > 0; and total mismatches > 2) after the first selection step (Figure IV.3).

3.2. Successful selection of 50 HLA-class I and 24 HLA-class II *Leishmania*-specific epitopes

The final epitope list containing 480 epitopes was ranked according predicted IC50, and selections were performed per supertype – most promiscuous and strongest binders per supertype. Finally, between 3 up to 6 epitopes per supertype were selected, for a total of 50 HLA-class I epitopes (Figure IV.4).

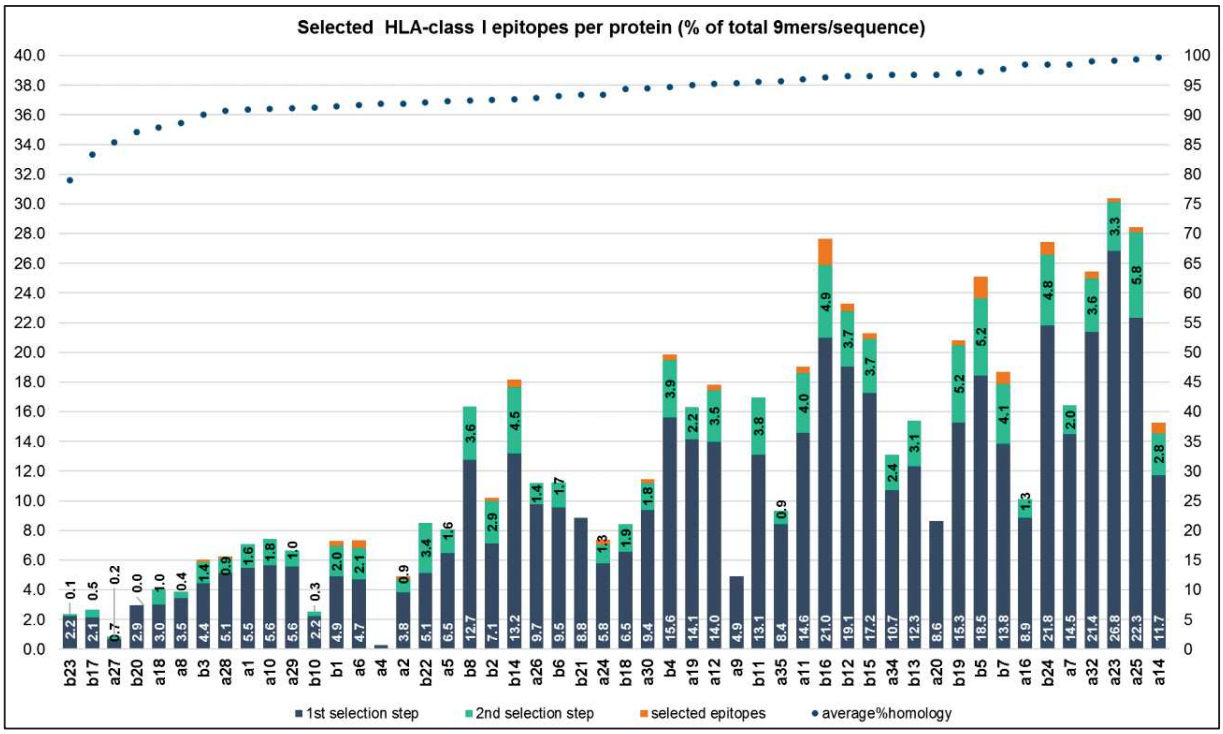


Figure IV. 4 Stepwise HLA-class I binding prediction results per protein antigen and protein conservation. *The first selection step (in blue) includes epitopes under allele-specific IC50 cut-off values, 100% conserved among species and predicted by at least 2 algorithms. These epitope lists were reduced in a second selection step (in green) with the application of filtering criteria (promiscuity>0; prediction by all 3 used algorithms; hydrophobicity<0,5; total mismatches > 1; anchor mismatches > 0). Epitopes were selected according to best predicted IC50 values per supertype (orange).*

The 50 HLA-class I restricted peptides come from 23 different protein antigens (11 from Set A; 12 from Set B). 20 epitopes are from Set A antigens and 30 are from Set B antigens.

All peptides are predicted to be strong binders, with average predicted IC50 of 40 nM (minimum 3,5 nM; maximum 111,5 nM). Most peptides are not predicted to be promiscuous, 6 epitopes are predicted to be promiscuous to 2 supertypes, and 2 epitopes are promiscuous to 3 allele supertypes.

The final HLA-class II epitope list containing 121 unique cores was ranked according predicted promiscuity, and selections were performed per allele. The most promiscuous epitopes were selected (20), and 4 other epitopes were selected according to best predicted IC50 for remaining alleles (not represented in the overall promiscuity prediction) (Figure IV.5).

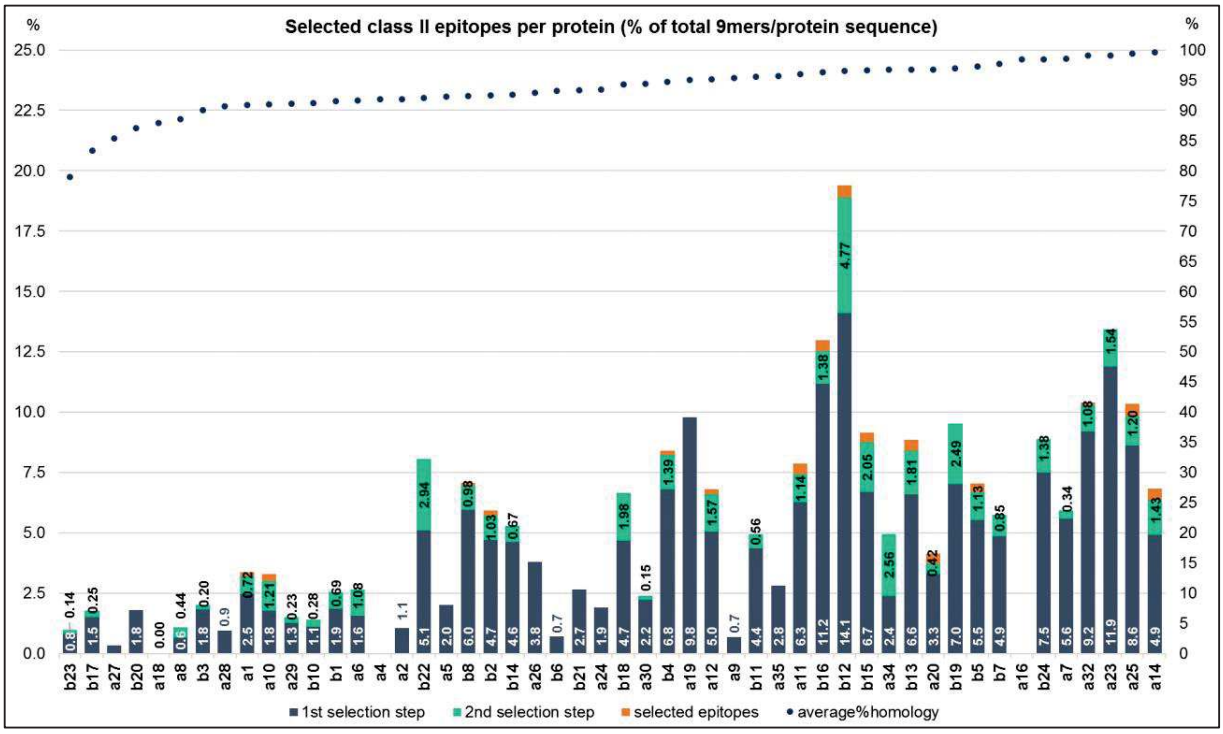


Figure IV. 5 Stepwise HLA-class II binding prediction results per protein antigen and protein conservation. *The first selection step (in blue) includes the top 10% of predicted epitopes, 100% conserved among species and predicted by at least 2 algorithms. These epitope lists were reduced in a second selection step (in green) with the application of filtering criteria (promiscuity>0; prediction by 2 algorithms; hydrophobicity<0,5; total mismatches > 3). Epitopes were selected according to predicted promiscuity (orange).*

The 24 HLA-class II-restricted epitopes come from 15 different protein antigens (7 from Set A; 8 from Set B). 12 epitopes originate from Set A antigens and 12 from Set B antigens. The average prediction percentile is 1,49% (minimum 0.01%; maximum 7.32%; median 0.53). These epitopes are predicted to be promiscuous, from 3 to 13 alleles. Between 4 to 18 epitopes per allele were selected.

For HLA-class II binding predictions, the total epitope size is 15-mer, however, the determining motif for HLA binding is the epitope core or register. The predictive power of different algorithms correlates better with the ability for core prediction (4). By comparing core predictions,

we compare different predictors and consider all potential cores within a 15-mer peptide, selecting the best core/full_peptide combination according to predicted IC50.

3.3. HLA-class I peptides are mostly water-soluble

From the 50 HLA-class I peptides ordered, one was never successfully synthesized (B44_2):

peptide	Charge	Hydrophobicity	pI	Protein	Supertype	IC50	Mismatches (Total/Anchor)
B44_2	0 (neutral)	-0.1999	5,99	b16	A2+B44+B44	N_18,6	4/1

Upon peptide solubilisation, 43 peptides are soluble in water; 5 soluble in sodium bicarbonate 0.5M solution; and 1 soluble only in DMSO.

24 HLA-class II peptides were directly and successfully solubilised in 10% DMSO.

3.4. Selected HLA-class I peptides are not toxic to human cells

No toxicity towards human PBMC was detected for any of the 49 HLA-class I peptides tested, regardless of peptide concentration (Figure IV.6).

Peptide toxicity assays confirm the synthetic peptides can be used for *in vitro* cellular stimulation of human peripheral immune cells, in subsequent experimental validation steps. Peptide toxicity assays were not performed with class II peptides because of reduced peptide availability.

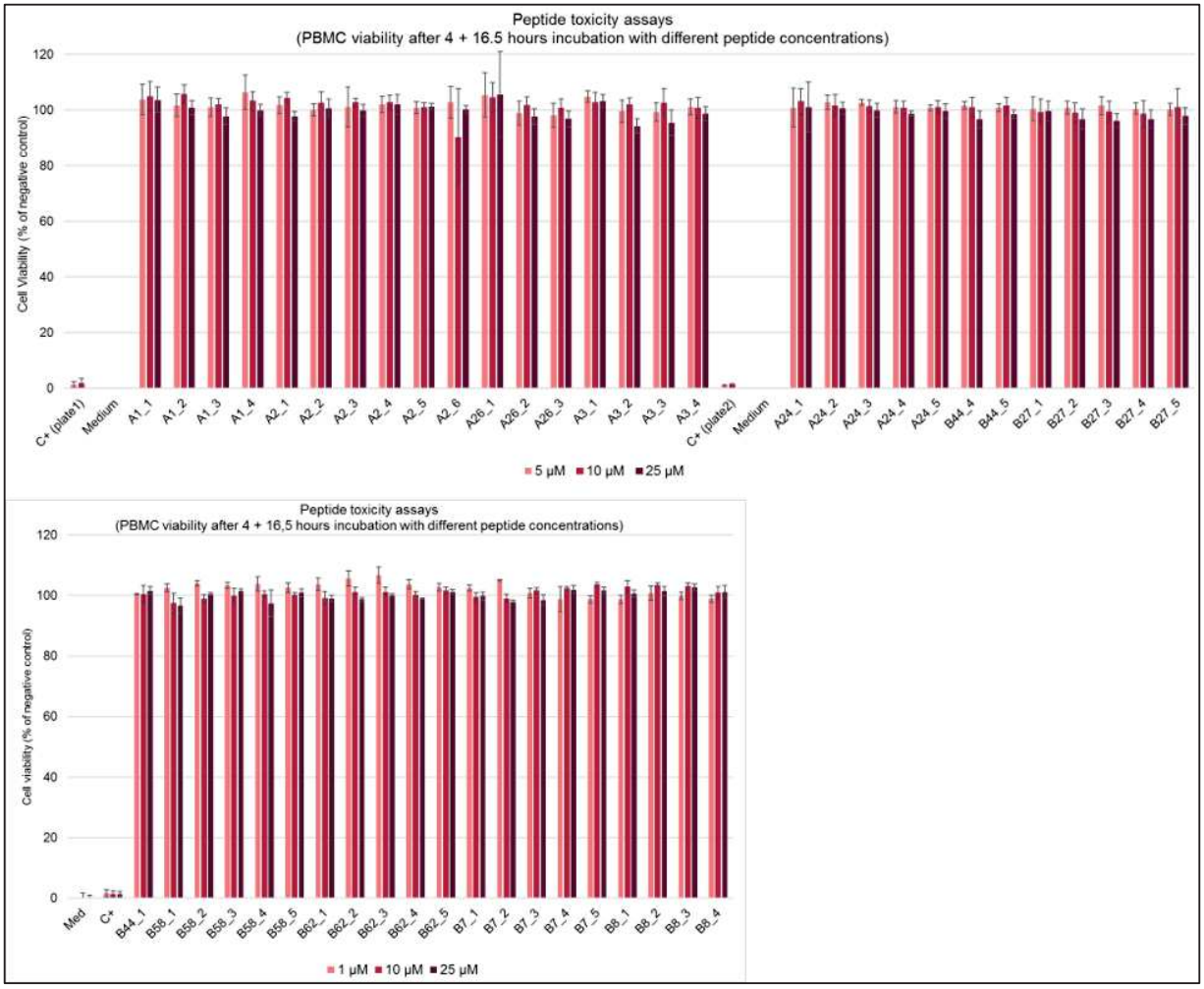


Figure IV. 6 Peptide toxicity assay results. *Peptide toxicity was assessed for 49 9-mer peptides in human total PBMC at three concentrations (1 or 5 µM, 10 µM and 25 µM) with PrestoBlue® Viability dye after 20-hour incubation.*

4. Discussion

Experimental validation steps limit the number of peptide candidates to test, meaning that these candidates must be carefully selected to increase the chances of selecting immunogenic peptides. The addition of an initial *in silico* screening in the development pipeline greatly diminishes costs associated with epitope mapping experiments by decreasing the number of peptides to test. Nevertheless, current *in silico* epitope prediction tools still present some shortcomings (Figure IV.8). The lack of gene and protein annotations as well as an underrepresentation of protozoan-derived epitopes in the databases renders sequence-based prediction less accurate (35). Similarly, protozoan proteins with different physicochemical composition are not represented in the databases (35). *Leishmania*-specific epitopes are underrepresented in epitope databases. To date, the IEDB database contains a total of 538,374 peptidic epitopes, which include only 965 *Leishmania* epitopes (IEDB search: any epitopes + *Leishmania* ID 5658, February 2019 annotation). Of these, only 555 are T-cell epitopes, and if we consider positive assays only this number decreases to 379, corresponding to less than 1% of total 448402 T-cell epitopes with positive results in the database.

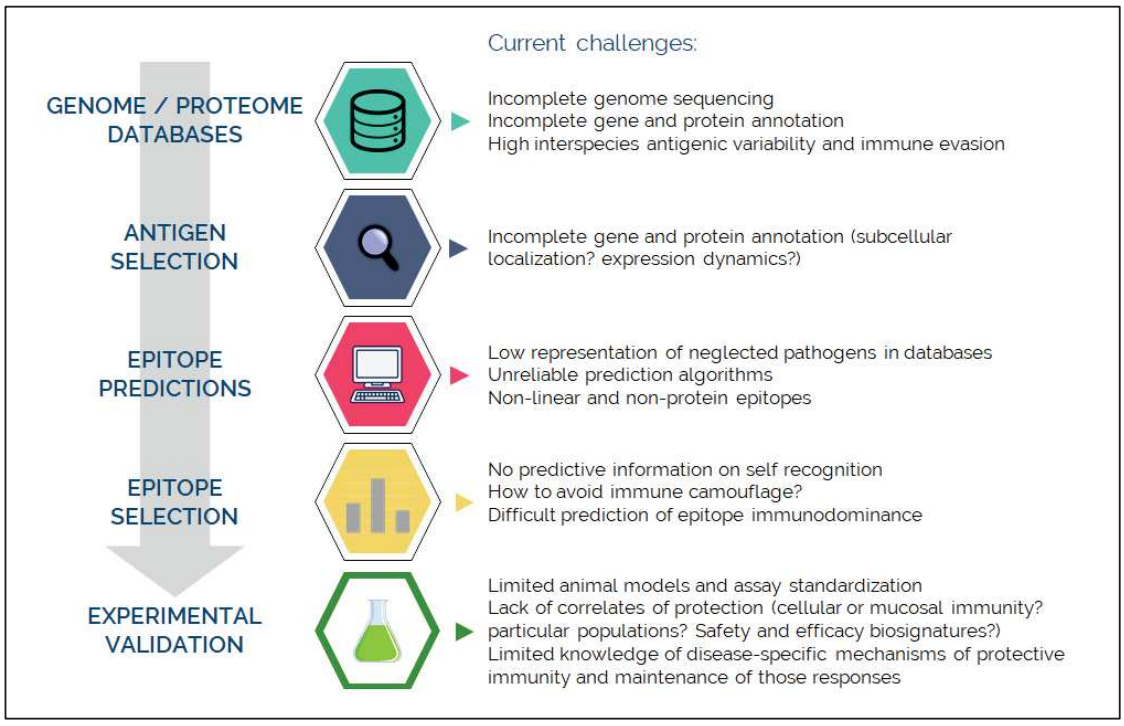


Figure IV. 7 Current challenges affecting reverse vaccinology approaches for vaccine design.

The term reverse vaccinology was first described by Rappuoli R. (31) to designate approaches that screen entire genomes or proteomes to filter out proteins of interest (antigens) using bioinformatic tools. Subsequently, particular features associated with antigenicity are searched, namely, subcellular localisation or expression timing.

At the time of the present analysis, only four (true) reverse vaccinology studies had been performed for vaccine development against *Leishmania*, with modest success in identifying strong epitopes (54,55). Briefly, most studies used BLASTp alignment to eliminate peptides or proteins with significant homology to host proteins (human or mouse). Also, most studies used sequence-based HLA-binding predictions for T-cell epitope discovery.

In the first study, in 2009, Herrera-Najera et al analysed the complete *L. major* proteome (8272 annotated proteins) and performed class I epitope predictions for mouse alleles. The first epitope prediction round used the algorithm RANKPEP based on position-specific score matrices (PSSM), subsequently, proteins with top scoring peptides were re-analysed using multiple HLA-binding prediction algorithms to generate consensus predictions (SYFPEITHI, BIMAS, ProPred-I, and MAPPP which are PSSM-based; ANNPred; SVMHC; ComPred which is combines ANN and QM; and Predep, a structure-based algorithm). The last analysis step involved a BLASTp analysis against human and mouse proteins, wherein epitopes showing over 80% of sequence identity were discarded, and epitopes conserved among kinetoplastids were prioritised. However, predictions were only performed for mouse alleles (H-2K^{d} andH-2D^{d}). Also, proteasomal cleavage filters were applied, potentially introducing a bias in epitope prediction. Experimental validation was performed in mice. This approach eliminated the known candidates GP63, LACK, histone 2B, LmSTI1, TSA, CPb, NH36 or beta-tubulin in the first selection step.

Similarly, Guerfali FZ et al also used the complete *L. major* proteome for epitope prediction. Again, predictions were performed only for mouse alleles (BALB/c and C57BL/6 alleles) and with redundant PSSM-based algorithms (SYFPEITHI, BIMAS, RANKPEP). BLASTp analysis was also used to exclude epitopes with high levels of similarity with host proteins, but only 100% identical sequences were excluded. Moreover, SignalP predictions were used to select secreted protein antigens for epitope prediction, which is extremely unsensitive for *Leishmania* excreted-secreted proteins since, as shown by the literature and in chapter II of the present thesis, only up to 10% of proteins in the secretome are classically secreted.

In 2012, John L. et al retrieved total protein sequences from *L. major* and *L. infantum* and 8122 common proteins were identified through BLASTp analysis. Subcellular localisation analysis was performed with PSORT and TMHMM to select cell surface and secreted proteins (cytoplasmic

proteins, and proteins with more than one transmembrane domain were excluded), and BLASTp analysis was used to select non-homologous proteins to human and mouse. Finally, epitopes were predicted through HLA-binding algorithms for class I alleles (BIMAS, SYFPEITHI, ProPred1 predictions for human alleles HLA_0201, HLA-A2, HLA-A 0205, HLA-Cw 0602, HLA- A2.1, HLA-A3, HLA-B14, HLA-B 5401, and HLA-B 5102) and for class II alleles (ProPred predictions for human alleles DRB1_0101, DRB1_0102, DRB1_1101, DRB1_1104, DRB1_1501, DRB1_1502, DRB1_0402, DRB1_0404, DRB1_0405, DRB1_1301, and DRB1_1302). All predicted peptides were again 'blasted' against the human and mouse proteomes, 19 peptides showed no similarity and were selected but not validated experimentally.

Overall, these RV studies focus on mouse alleles instead of human alleles (except John L et al), so extrapolation for human immune responses is very limited. Also, the filtering criteria used (proteasomal cleavage, SignalP and PSORT predictions) are likely to reject important *Leishmania*-specific proteins and/or epitopes. Discovered peptides were either validated in the mouse model or not at all.

In 2016, Freitas e Silva et al performed a reverse vaccinology approach followed by peptide validation in human samples. This study is detailed in chapter I - Table I.4 (*Leishmania*-specific peptide vaccine candidates validated using human samples) and describes the most complete RV approach. It includes the proteomes of three *Leishmania* species (*L. braziliensis*, *L. major*, and *L. infantum*), it uses one of the best available HLA-binding prediction algorithms (NetMHC), BLASTp analysis and structure-based epitope prediction. However, the proposed pipeline uses intensive computational analysis, namely structural modelling to HLA molecules, which is difficult to reproduce due to its extensive computational resource requirements.

We believe the successful identification of immunogenic epitopes depends on both the quality of used immunoinformatic algorithms and on the rationality of epitope selection criteria. These criteria should be permissive enough not to falsely reject immunogenic epitopes, and restrictive enough to effectively filter HLA-binding prediction data. Hence, there is an optimal balance between the used tools and the criteria chosen to filter their results, which sought to optimise. Overall, our proposed reverse vaccinology pipeline uses proteomic information from six *Leishmania* species, the most clinically relevant ones (*L. braziliensis*, *L. major*, *L. infantum, L. donovani*, *L.amazonensis* and *L. tropica*), it makes use of BLASTp analysis and the best available HLA-binding prediction algorithms. Because our starting sample is the *Leishmania* secretome, an optimal antigenic source, no subcellular localization filters need to be applied.

The best sequence-based HLA-binding prediction tools currently available were used in this study (24–27). Structure-based epitope prediction tools are also available, which combine binding data, sequence information, known crystallographic data of pMHC complexes, and computational modelling. However, structure-based prediction tools only apply to the few alleles with crystallographic data and PDB files (Protein Data Bank), thereby excluding alleles and affecting vaccine coverage or introducing bias in predictions for other alleles (56). Moreover, these tools require great computational processing power and time. Altogether, it is more practical and accurate to use sequence-based pan-specific methods and based on machine-learning algorithms for epitope prediction.

Some previously described *Leishmania* protein antigens were excluded from epitope prediction analysis. The antigen KMP11 was removed due to ambiguous results showing evidence of induced Th2 responses (57). Also, the LACK antigen was excluded since it may have an important role in parasite immunomodulation of host immune responses. Interestingly, studies have shown that LACK-specific TCD4^{+} cells are present in individuals never exposed to *Leishmania* (4,5). These cells are primed in gut-associated lymphoid tissues by cross-reactive microbial antigens and are able to quickly secrete IL-4 (6). Finally, the promastigote surface antigen was not included in the present epitope prediction analysis because this analysis previously performed by our team, both soluble and membrane-associated isoforms (58–60).

Furthermore, after performing *in silico* epitope prediction, the ability to compare different data sources and to synergistically combine diverse algorithms in the context of epitope prediction remains challenging. To this end, the selection script developed in R greatly assisted in data integration and epitope selection.

The binding affinity and peptide conservation filters applied to the result table allowed a 90% reduction of the initial epitope list. Despite this broad selection, testing over 2100 9-mer peptides or over 840 15-mer peptides is still an issue for the experimental validation assays and experiment costs. To further reduce this list, peptides were selected by relative comparison according to the predicted HLA restriction, promiscuity and binding affinity to select the best epitopes per supertype or allele. Binding affinity is a key characteristic of peptide immunogenicity and the general cut-off value of 500

nM has been extensively used in T-cell epitope selection. Allele-specific binding information allows the establishment of specific thresholds which correlate well with epitope prediction (61).

Homology was used firstly as a positive selection filter for highly conserved epitopes among *Leishmania* species. Logically, epitopes from proteins that show higher homology are more likely to be 100% conserved and, therefore, more likely to be selected. However, the 52 antigens processed are generally conserved, with between 60% up to 99% homology levels. Still, because selections were performed according to predicted IC50, there are conserved epitopes selected from less conserved proteins – e.g. B3, A28 proteins for HLA class-I epitopes, and A1, A10 for class-II epitopes.

Homology to host was then used as negative selection filter, not to exclude prediction data, but for epitope ranking and description. In this case, epitopes were aligned with human RefSeq proteome, and position-specific mismatches were counted. This allowed us to establish a low stringency threshold for homology with human proteins, (at least 2 amino-acids in 9-mer peptides or at least 3 amino-acid mismatches for 15-mer peptides). Although the mechanisms eliciting self-recognition and autoimmune responses are not fully understood, and therefore difficult to predict, the selection of epitopes with lower homology to host proteins is likely to decrease the chances of unwanted cross-reactivity.

The hydrophobicity filter, although not sensitive to immunogenicity, allows the discovery of soluble epitopes, which can be easily produced, solubilised and tested *in vitro*. After peptide synthesis, a vast majority of class I peptides was found to be water soluble, which facilitates handling. Solubility assays were not performed for class II peptides because of reduced peptide availability – HLA-class II peptides were directly solubilised in 10% DMSO to improve peptide stability upon freezing.

After peptide selection, the selection script was adapted to accommodate the steps performed in the spreadsheet analysis software and will be published under the title “Exploring -omics datasets for epitope-based development of vaccines and therapeutics – SILVI, an open-source pipeline for T-cell epitope selection”. This publication aims to make SILVI available to the community and, hopefully, assist future studies in epitope selection analysis. In the manuscript, an example protein was included (Hepatitis C Virus Genome polyprotein P26664) to illustrate the workflow and SILVI’s potential in assisting epitope selection. When compared to validated epitopes in the IEDB, the application of low stringency filters (promiscuity > 0 and scoreN under 1000 nM) to class I epitope predictions results in 81% sensitivity and 82% specificity (out of 614 peptides selected, 100 are P26664 validated epitopes in the IEDB). Notably, these results are not improved after proteasomal cleavage analysis.

Finally, the final epitope selections include proteins from both antigen datasets (Sets A and B). For the HLA-class I epitope list, Set A proteins correspond to 40%, with 20 epitopes compared to 30 from Set B proteins. For HLA-class II epitope lists, both datasets are equally represented (12 each). It is noteworthy that the new antigen candidates in Set B are equally or even potentially more immunogenic than known protein candidates.

We successfully selected 50 HLA-class I- and 24 HLA-class II-restricted epitopes, through an epitope selection pipeline using high-performing HLA-binding predictions (27), and homology alignments (53). Based on this information, together with conservation among species, physico-chemical properties and target population HLA restriction, prediction data was filtered, thereby reducing the number of peptides to test experimentally while increasing the chances of identifying immunogenic peptides.

The final peptide list includes high affinity binders with high immunogenicity potential according to *in silico* predictions. The synthetic peptides, and respective predicted immunogenicity, will be validated with experimental assays, detailed in Chapter V.

5. Bibliography

1. Murphy K, Weaver C. Glossary. In: Janeway's Immunobiology. 9th ed. Garland Science; 2017. p. 818–54.
2. Calis JJA, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, et al. Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. Asquith B, editor. PLoS Comput Biol. Public Library of Science; 2013 Oct 24;9(10):e1003266.
3. Gaddis DE, Fuller MJ, Zajac AJ. 6 CD8 T-cell Immunodominance, Repertoire, and Memory. In: Frelinger JA, editor. Immunodominance: The Choice of the Immune System. 1st ed. WILEY-VCH; 2006. p. 109–35.
4. Singh SP, Mishra BN. Major histocompatibility complex linked databases and prediction tools for designing vaccines. Hum Immunol. 2016;77(3):295–306.
5. Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, et al. The immune epitope database (IEDB) 3.0. Nucleic Acids Res. Oxford University Press; 2015 Jan 28;43(D1):D405–12.
6. Lund O, Nielsen M, Lundegaard C, Kesmir C, Brunak S. Immunological Bioinformatics. Methods. Cambridge, MA: MIT Press; 2005.
7. Tomar N, De RK. Immuno-informatics: A Brief Review. 2nd ed. Tomar N, De RK, editors. Vol. 1184, Immunoinformatics. Methods in Molecular Biology (Methods and Protocols). New York, NY: Humana Press; 2014.
8. Liljeroos L, Malito E, Ferlenghi I, Bottomley MJ. Structural and Computational Biology in the Design of Immunogenic Vaccine Antigens. J Immunol Res. 2015;156241.
9. Tung C-W. Chapter 6 - Databases for T-Cell Epitopes. In: De RK, Tomar N, editors. Immunoinformatics. New York, NY: Humana Press; 2014. p. 123–34.
10. Salimi N, Fleri W, Peters B, Sette A. Design and utilization of epitope-based databases and predictive tools.
11. Fleri W, Vaughan K, Salimi N, Vita R, Peters B, Sette A. The Immune Epitope Database: How Data Are Entered and Retrieved. J Immunol Res. 2017;2017.
12. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res. 2019 Jan 8;47(D1):D339–43.
13. Robinson J, Halliwell JA, McWilliam H, Lopez R, Marsh SGE. IPD—the Immuno Polymorphism Database. Nucleic Acids Res. Narnia; 2012 Nov 23;41(D1):D1234–40.
14. Robinson J, Halliwell JA, McWilliam H, Lopez R, Parham P, Marsh SGE. The IMGT/HLA

database. Nucleic Acids Res. Narnia; 2012 Oct 17;41(D1):D1222–7.

15. Brusic V, Rudy G, Harrison LC. MHCPEP: a database of MHC-binding peptides. Nucleic Acids Res. 1994 Sep;22(17):3663–5.
16. Rammensee H-G, Bachmann J, Emmerich NN, Bachor OA, Stevanovic S. SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics. 1999;50:213–9.
17. Lata S, Bhasin M, Raghava GPS. MHCBN 4.0: A database of MHC/TAP binding peptides and T-cell epitopes. BMC Res Notes. 2009 Jan;2:61.
18. Backert L. Immunoinformatics and epitope prediction in the age of next-generation sequencing and personalized medicine. Genome Med. Genome Medicine; 2015;1–12.
19. Lundegaard C, Hoof I, Lund O, Nielsen M. State of the art and challenges in sequence based T-cell epitope prediction. Immunome Res. 2010 Jan 3;6(Suppl 2):S3.
20. V. Desai D, Kulkarni-Kale U. Chapter 19 - T-Cell Epitope Prediction Methods: An Overview. In: De RK, Tomar N, editors. Immunoinformatics Methods in Molecular Biology (Methods and Protocols). 2nd ed. New York: Humana Press; 2014.
21. Morrow WJW, Sheikh NA, Schmidt CS, Davies DH. Vaccinology: Principles and Practice. Vaccinology: Principles and Practice. Wiley-Blackwell; 2012.
22. Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, et al. Peptide binding predictions for HLA DR, DP and DQ molecules. BMC Bioinformatics. 2010;11(1):568.
23. Zhang L, Udaka K, Mamitsuka H, Zhu S. Toward more accurate pan-specific MHC-peptide binding prediction: A review of current methods and tools. Brief Bioinform. 2012 May 1;13(3):350–64.
24. Lin HH, Zhang GL, Tongchusak S, Reinherz EL, Brusic V. Evaluation of MHC-II peptide binding prediction servers: applications for vaccine research. BMC Bioinformatics. 2008;9 Suppl 12:S22.
25. Wang P, Sidney J, Dow C, Mothé B, Sette A, Peters B. A Systematic Assessment of MHC Class II Peptide Binding Predictions and Evaluation of a Consensus Approach. PLoS Comput Biol. 2008;4(4):e1000048.
26. Peters B, Bui H-H, Frankild S, Nielsen M, Lundegaard C, Kostem E, et al. A Community Resource Benchmarking Predictions of Peptide Binding to MHC-I Molecules. PLoS Comput Biol. Public Library of Science; 2006;2(6):e65.
27. Trolle T, Metushi IG, Greenbaum JA, Kim Y, Sidney J, Lund O, et al. Automated benchmarking of peptide-MHC class I binding predictions. Bioinformatics. 2015;31(13):2174–2181.
28. Mallios RR. A consensus strategy for combining HLA-DR binding algorithms. Hum Immunol.

Elsevier; 2003 Sep 1;64(9):852–6.

29. Castelli M, Cappelletti F, Diotti RA, Sautto G, Criscuolo E, Dal Peraro M, et al. Peptide-based vaccinology: experimental and computational approaches to target hypervariable viruses through the fine characterization of protective epitopes recognized by monoclonal antibodies and the identification of T-cell-activating peptides. Clin Dev Immunol. 2013 Jan;2013(ID 521231).
30. Ferrante A, Gorski J. Cooperativity of hydrophobic anchor interactions: evidence for epitope selection by MHC class II as a folding process. J Immunol. 2007;178(11):7181–9.
31. Sette A, Rappuoli R. Reverse Vaccinology: Developing Vaccines in the Era of Genomics. Immunity. 2010;33(4):530–41.
32. Aebischer T. Leishmania spp. Proteome Data Sets: A Comprehensive Resource for Vaccine Development to Target Visceral Leishmaniasis. Front Immunol. Frontiers; 2014 Jan 10;5:260.
33. Rock KL, Shen L. Cross-presentation: underlying mechanisms and role in immune surveillance. Immunol Rev. 2005 Oct 1;207(1):166–83.
34. Vivona S, Gardy JL, Ramachandran S, Brinkman FSL, Raghava GPS, Flower DR, et al. Computer-aided biotechnology: from immuno-informatics to reverse vaccinology. Trends Biotechnol. 2008;26(4):190–200.
35. Resende DM, Rezende AM, Oliveira NJ, Batista ICA, Corrêa-Oliveira R, Reis AB, et al. An assessment on epitope prediction methods for protozoa genomes. BMC Bioinformatics. 2012;13(1):309.
36. De Groot AS, Moise L, Liu R, Gutierrez AH, Tassone R, Bailey-Kellogg C, et al. Immune camouflage: Relevance to vaccines and human immunology. Hum Vaccin Immunother. 2015;10(12):3570–5.
37. Noazin S, Modabber F, Khamesipour A, Smith PG, Moulton LH, Nasseri K, et al. First generation leishmaniasis vaccines: A review of field efficacy trials. Vaccine. Elsevier; 2008 Dec 9;26(52):6759–67.
38. Weber CA, Mehta PJ, Ardito M, Moise L, Martin B, De Groot AS. T cell epitope: Friend or Foe? Immunogenicity of biologics in context. Adv Drug Deliv Rev. 2009;61(11):965–76.
39. Settleman J, Sawyers CL, Hunter T. Challenges in validating candidate therapeutic targets in cancer. Elife. eLife Sciences Publications Limited; 2018 Feb 8;7:e32402.
40. Kaufmann SH, Juliana McElrath M, Lewis DJ, Del Giudice G. Challenges and responses in human vaccine development. Curr Opin Immunol. 2014;28:18–26.
41. Hagan T, Nakaya HI, Subramaniam S, Pulendran B. Systems vaccinology: Enabling rational

vaccine design with systems biological approaches. Vaccine. 2015 Apr 6;33(40):5294–301.

42. Yewdell JW, Bennink JR. Immunodominance in major histocompatibility complex class I–restricted T lymphocyte responses. Annu Rev Immunol. 1999;17:51–88.
43. Lazarski CA, Chaves FA, Jenks SA, Wu S, Richards KA, Weaver JM, et al. The kinetic stability of MHC class II:peptide complexes is a key parameter that dictates immunodominance. Immunity. 2005 Jul;23(1):29–40.
44. Sant AJ, Chaves FA, Jenks SA, Richards KA, Menges P, Weaver JM, et al. The relationship between immunodominance, DM editing, and the kinetic stability of MHC class II:peptide complexes. Immunol Rev. 2005 Oct 1;207(1):261–78.
45. Weiskopf D, Angelo MA, De Azeredo EL, Sidney J, Greenbaum JA, Fernando AN, et al. Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8 + T cells. 2013;
46. Greenbaum J, Sidney J, Chung J, Brander C, Peters B, Sette A. Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. Immunogenetics. 2011 Jun;63(6):325–35.
47. Manczinger M, Kemény L. Peptide presentation by HLA-DQ molecules is associated with the development of immune tolerance. PeerJ. PeerJ, Inc; 2018;6:e5118.
48. Koren E, De Groot AS, Jawa V, Beck KD, Boone T, Rivera D, et al. Clinical validation of the "in silico" prediction of immunogenicity of a human recombinant therapeutic protein. Clin Immunol. 2007;124(1):26–32.
49. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J Immunol. 2017 Nov 1;199(9):3360–8.
50. Moutaftsi M, Peters B, Pasquetto V, Tscharke DC, Sidney J, Bui HH, et al. A consensus epitope prediction approach identifies the breadth of murine TCD8+-cell responses to vaccinia virus. Nat Biotechnol. 2006 Jul 11;24(7):817–9.
51. Andreatta M, Karosiene E, Rasmussen M, Stryhn A, Buus S, Nielsen M. Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification. Immunogenetics. NIH Public Access; 2015 Nov;67(11–12):641–50.
52. Wu S, Zhu Y. ProPAS: standalone software to analyze protein properties. Bioinformation. Biomedical Informatics Publishing Group; 2012;8(3):167–9.
53. McGinnis S, Madden TL. BLAST: At the core of a powerful and diverse set of sequence analysis

tools. Nucleic Acids Res. 2004 Jul 1;32:W20–5.

54. John L, John GJ, Kholia T. A Reverse Vaccinology Approach for the Identification of Potential Vaccine Candidates from Leishmania spp. Appl Biochem Biotechnol. 2012;167(5):1340–50.
55. Herrera-Najera C, Piña-Aguilar R, Xacur-Garcia F, Ramirez-Sierra MJ, Dumonteil E. Mining the Leishmania genome for novel antigens and vaccine candidates. Proteomics. WILEY-VCH Verlag; 2009 Mar;9(5):1293–301.
56. He L, Zhu J. Computational tools for epitope vaccine design and evaluation. Curr Opin Virol. Elsevier; 2015 Apr 1;11:103–12.
57. Agallou M, Athanasiou E, Koutsoni O, Dotsika E, Karagouni E. Experimental validation of multi-epitope peptides including promising MHC class I- and II-restricted epitopes of four known Leishmania infantum proteins. Front Immunol. 2014;5(June):1–16.
58. Chamakh-Ayari R, Bras-Gonçalves R, Bahi-Jaber N, Petitdidier E, Markikou-Ouni W, Aoun K, et al. In vitro evaluation of a soluble Leishmania promastigote surface antigen as a potential vaccine candidate against human leishmaniasis. PLoS One. 2014;9(5):1–12.
59. Petitdidier E, Pagniez J, Papierok G, Vincendeau P, Lemesre J-L, Bras-Gonçalves R. Recombinant Forms of Leishmania amazonensis Excreted/Secreted Promastigote Surface Antigen (PSA) Induce Protective Immune Responses in Dogs. Louzir H, editor. PLoS Negl Trop Dis. World Intellectual Property Organization; 2016 May 25;10(5):e0004614.
60. Bras-Gonçalves R, Petitdidier E, Pagniez J, Veyrier R, Cibrelus P, Cavaleyra M, et al. Identification and characterization of new Leishmania promastigote surface antigens, LaPSA-38S and LiPSA-50S, as major immunodominant excreted/secreted components of L. amazonensis and L. infantum. Infect Genet Evol. Elsevier; 2014 Jun 1;24:1–14.
61. Paul S, Weiskopf D, Angelo MA, Sidney J, Peters B, Sette A. HLA class I alleles are associated with peptide binding repertoires of different size, affinity and immunogenicity. 2014;191(12):6–10.

CHAPTER V

EXPERIMENTAL VALIDATION OF SYNTHETIC PEPTIDES

Objectives: to assess T-cell-dependent peptide immunogenicity and discover *Leishmania*-specific epitopes restricted to HLA-class I and -class II alleles for vaccine development, using immunoassays exclusively with human samples.

Aims:

- to evaluate peptide *in vitro* immunogenicity in both naive and healed immune backgrounds;
- to validate the *in silico* epitope predictions;
- to select the most immunogenic peptides that will proceed for multi-epitope peptide design

1. Introduction

Immunogenicity testing is a key aspect in the development of vaccines, immunotherapies or biological therapeutics, from both a regulatory standpoint and a preclinical development perspective (1). Researchers may seek unwanted immunogenic regions within therapeutic molecules, for instance, therapeutic antibodies, or, on the other hand, immunogenic epitopes to induce specific protective immune responses through vaccination or immunotherapy. In either case, immunoassays must be performed, during clinical development and post-marketing surveillance, to identify and assess the ability of a given molecule to induce humoral or cellular immune responses in exposed or naive individuals.

The mechanisms of immunogenicity include T cell-dependent and T cell-independent responses. In the context of T-cell dependent immune responses, T-cell epitope mapping can be performed initially through *in silico* methods, subsequently validated with *in vitro* and *in vivo* assays, to describe Th1, Th2, cytotoxic, or regulatory epitopes. The most common marker associated with Th1-cell activation is IFN-γ, however, a maximum number of parameters should be included in the characterisation of induced responses.

There are currently several methods available to assess immunogenicity and T-cell responses *in vitro*, summarised in Figure V.1. Importantly, several tools exist to assess the magnitude and breadth of specific responses, as well as to describe cell function according to cytokine production or cell phenotype.

A

Objective	Assay	Readout
Cytokine production	**ELISA**: IFN-γ or IL-10 production	magnitude of response (amount of cytokine produced); type and quality of response
	CBA (cytokine bead assay): Multiple cytokine production (IFN-γ / TNF-α / Granzyme-B / IL-2 / IL-4 / IL-12 / IL-17 etc.)	magnitude of response (amount of cytokine produced); type and quality of response + multi-analyte and high-throughput
	ELISpot: IFN-γ production	Magnitude of response (number of cytokine-producing cells); very sensitive (up to 1 specific cell per million)
T cell phenotype	Expression of **surface markers on specific T cells** (CD3 / CD4 / CD3 / CD45RO / CD45RA / CCR7)	Phenotype of responder cells
	ICS - Intracellular Cytokine Staining (IFN-γ / TNF-α / Granzyme-B / IL-2)	Link between cytokine secretion and cell phenotype; accurately measures polyfunctionality
T cell proliferation	**CFSE staining** (carboxyfluorescein succinimidyl ester)	Cell proliferation in response to stimulation: easy to perform; co-staining can distinguish Tregs and T effector phenotypes (Th1, Th2, Th17, Th22)

B

CFSE staining < ELISA < ELISpot < CBA < surface marker staining < ICS

Figure V. 1 Summary of the current toolbox of cell-based assays to assess T-cell-dependent immunogenicity, *adapted from (2). A) In vitro T cell assay methods and respective application and readout. B) Advantages and disadvantages regarding the different methods.*

The CEF peptide pools are commonly used as positive control for peptide T cell activation in assays using total PBMC, both for HLA class I and class II peptide presentation (3). These pools contain well-defined epitopes capable of inducing memory cellular responses, from Cytomegalovirus (CMV), Epstein-Barr Virus (EBV) and Influenza virus (Flu) (CEF), which cause chronic infection with multiple reactivations or repeated acute infections.

There is, however, a lack of standardisation in immunogenicity testing protocols, specifically regarding the methods used, and the evidence needed regarding type, quantity and quality of the observed immune responses. For example, to address these issues in clinical studies using flow cytometry-based immunophenotyping assays, the Human Immune Phenotyping Consortium (HIPC) promoted standardisation guidelines, so that data could be compared across sites and studies (4).

For all mentioned techniques, the number of available cells is a limiting factor. Moreover, the time, cost and workload needed to perform the assays can be restrictive (5) (Figure V.1). High- or

medium-throughput approaches allow for the reduction of the number of cells needed and overall experiment costs (6).

Although *in vitro* lymphocyte cultures do not allow assessment of cell migration and systemic interactions, they provide extremely useful information on antigen recognition and cellular activation. Still, the cell frequency in peripheral blood differs greatly between naive and memory T cell populations.

Naive T cells specific to any given antigen have been shown to constitute only 0.01 to 1% of total T cells – 1 in 10^5 up to 10^6 T cells (Figure V.2). Memory cells, on the other hand, constitute more than 1% of total T cells in peripheral blood (Figure V.2). These differences in cell frequency imply that detection of induced responses in both repertoires must be performed differently.

Recall responses in immunized or protected individuals can be readily detected using total PBMC since specific memory cells are present at high frequencies in circulation (Figure V.2). Usually, short-term cultures (5 to 10 days) are performed in triplicates of 2.10^5 total PBMC, for a total of 6.10^5 cells tested per condition.

On the other hand, naive T cells clones are rare, which constitutes a limiting issue for detection sensitivity. A method based on peptide:MHC tetramers uses four specific peptide:MHC complexes bound to a single molecule of fluorescently labelled streptavidin (13). It has proven extremely useful to identify populations of antigen-specific T cells in the naive repertoire, for example, describing the number of T cell precursors specific to ovalbumin and viral-specific epitopes (14). However, this technique can only evaluate one epitope per experiment, it is not applicable to large numbers of MHC alleles and requires previous knowledge about the peptide sequences. Also, it is very expensive and, therefore, unsuited for epitope mapping or peptide screening experiments (12).

A suitable method to prime naive T cells *in vitro* is to use antigen-pulsed autologous DC, and several rounds of stimulation inducing T cell activation, proliferation and effector phenotypes (7–11). Cellular amplification allows the analysis and measurement of antigen-specific responses mediated by a single precursor cell or population (12). Also, the use of purified cell populations, enriched in lymphocytes, helps increase the level of detection of such responses (2).

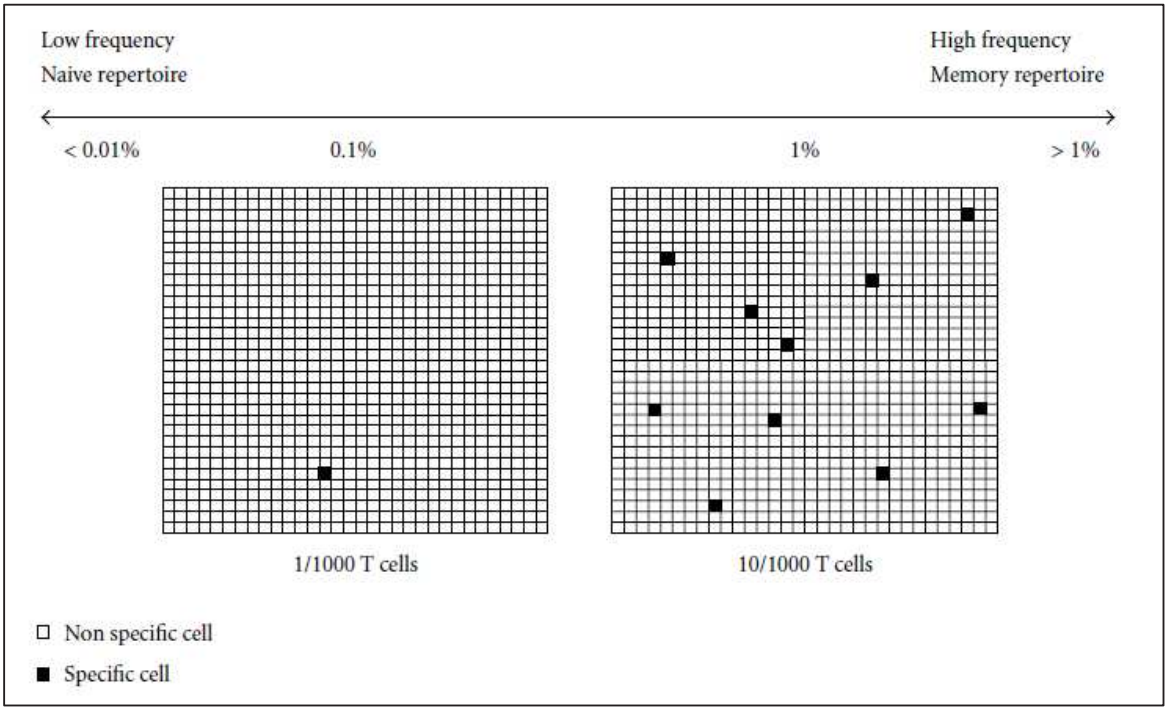


Figure V. 2 Specific cell frequency in different immune repertoires (5).

Precursor frequency experiments use limiting dilution conditions to determine the relative frequency of antigen-reactive cells in each population. Cell suspensions are distributed with the goal of isolating one antigen-specific cell in each well of a 96-well plate. Specific T cells are then amplified through re-stimulations with peptide-pulsed antigen-presenting cells. Due to the low frequency of specific cells in circulation, not all wells seeded will contain specific T cells that will proliferate in response to antigen recognition. Therefore, a minimum of 2 million cells should be tested to maximize the chances of detecting specific naive T cells. Wells or cell lines with one or more specific T cell will be considered positive, whereas negative wells never contained antigen-specific T cells. Using this assay, it is possible to estimate the specific cell frequency through applying the Poisson distribution formula to the number of total positive and negative wells. This approach also reduces inter-assay variability. Yet, if there is a high number of peptides to test, this approach can be difficult to perform as it is extremely laborious. Methods using batch cell stimulation can provide a viable alternative, so long as response specificity is confirmed.

2. Experimental strategy:

Immunoassays will be performed exclusively in human cells and using two immune backgrounds or status relative to *Leishmania* infection – naive and healed individuals (Figure V.3). Most healed individuals after *Leishmania* infection possess specific memory T cells, responsible for immunity against reinfection. Importantly, peptides that successfully induce T cell activation in both backgrounds are promising vaccine candidates (see Chapter I). These candidates are associated with protective recall responses and will be able to induce long-lasting memory responses in naive individuals, the objective of prophylactic vaccines.

The primary focus is the experimental validation of *Leishmania*-specific peptides, in order to validate *in silico* predictions and pre-select the most immunogenic peptides (Figure V.3). To this end, the *in vitro* immunoassays with naive human samples were optimised, healed donor samples were collected from endemic areas in Tunisia, and the 49 HLA-class I- and 24 HLA-class II-restricted synthetic peptides were screened in both immune backgrounds. The main experimental readout for the proposed immunoassay screenings will be IFN-γ production, assessed by IFN-γ ELISpot assays.

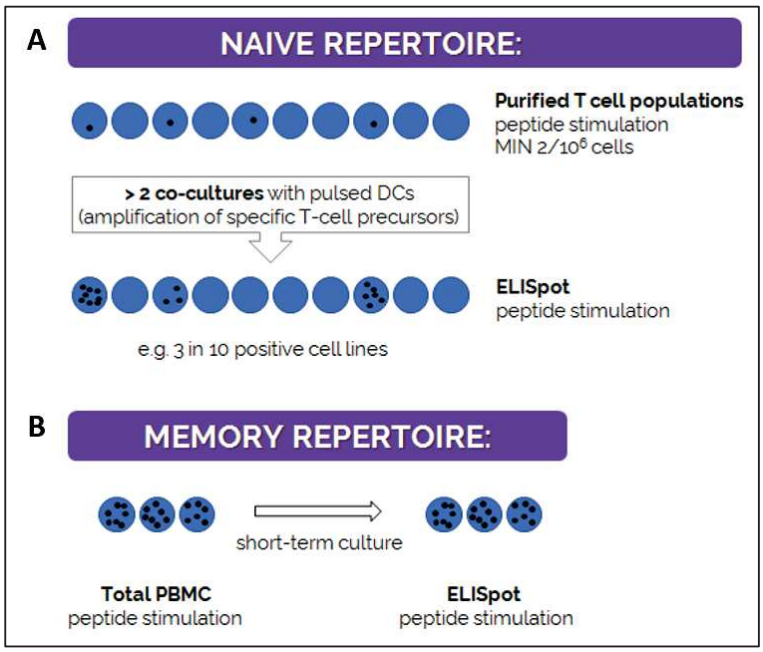


***Figure V. 3 Proposed* in vitro *T-cell assays to assess immunogenicity of synthetic peptides containing HLA-class-I and class-II* Leishmania*-specific epitopes**. *A) General view of* in vitro *assays with purified T cells from naive donors (naive repertoire). B) General view of the protocol used to assess immunogenicity in cells from healed individuals (memory repertoire).*

3. Methods

3.1. Ethics statement

The recruitment and sampling collection of different groups of volunteers were done in accordance with Good Clinical Practice (GCP).

In France, the *Etablissement Français du Sang* (EFS) is an established scientific partner of the *Institut de Recherche pour le Développement* (IRD), and performed the recruitment of blood donors and sample collection at the human blood bank of Toulouse.

In Tunisia, the recruitment of healed donors was based on the recommendations and approval of the local ethical committee from the *Institut Pasteur de Tunis* (IPT), the *Comité d'éthique de l'Institut Pasteur de Tunis* (*convention de collaboration* N°305256/00). A written informed consent was obtained from all subjects involved in this study.

3.2. Preparation of total soluble *Leishmania* promastigote antigens (TSLA)

All antigen extracts were prepared from promastigote stationary phase parasite cultures of *L. infantum*. Briefly, TSLA were obtained from washed parasites in 1x phosphate-buffered saline (PBS), centrifuged at 1000×g/10 min at 4°C and supernatants were removed. The pellets were resuspended in lysis buffer (50 mM Tris/5 mM EDTA/HCl, pH7. 1 mL/1×10^9 parasites), subjected to three rapid freeze/thaw cycles followed and to three sonication pulses of 20 seconds/40W. Samples were centrifuged at 5000×g for 20 min at 4°C, and supernatants were collected, aliquoted and stored at −80°C until use. Protein quantification was performed using Bradford method.

3.3. Synthetic Peptides

HLA-class I 9-mer peptides were synthesized by the Peptide Synthesis Platform from *the Institut de Biologie Paris-Seine* (*UMPC-Sorbonne Universités*, Paris, France), and HLA-class II 15-mer peptides by Proteogenix (Schiltigheim, France). The first peptide stocks were stored as 200 µM solutions in sterile MilliQ water, or Sodium Bicarbonate 0.2 mM (used in Tunisia and MN01-02), and the second peptide stocks were stored as 500 µM solutions in 10% DMSO and sterile MilliQ water (MN03-04).

The CEF-I control peptide pool is comprised of 23 well-defined peptides derived from Cytomegalovirus (CMV), Epstein Bar virus (EBV) and Influenza viruses and is widely used as a positive control for CD8 T cell activation, designed to stimulate T cells with a broad array of HLA types. CEF-I stock solution (Mabtec) was diluted 1:100 in culture medium for a final well concentration of 2 µg/mL.

3.4. Matrix-based peptide pools

A matrix-based pool testing strategy was chosen to validate and select the most immunogenic epitopes among the initial 48 class I and 24 class II peptides (Figure V.4). Briefly, peptides are attributed a random peptide number (pi1 to pi48 for HLA-class-I peptides, and pii1 to pii24 for HLA-class II peptides) and arranged in a matrix. Each peptide is included in two independent pools (1 vertical and 1 horizontal), and double-positive peptides are considered immunogenic.

A HLA-class I peptide pools:

Pools 1 to 7

Pools A to G	1	2	3	4	5	6	7
A	pi1	pi2	pi3	pi4	pi5	pi6	pi7
B	pi8	pi9	pi10	pi11	**pi12**	pi13	pi14
C	pi15	pi16	pi17	pi18	pi19	pi20	pi21
D	pi22	pi23	pi24	pi25	**pi26**	pi27	pi28
E	pi29	pi30	pi31	pi32	pi33	pi34	pi35
F	pi36	pi37	pi38	pi39	pi40	pi41	pi42
G	pi43	pi44	pi45	pi46	pi47	pi48	pi49

B HLA-class II peptide pools:

Pools 1 to 5

Pools A to E	1	2	3	4	5
A	pii1	pii2	pii3	pii4	pii5
B	pii6	pii7	pii8	pii9	pii10
C	pii11	pii12	pii13	pii14	pii15
D	pii16	pii17	**pii18**	pii19	**pii20**
E	pii21	pii22	pii23	pii24	

Figure V. 4 Matrix-based pool design. *A) 49 peptides restricted to HLA-class I alleles (pi) were randomly distributed in 14 pools, pools 1 to 7 and pools A to G. B) 24 peptides restricted to HLA-class II alleles (pii) were randomly distributed in 10 pools, pools 1 to 5 and pools A to E* (15). *Pools highlighted in yellow are example positive pools, leading to the identification of 2 immunogenic peptides.*

3.5. Immunoassays with samples from naive donors

a) Naive donor bank

Buffy coat samples from naive donors are collected in Toulouse by the EFS and shipped to Montpellier. Each naive donor was given an internal sample identification code – MPLn. PBMC were purified using Ficoll-Paque PLUS density gradients (GE Healthcare, Uppsala, Sweden). Briefly, buffy coat samples are diluted in PBS 1x 2mM EDTA, gently layered on top of the Ficoll gradient and centrifuged for 30minutes at 400g. After centrifugation, the interface ring layer is collected and washed in PBS 1x 2mM EDTA. Red blood cell lysis was performed with homemade ACK lysis buffer – 0.15 M ammonium chloride (NH_4Cl), 10 mM potassium bicarbonate ($KHCO_3$), 0.1 mM Na_2EDTA. Total PBMC are seeded in T175 flasks for 2 hours at 37°C for monocyte adherence. Non-adherent PBMC (PBMC NA) are collected and cryopreserved at -150°C for future use.

b) Generation of monocyte-derived dendritic cells

Monocyte-derived dendritic cells (MoDC) were generated from plastic-adherent PBMCs (from 500 million to 1 billion total PBMC), after 5-day culture in AIM-V medium (Invitrogen) supplemented with 1000 U/mL of interleukin-4 (rhIL-4, Miltenyi Biotec) and 1000 U/mL of granulocyte macrophage colony-stimulating factor (rhGM-CSF, Miltenyi Biotec).

To evaluate the *in vitro* generation of monocyte-derived immature DC, the differentiation status of MPL9 and MPL10's immature DC was assessed with the Mo-DC Differentiation Inspector human (Miltenyi Biotec, ref. 130-093-567), according to manufacturer's instructions. It comprises monoclonal antibodies recognizing CD14, CD83, and CD209 and corresponding isotype controls. After staining, samples were acquired by flow cytometry (FACSCanto, Becton Dickinson) and analysed with the BD FACSDiva™ software (version 6.2).

c) Cell counting

Total PBMC were counted manually using a KOVA™ Glasstic™ slide, after Trypan Blue 0,4% staining (cell suspensions diluted in PBS1x).

Purified T cells were counted using the LUNA-FL™ Automated Cell Counter (Logos Biosystems, Annandale, VA), after Trypan Blue 0,4% staining (cell suspension diluted in PBS1X, 10 µL are added to the LUNA slide, focus is manually adjusted, and cells/mL are counted in the range $5x10^4$ to $1x10^7$ cells/mL).

d) Immunoassays with samples from naive donors (MN01-MN04)

A total of four optimisation experiments with naive donor samples were performed, summarised in Table V.8.

Table V. 1 T cell amplification assays performed. *Four different protocols were tested and optimised: in experiment MN01, the Wolfl protocol described in* (16) *was used. In experiment MN02, the T cell amplification assay was adapted from the protocol for TCD4 cells described in* (8,9)*. In experiment MN03, this protocol was further optimised through the addition of a third co-culture and different peptide concentrations, tested in two different formats (96-well plate format and 48-well batch stimulation format). Finally, experiment MN04 used a 48-well format, batch stimulation assay, with the long peptide protocol for APC stimulation (immature dendritic cells, matured and pulsed for 16 hours).*

	MN01: **simplified Wolfl protocol**	MN02: **T cell amplification assay**	MN03: **T cell amplification assay**	MN04: **T cell amplification assay**
Cell population	Naive TCD8	Total TCD8	Total TCD8	Total TCD8 (48-well plate format)
Antigen-Presenting Cell (ratio APC:T cell)	Total PBMC (1:4)	mDC (1:10)	mDC (1:10)	mDC and iDC (1:10)
#co-cultures	3	2	3	3
Peptide pools	Matrix-based (14)	Matrix-based (14)	Subgroup (2)	poolGOOD (1)
[Peptide]	1 µM	5 µM	1 / 2,5 / 5 µM	2,5 µM
culture time	21-23 days	14-16 days	21-23 days	21-23 days
Readout	Elispot (1)	Elispot (1)	Elispot (1)	Elispot (1)

- **Naive T cell assay adapted from Wolfl et al (MN01):**

Naive T $CD8^+$ cells were isolated from total PBMC from donor MPL3, in a two step-procedure, firstly by depletion of non-naive T cells and NK cells (cocktail of biotin-conjugated human monoclonal antibodies against CD45RO, CD56, CD57, CD244), and, subsequently, by positive selection of $TCD8^+$ naive cells using an anti-CD8 monoclonal antibody coupled to magnetic microbeads (Miltenyi CD8 MicroBeads), as recommended by the manufacturer (Naive $CD8^+$ T cell Isolation Kit, ref. 130-093-244, Miltenyi Biotec). Cell purity was not assessed due to technical problems with the flow cytometer.

Naive $TCD8^+$ cells were plated in 96-well plate, 200.000 per well (200k/w), and 10 wells per condition (2 million TCD8 cells), in CellGro medium (CellGenix #0020801-0500) supplemented with Penicillin/Streptomycin 1X and 5% human serum type AB (SAB, Lonza lot#0000166597) (=**complete CellGro**). Autologous PBMC (50.000 per well) were seeded in 96-well plates in complete CellGro medium (for a APC:TCD8 1:4 ratio). PBMC were loaded with peptide pools (1 µM each

peptide), medium alone or CEF peptide control pool (2 µg/mL) and incubated for 4 hours at 37°C/5% CO_2. Pulsed PBMC were washed and irradiated 6 minutes under 50 Hz, 312 nm UV light (VL-6.M lamp, Fischer BioBlock Scientific). Naive T CD8+ cells were added to pulsed PBMC and cultured in complete CellGro medium supplemented with human recombinant (rh) IL-7 (500 U/mL, Miltenyi Biotec) and rhIL-15 (75 U/mL, Miltenyi Biotec). Medium was changed after 3 days. Naive T CD8+ cells were re-stimulated on day 7 and day 14, for a total of 3 co-cultures in the same conditions (d0, d7, d14).

The protocol recommended by Wolfl et al was adapted to use total PBMC as APC. The protocol compares several cell types as APC, including total PBMC, and suggests the use of monocyte-derived dendritic cells. Monocyte-derived DCs were not generated for donor MPL3 so the adapted protocol uses total PBMC, pulsed with peptides for 4 hours, without any differentiation or maturation factors.

IFN-γ production was assessed on day 21 by IFN-γ enzyme-linked immunospot assay (ELISpot). All cells from culture wells were seeded in the ELISpot plate and compared with unstimulated cells from the same individual. Naive TCD8+ cell lines were considered positive when a spot count was twofold higher in the presence of the peptides than in their absence, with a minimal difference of 44 spots, the highest NS spot count (STIM spot count > 2xNS spot count average+44 spots).

- **T cell amplification protocol_v1 (MN02):**

The T cell amplification protocol using total TCD8+ cells from the naive repertoire was adapted from the protocol for TCD4 cells developed by the Bernard Maillere Lab (*Laboratoire d'immunochimie de la réponse immune cellulaire*, *Institut de Biologie et technologies*, CEA – Saclay)(8,10,17). This protocol uses magnetically purified total TCD8 cells which are co-cultured with autologous peptide-pulsed mature DC.

Immune cell populations (monocytes, TCD8+ and TCD4+ cells) from MPL9 buffy coat sample were purified differently from other donors. This sample was used to test a semi-automatic cell separator and sequential cell isolation and assess the protocol feasibility for the generation of the naive donor bank. Monocytes, TCD8+ and TCD4+ cell populations were magnetically isolated with the use of a MultiMACS™ Cell24 Separator Plus (Miltenyi Biotec). Briefly, the buffy coat sample was diluted in PBS1x 2mM EDTA and the distinct cell populations were sequentially purified: i) CD14+ monocytes through positive selection with the StraightFrom®Buffy Coat CD14 MicroBead Kit, human (Miltenyi Biotec, ref 130-114-976); ii) total TCD8+ cells through positive selection with the StraightFrom® Buffy Coat CD8 MicroBead Kit, human (Miltenyi Biotec, ref. 130-114-978); and,

finally, iii) total TCD4+ cells through positive selection with the StraightFrom® Buffy Coat CD4 MicroBead Kit,human (Miltenyi Biotec, ref. 130-114-980).

Monocytes were cultured in T175 flasks for the generation of monocyte-derived DC (see 3.5.b) and, subsequently, immature DCs were cryopreserved. All T cell populations were cryopreserved and kept at -150°C until use. Cell purity was assessed by surface marker staining with the following antibodies: mouse anti-CD8-PE antibody (BD Biosciences); mouse anti-CD4-PeCy7 antibody (BD Biosciences); mouse anti-CD14-FITC antibody (BD Biosciences); and mouse anti-CD45-PerCPCy5 (BD Biosciences).

Three cryovials containing total TCD8+ cells from donor MPL9 were gently thawed (vials with 50 million, 25 million and 12.5 million TCD8+ cells). Total TCD8+ cells were seeded in 96-well plate, 200.000 cells per well in Iscove Modified Dulbecco medium (IMDM) supplemented with 10% human AB serum, 1x non-essential amino-acids (Invitrogen), 50 U/mL of penicillin and 50 µg/mL of streptomycin (Invitrogen) (**=IMDMc**), for a total of 10 cell lines (2 million TCD8+ cells) per condition. Autologous DCs (20000 per well) were seeded in 96-well plates and pulsed with the different stimuli using the **short peptide protocol** – immature DCs are incubated in AIMV medium (Invitrogen) and matured with lipopolyssacharide (LPS, 1 µg/mL) and Resiquimod (R848, 10 µg/mL) for 16 hours, and then pulsed with either medium alone, peptide pools or positive controls for 4 hours at 37°C/5% CO_2. After peptide stimulation, APC were washed and irradiated 6 minutes under 50 Hz, 312 nm UV light (VL-6.M lamp, Fischer BioBlock Scientific).

TCD8 cell lines were generated by adding total TCD8+ cells to the pre-pulsed DC wells and cultured in IMDMc supplemented with rhIL-21 (30 U/mL, Miltenyi Biotec) for cell priming. Medium was changed after 3 days to IMDMc supplemented with rhIL-7 (500 U/mL, Miltenyi Biotec) and rhIL-15 (75 U/mL, Miltenyi Biotec). Cells were stimulated on day 0 and 7, for a total of 2 co-cultures, plus the ELISpot stimulation.

The specificity of the TCD8+ cell lines was analysed with ELISpot assays: 20.000 T cells from each cell line were incubated with autologous iDCs alone (cNS control) or with iDCs previously loaded with peptide pools (1 culture well = 2 ELISpot wells with 20.000 cells each). TCD8 cell lines were considered specific when a spot count was twofold higher in the presence of the protein than in its absence, with a minimal difference of 25 spots (STIM spot count > 2xNS spot count average+25 spots).

- **T cell amplification protocol_v2 (MN03):**

Three cryovials containing total TCD8+ cells from donor MPL9 were gently thawed (vials with 50 milion, 25 million and 12,5 million cells each). Total TCD8 cells were counted and cultured in IMDMc medium in: i) 96-well plates, 200.000 TCD8 cells per well and 10 cell lines per condition, or ii) 48-well plates, 1 million TCD8 cells per well and 2 wells per condition. In both formats, TCD8 cells were co-cultured at 1:10 ratio with DC cells stimulated using the **short peptide protocol** with mature monocyte-derived DCs, and a total of 3 co-cultures were performed (d0, d7, d14) plus the ELISpot stimulation. Different peptide concentrations were tested – 1, 2,5 and 5 µM per peptide in each pool. After peptide stimulation, APC were washed and irradiated 6 minutes under 50 Hz, 312 nm UV light (VL-6.M lamp, Fischer BioBlock Scientific).

For the IFN-γ ELISpot, and for all peptide concentrations, total TCD8 cells were incubated with autologous DCs alone (control) or with DCs previously loaded with peptide pools at a ratio of 1:10 (APC:TCD8). For cell lines cultured in 96-well plates, 20000 TCD8 cells were seeded in the ELISpot plate. For cells cultured in 48-well plates, 50.000, 100.000 or 150.000 TCD8 cells were seeded in the ELISpot plate. Responses were considered specific when a spot count was twofold higher in the presence of the protein than in its absence, with a minimal difference of 25 spots (STIM spot count > 2xNS spot count average+25 spots).

- **T cell amplification protocol_v3 (MN04):**

Total TCD4 and TCD8 cells were isolated using magnetic microbeads according to the manufacturer's instructions. Briefly, cryopreserved non-adherent PBMCs (PBMC NA) from healthy donor MPL10 were gently thawed. Total $TCD4^+$ lymphocytes were isolated by positive selection using an anti-CD4 monoclonal antibody coupled to magnetic microbeads (Miltenyi Biotec, 'CD4 Microbeads kit' ref 130-045-101), as recommended by the manufacturer, and cryopreserved at -150°C for future use. Total $TCD8^+$ cells were isolated from the flow through by negative selection using magnetic microbeads (Miltenyi Biotec, '$CD8^+$ T cell isolation kit' ref 130-096-495) according to manufacturer instructions.

In this experiment, only the 48-well format (batch stimulation) was used, 1 million TCD8 cells were seeded per well and 2 wells per condition (2 million TCD8 tested). A total of 3 co-cultures were performed (d0, d7, d14) and monocyte-derived DCs stimulated with the i) **short peptide protocol** (see above), and ii) **long peptide protocol** (simultaneous maturation with LPS/R848 and peptide

stimulation during 16h at 37°C). After peptide stimulation, APC were washed and irradiated 6 minutes under 50 Hz, 312 nm UV light (VL-6.M lamp, Fischer BioBlock Scientific).

The selected donor was MPL10, and a tailored peptide pool was designed according to the HLA-typing results. The peptide pool consists of 3 peptides per supertype (best predicted IC_{50}), for a total of 12 peptides (**poolGOOD**). Autologous monocyte-derived DCs were stimulated with 2,5 µM per peptide.

TCD8 cells were considered specific when a spot count was twofold higher in the presence of the protein than in its absence, with a minimal difference of 25 spots (STIM spot count > 2xNS spot count average+25 spots).

3.6. Immunoassays with samples from Healed Donors

a) Healed donor samples from a *Leishmania*-endemic area in Tunisia

Human donor groups (cured individuals and healthy individuals with no history of leishmaniasis) were recruited from endemic areas for CL, based on the following defined inclusion criteria: i) individuals living in the Gadarif Region in Tunisia, endemic foci to *L. major* transmission, and who have not moved away in the last 10 years; and ii) the presence of typical scars for cured CL group. A complete medical questionnaire was completed during examination. Adult healthy individuals recruited in low or non-endemic areas (Tunis), and no or low IFN-γ response to SLA (<100 pg/ml) will be considered as non-immune/naive. Exclusion criteria were immunosuppressive diseases other than leishmaniasis, long term treatment and pregnancy. Each donor was assigned an internal sample identification code – TUNn.

Heparinized blood was collected from a total of 20 healed donors and 10 healthy controls (1st healed series n=10, 2nd healed series n=10, naive series n=10).

b) Total PBMC stimulation assays

Total PBMC were isolated from blood by density centrifugation through Ficoll-Hypaque (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Whole blood samples were transported at room temperature, in a cooler to keep temperature stability during the car trip from Gafsa Hospital to the Institut Pasteur of Tunis. Upon arrival in the lab, the cells were cultured in RPMI 1640 supplemented with 10% heat inactivated Human AB serum (SAB, Lonza), 100 IU/mL penicillin, 100 µg/mL streptomycin, 2 mM L-glutamin, 1 mM sodium pyruvate and 1X non-essential amino acids. Briefly,

cells were plated in 96-well plates (TPP, Switzerland), 200.000 PBMC/well, and were kept with media alone (NS) or stimulated with: i) Phytohemagglutinin 10 µg/mL (PHA, Sigma-Aldrich); ii) TSLA 10 µg/mL; iii) CaniLeish® Antigen 10 µg/mL (GMP-produced *Li*ESAp lyophilised without adjuvant) as positive controls; and iv) peptide pools (5 µM per peptide). Cells were cultured at 37°C, 5% CO_2 for 10 days. On day 1, 4 and 6, human recombinant IL-2 (100 U/mL, R&D systems) was added to the wells.

The peptide pool conditions tested were: i) one pool containing class I 9-mer peptides (pool ALL_I with 48 peptides); ii) fourteen matrix-based class I peptide pools 1 to 7 and pools A to G (containing 6 or 7 9-mer peptides each); iii) one pool containing class II 15-mer peptides (pool ALL_II with 24 peptides); iv) and ten matrix-based class II peptide pools 1 to 5 and pools A to E (containing 4 or 5 15-mer peptides each).

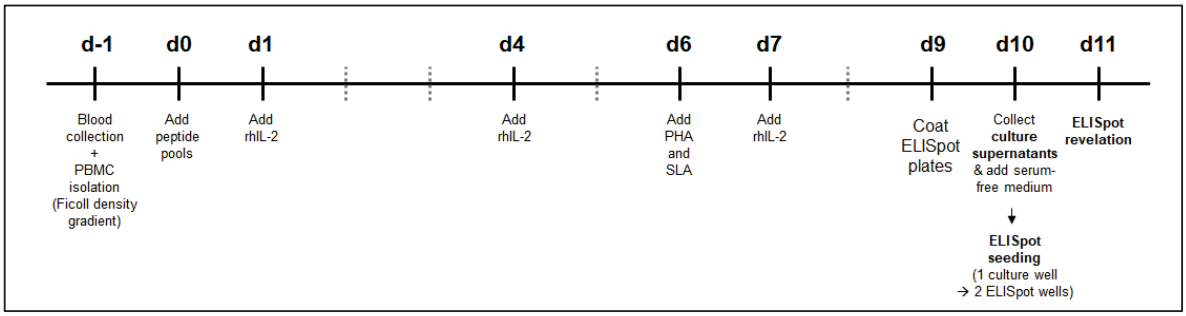


Figure V. 5 Experimental planning for immunoassays with samples from healed donors.

3.7. Interferon-γ (IFN-γ) enzyme-linked immunospot (ELISPOT)

Briefly, 96-well polyvinylidene difluoride (PVDF) plates (MSIP, Millipore) were coated overnight at 4°C with 2,5 ug/mL anti-human IFN-γ monoclonal antibody (mAb 1-D1K; Mabtech). Cells were cultured overnight (16-20h) with positive controls, peptide pools, or with culture medium as negative control, in triplicate. IFN-γ secretion was detected by addition of biotinylated anti-human IFN-γ mAb (7-B6-1; Mabtech) in PBS 1% bovine serum albumin (BSA) for 2 hours at 37°C. After washing, extravidin-phosphatase solution (dilution 1:3000 in PBS 1% BSA; Sigma-Aldrich) was added to the plate for 1 hour at room temperature, and spots were revealed with BCIP substate solution (tablets, Sigma-Aldrich) for 10 min (maximum revelation time of 20 min). IFN-γ spots were quantified using an AID Immunospot analyzer (C.T.L.).

3.8. Statistical Analysis

Statistical analysis was performed using GraphPad Prism version 6 software (GraphPad). For healed experiment series, the non-parametric Mann-Whitney test (one-tailed with confidence interval 95%) was applied to compare independent sample groups (STIM vs NS). A p-value ≤ 0.05 was considered statistically significant.

3.9. HLA-typing

All HLA-typing services were provided by DKMS Life Science Lab, an affiliated company with DKMS German Bone Marrow Donor Centre. High resolution HLA typing (99%) is performed by Next-Generation Sequencing (Long Range Sequencing Service). Genomic DNA was extracted using the QIAamp DNA Blood Mini Kit (Qiagen, ref 51104) and shipped to DKMS Lab in Germany. Sequencing was performed for exons 2 and 6 for 6 HLA loci (HLA-A, -B, -C, -DRB1, -DQB1, and -DPB1) leading to the 6-digit identification of twelve HLA alleles per donor.

In order to compare allele frequencies between the recruited donors and country-specific populations, allele frequency data was retrieved from the Allele Frequency Net Database (AFND). The AFND provides a central source, freely available to all, for the storage of allele frequencies from different polymorphic areas in the Human Genome. Currently, collected data are in allele, haplotype and genotype format. In the present analysis, only HLA studies were considered (18). Allele frequency information with two-digit level of resolution for the French Population was retrieved from the AFND (Table V.2). Allele frequency information with two-digit level of resolution for the Tunisian Population was retrieved from the AFND (Table V.3).

Allele frequency refers to the total number of copies of the allele in the population sample (Alleles/2n) in decimal format. Some studies present results as percentage of individuals that have the allele (% ind), and it refers to the percentage of individuals who have the allele in the population (heterozygous allele count / n). Since Tunisian studies use mostly allele frequency to present results, for consistency purposes, only studies with results expressed as allele frequency were included.

Table V. 2 Description and size of 17 French population studies from AFND with HLA data considered for HLA allele frequency analysis

Population (Studies)	Ethnicity	Study	Sample Size	HLA Loci
France Bordeaux	Mixed	HLA	990	-A, -B, -C, -DRB1
France Ceph	Caucasoid	HLA	124	-DPA1, -DPB1, -DQA1, -DQB1
France Corsica Island	Caucasoid	HLA	100	-A, -B, -C
France Grenoble	Mixed	HLA	1	-A, -B, -C, -DRB1
France Grenoble, Nantes and Rennes	Caucasoid	HLA	6,094	-A, -B, -C, -DRB1
France Lille	Caucasoid	HLA	95	-DPB1, -DQA1
France Lyon	Caucasoid	HLA	4,813	-A, -B, -C, -DRB1
France Marseille	Mixed	HLA	1	-A, -B, -C, -DRB1
France Reims	Caucasoid	HLA	102	-A, -B
France Rennes	Caucasoid	HLA	200	-DPB1, -DQB1, -DRB1
France Rennes pop 2	Caucasoid	HLA	148	-DPB1
France Rennes pop 3	Caucasoid	HLA	200	-A, -B, -DQB1, -DRB1
France Rennes pop 4	Mixed	HLA	1	-A, -B, -C, -DRB1
France South	Caucasoid	HLA	350	-DQA1, -DQB1, -DRB1
France Southeast	Caucasoid	HLA	130	-A, -B, -C, -DPB1, -DQB1, -DRB1
France West	Caucasoid	HLA	100	-C, -DQB1, -DRB1
France West Breton	Caucasoid	HLA	150	-DPB1, -DQA1, -DQB1, -DRB1

Table V. 3 Description and size of 8 Tunisian population studies from AFND with HLA data considered for HLA frequency analysis

Population (Studies)	Ethnicity	Study	Sample Size	HLA Loci
Tunisia	Arab	HLA	100	-A, -B, -C, -DPB1, -DQA1, -DQB1, -DRB1
Tunisia Gabes	Mixed	HLA	95	-A, -B, -DQB1, -DRB1
Tunisia Gabes Arab	Arab	HLA	96	-DQB1, -DRB1
Tunisia Ghannouch	Arab	HLA	82	-A, -B, -DQB1, -DRB1
Tunisia Jerba Berber	Berber	HLA	55	-DQB1, -DRB1
Tunisia Matmata Berber	Berber	HLA	81	-DQB1, -DRB1
Tunisia pop 2	Arab	HLA	111	-DQB1, -DRB1
Tunisia pop 3	Arab	HLA	104	-A, -B, -DQB1, -DRB1

For the analysis of HLA allele frequency in different world populations affected by leishmaniasis, data was collected from the AFND for the following world regions: South and Central America, North Africa, Western Asia, South Asia (Appendix V.1). Data were retrieved from AFND through a 'HLA allele freq (Classical)' search with the following filters: i) region name; ii) level of resolution=2-digits; iii) Population standard=Gold and Silver (19); iv) show frequencies=only positives (Table V.4).

Table V. 4 Populations with HLA allele frequency data from four world regions most affected by leishmaniases. *Data from AFND, world regions plus filters (level of resolution: 2 digits + population standard: Gold and Silver only + only positive results). "records" are all combinations of 2-digit allele/population per world region.*

World Region	countries included	FULL results (records)	records WITH FILTERS	allele list (filtered results rows)
Western Asia	Armenia, Gaza, Georgia, Iran, Jordan, Iraq, Israel, Lebanon, Oman, Turkey, Saudi Arabia, UAE	8976	66	94
South and Central America	Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Guatemala, Ecuador, Jamaica	13700	411	96
North Africa	Algeria, Ethiopia, Morocco, Sudan, Tunisia	6210	53	83
South Asia	India, Bangladesh, Pakistan, Sri Lanka	4707	325	85

4. Results - Immunoscreenings with samples from naive donors

4.1. The naive donor bank and respective HLA-typing results

A total of 21 buffy coat samples from naive individuals were processed (MPL1-21) (Table V.9). HLA-typing was performed for donors MPL3-21 included in the naive donor bank (Table V.10). Individual HLA-class I alleles sharing peptide binding properties were grouped in allele groups (supertypes) according to the classification by Sidney J et al 2008 – supertypes A01, A02, A03, A24, A26, B07, B08, B27, B44, B58, B62 (cf chapter IV). Whenever specific alleles possess B and F pocket specificities that are shared by different allele groups they can be classified, for example, as A01A24 or remain unclassified.

Table V. 5 Naive donor bank description. *Buffy coat samples from 21 individuals from the French blood bank EFS (Établissement Français du Sang), in Toulouse, France. EFS code = collection site code/internal sample ID. Total PBMC were recovered using Ficoll density gradient.*

Donor ID	EFS code	Gender	Blood type	Volume (mL)	Total cells recovered
MPL1	B3111/72170830625	M	0-	66	$3{,}25x10^8$ total PBMC
MPL2	B3111/72170938740	F	0+	74	$5{,}59x10^8$ total PBMC
MPL3	B3111/72171040510	F	0-	74	$6{,}21x10^8$ total PBMC
MPL4	B3111/7217182530-	M	0-	75	$5{,}44x10^8$ total PBMC
MPL5	B3111/72171817414	F	A+	73	$5{,}62x10^8$ total PBMC
MPL6	B3111/72171817385	M	A+	65	$5{,}96x10^8$ total PBMC
MPL7	B3111/7217184683-	M	A+	70	$8{,}32x10^8$ total PBMC
MPL8	B3111/72171846995	M	0+	67	$1{,}11x10^9$ total PBMC
MPL9	B3111/72171970849	M	0+	69	$2.3x10^8$ monocytes/$1.5x10^8$ TCD8/$5.2x10^8$ TCD4
MPL10	B3111/72172236662	M	0+	72	$1{,}611x10^9$ total PBMC
MPL11	B3111/72171576825	M	0+	72	$6{,}682x10^8$ total PBMC
MPL12	B3111/72172340953	M	0+	88	$9{,}824x10^8$ total PBMC
MPL13	B3111/7217234097-	M	0+	71	$1{,}026x10^9$ total PBMC
MPL14	B3111/72172272292	M	A+	77	$1{,}517x10^9$ total PBMC
MPL15	B3111/72172234851	F	0+	73	$5{,}571x10^8$ total PBMC
MPL16	B3111/72172234894	M	A+	72	$6{,}653x10^8$ total PBMC
MPL17	B3111/72172166289	F	A+	66	$7{,}188x10^8$ total PBMC
MPL18	B3111/72172203158	M	0-	71	$9{,}861x10^8$ total PBMC
MPL19	B3111/72172171731	M	0+	71	$1{,}035x10^9$ total PBMC
MPL20	B3111/72172611397	M	A+	68	$1{,}104x10^9$ total PBMC
MPL21	B3111/72172611901	M	0+	70	$9{,}756x10^8$ total PBMC

Table V. 6 HLA-typing results for donors MPL3 to MP21. *A) Six loci corresponding to HLA-class I and class II polymorphic regions (HLA-A, -B, -C, -DRB1, -DPB1, -DQB1) were sequenced by NGS (exons 2 and 3), leading to the 6-digit identification of the alleles expressed by each donor. Allele supertype classification by Sidney et al 2008* (20)*. B) Ambiguous identifications are given a multiple allele code (MAC, https://hml.nmdp.org/MacUI/).*

A

Donor ID	HLA-A1	HLA-A1 supertype	HLA-A2	HLA-A2 supertype	HLA-B1	HLA-B1 supertype	HLA-B2	HLA-B2 supertype	HLA-C1	HLA-C2	HLA-DRB1	HLA-DRB2	HLA-DQB1	HLA-DQB2	HLA-DPB1	HLA-DPB2
MPL3	23:01:01G	A24	23:01:01G	A24	44:03:01G	B44	49:01:01G	unclassified	04:01:01G	07:01:01G	11:04:01	12:JV	03:BGCWT	03:BGCWT	04:01:00	04:01:00
MPL04	01:01:01G	A01	03:01:01G	A03	35:03:01G	B07	08:01:01G	B08	07:01:01G	04:01:01G	03:01:01	01:01:01	05:01	02:01:01	03:FYKD	02:AHZAM
MPL05	33:03:01G	A03	03:01:01G	A03	35:03:01G	B07	07:02:01G	B07	04:01:01G	07:02:01G	03:01:01	15:01:01	06:02:01	02:01:01	03:FYKD	04:ADCGE
MPL06	01:01:01G	A01	11:01:01G	A03	40:01:01G	B44	08:01:01G	B08	07:01:01G	03:04:01G	03:01:01	04:01:01G	02:01:01	03:01:01	01:AETTG	04:ADCGE
MPL07	31:01:02G	A03	03:01:01G	A03	07:02:01G	B07	07:02:01G	B07	07:02:01G	07:02:01G	16:01:01	07:DFRJ	05:02:01	02:GKDU	02:01	04:01
MPL08	11:01:01G	A03	68:02:01G	A02	35:01:01G	B07	14:02:01G	B27	08:02:01G	04:01:01G	13:03:01	01:03:01	05:01	03:01:01	02:01	02:01
MPL09	29:02:01G	A01 A24	68:02:01G	A02	18:01:01G	B44	44:03:01G	B44	05:01:01G	04:01:01G	11:02:01	07:DFRJ	06:09:01	02:GKDU	10:01:01G	03:FNVX
MPL10	03:01:01G	A03	02:05:01G	A02	15:03:01G	B27	07:02:01G	B07	12:03:01G	07:02:01G	15:01:01	07:DFRJ	06:02:01	03:03:02	15:01:01	04:01
MPL11	01:01:01G	A01	02:01:01G	A02	58:01:01G	B58	44:03:01G	B44	07:01:01G	04:01:01G	07:DFRJ	07:DFRJ	02:GKDU	02:GKDU	04:01	17:01:01
MPL12	03:01:01G	A03	26:01:01G	A26	51:01:01G	B07	13:02:01G	unclassified	02:02:02G	06:02:01G	09:01:02G	07:DFRJ	03:03:02	02:GKDU	17:01:01	04:01
MPL13	03:01:01G	A03	32:01:01G	A01	07:02:01G	B07	44:02:01G	B44	05:01:01G	07:02:01G	15:01:01	15:01:01	06:02:01	06:02:01	14:01:01	05:RGPW
MPL14	02:01:01G	A02	26:01:01G	A26	44:03:01G	B44	44:03:01G	B44	04:01:01G	04:01:01G	14:54:01	07:DFRJ	05:AWXGU	02:GKDU	06:01:01	05:RGPW
MPL15	01:01:01G	A01	03:01:01G	A03	35:03:01G	B07	08:01:01G	B08	07:01:01G	12:03:01G	04:04:01	03:01:01	03:02:01	02:01:01	04:01	02:01
MPL16	23:01:01G	A24	26:01:01G	A26	07:02:01G	B07	50:01:01G	B44	06:02:01G	07:02:01G	04:05:01	15:01:01	06:02:01	03:02:01	15:01:01	04:01
MPL17	11:01:01G	A03	32:01:01G	A01	07:02:01G	B07	40:01:01G	B44	03:04:01G	07:02:01G	04:01:01G	13:01:01	03:03:02	03:02:01	14:01:01	06:01:01
MPL18	30:02:01G	A01	02:01:01G	A02	07:02:01G	B07	18:01:01G	B44	05:01:01G	07:02:01G	03:01:01	07:DFRJ	02:01:01	02:GKDU	11:AWXGR	02:02:01
MPL19	24:02:01G	A24	03:01:01G	A03	35:01:01G	B07	14:02:01G	B27	15:05:01G	01:02:01G	01:01:01	01:02:01G	05:01	05:01	17:JECX	04:ADCGE
MPL20	01:01:01G	A01	11:01:01G	A03	35:01:01G	B07	14:02:01G	B27	08:02:01G	04:01:01G	13:01:01	04:01:01G	06:03:01	03:02:01	04:01	04:01
MPL21	24:02:01G	A24	66:01:01G	A03	18:01:01G	B44	41:02:01G	B44	05:01:01G	17:01:01G	13:03:01	03:01:01	02:01:01	03:01:01	04:HJMR	04:FNVS

B

Multiple Allele Code (MAC)	Decode result (shows the allele list for a MAC at a locus and allele-family)
DRB1*07:DFRJ	DRB1*07:01/DRB1*07:79
DRB2*07:DFRJ	DRB2*07:01/DRB2*07:79
DQB1*02:GKDU	DQB1*02:02/DQB1*02:97
DQB2*02:GKDU	DQB2*02:02/DQB2*02:97
DPB2*04:FNVS	DPB2*04:02/DPB2*105:01
DPB1*11:AWXGR	DPB1*11:01/DPB1*654:01
DPB1*17:JECX	DPB1*17:01/DPB1*131:01
DPB1*01:AGXDR	DPB1*01:01/DPB1*417:01/DPB1*462:01
DPB2*04:HJMR	DPB2*04:01/DPB2*126:01
DPB2*04:ADCGE	DPB2*04:01/DPB2*350:01
DPB2*04:AHPJH	DPB2*04:02/DPB2*463:01
DPB2*17:JECX	DPB2*17:01/DPB2*131:01

HLA typing results for the 19 individuals describe the diploid loci for HLA-A, -B, -C, -DRB1, -DQB1 and -DPB1 genes corresponding to 12 alleles per donor. The only HLA class I genes analysed were the HLA-A and -B genes (4 gene loci per donor), and homozygous alleles were excluded from allele counts. The only HLA class II genes analysed were the HLA-DRB1 and -DPB1 genes (4 loci) and homozygous alleles were excluded from allele counts.

HLA-typing results for the 19 naive donors included in the naive donor bank show the most frequent HLA-A allele is HLA-A*03, with 42,1% of individuals expressing an A*03 allele (8 allele counts) (Figure V.6). Almost 80% of individuals express an allele from the A03 supertype (15 in 19 donors), which also includes the alleles HLA-A*11:01:01G, A*31:01:02G, A*33:03:01G, A*66:01:01G (Table V.10).

The most frequent HLA-B allele expressed by the naive donors is HLA-B*07, with 36,8% of individuals expressing HLA-B*07 alleles (7 allele counts) (Figure V.6). Almost 75% of individuals express an allele from the B07 supertype (14 in 19 donors), which also includes the B*35 alleles, namely HLA-B*35:01:01G (Table V.10). The individuals expressing alleles that belong to the B07 supertype include 36,8% and 31,6% of individuals who express HLA-B*07 and -B*35 alleles, respectively. Even if these alleles share peptide binding specificities, the high frequency of B*35 alleles in the human populations could justify a separate allele group to minimise bias in epitope selection.

Another common HLA class I allele expressed by naive donors is the HLA-B*44, with 26,3% individuals expressing HLA-B*44 alleles (5 allele counts) and 63,2% of individuals expressing an allele from the B44 supertype (12 in 19 donors) (Figure V.6). The other main alleles belong to supertypes A01 and A02, with 8 and 6 individuals expressing alleles from these supertypes, respectively (Figure V.6). Three of the donors are homozygous for one of the HLA class I loci: donor MPL3 (HLA-A*23:01:01G, A24), donor MPL07 (HLA-B*07:02:01G, B07), and donor MPL14 (HLA-B*44:03:01G, B44).

The HLA-class I supertype B62, with B*15:01 as the representative allele, is not represented in the donor bank, and the supertypes B58 (B*58:01) and A01A24 are only represented once. The only B*15 allele expressed is B*15:03:01G (donor MPL10) which is classified in the B27 supertype.

The supertype A01A24, according to Sidney J et al 2008, includes alleles with specificities that are compatible with multiple supertypes (i.e HLA-A*29:02) – B pocket specificity for small, aliphatic and aromatic residues, and F pocket specificity for aromatic and large hydrophobic residues. In this case, these features are compatible with both the A1 supertype (B pocket specificity for small and aliphatic residues, and F pocket specificity for aromatic and large hydrophobic residues), and the A24

supertype (B pocket specificity for aromatic and aliphatic residues, and a F pocket specificity for aromatic, aliphatic and hydrophobic residues).

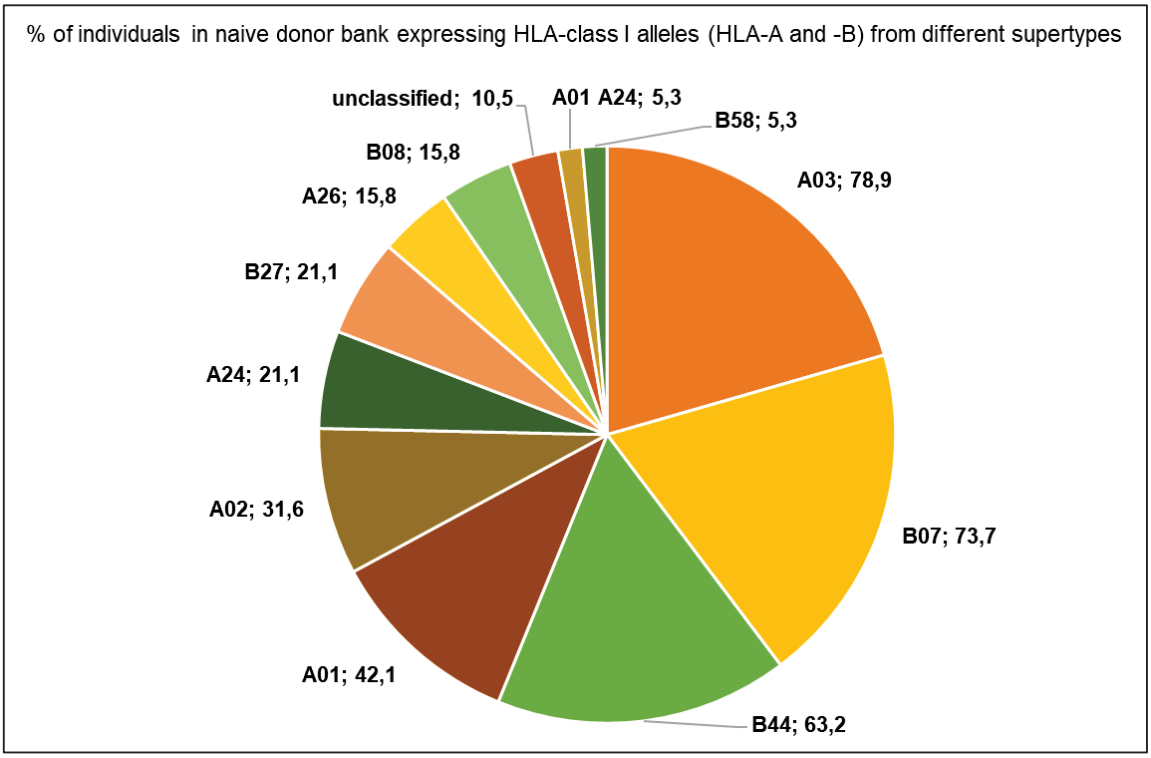


Figure V. 6 Percentage of individuals in the naive donor bank expressing HLA class I alleles according to allele supertypes*. Results for HLA-class I alleles (HLA-A and -B genes) grouped in supertypes (allele groups with shared binding properties).*

4.2. HLA allele frequency of the naive donor bank matches the French population

Studies from the AFND describing allele frequency for HLA-A and -B alleles in the French population show that the most prevalent HLA-A class I alleles are, by descending order, HLA-A*02, HLA-B*44, HLA-A*01, HLA-A*03, HLA-B*35, HLA-A*24, HLA-B*07, and HLA-B*08 (Figure V.7).

The overall allele frequencies in the French population are similar to the results obtained for the naive donors included in the bank, as the most prevalent HLA-A genes expressed are also HLA-A*02, HLA-A*01, HLA-A*03, HLA-A*24, and HLA-B genes expressed are HLA-B*44, HLA-B*35 and HLA-B*07 (Figure V.7).

Naive donors express mostly HLA-A*03 alleles (0,2105 frequency / 42,1% of individuals), HLA-B*07 alleles (0,1842 frequency / 36,8% of donors), HLA-B*35 alleles (0,1579 frequency / 31,6%

of individuals), HLA-B*44 alleles (0,1316 frequency / 26,3% of individuals), HLA-A*01 alleles (0,1316 frequency / 26,3% of individuals), and HLA-B*08 alleles (0,079 frequency / 15,8% of individuals) (Figure V.7).

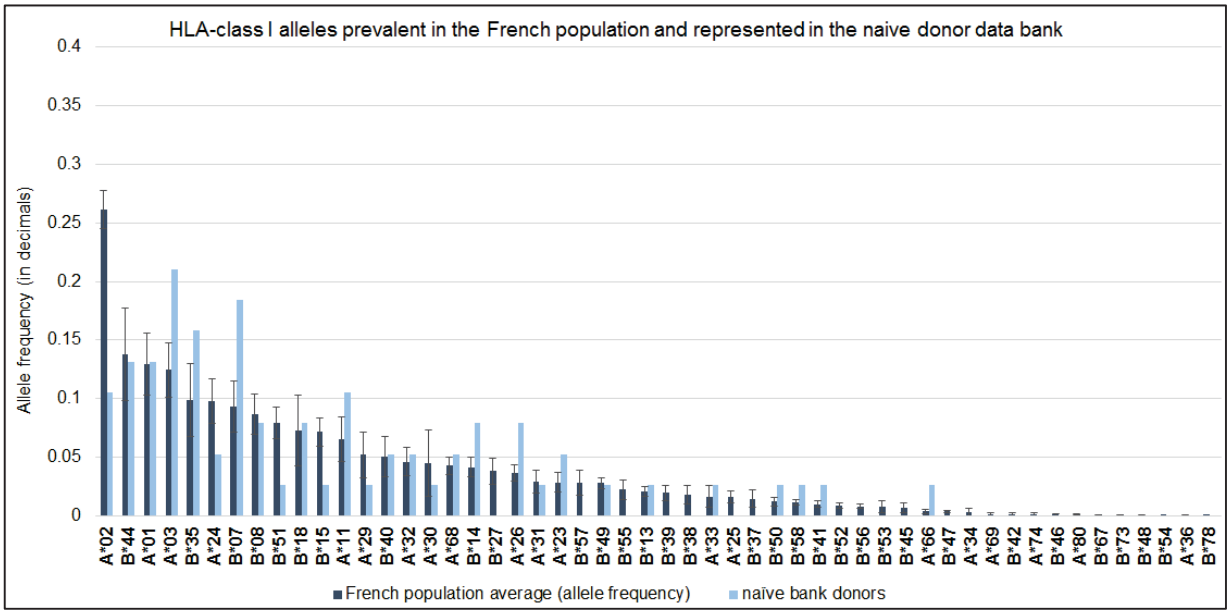


Figure V. 7 HLA class I allele frequency in the French population and in the naive donor bank. *Allele frequency for HLA-A and -B genes (allele / 2n).*

Similar results were obtained for HLA-class II alleles, wherein the most prevalent alleles in the French population correspond to the most common alleles expressed by the naive donors. The main HLA class II alleles, for HLA-DRB1 and -DPB1 genes, are HLA-DPB1*04, HLA-DPB1*02, HLA-DRB1*04, HLA-DRB1*07, HLA-DRB1*13, HLA-DRB1*15, HLA-DRB1*11, HLA-DRB1*01 and HLA-DRB1*03 (Figure V.8).

The most expressed HLA-class II allele is HLA-DPB1*04 with 68,4% of all donors expressing this allele (allele count 13, frequency 0,342). The most frequent HLA-DRB1 allele, and the second most frequent DRB allele overall, is HLA-DRB1*07 expressed by 36,8% of individuals (allele count 7, frequency 0,184).

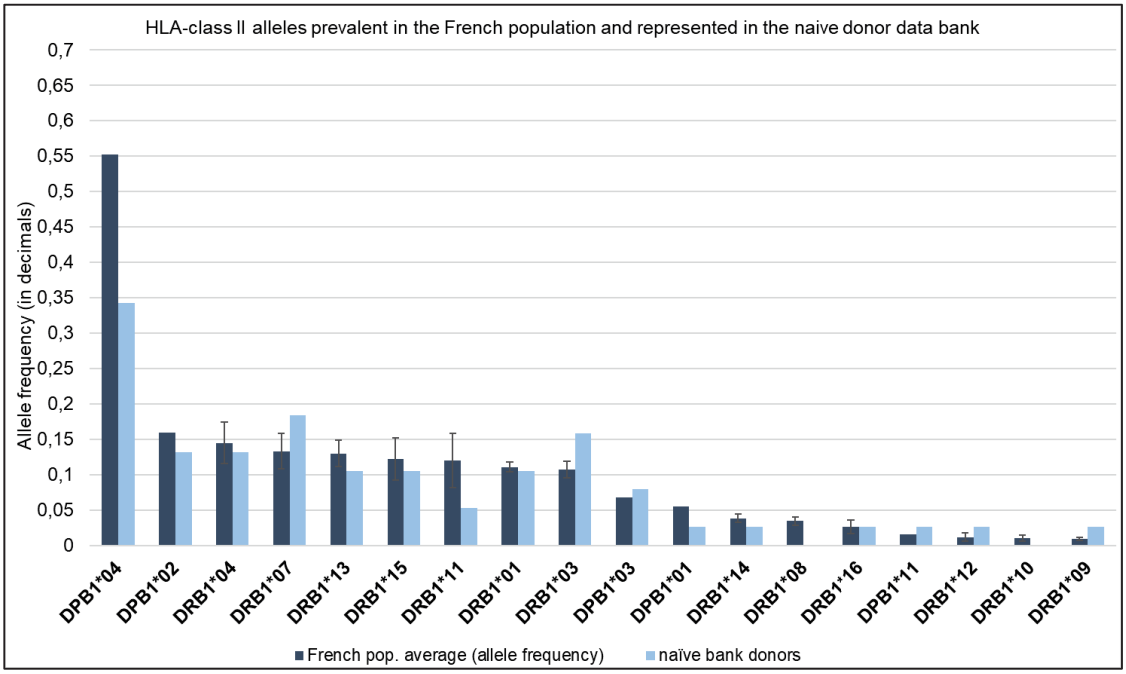


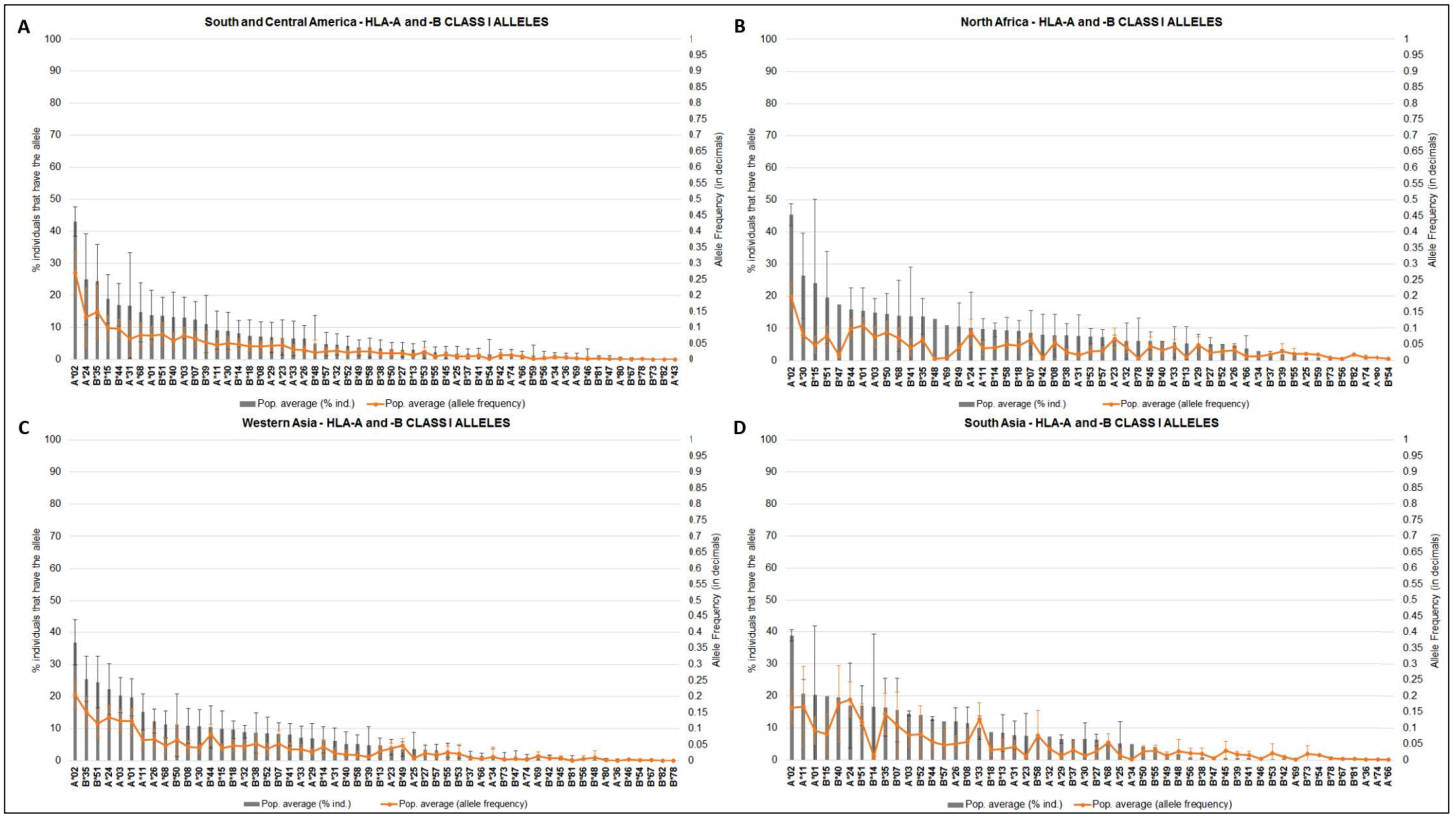
Figure V. 8 HLA class II allele frequency in the French population and in the naive donor bank. *Allele frequency for HLA-DRB1 and -DPB1 alleles (allele / 2n).*

4.3. HLA allele frequency of the naive donor bank includes the most prevalent alleles in the populations affected by leishmaniasis

The regions included in the AFND that match the populations most at risk for *Leishmania* infection are South and Central America, North Africa, Western Asia, South Asia. All positive HLA frequency studies (with Gold and Silver population standards) were included, which present results either as percentage of individuals expressing a given allele and/or HLA frequency in decimals.

The most common HLA-class I alleles across all four world regions, for HLA-A genes are HLA-A*02, -A*24, -A*68, -A*01 and -A*03. As for HLA-B genes the most common alleles are HLA-B*35, -B*44, -B*51, -B*07. All these class I alleles are consistently expressed by over 10% of the population in these regions (Figure V.9).

Figure V. 9 Average HLA-class I allele frequencies according to world region. *Averaged results from the population studies considered (%individuals and allele frequency) for HLA-class II alleles HLA-A and HLA-B. A) South and Central America populations. B) North Africa population. C) Western Asia populations. D) South Asia populations.*



The most frequent HLA-class II alleles across all four world regions, expressed by over 10% of individuals in all these populations are HLA-DRB1*03, -DRB1*04, -DRB1*07, -DRB1*08, -DRB1*11, -DRB1*13, -DRB1*15. The allele HLA-DRB1*01 is less prevalent is the western and south asian populations. Some alleles are often not included in the allele frequency studies, as for example the HLA-DPB1 locus, which impedes comparisons among populations.

Typing of the HLA-DQB1 loci was performed for the individuals in the naive donor bank (Table V.10), and data exists for some of the selected populations, however, these were not included in the analysis since epitope prediction was not performed to HLA-DQ alleles.

HLA allele frequencies for these four world regions match the most prevalent alleles in the naive donor bank, suggesting the allele frequencies of naive donors are representative of these human populations.

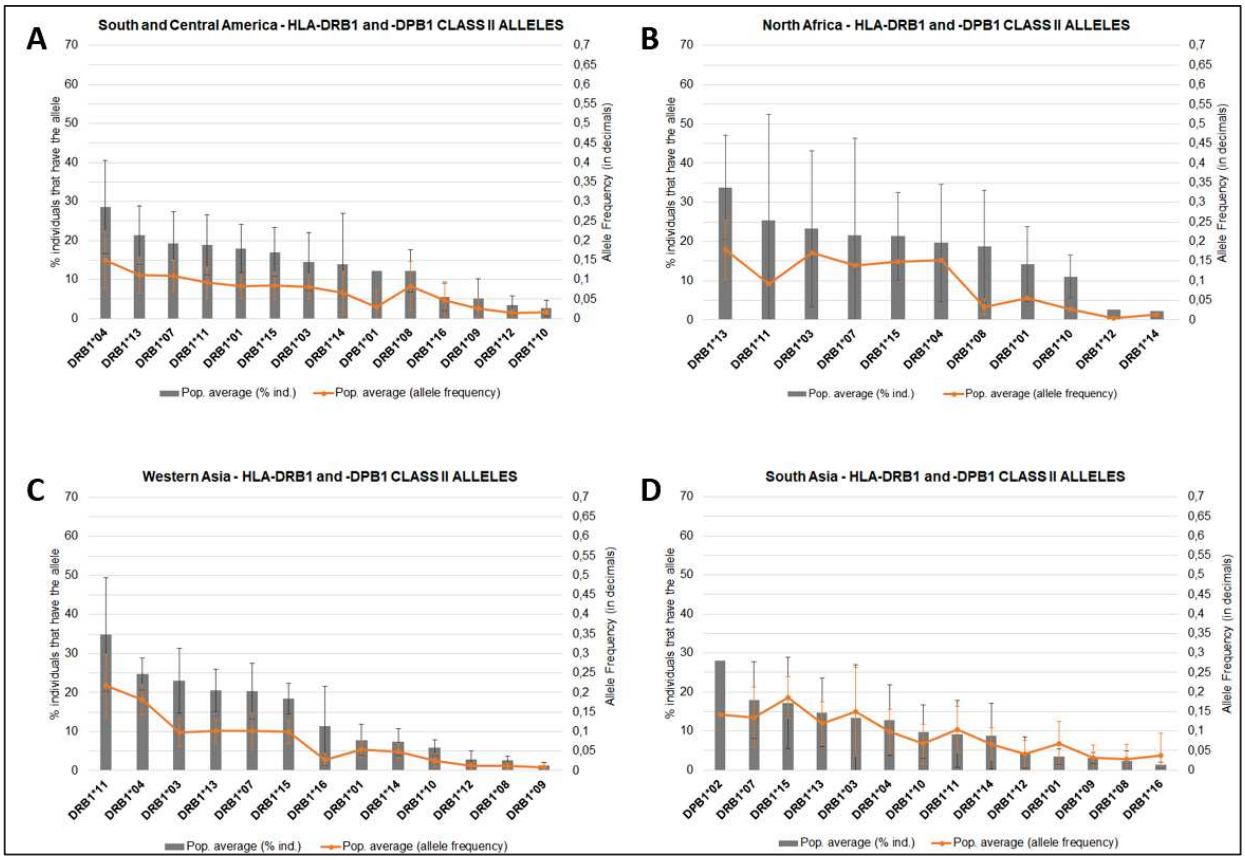


Figure V. 10 Average HLA-class II allele frequencies according to world region. *Averaged results from the population studies considered (%individuals and allele frequency) for HLA-class II alleles HLA-DRB1 and HLA-DPB1. A) South and Central America populations. B) North African populations. C) Western Asian populations. D) South Asia populations.*

4.4. Monocyte-derived dendritic cells are efficiently generated *in vitro*

The surface markers CD14, CD209 (DC-SIGN) and CD83 were used to determine the phenotype of monocyte-derived dendritic cells, after differentiation from monocytes and, for MPL10, after maturation. CD14 is highly expressed by monocytes, and functions as a co-receptor for LPS and LPS-binding protein (LBP). $CD14^+$ monocytes can differentiate into DCs in a IL-4/GM-CSF cytokine environment. CD209 (DC-SIGN) is a C-type lectin and enables TCR engagement by stabilisation of the DC-T-cell contact zone, and expression is limited to tissue DCs and monocyte-derived DCs. CD83 is regarded as a maturation marker for DCs, it regulates antigen expression presentation and is only expressed by DCs, Langerhans cells and B cells. Hence, immature DCs are $CD14^-CD209^+CD83^-$ and mature DCs are $CD14^- CD209^+CD83^+$.

In vitro monocyte differentiation into immature DC induced by IL-4 and GM-CSF is highly efficient, resulting in the recovery of 99,1% for MPL9 and 98,1% for MPL10 $CD14^-$ cells, excluding debris and contaminant cells (Figure V.11).

For MPL9, immature DCs as defined by the expression of CD209 and the absence of CD83 are 42,8% of total $CD14^-$ cells. Some mature DC expressing CD83 are found (2,3%) (Figure V.11 panel A).

For donor MPL10, the immature and mature DC populations were compared. Results show that after *in vitro* monocyte differentiation, immature DCs represent 17,9% of total $CD14^-$ cells, with 2,3% also expressing CD83 (Figure V.11 panel B). For mature DCs, after overnight maturation with LPS and Resiquimod, the percentage of mature DC ($CD14^-CD209^+CD83^+$) increases to 28,6%, while 8,3% of cells remain in an immature state ($CD14^-CD209^+CD83^+$) (Figure V.11 panel C).

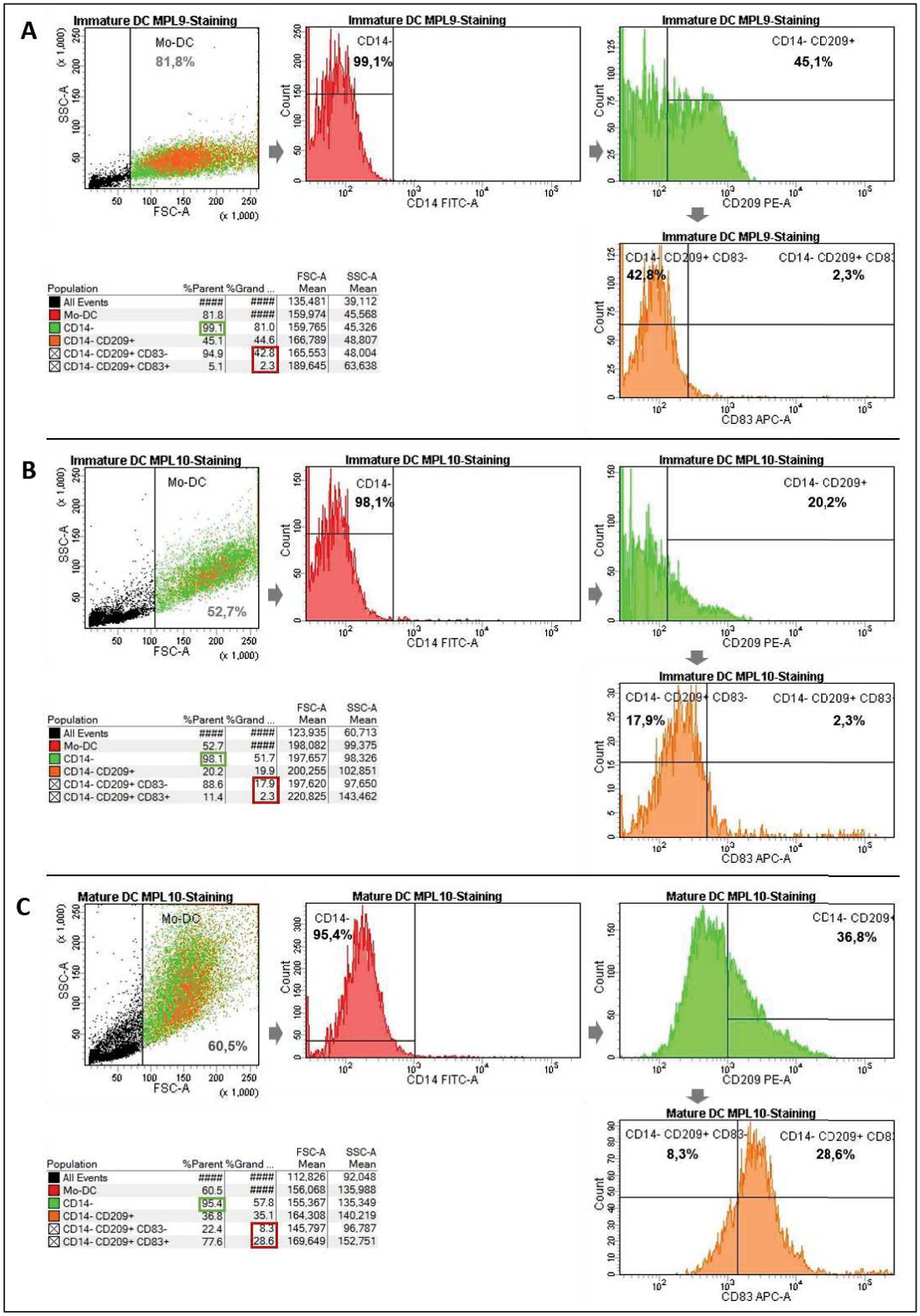


Figure V. 11 Monocyte-derived DC phenotype analysis for donors MPL9 and MPL10. *DC were stained in independent experiments with Mo-DC Inspector kit, human (Miltenyi) that detects CD14, CD209 (DC-SIGN), and CD83 expression. A) immature DC from donor MPL9. B) immature DC from donor MPL10. C) mature DC from donor MPL10.*

4.5. *Leishmania*-specific TCD8^{+} cells are present in the naive repertoire

The first peptide immunoscreening performed tested all 49 HLA-class I peptides in matrix-based pools using the protocol adapted from Wolfl et al, with magnetically isolated naive TCD8+ cells from donor MPL3 (experiment MN01).

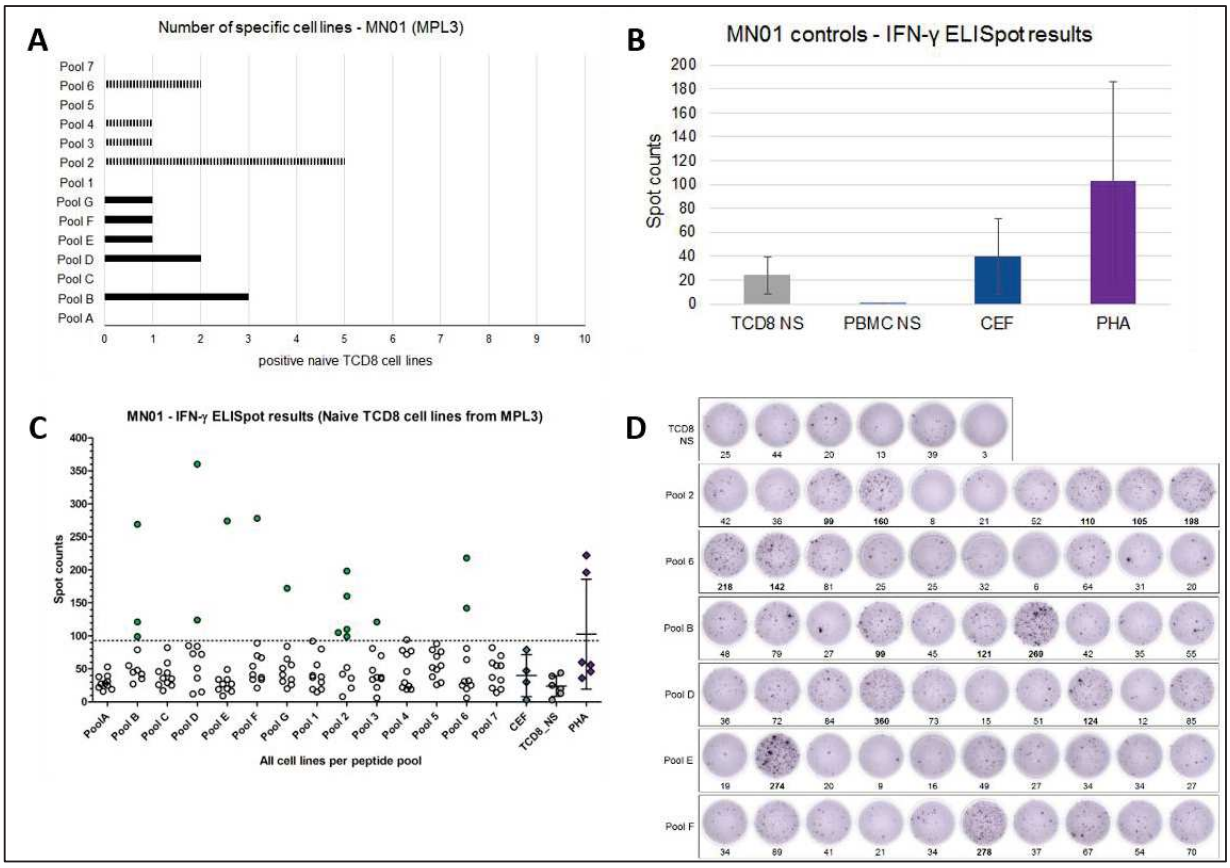


Figure V. 12 IFN-γ ELISpot results for the MN01 experiment: T cell amplification assay with naive TCD8 cells from naive donor MPL3*. A) number of positive cell lines per condition (14 class I peptide pools). B) control results (TCD8 NS, PBMC NS, CEF and PHA). C) Spot counts per cell line and the threshold mark for well positivity. D) ELISpot well pictures and respective spot counts.*

In experiment MN01, we successfully generated and detected specific T cell lines against *Leishmania*-specific peptide pools (Figure V.12 panel A). A total of 17 positive cell lines (out of 140) specific to 9 different peptide pools were generated from donor MPL3. Results show that 5 (out of 14) peptide pools generated 1 positive cell line, and 4 pools generated between 2 and 5 specific cell lines.

ELISpot results for negative controls (never stimulated TCD8 cells) showed overall low background levels (NS spot counts: 25, 20, 44, 13, 39, 3) for an average of 24 spots (±15,5). The

highest NS value (44) was included in the empirical test to determine well positivity (2*NS average+44 spots). The PHA solution used in this experiment is thought to have been from on old stock, probably degraded, as only 2 wells showed high spot counts, and still much lower than expected (PHA spot counts plate A: 222, 196, 60, 56; plate B: 46, 36) (Figure V.12 panel B).

In this experiment, the CEF-stimulated naive TCD8 cells underwent three co-cultures plus the ELISpot stimulation. The spot counts of the CEF peptide pool are quite variable with some low values (47, 79, 30, 3), which is expected since this assay uses a purified population of naive T cells, and very few memory cells are present (Figure V.12 panel B). Memory T cells are responsible for long-term responses against the CEF peptide pool (CMV-, EBV-, influenza-specific epitopes), so low spot counts can be expected when using purified naive T cell populations. $CD45RA^+$ cell enrichment after magnetic purification was not assessed due to technical problems with the flow cytometer, even though staining was performed. Although there are specific naive T cells in naive individuals against the CEF peptides which can be stimulated an amplified *in vitro*, only 600.000 naive TCD8+ cells (triplicate wells) were tested. Still, it is noteworthy that some *Leishmania*-specific peptide pools induced much higher number of IFN-γ-producing cells than the CEF control pool.

A

Positive wells >1 (spot count > 2*NS mean+44)	
Pool 2	**5**
Pool 6	**2**
Pool B	**3**
Pool D	**2**

4 positive pools:

Peptide pools	1	2	3	4	5	6	7
A	B58_2	A24_2	B62_3	A2_6	A26_3	A1_2	A26_2
B	B27_1	**B7_3**	B8_2	A2_1	B58_4	**A24_3**	A24_4
C	B62_4	B27_2	A3_4	B7_1	A1_3	B58_1	B62_1
D	A1_1	**B8_3**	A3_3	A24_5	B44_5	**B58_3**	B27_4
E	A24_1	B62_2	A3_1	A26_1	B27_5	A2_2	B44_1
F	A2_4	B8_4	B7_2	A2_3	B7_5	B27_3	B44_4
G	B58_5	A3_2	B8_1	A2_5	A1_4	B7_4	

	double positive peptides
pi9	B7_3
pi13	A24_3
pi23	B8_3
pi27	B58_3

B

Donor ID	HLA-A1	HLA-A1 supertype	HLA-A2	HLA-A2 supertype	HLA-B1	HLA-B1 supertype	HLA-B2	HLA-B2 supertype
MPL03	23:01:01G	A24	23:01:01G	A24	44:03:01G	B44	49:01:01G	Unclassified (B44)

***Figure V. 13 Four peptide pools successfully generated over 2 positive cell lines**. A) Cross-matching results with matrix peptide pool constitution leads to the identification of 4 double-positive, hence immunogenic peptides. B) Naive donor MPL3's HLA-typing results for 6 HLA loci (HLA-A, -B, -C, -DRB, -DP, -DQ)*

Peptide pool 2 induced 5 positive cell lines out of 10, the highest number overall, ranging from 99 to 198 spots. Pool B had the second highest number of positive lines, 3 in 10, with spots counts ranging from 99 to 269. Pool D and Pool 6 induced 2 positive cell lines each, with spot counts of,

respectively, 124/360, and 142/218. Pools E and F induced only 1 cell line but with very high spot counts of 274 and 278, respectively (Figure V.12 panels C and D).

If we consider a minimum of 2 positive cell lines, 4 immunogenic peptide pools are identified – pools B and D, and pools 2 and 6 (Figure V.13 panel A). Cross-matching these results with the pool matrix, four double-positive immunogenic peptides are identified – B7_3, A24_3, B8_3, B58_3 (Figure V.12 panel A).

The donor MPL3 expresses the alleles HLA-A*23:01:01G (A24), HLA-A*23:01:01G (A24), HLA-B*44:03:01G (B44), HLA-B*49:01:01G (unclassified but probable B44) (Figure V.12 panel B). Considering this information, only one peptide (A24_3) corresponds to the donor's HLA typing, implying peptides are more promiscuous than predicted by *in silico* algorithms.

If we consider all peptide pools with at least 1 positive well, there are 9 positive matrix-based peptide pools leading to the identification of 20 double-positive peptides. However, from these 20 double positive peptides, only 2 correspond to the donor's HLA-typing. The number of restricted peptides is not significantly increased as these do not include more donor-specific supertypes. Interestingly, peptide pool E which generated one cell line with high spot count (274) contains both A24- and B44-restricted peptides.

This assay constitutes preliminary evidence of the presence of T-cell precursors specific to *Leishmania* epitopes in naive individuals' immune repertoire. Although the magnitude of response was assessed for *in vitro* stimulated cell lines, specificity was not absolutely confirmed because spot counts were compared to NS wells (never stimulated autologous $TCD8^+$ cells). An ELISpot well containing culture-stimulated cells but unstimulated during the ELISpot (cNS) will demonstrate that IFN-γ production is dependent on peptide-stimulus and, therefore, specific.

The remaining issues with this experiment are i) that no cell counts were performed before seeding in ELISpot plate (in this experiment 1 culture well corresponds to 1 ELISpot well) which may introduce bias when comparing wells; ii) the use of an empirical rule for well comparison, because there are no triplicates no statistical test can be applied; iii) *Li*ESAp or CaniLeish® antigen, the original antigenic mixture, were not included as positive controls; iv) finally, the high number of purified naive TCD8 cells needed per experiment and per individual (MINIMUM 200.000*10 wells/condition) is logistically unviable.

In this experiment total PBMC were used as APC, without supplementation with differentiation or maturation factors. We expect the antigen presentation protocol can be further

optimised using professional APC, such as monocyte-derived DCs for short peptide presentation in HLA class-I molecules.

4.6. Total TCD8$^+$ cell *in vitro* amplification needs at least three stimulation rounds

In experiment MN02, total TCD8 cells from donor MPL9 were used, and the same matrix-based peptide pools were tested at 5 µM concentration per peptide. Cell populations from MPL9 were isolated differently from other donors using the MultiMACS™ Cell24 Separator Plus (see Material and Methods). Surface marker phenotyping shows that CD8$^+$CD45$^+$ cells were present at 88,2% purity (66,4% of total events excluding debris) (data not shown). Also, CD14$^+$CD45$^+$ monocytes were isolated with 93,9% purity (89,4% of total event excluding debris) and CD4$^+$CD45$^+$ cells were isolated with 96% purity (84,7% of total events excluding debris) (data not shown).

The protocol used was adapted from analogous TCD4$^+$ cell amplification assays (8,21), and uses the total TCD8$^+$ cell population and mature monocyte-derived DCs (mDC) as APC. Since the class I peptides are only 9-mer in length, and do not necessarily need processing, a short peptide stimulation protocol was adopted, where monocyte-derived immatured DCs are firstly matured with LPS and TLR-7/8 agonist Resiquimod/R848, and only after stimulated with peptides for 4 hours.

In this experiment, no specific cell lines were generated against any of the pools at 5 µM peptide concentration (Figure V.14 panel A) after two co-cultures.

Negative control background levels from never stimulated TCD8$^+$ cells remained low (NS spot counts 15/6/20/20/17, average 16). As expected, since in this experiment total TCD8 cells were used, the CEF control pool results were higher than MN01, even with just one co-culture plus the ELISpot stimulation (CEF spot counts 138/49/153/51, CEF MN02 Elispot2: 34, 238 (Too Numerous To Count), 39, 186 (TNTC), with only 20.000 TCD8+ cells seeded in the ELISpot wells.

PHA stimulation results (spot counts 366, TNTC, 404, TNTC) and average cell viability of 77% (± 9.9%) assessed before ELISpot seeding, confirm cells were viable and responsive (Figure V.13 panel B). We estimate viability is higher than calculated by LUNA reader, for there is some cellular debris from DC cells and SAB used for the co-cultures, considered as dead cells by the automatic counter (a cell size filter was not applied).

The positive controls *Li*ESAp and CaniLeish® antigen were tested in this experiment, however, the observed negative results remain inconclusive since the DC stimulation protocol was not adapted for protein processing. The short peptide stimulation protocol was used for all conditions and we

estimate a 4-hour incubation time of matured DCs is not sufficient to process and present the full protein antigens, and subsequent T-cell activation is abrogated (Figure V.14 panel B).

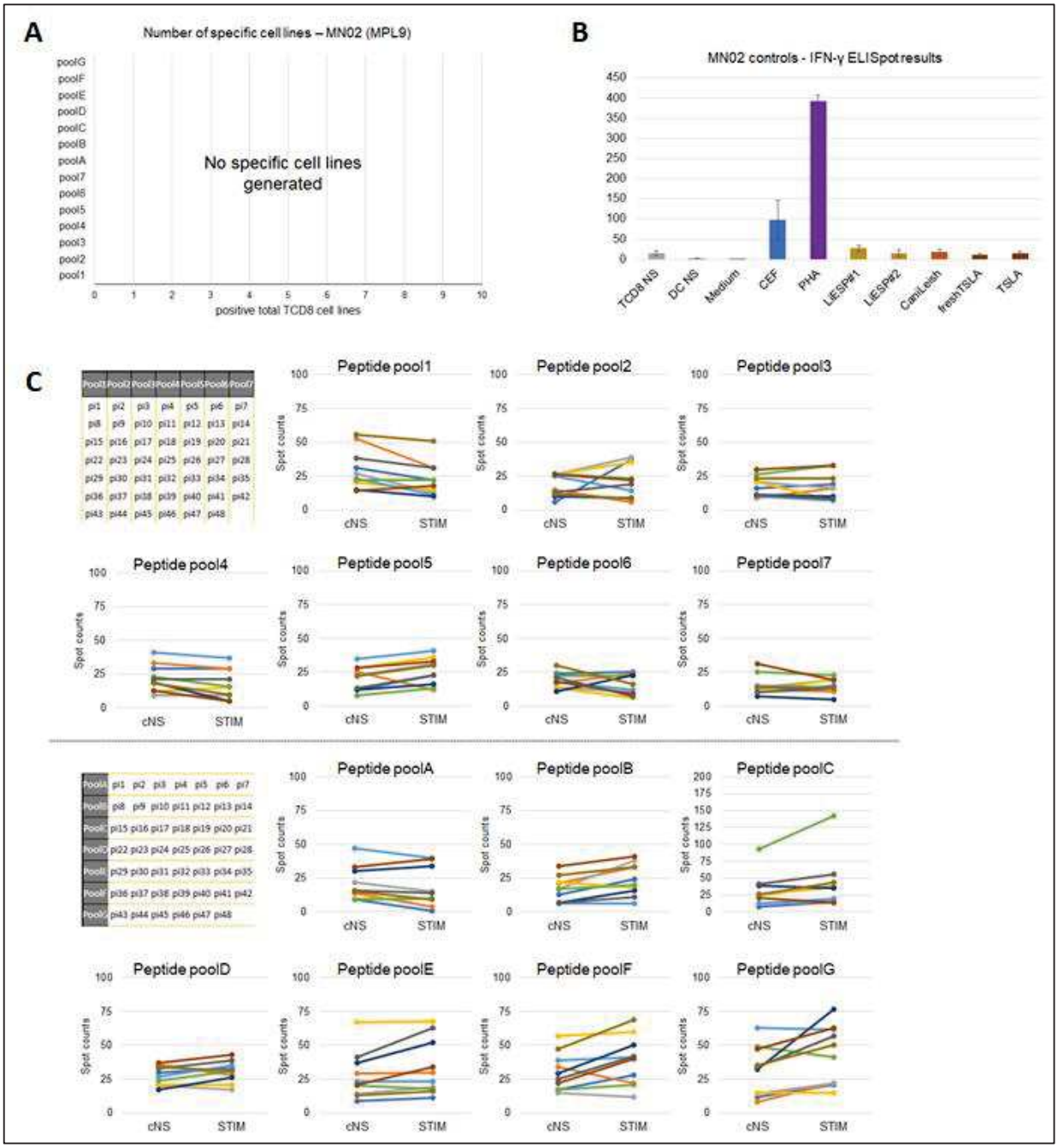


Figure V. 14 IFN-γ ELISpot results for the MN02 experiment (96-well format), T cell amplification assay with total TCD8+ cells from donor MPL9. *A) Number of specific T cell lines generated per condition. B) Spot counts for control conditions. C) Spot counts per condition showing individual cell lines; cNS, culture-stimulated cells but unstimulated in the ELISpot plate.*

In this experiment, response specificity per cell line was confirmed with culture-stimulated cells but unstimulated during the ELISpot (cNS). Overall, observed spot count values in stimulated wells (STIM) are sometimes higher than cNS, but not enough to consider as positive responses (Figure V.14 panel C). Otherwise, spot counts are similar to respective cNS unstimulated wells. Also, STIM and cNS are mostly similar to background levels for negative controls (never stimulated TCD8 cells – NS) but much more variable. All conditions and all cell lines show some variability in background levels (cNS spot counts), and even some cell lines present high background levels (e.g. pools 1, C, E, F and G) (Figure V.14 panel C). For peptide pool C, one cell line has cNS spot count of 93 and STIM spot count of 142, which if we compared with the never stimulated TCD8+ negative control would be considered a positive cell line (Figure V.13 panel C). This can be due to approximation errors when averaging cell counts to calculate 20.000 seeded cells in the ELISpot (3 in 10 wells measured per condition). Still, pools that successfully generated more than one cell line in MN01 – pools B, D, 2 and 6 – did not induce IFN-γ production/positive cell lines (maximum spot counts were 41, 43, 39 and 25, respectively).

The use of statistical analysis to assess positive responses is more reliable and recommended as they are based on a theoretical background, have a universal application across experiments, and even more suitable to detect weaker responses. Furthermore, the primary objective of the peptide immunoscreenings is to exclude non-immunogenic and unstable peptides, before inquiring about precursor frequency. Therefore, the following adaptations were made to the protocol: i) an additional total TCD8:DC co-culture (3 instead of 2); ii) testing different peptide concentrations and fresh aliquots diluted in 10% DMSO; and iii) testing a different assay format that allow the application of statistical tests to determine well positivity, the counting of all culture wells and show results as SFC/million, a widely accepted way to present ELISpot results.

4.7. *In vitro* amplification of total TCD8$^+$ cells depends on the antigen-presenting cells and respective stimulation protocol

In experiment MN03, total TCD8 cells from donor MPL9 were used, and all peptides were tested in two pools at different peptide concentrations. At the time of this experiment, HLA-typing results of MPL9 donor were not yet available to design one personalised peptide pool, so all class I peptides were tested in two pools, containing 22 peptides with predicted restriction to HLA-A genes (pool22), and 27 peptides with predicted restriction to HLA-B genes (pool27). Furthermore, two assay formats were tested: 96-well plate format, establishment of cell lines and empirical rule to determine positive responses, and the 48-well plate format (batch stimulation), that allows for

statistical determination of well positivity and cell counting before seeding in the ELISpot plates, while still testing the same number of cells.

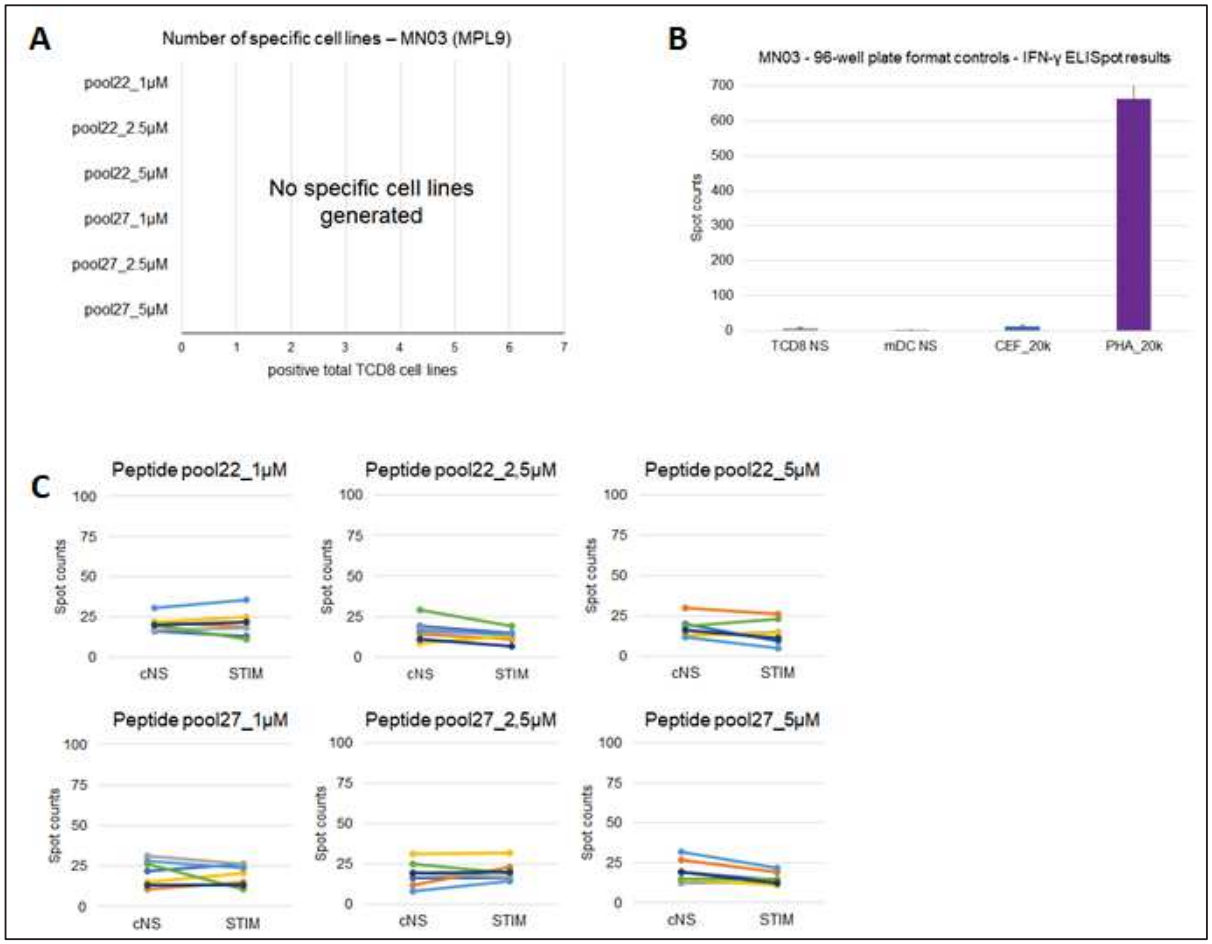


Figure V. 15 IFN-γ ELISpot results for the MN03 experiment (96-well format), T cell amplification assay with total TCD8 cells from donor MPL9*. A) Number of specific cell lines and IFN-γ ELISpot spot counts for 96-well plate format control conditions; B. all results per cell line, 7 seven lines per condition (1, 2.5 and 5µM peptide concentrations).*

The addition of a third re-stimulation or different peptide concentrations did not improve results obtained with the 96-well format T cell amplification assay, and no specific cell lines were generated (Figure V.15 panel A). With this format, positive control CEF stimulation also yielded very low spot counts (16/8/11), when cells are responsive, since PHA-stimulation yielded high IFN-γ production (PHA_20k spot counts 784, 693, 715 in plate A, and 595, 589, 608 in plate B).

Background levels of culture-stimulated cells but unstimulated in the ELISpot wells (cNS) of all cell lines are less variable than in the previous experiment MN02, but spot counts of stimulated cells are very similar to background (Figure V.15 panel C).

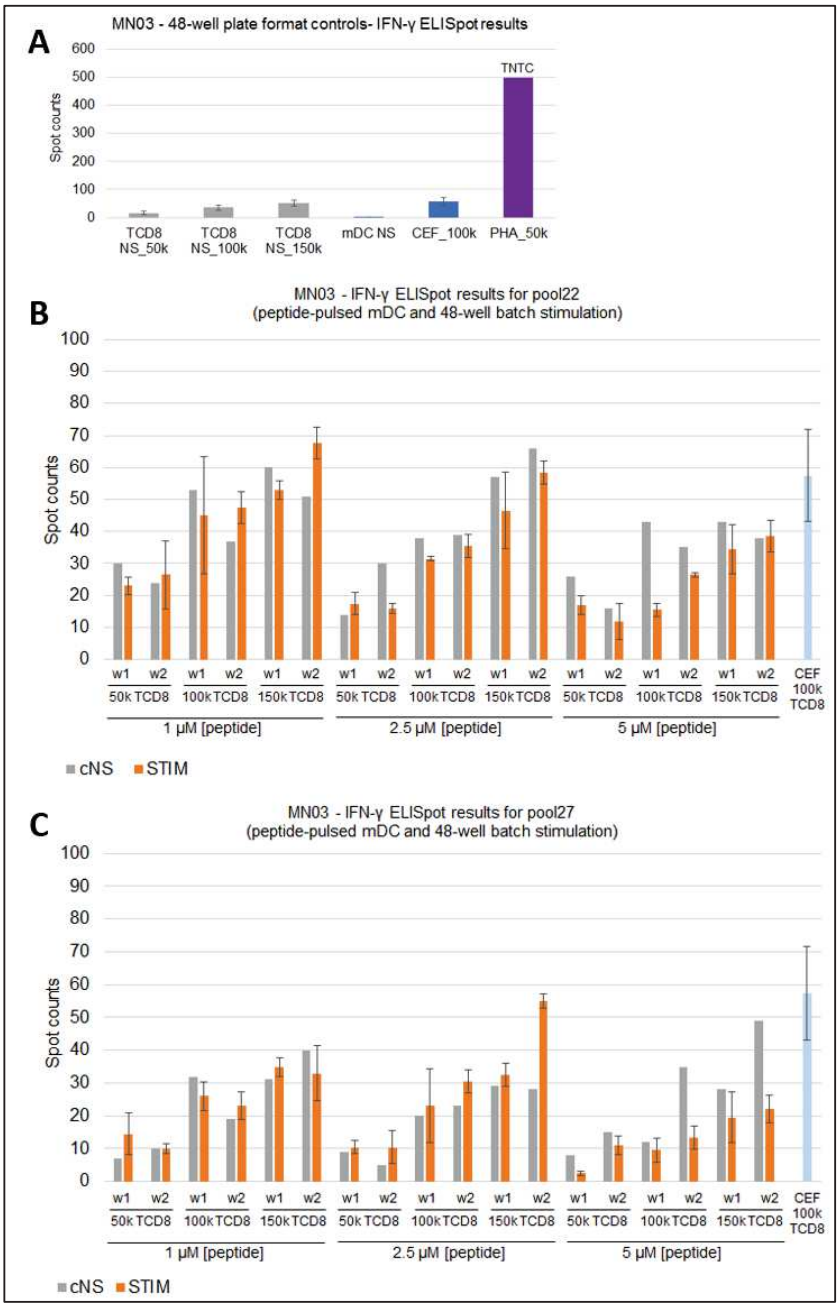


Figure V. 16 IFN-γ ELISpot results for the MN03 experiment (48-well format), T cell amplification assay with total TCD8 cells from donor MPL9*. A) IFN-γ ELISpot results for control conditions. B) Pool22 IFN-γ ELISpot results for all conditions tested: 1, 2.5 and 5µM peptide concentrations and 50.000 (50k), 100.00 (100k), or 150.000 (150k) TCD8 cells per ELISpot well. C) Pool27 IFN-γ ELISpot results for all conditions tested: 1, 2.5 and 5µM peptide concentrations and 50.000 (50k), 100.00 (100k), or 150.000 (150k) TCD8 cells per ELISpot well.*

Likewise, in the 48-well batch stimulation format, no differences were found in IFN-γ productions were observed between unstimulated (cNS) and stimulated cells (STIM). Nevertheless, better results were obtained with the CEF peptide control pool (CEF_100k spot counts 72/72/62 in plate A, and 36/47/56 in plate B, large spots), and high IFN-γ responses PHA-stimulated wells (PHA_50k spot counts: TNTC) (Figure V.16 panel C).

The reference number of cells seeded in the ELISpot plate is 10% of the culture well, in accordance with the 96-well plate format assay (200.000 TCD8 per culture well and 20.000 TCD8 seeded in the ELISpot well). To optimise response detection in the 48-well batch stimulation assay, we expect to use 100.000 TCD8 cells per ELISpot well (10% of 1 million TCD8 cells per culture well), but different cell numbers were tested – 50.000 TCD8 (50k), 100.000 TCD8 (100k), 150.000 TCD8 (150k) per ELISpot well. Response specificity for all cell densities was assessed by including culture-stimulated wells unstimulated in the ELISpot wells (cNS).

For some conditions, in the 48-well assay format, stimulated conditions had fewer spot counts than unstimulated cNS wells (Figure V.16 panels B and C). Also, there appears to be a dose-dependent effect between peptide concentration and decreased spot counts in stimulated wells, very noticeable at 5 µM per peptide (Figure V.15 panels B and C). This effect is probably not due to peptide toxicity since cell viability was assessed, but probably to peptide binding competition, as each pool contains a large number of peptides (22 and 27 peptides).

From this experiment, we realized that stimulation of monocyte-derived matured DCs is not optimal and should be further optimised. Even if class I peptides are 9-mer in length, apparently these were not successfully presented by MHC molecules and were unable to prime specific T cells *in vitro* under these conditions. Furthermore, the cells used in the MN01 experiment were highly immature and were successful for T cell activation and cell line generation. Therefore, two different protocols for monocyte-derived DC stimulation were adopted: i) **short peptide protocol,** immature DC are matured (with LPS and Resiquimod/R848) for 16 hours, and then pulsed with peptides for 4 hours at 37°C/5% CO_2; and ii) **long peptide protocol**, immature DC are matured (with LPS and Resiquimod/R848), and simultaneously pulsed with peptides for 16 hours at 37°C/5% CO_2.

4.8. IFN-γ production by *Leishmania*-specific total TCD8^{+} cells was successfully induced and detected with adequate DC stimulation

In experiment MN04, total TCD8 cells from donor MPL10 were magnetically isolated and co-cultured with mDC (short peptide protocol) or iDC (long peptide protocol).

Since we week to validate the proposed protocol for DC stimulation, only 12 peptides were used in one peptide pool, customised according to the naive donor's HLA alleles, instead of testing all 49 HLA-class I peptides. Naive donor MPL10 expresses the HLA-class I alleles HLA-A*03:01:01G, HLA-A*02:05:01G, HLA-B*15:03:01G, HLA-B*07:02:01G, belonging to supertypes A3, A2, B27 and B7, respectively (Figure V.17). Accordingly, the peptide pool tested (poolGOOD) includes 3 peptides per supertype, each at a 2,5 µM concentration – peptides A3_1, A3_2, A3_3, A2_1, A2_2, A2_6, B27_2, B27_3, B27_5, B7_1, B7_2, B7_3 (Figure V.17).

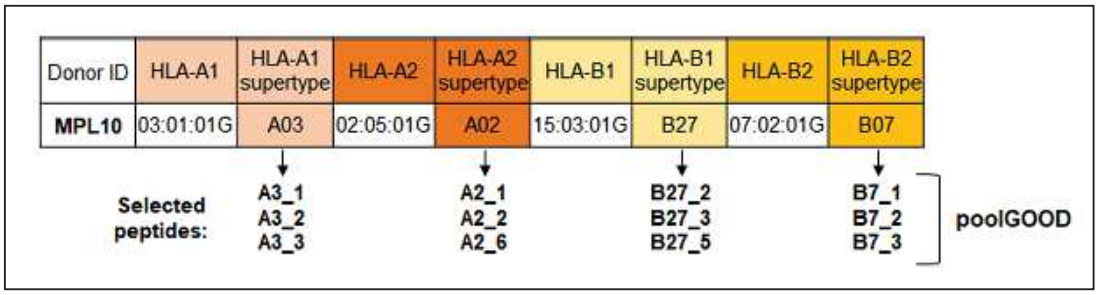


Figure V. 17 MPL10 HLA-typing for HLA-class I loci and peptides selected to compose the peptide pool.

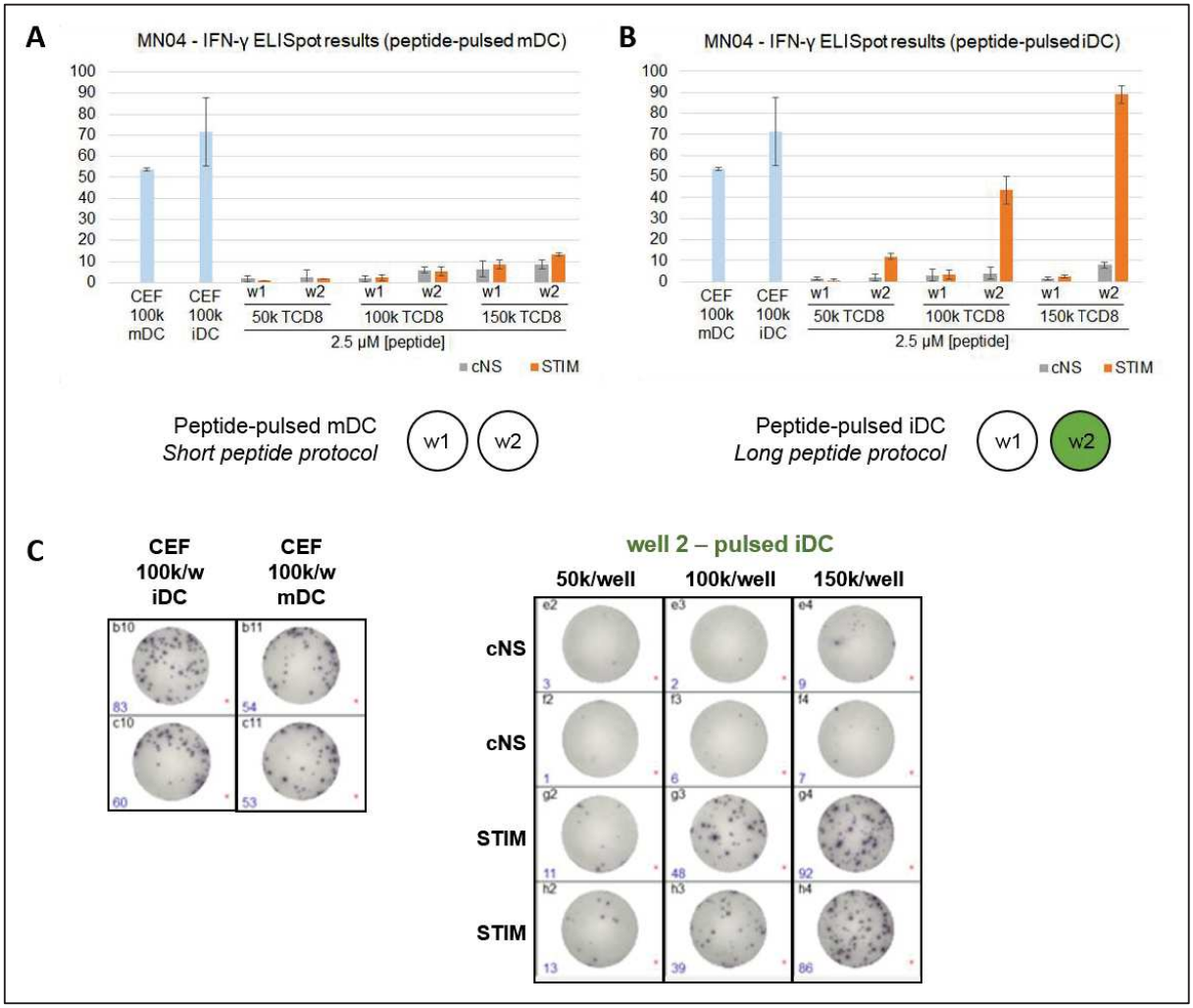


Figure V. 18 IFN-γ ELISpot results for experiment MN04, T cell amplification assays with total TCD8 T cells from donor MPL10. *A) TCD8+ amplified with the short peptide protocol (mDC stimulation). B) TCD8+ cells amplified with the long peptide protocol (iDC stimulation). C) ELISpot well images and respective spot counts for CEF and positive well.*

Again, to confirm optimal cell density for the detection of the induced response after stimulation in the 48-well plates, different TCD8⁺ cell numbers were seeded in the ELISpot plate in duplicate wells, similarly to experiment MN03. For the same number of seeded cells (100.000 TCD8 cells/well), iDC stimulation (long peptide protocol) with CEF control peptide pool yielded better results than mDC stimulation (short peptide protocol) – CEF 100k/iDC spot counts 60 and 83 and CEF 100k/mDC spot counts 53 and 54 (Figure V.18). Background levels of cNS wells are low, maximum spot count of 10, but are proportional to cell density (Figure V.18).

In this experiment, stimulated cells with the short peptide protocol show low spot counts regardless of cell density, with a maximum of 13 spot counts (Figure V.18 panel A). On the contrary,

successful amplification and detection of specific TCD8 cells is observed when stimulated with pulsed iDC, (long peptide protocol) (Figure V.18 panel B). The observed responses are highly specific, as cNS wells (culture-stimulated cells unstimulated in the ELISpot) show very low spot counts, meaning IFN-γ-producing cells are primed through peptide-specific stimulus and become activated.

5. Discussion - Immunoscreenings with naive donor samples

The immunoassays with naive donor samples currently still present several challenges. The main limitation is the number of cells needed per condition, since specific cells are rare. These experiments are also extremely time-consuming (with over 20-day cultures), and resource-consuming due to reagents' costs. Also, *in vitro* cellular amplification occurs under non-physiological conditions, so analysis should be performed with care. Finally, a high number of individuals should be tested to assess the peptides' overall immunogenicity in the naive repertoire. Specific approaches were adopted to mitigate some of these issues, namely: i) the generation of a naive donor bank and respective HLA-typing; and ii) the use of matrix-based peptide pools.

Peptide immunogenicity testing through immunoassays with samples from naive donors is rather innovative in the field of peptide-based vaccine development. Several laboratories, namely the Bernard Maillere Lab (CEA – Saclay), have developed techniques of *in vitro* cellular amplification that allow the analysis of specific T cell precursors in the naive repertoire (8–10,17,22–24). Immunogenicity testing is often focused on samples from exposed/healed individuals or animal models. We believe immunogenicity testing in the naive repertoire is essential to assess the prophylactic potential of peptide vaccine candidates, since there is a minimum threshold of specific naive T cells in circulation needed to generate specific effector and memory T cells. Also, the TCR-specificity of responding naive T cells correlates with the specific memory T cells present in immune individuals. We propose to assess and compare the peptide immunogenicity profiles in the naive and memory repertoires to select immunogenic peptide candidates to use in a final vaccine formulation against leishmaniasis. Overall, this approach allows for the fast validation of peptide vaccine candidates with increased confidence in late-stage human *in vivo* immunogenicity and vaccine effectiveness.

5.1. Naive donor bank

The generation of a naive donor bank and respective HLA-typing results allows the optimization of samples to use in the immunoscreening experiments. Donors can be selected according to HLA-typing results to maximise HLA coverage while testing a minimal number of individuals.

The naive donor bank was generated with samples from the French Blood Bank in Toulouse. The HLA-typing results for HLA-A, -B, -DRB1, and -DPB1 alleles expressed by the naive donors match that of the French population. The HLA diversity represented in the naive donor bank also corresponds to the most common alleles expressed by populations in South and Central America, North Africa, Western Asia and South Asia. These regions include the populations most at risk of developing leishmaniasis and are the target populations for an anti-*Leishmania* vaccine. Unquestionably, the naive donor bank is an extremely small population compared with the populations considered for the allele frequency studies, but the most frequent alleles are nonetheless represented.

Ultimately, we aim at increasing the sample size of the naive donor bank and increase the number of represented alleles, particularly those which are currently underrepresented. However, the most prevalent HLA alleles in the world population, and particularly in populations in endemic leishmaniasis areas, already represented in the naive donor bank will provide a great HLA coverage for future experiments.

A peptide-based vaccine formulation that includes peptides restricted to nine HLA-class I alleles (HLA-A*02, -A*24, -A*68, -A*01, -A*03, -B*35, -B*44, -B*51, -B*07) will be effective in these world regions. Regarding HLA-class II alleles, although these are expected to be more promiscuous than class I peptides, peptides restricted to HLA-DRB1*03, -DRB1*04, -DRB1*07, -DRB1*08, -DRB1*11, -DRB1*13, and -DRB1*15 should be included in the final vaccine formulation.

5.2. Peptide immunogenicity testing with samples from naive donors

The naive T cell assay (experiment MN01) confirmed the presence of *Leishmania*-specific TCD8 cells in the naive repertoire, with 4 peptide pools generating at least 2 positive cell lines (in 10 tested per condition, i.e. 2 million T CD8+ cells). This assay shall be repeated in the future, after pre-selecting peptides, with fewer peptides further along the peptide validation pipeline and with some protocol changes (i.e. cell counting before ELISpot seeding, ELISpot triplicates and higher number of cell lines).

The protocol used in experiment MN01 was adapted from Wolfl et al, to a simplified version using total PBMC as APC. While the authors recommend the use of monocyte-derived dendritic cells, at the time of this experiment it was not possible to induce DC differentiation *in vitro*, and the use of autologous total PBMC as APC is commonly done as well (25). The B cells and monocytes present in this population are for the most part in an immature state.

The donor MPL3 is homozygous for the HLA-A loci (HLA-A*23:01:01G alleles classified as A24 supertype) and expresses HLA-B*44:03:01G (B44 supertype) and HLA-B*49:01:01G, an unclassified allele. Considering four double-positive immunogenic peptides that generated more than 2 specific cell lines, only one (A24_3) matches MPL3's HLA typing results. Interestingly, the allele B*49:01 has ambiguous binding pocket specificities. This allele has a B pocket specificity matching to the B44 supertype, and F pocket binding like the B*38:01 allele (hydrophobic), so it likely this allele shares binding specificities with other alleles in the B44 supertype, and perhaps even other supertypes.

Surprisingly, when the total TCD8 stimulation assay protocol was adopted, the use of mature DCs with the short peptide protocol no longer stimulated T cell proliferation and IFN-γ production.

The number of co-cultures performed (MN02: 2 co-cultures, MN03: 3 co-cultures), the peptide concentration (1, 2,5 or 5 μM), or the assay format do not seem to be as important as the DC stimulation protocol, since IFN-γ responses were not detected in both MN02 and MN03 experiments. Many factors may contribute to this and apparently these *Leishmania*-specific peptides, even if short (9-mer) may need intracellular processing. Furthermore, we estimate that, even for short peptides, optimal DC stimulation requires peptide processing and presentation with simultaneous DC maturation through TLR activation, due to peptides' intrinsic weak immunogenicity.

In experiment MN04, when immature DCs were again used as APC, and stimulated during 16 hours with both peptides and maturation factors, specific IFN-γ production was induced and detected. These results suggest APC's immature state and the simultaneous peptide/TLR stimulation are key for proper peptide presentation and T cell activation.

During *in vivo* responses, immature DC uptake antigen in the periphery and migrate to the lymph nodes as mature cells where they prime responding T cells. Immature DCs are phagocytic and process and present exogenous antigens. TLR stimulation of immature DCs induces cell maturation process consisting of the upregulation of MHC and costimulatory molecules (CD40, CD80 and CD86) and migration to the lymph nodes. Mature DC become powerful APC, able to prime naive T cells and influence T-cell differentiation. TLRs are therefore a unique link between pathogen recognition and induction of T cell responses. The maturation conditions and stimulation by different pattern recognition receptors (PRR) influence DC's costimulatory molecules and the ability to produce inflammatory cytokines, therefore impacting the observed T cell responses (26,27).

Notably, both immature and mature DC populations generated from donors MPL9 and MPL10 include cells with intermediary phenotypes, not expressing any of the stained markers (CD14-CD209-

$CD83^-$). To describe these populations other markers should be included in the analysis. Also, some contaminant cells (probably non-adherent PBMC) and debris due to cell thawing prior to staining are observed (18 to 47% of total events). Nevertheless, the differentiation of monocytes into DCs is successful as $CD14^-$ cells correspond to over 98% of all cells.

For MPL9 over 40% of total $CD14^-$ cells are immature DCs ($CD14^-CD209^+CD83^-$). The percentage of immature DCs after differentiation for MPL10 is much lower (18% of total $CD14^-$ cells). These results may result from a shorter differentiation time for monocytes from MPL10 (four-day stimulation instead of five). However, overnight maturation with LPS and Resiquimod greatly increased the number of mature DCs (from 2,3% to 28,6% of $CD14^-CD209^+CD83^+$ cells).

The experiments MN02-04 used monocyte-derived DC as antigen-presenting cells. These were co-cultured with $TCD8^+$ cells at a 1:10 ratio. Although no optimisation experiments were performed to improve this ratio, there is evidence that optimal APC:T-cell ratio can go up to 1:160, or be as low as 1:2, with lower ratios having potentially negative impact on T cell stimulation (28). The intermediary 1:10 ratio was adopted, same as in the total $TCD4^+$ cell amplification protocol.

After experiment MN01, a total TCD8 amplification assay protocol was adopted due to the high number of conditions to test and to decrease the experiment-associated costs. Even though naive donors are expected to have never been exposed to leishmaniasis, there is an increased risk of triggering cross-reactive responses, potential cross-specific memory T cells in circulation against other pathogens or antigens. To assure no recall responses are being detected, short term total PBMC stimulation experiments will be performed with both TSLA and peptide pools to confirm, respectively, the naive donor status (TSLA-negative) and IFN-γ production by amplified specific naive T cells and not cross-specific memory cells.

Peptide concentrations in the peptide pools tested ranged from 1 to 5 μM. While the nature of the induced response to change remains unchanged regardless of the peptide concentration, a positive effect on T cell activation is expected in the range of 1-10 μg/mL of a 9-mer peptide (29). Nevertheless, the peptide pool composition may also influence the observed responses. Response specificity against a single peptide may be decreased if pools contain many peptides, and likewise, pool sensitivity may be increased if there are multiple positive peptides. On the other hand, peptide competition and low frequency responses may hinder response detection. We expect that the peptides tested present binding competition since they are all predicted to be strong binders. In

experiment MN03, peptide pools with over 20 peptides each were used and stimulated cells at 5 µM show decreased spot counts when compared with the other tested concentrations. Therefore, it is possible that increased peptide numbers and/or increased peptide concentration inhibit T cell activation, either through peptide competition or due to the absence of high avidity specific T cells. We believe peptide competition is higher in the tested peptide pools than expected for traditional epitope mapping experiments, where only a few immunogenic regions are found within a full protein, and many negative peptides are included in the peptide pools. The matrix-based pools contain 7 peptides, and the poolGOOD used in MN04 contains 12 peptides at 2,5 µM each. Peptide pools containing up to 12 peptides at 2,5 µM do not seem to inhibit T cell activation, so this peptide concentration will be adopted for future experiments.

Peptide pools are an efficient and sensitive way to optimise epitope mapping experiments and even to detect low level responses while decreasing experiment-associated costs and sample size. The use of matrix-based pools provides additional advantages regarding the number of assays performed and cryovials needed (compared with single peptide testing or mini-pool peptide testing). However, some unclear results regarding double positive peptides argue for the need of a second ELISpot to confirm peptide specificity (two-stage matrix-based pool approach).

The CEF control peptide pool provides a good control for peptide processing and presentation to total TCD8+ cells. Although results were quite variable (from 16 spot counts up to TNTC), they were consistently positive and with large spots. Evidently, the observed responses are mediated by memory T cells instead of naive T cells. While CEF-induced responses do not necessarily provide evidence for amplification of specific T cell populations through clonal expansion, they still provide evidence that antigen presentation mechanisms by peptide-pulsed APC are effective in priming specific T cells *in vitro*. True positive control peptides for the naive T cell repertoire should be included, however, universal control peptides for naive T cell stimulation do not exist. A few peptides were so far described to have specific naive T cells in the naive repertoire. The most studied are peptides from the Melan-A/MART-1 antigen (Melanoma-Associated Antigen) (30,31). Still, these are restricted to HLA-A*02 alleles and have an unusually high precursor frequency. We propose and expect to find the most immunogenic peptides with specific naive T cells through the optimisation of T cell amplification assays, which have been shown to be effective in detecting specific T cells in the naive repertoire.

In experiment MN04, which used the long peptide stimulation and a tailored peptide pool containing 12 peptides (2,5 µM each), we successfully amplified and detected *Leishmania*-specific T cells from the naive repertoire. Only one culture well was positive (1 million TCD8 cells seeded) among two stimulated wells (2 million TCD8 cells tested), as expected for low frequency specific cells.

The 48-well batch stimulation assay format allows for the successful amplification of specific T cells present in the naive repertoire, as well as practical advantages that permit cell counting, ELISpot well triplicates, and therefore the application of statistical tests for well positivity determination. Although, in this experiment, only two wells per condition were seeded in the ELISpot plate due to the optimisation of peptide concentration and cell density.

We aim to use the 48-well batch stimulation assay with a higher number of donor samples (n=10) to screen the synthetic peptides and select the most immunogenic ones. Possibly, the number of total TCD8 cells tested per condition can be increased (3 or 4 million total TCD8 per condition).

However, the 48-well batch stimulation assay format eliminates the application of the Poisson distribution formula for the calculation of the specific T cell precursor frequency. Hence, we propose to use the batch stimulation assays to pre-select the (most) immunogenic peptides and, subsequently, to perform precursor frequency analysis experiments with a fewer number of pre-validated immunogenic peptides, and possibly bypassing polyclonal activation with peptide pools but performing co-cultures with DC pulsed with single peptides.

Because naive T cell amplification assays were performed in a limited number of individuals with all peptides, comparison is very limited. However, MN01 and MN04 clearly evidence the existence of a minimum number of naive T cells specific to *Leishmania* peptides in naive donors, hence, validating their predicted immunogenicity and the feasibility of using such peptides in an effective vaccine formulation. The naive donor bank, which will be augmented with additional naive donor samples, and the *in vitro* stimulation protocol, which is now optimised, will allow to screen all peptides and select the most immunogenic by serial routine experiments.

6. Results - Immunoscreenings in healed individuals from endemic areas (Tunisia)

The first peptide immunoscreenings performed with healed donors' samples tested all 49 class I peptides and 24 HLA-class I peptides in matrix-based pools, using a short term total PBMC stimulation protocol.

Two series of blood samples were received, both including 10 samples from healed donors – first series TUN1 to TUN10, and second series TUN11 to TUN20. In the first series (TUN1 to 10), one blood sample was hemolysed (TUN2) so no PBMC were purified. IFN-γ ELISpot was performed for both series and culture supernatants were collected and stored at -80°C. The series of blood samples with naive donors was performed in the Institut Pasteur Tunis, no IFN-γ ELISpot was performed for these samples.

Peptide immunogenicity or the capacity to prime *in vitro* specific recall responses was successfully assessed for 9 healed donors (see Appendix V.2 for individual data).

6.1. Healed donors' personal information and HLA-typing results

The average age for the first healed donor series was 43 years-old, ranging from 26 to 64 years-old (Table V.7). Male healed donors are much more frequent than female donors in this series, with only one woman in nine donors (TUN4). The presence of CL lesions was inquired in the medical questionnaire. On average, healed donors in the first series presented 2 cutaneous lesions, between one up to five lesions per donor. Most cutaneous lesions were localised to the legs or arms, only a few CL lesions were localised in the face (two donors, 11%) (Table V.7).

HLA-typing results for the first healed donor series was performed (Table V.8)

Table V. 7 Personal information and medical history from recruited healed donors (TUN1-TUN20).

Donor ID	Blood collection		Personal information							Medical history		Cutaneous Leishmaniasis medical history								
	Collection date	Collection volume (mL)	Gender	Date of birth	Always lived in this endemic area?	Years in endemic area	Endemic area	Profession	Literary habilitations	Relevant clinical history (including HIV+)	Current medical treatment	Scar(s)?	Scar localization	Mucous symptoms?	Medical Report	Parasitological confirmation	Parasitological confirmation date	Parasitological confirmation method	Treatment?	Treatment with	Treatment duration (months)
TUN1	20/09/2017	20	M	22/06/1984	yes	21	Lela Ksar GAFSA	hairdresser	High school	none	none	1	right leg	no	Hopital Regional GAFSA	yes		smear	yes	arab traditional treatment	6
TUN2	20/09/2017	20	M	03/04/1984	yes	11	Lela GAFSA	factory worker/other	High school	none	none	1	left lower leg	no	Hopital Regional GAFSA	yes		smear	yes	cryotherapy+antibiotherapy	0.5
TUN3	20/09/2017	20	M	21/10/1986	yes		Ras el Kef GAFSA		High school + 'Bac'	none	none	2	left lower leg+left foot	no		yes	08/08/2014	smear	yes	cryotherapy+antibiotherapy	
TUN4	20/09/2017	20	F	23/09/1981	yes		Alim Sened GAFSA		Superior education	none	none	2	nose+left wrist	no	Hopital Regional GAFSA	yes	01/10/2007	smear	yes	arab traditional treatment	1.5
TUN5	20/09/2017	20	M	19/12/1968	yes		Lela Ksar GAFSA / Gharien Lybie	employee	Superior education	none	none	1	left lower leg	no		yes	08/12/2015	smear+PCR	yes	15-day hospitalisation: Glucantime® IM (60mg/Kg)+cryotherapy	0.5
TUN6	20/09/2017	20	M	16/01/1991	yes		Lela Ksar GAFSA		High school	none	none	1	right lower leg	no	Hopital Regional GAFSA	yes	25/11/2016	smear	yes	cryotherapy	
TUN7	20/09/2017	20	M	15/04/1966	yes		Ajama Ksar GAFSA	farmer	High school	none	none	5	right arm+right leg (5 lesions)	no		yes		smear+PCR	yes	cryotherapy+antibiotherapy	
TUN8	20/09/2017	20	M	26/10/1972	yes		Ajama Ksar GAFSA	mechanic	Primary school	none	none	2	abdomen+ right foot	no	Hopital Regional GAFSA	yes	19/12/2014	smear	yes	17-day hospitalisation: Glucantime® IM (60mg/Kg)+cryotherapy	0.5
TUN9	20/09/2017	20	M	05/11/1953	yes		Ajama Ksar GAFSA	engineer	Superior education	none	none	1	left leg	no		yes	10/05/2005	smear+PCR	yes	cryotherapy+antibiotherapy+ local therapy (paromomycin-based topical cream)	
TUN10	20/09/2017	20	M	20/09/1953	yes		Gafsa centre ville	hairdresser	High school	none	none	1	face	no	Hopital Regional GAFSA	yes		smear	yes	cryotherapy (2 applications in a 14-day interval)	1

Table V.7 (continued)

Donor ID	Blood collection		Personal information							Medical history		Cutaneous Leishmaniasis medical history									
	Collection date	Collection volume (mL)	Gender	Date of birth	Always lived in this endemic area?	Years in endemic area	Endemic area	Profession	Literary habilitations	Relevant clinical history (including HIV+)	Current medical treatment	Scar(s)?	Scar localization	Mucous symptoms?	Medical Report	Parasitological confirmation	Parasitological confirmation date	Parasitological confirmation method	Treatment?	Treatment with	Treatment duration (months)
TUN11	03/10/2017	20	F	12/30/1996	yes	17	Souk el Fajer	employee	Superior education	none	none	1	left knee	no	Hopital Regional dispensaire Sidi Bouyehir	yes	07/10/2002	smear	yes	Glucantime® intralesional (2 times per week)	1
TUN12	03/10/2017	20	F	16/04/1969	yes		Cite Echebeb GAFSA	unemployed	High school	none	none	1	right arm	no		yes	08/08/2013	smear	yes	cryotherapy+antibiotherapy+ local therapy (paromomycin-based topical cream)	21 days
TUN13	03/10/2017	20	M	11/05/1998	yes		Cite Echebeb GAFSA		High school + 'Bac'	none	none	1	left hand (back)	no	Hopital Regional GAFSA	yes	08/08/2007	smear	yes	cryotherapy+local therapy (paromomycin-based topical cream)	
TUN14	03/10/2017	20	F	09/05/1972	yes	10	Lela GAFSA	factory worker	High school	none	none	1	leg	no		yes	09/09/1980	smear	yes	Glucantime® intralesional (2 times per week)	1
TUN15	03/10/2017	20	M	22/07/1995	yes		Lela GAFSA	carpenter	High school	none	gastric ulcer	4	both arms	no	Hopital Regional GAFSA	yes	07/08/2007	smear	yes	cryotherapy+ local therapy (paromomycin-based topical cream)	0.5
TUN16	03/10/2017	20	M	28/08/1993	yes		Lela GAFSA	factory worker	High school	none	none	2	both lower legs	no	Hopital Regional GAFSA	yes	03/08/2001	smear	yes	cryotherapy+antibiotherapy	
TUN17	03/10/2017	20	M	27/10/1968	yes	50	Lela GAFSA	employee	High school	none	none	6	both legs	no	Hopital Regional GAFSA	yes	19/09/1980	smear	yes	Glucantime® IL + traditional treatment	
TUN18	03/10/2017	20	F	01/08/1975	yes	42	Lela GAFSA	sans emploi	Primary school	none	at 12 years old: 15-day corticosteroid therapy due to chronic bronchitis and pneumothorax	2	both legs	no	Hopital Regional GAFSA	yes	08/06/1985	smear	yes	Glucantime® intralesional (6 injections)	
TUN19	03/10/2017	20	F	07/02/1978	yes	18	Souk el Fajer / Moulares	employee	Superior education	none	none	3	both lower legs	no	Hopital Regional Moulares	yes	10/09/1992	smear	yes	Glucantime® intralesional (2/week)	3 weeks
TUN20	03/10/2017	20	F	x/x/1967	yes	50	Lela GAFSA	unemployed	illiterate	none	breast nodule ablation (4 years ago)	2	both hands (back)	no		yes	07/10/2015	smear	yes	hospitalisation +cryotherapy+Glucantime® IM (60mg/Kg for 15 days)	0.5

Table V. 8 HLA-typing results for the first healed donor series (TUN1-TUN3 to TUN10). *A) Six loci corresponding to HLA-class I and class II polymorphic regions (HLA-A, -B, -C, -DRB1, -DPB1, -DQB1) were sequenced by NGS (exons 2 and 3), leading to the 6-digit identification of the alleles expressed by each donor. Allele supertype classification by Sidney et al 2008* (20)*. B) Ambiguous identifications are given a multiple allele code (MAC, https://hml.nmdp.org/MacUI/)*

A

Donor ID	HLA-A1	HLA-A1 supertype	HLA-A2	HLA-A2 supertype	HLA-B1	HLA-B1 supertype	HLA-B2	HLA-B2 supertype	HLA-C1	HLA-C2	HLA-DRB1	HLA-DRB2	HLA-DQB1	HLA-DQB2	HLA-DPB1	HLA-DPB2
TUN1	30:01:01G	A01 A03	02:05:01G	A02	50:01:01G	B44	42:01:01	B07	06:02:01G	17:01:01G	03:02:01	15:02:01G	04:02	06:01	01:AGXDR	04:AHPJH
TUN3	01:01:01G	A01	03:01:01G	A03	15:03:01G	B27	44:02:01G	B44	02:10:01G	16:04:01G	11:04:01	10:01:01	05:01	03:01:01	17:01:01	23:01:01
TUN4	23:01:01G	A24	02:22:01G	A02	58:01:01G	B58	44:03:01G	B44	03:02:01G	04:01:01G	07:DFRJ	07:DFRJ	03:03	02:GKDU	01:01:02G	17:JECX
TUN5	26:01:01G	A26	34:02:01	A03	07:02:01G	B07	08:01:01G	B08	07:01:01G	07:02:01G	03:01:01	15:01:01	06:02:01	02:01:01	01:AETTG	04:ADCGE
TUN6	01:01:01G	A01	29:02:01G	A01 A24	15:03:01G	B27	08:01:01G	B08	02:10:01G	07:01:01G	03:01:01	07:DFRJ	02:01:01	02:GKDU	11:AWXGR	04:01
TUN7	29:02:01G	A01 A24	32:01:01G	A01	35:03:01G	B07	44:03:01G	B44	16:01:01G	04:01:01G	11:01:01	15:01:01	06:02:01	03:01:01	04:FNVS	04:HJMR
TUN8	01:01:01G	A01	02:01:01G	A02	52:01:01G	B62	47:01:01G	unclassified	12:02:01G	06:02:01G	15:02:01G	07:DFRJ	03:03:02	06:01	09:01:01	04:01
TUN9	02:02:01G	A02	32:01:01G	A01	53:01:01G	B07	35:03:01G	B07	04:01:01G	04:01:01G	11:01:02	11:01:01	03:19:01	03:01:01	13:AKXDF	04:02
TUN10	02:05:01G	A02	02:02:01G	A02	35:08:01G	B07	44:03:02G	B44	07:01:01G	04:01:01G	04:03:01	11:01:02	06:02:01	03:02:01	17:JECX	04:ADCGE

B

Multiple Allele code (MAC)	Decode results (shows the allele list for a MAC at a locus and allele-family)
DRB1*07:DFRJ	DRB1*07:01/DRB1*07:79
DPB1*01:AGXDR	DPB1*01:01/DPB1*417:01/DPB1*462:01
DPB1*04:AHPJH	DPB1*04:02/DPB1*463:01
DPB1*17:JECX	DPB1*17:01/DPB1*131:01
DPB1*04:ADCGE	DPB1*04:01/DPB1*350:01
DPB1*11:AWXGR	DPB1*11:01/DPB1*654:01
DPB1*04:FNVS	DPB1*04:02/DPB1*105:01
DPB1*04:HJMR	DPB1*04:01/DPB1*126:01
DPB1*13:AKXDF	DPB1*13:01/DPB1*107:01/DPB1*518:01/DPB1*519:01
DPB1*17:JECX	DPB1*17:01/DPB1*131:01

The HLA-typing results for the healed donors, although the sample size is quite small, mostly match the most frequent alleles in the Tunisian population (Figure V.19 and V.20). The HLA-class I alleles HLA-A*02 are the most frequent. HLA-A*24 alleles, although very frequent in the Tunisian population, were not represented in the healed donor series. On the other hand, other alleles such as HLA-B*44, -B*35 and -A*01 were more frequent in the healed donor series (Figure V.19).

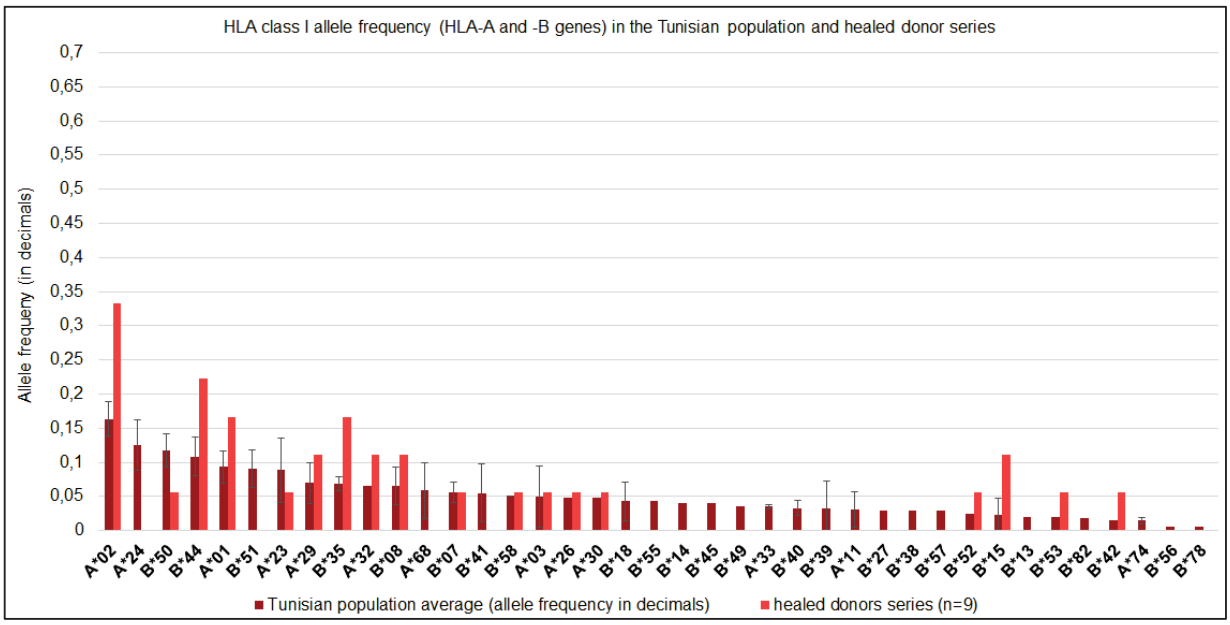


Figure V. 19 HLA class I allele frequency in the Tunisian populations and in the healed donor series. *Allele frequency for HLA-A and -B alleles (allele / 2n).*

The HLA-class II allele HLA-DPB1*04 is the most frequent in the Tunisian populations yet poorly represented in the healed donor series (Figure V.20). On the other hand, the second most frequent allele HLA-DRB1*07 is also well represented. Other alleles such as HLA-DPB1*03, -DPB1*06, -DRB1*11 and -DRB1*03 were more frequent in the healed donor series (Figure V.20).

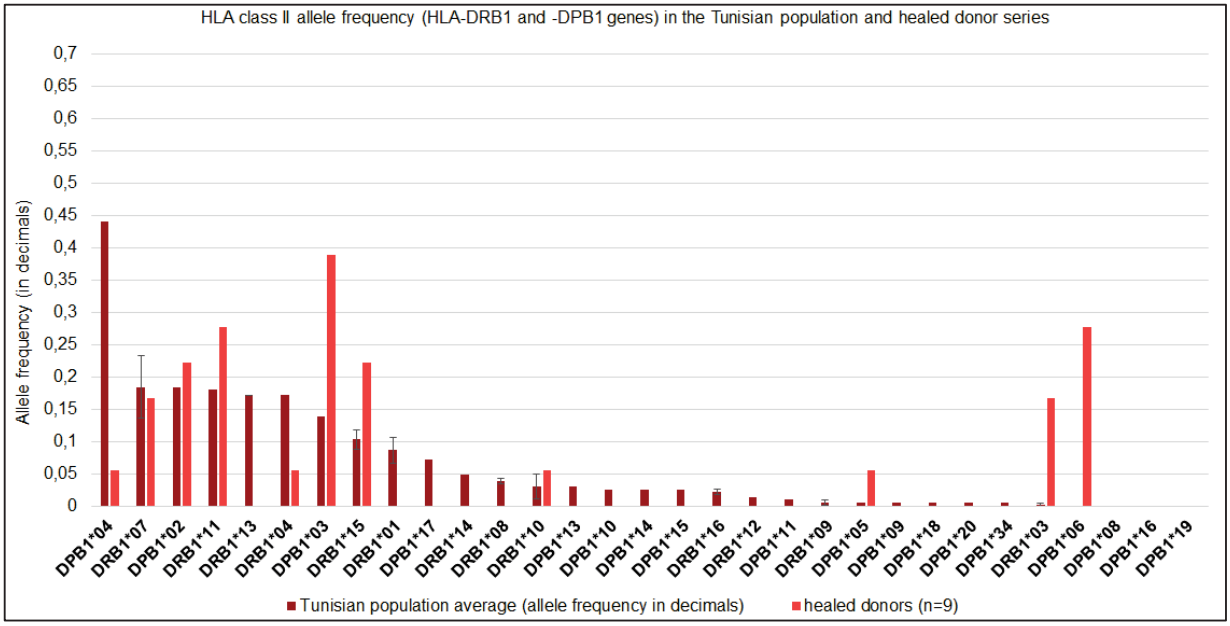


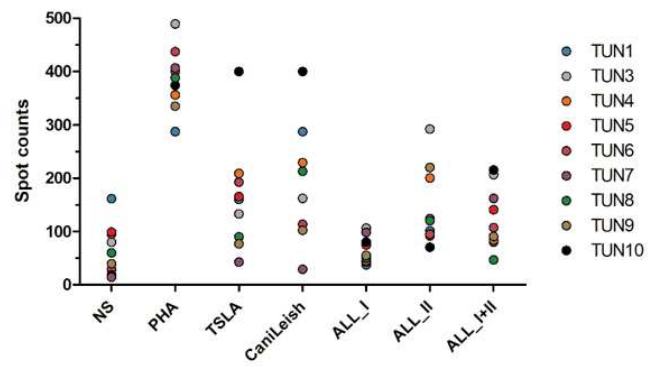
Figure V. 20 HLA class II allele frequency in the Tunisian populations and in the healed donor series. *Allele frequency for HLA-DRB1 and -DPB1 alleles (allele / 2n).*

6.2. Healed status validated by positive responses against TSLA

As expected, most donors show positive TSLA reactivity, defined by IFN-γ production by ELISpot (Figure V.19). Four donors (TUN3, TUN4, TUN5 and TUN10) have significantly higher IFN-γ responses against TSLA (TSLA-positive donors). Donor TUN7 has almost statistically significant differences (p-value=0,0571). Statistical significance was impossible to determine for donor TUN6 because only two ELISpot wells can be analysed due to an error during seeding (CaniLeish®-stimulated cells were stimulated with TSLA), but 192,5 spot count average is considerably higher than autologous NS spot counts (28,83 ±22,07). Two other donors (TUN9, TUN10) were considered TSLA-positive since ELISpot results show IFN-γ production during culture stimulation, evidenced by dark, highly saturated wells, even though spot counts are low or innumerable. In addition to TSLA, CaniLeish® was also used as positive control. Interestingly, CaniLeish® antigens induced vigorous IFN-γ responses in >75% of individuals (7 in 9 donors). Moreover, donors TUN1, TUN7, TUN8 and TUN9 had significantly higher responses against CaniLeish® but not TSLA (Figure V.21).

IFN-γ ELISpot results show donors have quite different background levels, from an average NS spot count of 14,5 (TUN7) up to an average spot count of 161,5 (TUN1) (Figure V.21).

A



B

Donor	NS	TSLA stim	CaniLeish® stim
TUN1	161.5	160,3	287.3
TUN3	80.2	**133***	**162.3***
TUN4	95.3	**209***	**229.3***
TUN5	98.8	**165,75***	**400 (TNTC)***
TUN6	28.8	192.5	113.5
TUN7	14.5	42.75 *(0.0571)*	**29.33***
TUN8	60.2	90.25	**213***
TUN9	39.7	77	**102.3***
TUN10	18.333	**400 (TNTC)***	**400 (TNTC)***

Figure V. 21 IFN-γ ELISpot results for controls for 1st healed series (n=9). *A) Spot counts for NS (never stimulated PBMC), for positive controls PHA, TSLA and CaniLeish®, and pools ALL_I (49 class I peptides), ALL_II (24 class II peptides) and ALL_I+II (all 73* Leishmania*-specific peptides). B) Table with averaged spot counts for all controls (* significantly different from autologous NS with the Mann-Whitney statistical test, p-value ≤ 0.05).*

6.3. IFN-γ responses against *Leishmania*-specific peptides

To evaluate recall responses against *Leishmania*-specific peptides, total PBMC were stimulated *in vitro* in the absence or presence of peptide pools containing HLA-class I or -class II peptides. Positive IFN-γ responses were induced by multiple peptide pools and by most donors – seven responder donors for HLA-class I peptide pools, and nine responder donors for HLA-class II peptide pools (Figure V.22). The number of responders and the number of peptide-specific responses induced for each donor were different according to the peptide pool (Figure V.22).

Overall, all HLA-class II pools induced IFN-γ production in at least one individual (II_pool5 and B), up to 7 individuals (II_pool4) (Figure V.22 panel B). On the other hand, three HLA-class I peptide pools did not elicit IFN-γ production in any of individuals tested (I_pool4, A and C), and the remaining 11 pools induced IFN-γ production in at least one donor (I_pool3 and 7), up to 4 donors

(I_pool1 and E) (Figure V.22 panel A). Individualised ELISpot results are included in Appendix V.2 (donors TUN1, TUN3-10).

For HLA-class II peptide pools, two donors had only one positive peptide pool so single immunogenic peptides were not identified (TUN1, II_poolE and TUN6, II_pool1). Likewise, for HLA-class I peptide pools, TUN6 had only one positive peptide pool (I_poolE) (Figure V.22). Results regarding classI peptide pools I_pool5 and I_pool6 are inconclusive for donor TUN7 since results are not available (no revelation antibody added to row C during ELISpot revelation).

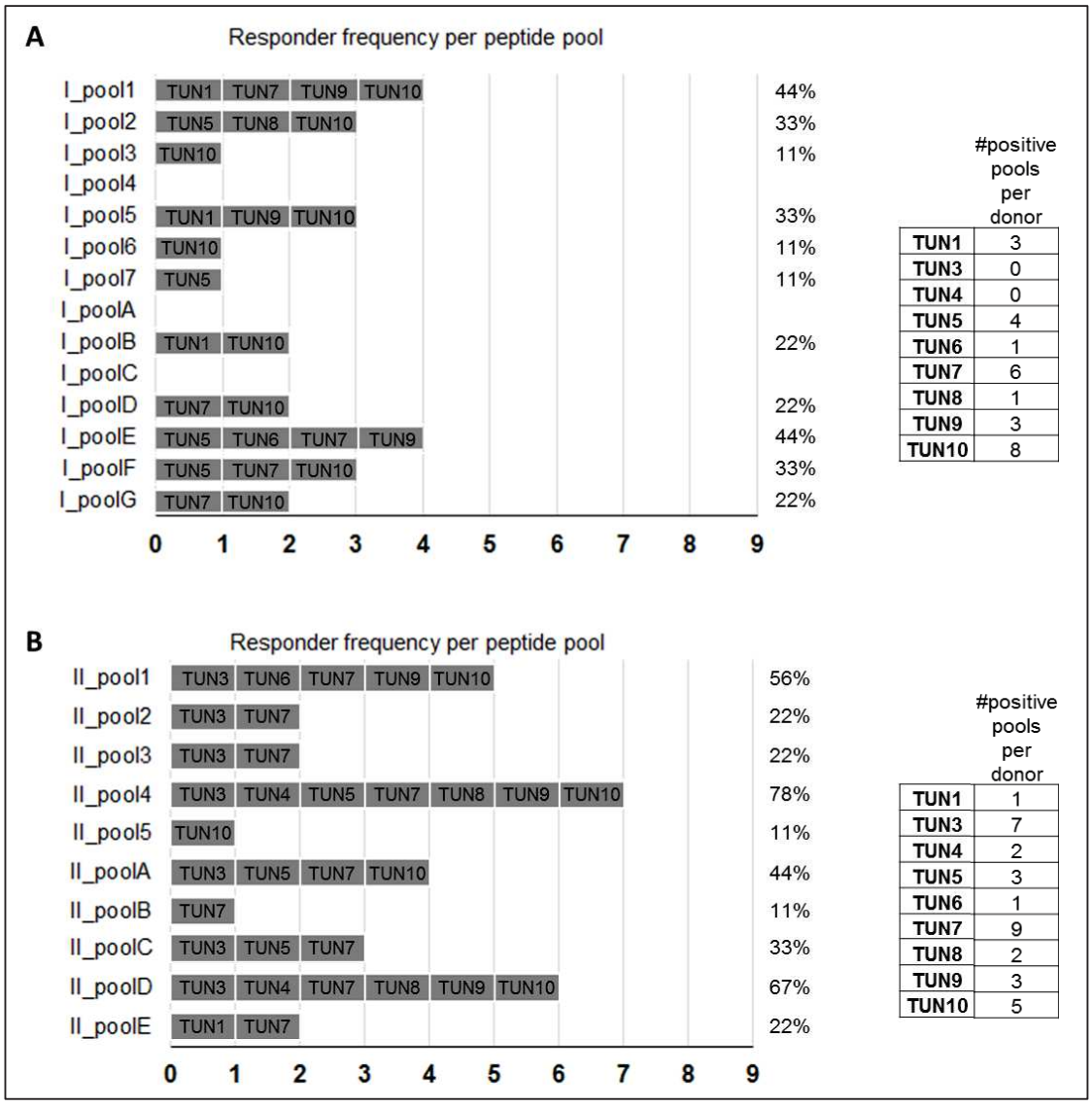


Figure V. 22 Responder frequency per peptide pool for the 1st healed series (n=9). *Number of responders, individuals with significantly increased IFN-γ production (responders) per peptide pool. A) Responder donors to class I peptide pools. B) Responder donors to class II peptide pools.*

TSLA-positive donor **TUN3** showed positive IFN-γ responses against HLA-class II peptide pools only, but against a high number of pools (7 out of 10: II_pool1, 2, 3, 4, II_poolA, C, and D) (Figure V.23 panel A). Twelve double-positive HLA-class II peptides were identified (14_b12, 16_b13, 3_a11, 22_b4, 7_a14, 12_a32, 21_b4, 17_b15, 4_a11, 5_a12, 9_a25, 13_b12). These peptides are predicted to be very promiscuous and bind to four up to eleven HLA-class II alleles.

For TSLA-positive donor **TUN4**, HLA-class I peptide pools did not induce significantly higher responses, but two positive HLA-class II pools did (II_pool4 and D). One double-positive immunogenic class II peptide was identified (pii19/13_b12) (Figure V.23 panel B). The **13_b12 peptide** is predicted to be highly promiscuous, predicted to bind to seven class II alleles: HLA-DPA1*01:03-DPB1*02:01, -DPA1*01:03-DPB1*04:01, -DPA1*02:01-DPB1*01:01, -DPA1*03:01-DPB1*04:02, -DPA1*03:01-DPB1*04:02, -DRB1*07:01, -DRB1*09:01, and -DRB1*13:02. These include alleles expressed by TUN4 namely, HLA-DRB1*07:01 and -DPB1*01:01 (Table V.8).

TSLA-positive donor **TUN5** had positive IFN-γ responses against four HLA-class I pools (I_pool2, 7, E, F) and three HLA-class II pools (II_pool4, A, C) (Figure V.23 panel C). The double-positive HLA-class I immunogenic peptides **B62_2**, **B8_4**, **B44_1** and **B44_4** were identified. Yet, the donor TUN5 expresses HLA alleles which belong to A26 (A*26:01:01G), A03 (A*34:02:01), B07 (B*07:02:01G), and B08 (B*08:01:01G) supertypes, meaning only one in four positive peptides match the donors's HLA typing (B8_4).Additionally, the double-positive immunogenic HLA-class II peptides **22_b4** and **17_b15** were identified in donor TUN5. Peptide 22_b4 is predicted to be promiscuous and bind to five class II alleles (HLA-DPA1*01:03-DPB1*02:01, -DPA1*01:03-**DPB1*04:01**, -DPA1*02:01-**DPB1*01:01**, -DPA1*03:01-DPB1*04:02, -DRB1*09:01). Peptide 17_b15 is predicted to be promiscuous and bind to seven class II alleles (HLA-DRB1*01:01, **-DRB1*03:01**, -DRB1*07:01, -DRB1*09:01, -DRB1*11:01, -DRB3*01:01, -DRB5*01:01, -DRB5*01:01). Some of these allele specificities match donor TUN5 typing results for HLA-class II loci -DRB1 and -DPB1, namely HLA-DRB1*03:01, -DPB1*04:01, and -DPB1*01:01 alleles (Table V.8).

TSLA-positive donor **TUN10** responded to seven positive HLA-class I pools (I_pool1, 2, 5, 6, B, D, G) and five positive HLA-class II pools (II_pool1, 4, 5, A, D) (Figure V.23 panel D). The double-positive immunogenic class I peptides identified were A1_1, A1_4, A3_2, A24_3, B7_3, B7_4, B8_3, B27_1, B44_5, B58_3, B58_4, B58_5. Given donor TUN10's HLA typing results, 5 in 16 peptides match the donor's HLA allele typing (**B7_3, B7_4, B7_5, B44_5, A2_4**) (Table V.8). The double-positive immunogenic HLA-class II peptides identified in donor TUN10 were **14_b12, 22_b4, 18_b15, 4_a11, 13_b12, 10_a25**. Again, all these peptides are predicted to be very promiscuous and bind to five up to eleven class II alleles.

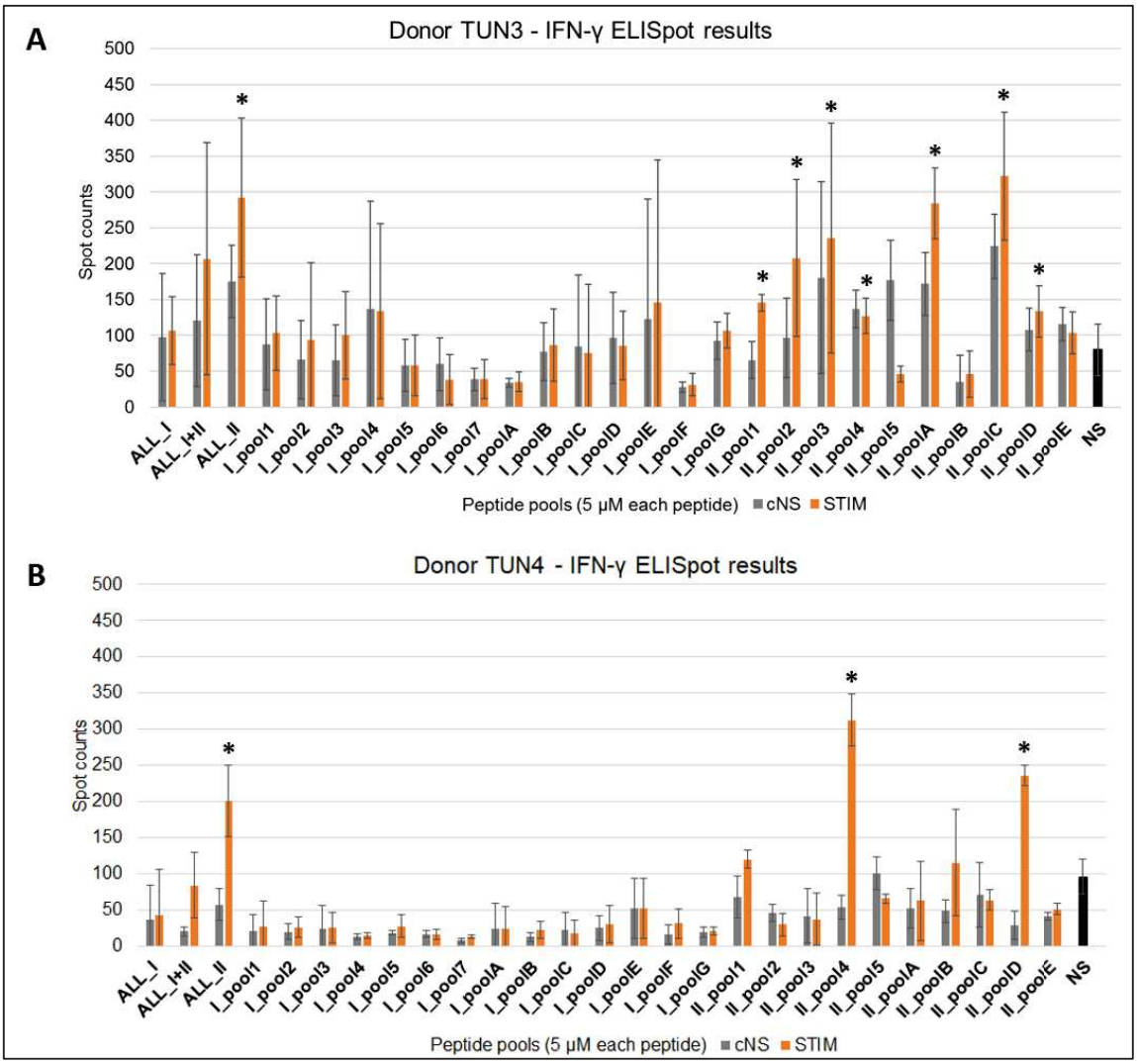


Figure V. 23 IFN-γ ELISpot results for TSLA-positive donors TUN3 (A), TUN4 (B), TUN5 (C), and TUN10 (D). *Highlighted pools (*) induced significantly higher IFN-γ production when compared with autologous NS (Mann-Whitney test, p-value < 0.5).*

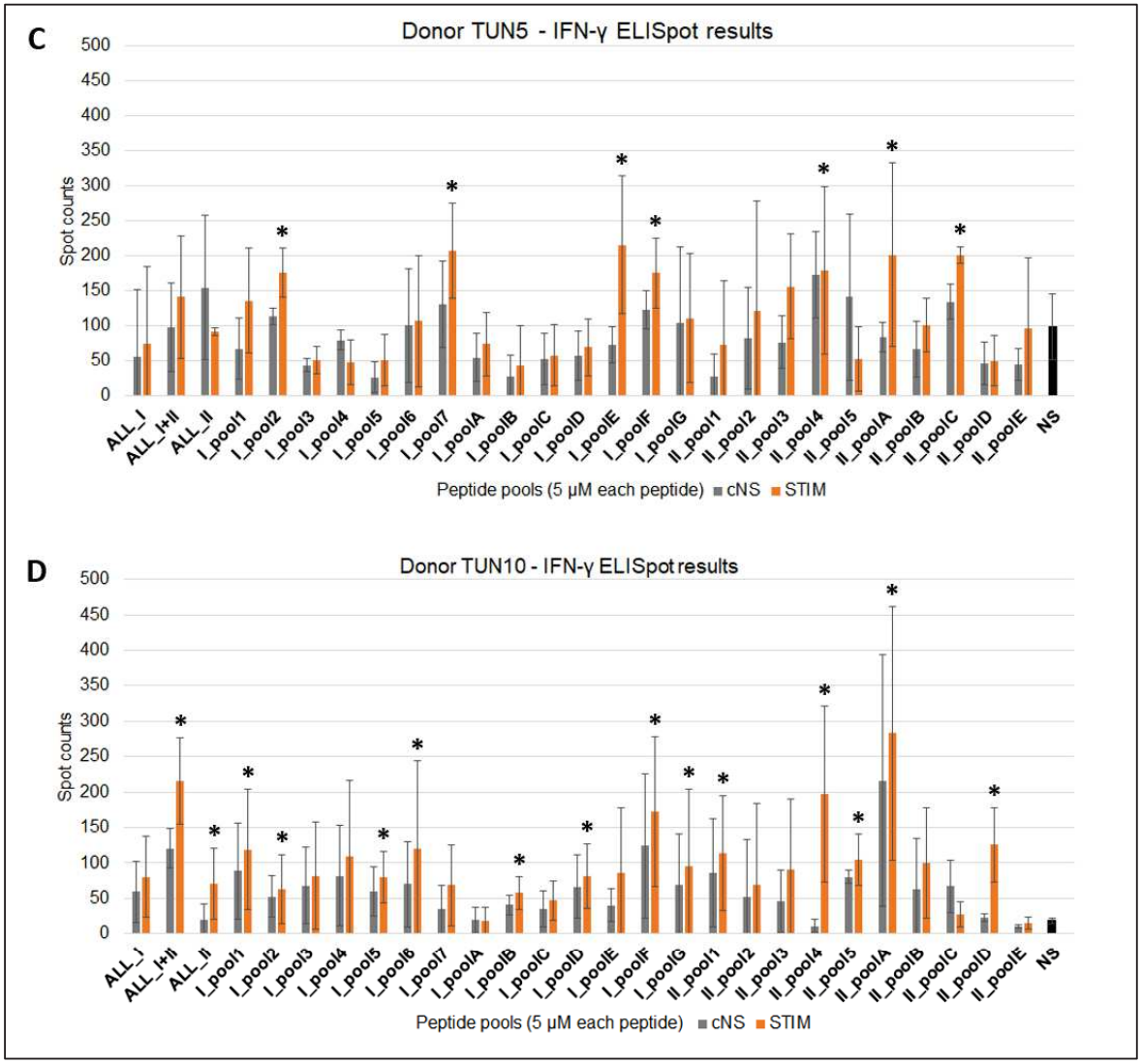


Figure V.23 (continued)

Donor **TUN7** is also highly responsive. This donor shows the lowest background values out of all the donors which may influence the observed results. Positive responses may be overestimated since this donor also shows the weakest responses to TSLA and CaniLeish®. Nevertheless, from the positive HLA-class I pools and from the 8 immunogenic peptides identified (A1_1; A3_3; A24_1; A3_1; A2_4; B7_2; B58_5 and B8_1), three peptides correspond to the donor's HLA-typing results who expresses HLA-class I alleles which belong to the A01 (A*32:02:01G), A01A24 (A*29:01:01G), B07 (B*35:03:01G) and B44 (B*44:03:01G) supertypes – peptides **A1_1, A24_1 and B7_2** (Table V.8) .

As expected, individuals with lower background correspond to higher magnitude of response, with TUN7 and TUN10 presenting the strongest IFN-γ responses (Figure V.24). Overall, HLA-class II

peptide pools induced stronger responses, with response magnitude ranging from 2-fold up to almost 24-fold higher responses by peptide stimulated cells (Figure V.24).

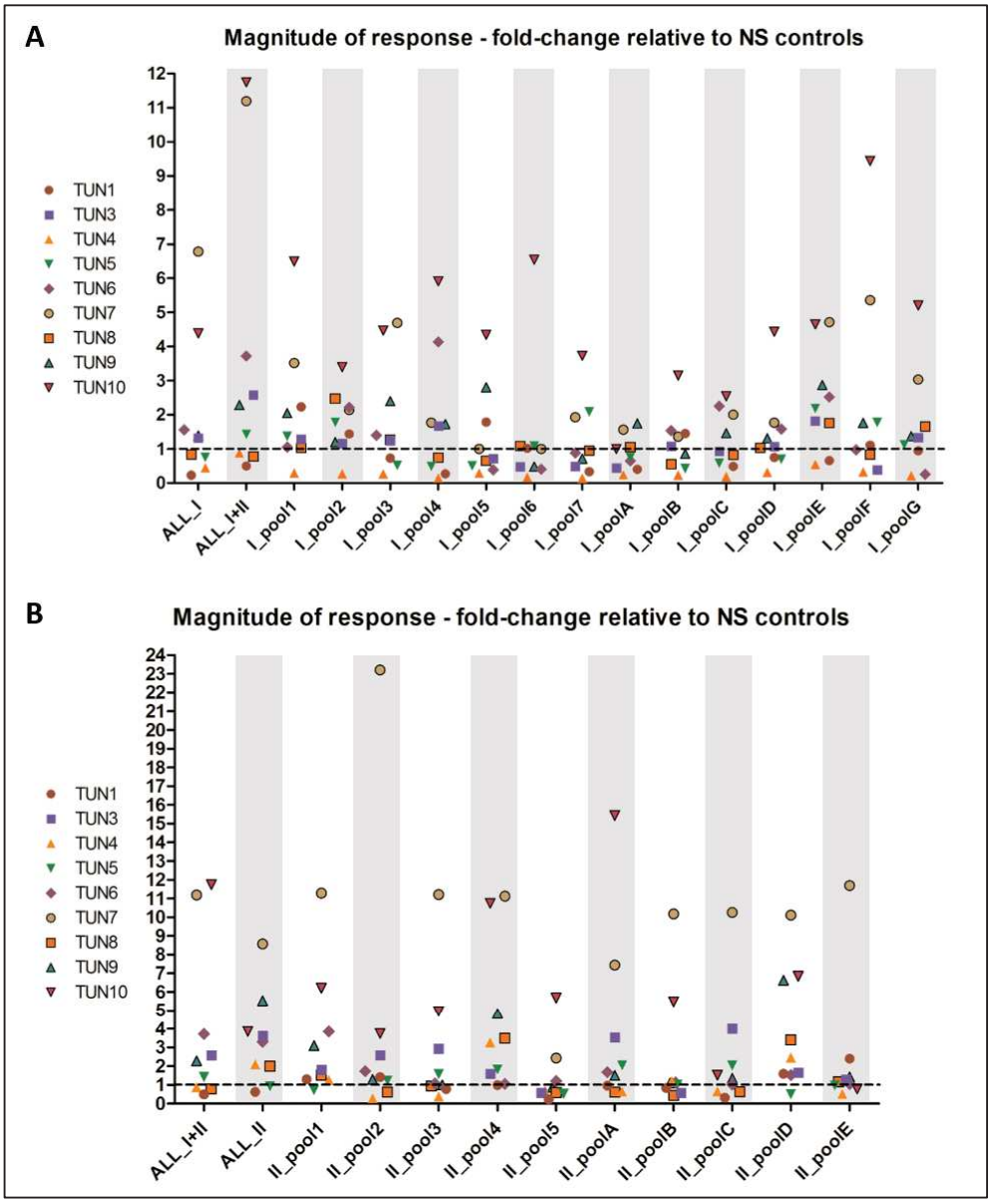


Figure V. 24 Magnitude of response against matrix-based peptide pools for all healed donors. Fold-change differences in IFN-γ production after peptide stimulation. *A) Pools containing HLA-class I peptides. B) Pools containing HLA-class II peptides.*

Cross-matching the matrix-based peptide pools with positive results for all donors, a total of 25 HLA-class I and 14 HLA-class II peptides were identified as double-positive immunogenic peptides (Figure V.25). Although donors expressed different HLA-DR and -DQ molecules, the immunoprevalent

HLA-class II-restricted peptide pii19 (**13_b12**) was immunogenic for 6 donors, becoming undoubtedly the most promising peptide candidate so far (Figure V.25 panel B). HLA-class II peptides pii4 (**22_b4**) and pii16 (**4_a11**) were immunogenic for 3 donors. Two HLA-class II peptides (pii14/**17_b15** and pii1/**14_b12**) and seven HLA-class I peptides (pi8/**B27_1**, pi12/**B58_4**, pi22/**A1_1**, pi29/**A24_1**, pi36/**A2_4**, pi37/**b4_B8**, pi43/**B58_5**) were immunogenic for two donors. The remaining nine double-positive HLA-class II peptides and eighteen HLA-class I peptides were immunogenic only for one donor (Figure V.25).

A

25 double-positive CLASS I peptides				
Matrix peptide nr	Proteincode_ supertype	Peptide name	#donors	Donor ID
pi8	b5_B27	B27_1	2	TUN1, TUN10
pi12	b3_B58	B58_4	2	TUN1, TUN10
pi22	b1_A1	A1_1	2	TUN7, TUN10
pi29	a6_A24	A24_1	2	TUN7, TUN9
pi36	b4_A2_1	A2_4	2	TUN7, TUN10
pi37	b4_B8	B8_4	2	TUN5, TUN10
pi43	b15_B58_B62	B58_5	2	TUN7, TUN10
pi35	b12_B44_1	B44_1	1	TUN5
pi42	b12_B44_2	B44_4	1	TUN5
pi24	b7_A3	A3_3	1	TUN7
pi30	a14_B62	B62_2	1	TUN5
pi31	a25_A3	A3_1	1	TUN7
pi38	a28_B7	B7_2	1	TUN7
pi45	b16_B8	B8_1	1	TUN7
pi9	b2_B7	B7_3	1	TUN10
pi13	b16_A24	A24_3	1	TUN10
pi23	a14_B8	B8_3	1	TUN10
pi26	a12_B44	B44_5	1	TUN10
pi27	a24_B58	B58_3	1	TUN10
pi40	b5_B7	B7_5	1	TUN10
pi41	a14_B27	B27_3	1	TUN10
pi44	a32_A3	A3_2	1	TUN10
pi47	a30_A1	A1_4	1	TUN10
pi48	a32_B7	B7_4	1	TUN10
pi33	a25_B27	B27_5	1	TUN9

B

14 double-positive CLASS II peptides			
Matrix peptide nr	Peptide name (peptide number_protein)	#donors	Donor ID
pii19	13_b12	6	TUN3, TUN4, TUN7, TUN8, TUN9, TUN10
pii4	22_b4	3	TUN3, TUN5, TUN10
pii16	4_a11	3	TUN3, TUN7, TUN10
pii14	17_b15	2	TUN3, TUN5, TUN7
pii1	14_b12	2	TUN3
pii2	16_b13	1	TUN3
pii3	3_a11	1	TUN3
pii5	18_b15	1	TUN10
pii11	7_a14	1	TUN3
pii12	12_a32	1	TUN3
pii13	21_b4	1	TUN3
pii17	5_a12	1	TUN3
pii18	9_a25	1	TUN3
pii20	10_a25	1	TUN10

C

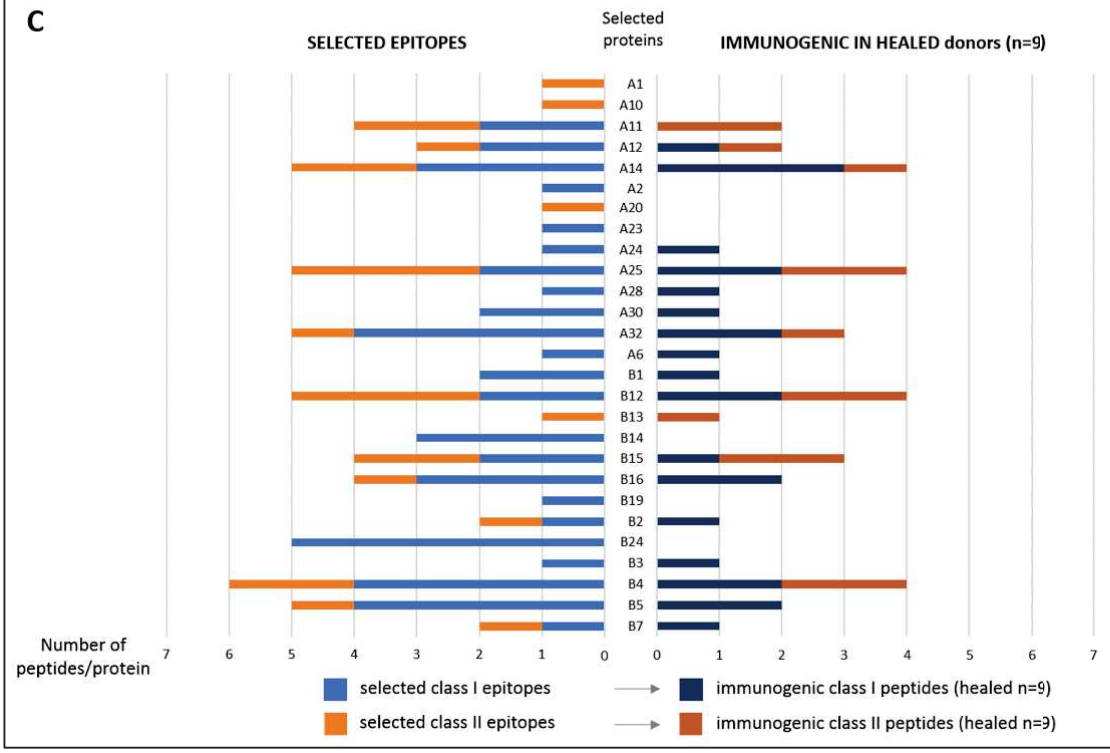


Figure V. 25 Double-positive immunogenic peptides (peptides present in two positive matrix-based pools). *A) 25 class I immunogenic peptides identified in 6 healed donors. B) 14 class II immunogenic peptides identified in 7 healed donors. C) Selected class I and class II peptides per protein antigen and peptides found to be immunogenic in samples from healed donors.*

Only a fraction of observed responses is specific to HLA-class I-restricted epitopes, as most of the reactions were induced by peptides not necessarily predicted to bind to the donors' HLA alleles. From the seven donors with positive responses to HLA-class I pools, only for TUN3 (1 in 4 peptides), TUN7 (3 in 8 peptides), and for TUN10 (5 in 16 peptides) correspond to the donor's HLA-typing. The other immunogenic HLA-class I peptides do not match the donors' allele specificity, and yet are not expected to be promiscuous. This observation agrees with the presence of unpredicted epitope promiscuity, the ability of peptides to bind to two or more different HLA alleles.

The analysis of cytokine production in response to *Leishmania*-specific peptides will help not only confirm, but also better characterise the immune responses and cellular functional activity induced by the peptide pools. Cytokine analysis of culture supernatants will be performed by Cytokine-Bead assays (CBA) for healed donors and the naive control donors, for which no results were presented here.

7. Discussion - Assays with Healed donor samples

Healed LCL donors, patients who developed active CL and either self-cured or received chemotherapy, generally possess strong and specific immune responses to *Leishmania* parasites and resistance to reinfection or relapse. In Tunisia, there are endemic areas for CL-causing *L. major* and *L. tropica* transmission. Historically, the Gafsa region was always an endemic area for CL, with typically sporadic and occasionally epidemic transmission (32). The Gafsa region in Tunisia is estimated to include a population of almost 324000 people, of which almost 147000 are at risk of developing leishmaniasis (33). Although parasitological confirmation was not performed for all donors, we expect *L. major* to be the species responsible for the primary infection.

The ELISpot results for the first series of healed donors are available (n=9) and presented here. The conclusion of the CBA analysis of these samples will allow the definite confirmation of the ELISpot responses and identified immunogenic epitopes, and more importantly, characterise in more detail the induced immune response by including a complete Th1/Th2/Th17 cytokine panel. The assessment of the production of the immunoregulatory cytokine IL-10 will be particularly informative. It has been described that Tunisian LST-positive individuals produce IL-10, also, IL-10 mRNA is detected in CL lesions caused by *L. major*. These observations further support the analysis of balanced Th1/Th2 responses in individuals with protective immune responses against *Leishmania*.

The second series of samples yielded no ELISpot results (data not shown), with few and faint spots in the control wells indicating very low cell viability before seeding in the ELISpot plates, although ELISpot revelation was successful after 20-minute incubation with substrate solution. These results are probably due to the change, at culture day 5, from commercial human AB serum to homemade serum. This change probably caused cell death, and cells were mostly unviable when seeded in the ELISpot plate. Still, the CBA analysis on the culture supernatants will still be performed to confirm this hypothesis.

The first immunoscreening assays performed with healed donor samples successfully identified IFN-γ responses specific to *Leishmania* peptides and have led to the identification of 25 HLA-class I and 14 HLA-class II immunogenic peptides. Overall, HLA-class II peptides appear to be more immunogenic as they induced higher IFN-γ production and in more individuals. An 'ELISA' effect in the ELISpot plates is observed at times – a dark background and high well saturation due to leftover IFN-γ from the culture wells. Yet, the AID reader successfully read and count all plates (Appendix V.2).

Four donors had significantly higher IFN-γ production after TSLA stimulation (TUN3, TUN4, TUN5, TUN10). Some individuals (n=5) did not have significantly increased responses after TSLA stimulation by ELISpot. However, for two of these donors (TUN9, TUN10), although spot count may be low or innumerable, an 'ELISA' effect is observed, where secreted IFN-γ diffuses into the supernatant and is absorbed on the ELISpot plate membrane producing a color carpet. This effect was observed probably because cells were not washed before seeding the ELISpot plate (plate centrifuge unavailable in the laboratory), a simple media change was performed. These individuals were considered as TSLA-positive since there is evidence of IFN-γ production during culture stimulation. Donor TUN1, who appears to not respond to TSLA stimulation, however, responds strongly to CaniLeish®. Curiously, the four TSLA-positive donors received different treatment regimens (respectively, cryotherapy plus antibiotherapy, arab traditional treatment, 15-day hospitalization with administration of intramuscular Glucantime® 60mg/Kg, or cryotherapy only). The time since diagnosis also differed greatly between donors (6 donors with date of parasitological confirmation), ranging from 1 to 12 years, and of 10 years for TUN4 and 2 years for TUN5. There appears to be no correlation with type of treatment or healing time and development of TSLA-specific immune responses in this healed donor sample series.

Two individuals did not respond to any HLA-class I peptide pool, and although there are peptides specific to all possible alleles present in the tested individuals. This may due to reduced peptide stability which prevents peptide processing or presentation and T cell activation, as class I peptides were solubilized in water. Peptide stability assays will be performed with pre-selected peptides.

Only one peptide concentration was tested for total PBMC stimulation in these experiments. A possible inhibitory effect on induced IFN-γ responses is observed when peptide pools contain 24 or 49 peptides (peptide pools ALL_I, ALL_II, and ALL_I+II). The 5 μM concentration per peptide in these pools may inhibit T cell responses or induce apoptosis or anergy. This effect may also be due to a high dilution of culture medium, since stock peptide solutions were at 200 mM, depriving the cells from essential nutrients. For this reason, new peptide stocks were prepared at 500 mM concentration.

To establish if the peptide pools significantly increased IFN-γ production in total PBMCs from healed donors, the Mann-Whitney non-parametric statistical test compared the triplicate STIM wells and the 6 autologous NS wells. A high variability in the ELISpot background is observed, as well as

some discrepancies in the ELISpot triplicates, possibly because red blood cell (RBC) lysis was not performed before seeding (reagents unavailable in the laboratory), and the number of contaminant RBC can affect successful T cell activation in a given well. Furthermore, dividing each culture well in two ELISpot seeding wells, may affect the minimum number of 100.000 PBMC per well resulting in failure to detect the specific T cells. Cell viability was also not assessed prior to ELISpot seeding, however, PHA responses were extremely high indicating the cells were viable and responsive.

The unstimulated ELISpot wells containing cells that were stimulated in the culture (cNS) have often an intermediary spot count value between the individuals' NS and the STIM wells. This may be because in a stimulated culture well there are cells in different activation stages, especially since culture is supplemented with rhIL-2, and some cells are still producing IFN-γ from the culture stimulation. In future experiments, cells should rested for at least 3 days after the last media change, washed and checked for viability before seeding in the ELISpot plate, and a minimum number of 200.000 PBMC seeded per well.

There does not seem to be a clear correlation between *in silico*-predicted HLA-class I restriction and individuals' HLA typing. There is increasing evidence of extensive promiscuity in HLA-class I antigen presentation which the *in silico* HLA-binding prediction algorithms do not account for (34,35). It is possible that *in silico* HLA-binding algorithms can predict immunogenic epitopes from antigen sequence data but perform less well in predicting HLA-restriction for promiscuous peptides. This argues for the selection of epitopes predicted to be promiscuous by the immunoinformatic tools available but, likewise, to expect unpredicted peptide HLA restriction.

There is, however, a good overlap regarding the most frequent alleles between the Tunisian population and the healed donors. Furthermore, if we consider the HLA coverage of the 25 immunogenic HLA-class I peptides, the most frequent alleles in the Tunisian and world population are covered (A*01, A*02, A*03, A*24, B*07, B*08, B*44), although only one peptide is restricted to A*02 alleles.

The cellular source of IFN-γ production was not assessed. $TCD4^+$ cells are the most common origin for IFN-γ production in response to TSLA, and probably to HLA-class II peptide pools, but this remains to be confirmed.

Finally, positive immune responses are observed against several *Leishmania*-specific peptide pools, as PBMC from healed donors produced IFN-γ in response to *in vitro* stimulation. Also, it is the first time IFN-γ responses from human PBMC after CaniLeish® stimulation are observed. This

preliminary observation needs further characterisation. As expected from the dog immune responses and observed here through *in vitro* human PBMC stimulation, the *Leishmania* promastigote secretome contains promising vaccine antigen candidates, which are also immunogenic in humans and associated with memory responses.

8. Conclusions

T cell activation *in vitro* occurs under non-physiological conditions, and factors influencing T cell activation should be controlled whenever possible, namely antigen-presentation, peptide dose, and stimulation conditions. Nevertheless, *in vitro* assays with human immune cells are still more conclusive on peptide immunogenicity than animal models. Cell-based immunoscreening assays to assess peptide immunogenicity *in vitro* are suitable and informative because they recreate the T cell responses induced after *in vivo* peptide administration, and allow the analysis of T cell populations involved. Furthermore, immunoscreening assays can be performed with cells from human donors of different immune status regarding leishmaniasis, allowing the evaluation of both correlations with protective responses and prophylactic potential of the peptide candidates. We expect to find peptides that are associated with protective T cell responses against *Leishmania* present in healed patients, as well as to find corresponding peptide-specific T cell precursors in the naive repertoire.

Although immunoscreenings in the naive repertoire were mostly optimisation experiments and only a few individuals were tested, preliminary evidence exists for peptides which are immunogenic in both immune backgrounds. Notably, the four double positive peptides identified in experiment MN01 from pool stimulations with more than one positive cell line (B7_3, A24_3, B8_3 and B58_3) were immunogenic in one healed donor sample (TUN10). The peptide B7_3 is also included in the poolGOOD (experiment MN04) which induced specific IFN-γ production by T cells from the naive repertoire (Table V.9).

Other peptides included in the poolGOOD induced both one positive naive T cell line (experiment MN01) and higher IFN-γ responses in healed donors: peptides A3_1, A3_3, B7_2 were immunogenic in donor TUN7, and the peptide A3_2 was immunogenic in donor TUN10. The peptide A2_1, which induced one positive naive T cell line (experiment MN01) was included in the poolGOOD (Table V.9).

Some of the peptides which induced one positive cell line in experiment MN01 were also immunogenic in healed donor samples. Namely, the peptide B8_1 was immunogenic in donor TUN7, also the peptides A3_2 and B7_4 were immunogenic for donor TUN10. The double-positive peptide B8_4 identified in experiment MN01 (one positive cell line) induced significant IFN-γ responses in TSLA-positive donors TUN5 and TUN10. Similarly, peptide B62_2 which induced one positive cell line in experiment MN01 was also identified in TSLA-positive donor TUN5 (Table V.9).

Some double-positive peptides inducing one positive T cell line in experiment MN01 did not induce significantly higher responses in healed donors (peptides A2_1, A2_2, A2_3, A2_5, B27_3, A26_1). Still, peptides B27_2, A2_1, and A2_2 were included in the poolGOOD that induced positive IFN-γ responses in naive donor MPL10 (Table V.9).

Table V. 9 Double-positive immunogenic peptides identified in the naive and memory repertoires.

HLA-class I peptides	Positive results in the naive repertoire	Positive results in the memory repertoire
A24_3, B8_3, B58_3	MN01 (2 positive cell lines)	1 healed donor (TUN7 or TUN10)
B7_3	MN01 (2 positive cell lines) + MN04 (poolGOOD)	1 healed donor (TUN10)
A3_2	MN01 (1 positive cell line) + MN04 (poolGOOD)	1 healed donor (TUN10)
A3_1, A3_3, B7_2	MN04 (poolGOOD)	1 healed donor (TUN7 or TUN10)
B8_4	MN01 (1 positive cell line)	2 healed donors (TUN5 and TUN10)
B8_1, B7_4, and B62_2	MN01 (1 positive cell line)	1 healed donor (TUN7, TUN10 or TUN5, respectively)
B27_2, A2_1, and A2_2	MN01 (1 positive cell line) + poolGOOD	-
B27_1, B58_4, A1_1, A24_1, A2_4, B58_5	-	2 healed donors (TUN1, TUN5, TUN7, TUN9, or TUN10)

Assays were not yet performed with TCD4+ cells from the naive repertoire, so no conclusion can be made on corresponding patterns for HLA-class II peptides. However, more healed individuals responded to the HLA-class II peptides, and 14 immunogenic HLA-class II peptides were identified. Some immunoprevalent peptides were identified which induced higher IFN-γ responses in over half the individuals tested (peptide 13_b12), and in three donors (22_b4 and 4_a11). Remarkably, two of these peptides come from new protein antigen candidates discovered through the reverse vaccinology approach. Due to intellectual property issues, peptide sequences and antigen names must remain confidential and cannot be divulged.

Overall, twelve HLA-class I 9-mer peptides (A24_3, A3_1, A3_2, A3_3, B58_3, B62_2, B7_2, B7_3, B7_4, B8_1, B8_3, B8_4) and five HLA-class II 15-mer peptides (13_b12; 22_b4; 17_b15; 4_a11; 14_b12) are, for now, the most promising candidates to advance for the multi-epitope peptide design.

We aim to re-test all the synthetic peptides in more immunoscreenings with samples from both naive and healed donors. Possibly, re-test the matrix-based pools and these pre-selected peptides, with some small protocol improvements. Also, we seek to include healed donors from more endemic areas, namely *L. braziliensis*, *L. infantum* and *L. donovani* endemic areas.

The present results provide evidence that the proposed peptide-based vaccine development pipeline, from the *in silico* epitope selection to the immunoscreenings exclusively in human samples, swiftly delivered some potentially interesting vaccine candidates.

Although still presenting some limitations, experimental validation with human samples from naive and healed individuals is conclusive regarding future *in vivo* immunogenicity of vaccine candidates, and it may accelerate clinical trial testing while preventing late-stage failure. This approach particularly benefits vaccine research against leishmaniases, since experience shows a development gap between promising vaccine candidates tested in animal challenge experiments and *de facto* candidate immunogenicity in humans living in endemic areas,

This evaluation model for preclinical development of peptide-based vaccines, especially in the context of NTDs, provides a powerful approach to fast-track the development and deployment of effective tools that will assist leishmaniasis control and improve Public Health programmes in affected areas.

9. Bibliography

1. Weber CA, Mehta PJ, Ardito M, Moise L, Martin B, De Groot AS. T cell epitope: Friend or Foe? Immunogenicity of biologics in context. Adv Drug Deliv Rev. 2009;61(11):965–76.
2. Jawa V, Cousens LP, Awwad M, Wakshull E, Kropshofer H, De Groot AS. T-cell dependent immunogenicity of protein therapeutics: Preclinical assessment and mitigation. Clin Immunol. 2013;149(3):534–55.
3. Currier JR, Kuta EG, Turk E, Earhart LB, Loomis-Price L, Janetzki S, et al. A panel of MHC class I restricted viral peptides for use as a quality control for vaccine trial ELISPOT assays. J Immunol Methods. Elsevier; 2002 Feb 1;260(1–2):157–72.
4. Finak G, Langweiler M, Jaimes M, Malek M, Taghiyar J, Korin Y, et al. Standardizing Flow Cytometry Immunophenotyping Analysis from the Human ImmunoPhenotyping Consortium. Nat Publ Gr. 2016;
5. Li Pira G, Ivaldi F, Moretti P, Manca F. High Throughput T Epitope Mapping and Vaccine Development. J Biomed Biotechnol. 2010;2010:1–12.
6. Li Pira G, Ivaldi F, Starc N, Landi F, Locatelli F, Rutella S, et al. Miniaturized and high-throughput assays for analysis of T-cell immunity specific for opportunistic pathogens and HIV. Clin Vaccine Immunol. American Society for Microbiology; 2014 Apr;21(4):488–95.
7. Wolfl M, Kuball J, Ho WY, Nguyen H, Manley TJ, Bleakley M, et al. Activation-induced expression of CD137 permits detection, isolation, and expansion of the full repertoire of CD8+ T cells responding to antigen without requiring knowledge of epitope specificities. Blood. 2007;110(1):201–10.
8. Delluc S, Ravot G, Maillere B. Quantitative analysis of the CD4 T-cell repertoire specific to therapeutic antibodies in healthy donors. FASEB J. 2011;25(6):2040–8.
9. Castelli FA, Leleu M, Pouvelle-Moratille S, Farci S, Zarour HM, Andrieu M, et al. Differential capacity of T cell priming in naive donors of promiscuous CD4+ T cell epitopes of HCV NS3 and Core proteins. Eur J Immunol. Wiley-Blackwell; 2007 Jun;37(6):1513–23.
10. Maillère B, Bats A-S, Castelli FA, Tartour E, Oudard S, Hans F, et al. The Tumor Antigen Cyclin B1 Hosts Multiple CD4 T Cell Epitopes Differently Recognized by Pre-Existing Naive and Memory Cells in Both Healthy and Cancer Donors. J Immunol. 2015;(195):1891–901.
11. Ho WY, Nguyen HN, Wolfl M, Kuball J, Greenberg PD. In vitro methods for generating CD8+ T-cell clones for immunotherapy from the naïve repertoire. J Immunol Methods. Elsevier; 2006 Mar 20;310(1–2):40–52.

12. Geiger R, Duhen T, Lanzavecchia A, Sallusto F. Human naive and memory CD4 + T cell repertoires specific for naturally processed antigens analyzed using libraries of amplified T cells. J Exp Med. 2009;206(7):1525–34.
13. Murphy K, Weaver C. Glossary. In: Janeway's Immunobiology. 9th ed. Garland Science; 2017. p. 818–54.
14. Obar JJ, Khanna KM, Lefrançois L. Endogenous Naive CD8+ T Cell Precursor Frequency Regulates Primary and Memory Responses to Infection. Immunity. Cell Press; 2008 Jun 13;28(6):859–69.
15. Fiore-Gartland A, Manso BA, Friedrich DP, Gabriel EE, Finak G, Moodie Z, et al. Pooled-Peptide Epitope Mapping Strategies Are Efficient and Highly Sensitive: An Evaluation of Methods for Identifying Human T Cell Epitope Specificities in Large-Scale HIV Vaccine Efficacy Trials. Lu S, editor. PLoS One. Public Library of Science; 2016 Feb 10;11(2):e0147812.
16. Wölfl M, Greenberg PD. Antigen-specific activation and cytokine-facilitated expansion of naive, human CD8+ T cells. Nat Protoc. 2014;9(4):950–66.
17. Castelli FA, Houitte D, Munier G, Szely N, Lecoq A, Briand JP, et al. Immunoprevalence of the CD4+ T-cell response to HIV Tat and Vpr proteins is provided by clustered and disperse epitopes, respectively. Eur J Immunol. 2008;38(10):2821–31.
18. González-Galarza FF, Takeshita LYC, Santos EJM, Kempson F, Maia MHT, Silva ALS da, et al. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. Nucleic Acids Res. Oxford University Press; 2015 Jan 28;43(D1):D784–8.
19. dos Santos EJM, McCabe A, Gonzalez-Galarza FF, Jones AR, Middleton D. Allele Frequencies Net Database: Improvements for storage of individual genotypes and analysis of existing data. Hum Immunol. 2016;77(3):238–48.
20. Sidney J, Peters B, Frahm N, Brander C, Sette A. HLA class I supertypes: a revised and updated classification. BMC Immunol. 2008 Jan;9:1.
21. Meunier S, Menier C, Marcon E, Lacroix-Desmazes S, Maillère B. CD4 T cells specific for factor VIII are present at high frequency in healthy donors and comprise naïve and memory cells. Blood Adv. American Society of Hematology; 2017 Sep 26;1(21):1842–7.
22. Hamze M, Meunier S, Karle A, Gdoura A, Goudet A, Szely N, et al. Characterization of CD4 T Cell Epitopes of Infliximab and Rituximab Identified from Healthy Donors. Front Immunol. 2017;8:500.
23. Besse MA, Levy Y, Schwartz O, Casartelli BN, Grygar C, Schneider A, et al. Hierarchy of CD4 T

Cell Epitopes of the ANRS Lipo5 Synthetic Vaccine Relies on the Frequencies of Pre-Existing Peptide-Specific T Cells in Healthy Donors. J Immunol. 2013;190:5757–63.

24. Maillere B. Comment on ' 'The Role of Naive T Cell Precursor Frequency and Recruitment in dictating Immune Response Magnitude" J Immunol Ref J Immunol Clin Immunol Eur J Immunol Cancer Immunol Immunother. 2013;190(56):2404–11.
25. Wullner D, Zhou L, Bramhall E, Kuck A, Goletz TJ, Swanson S, et al. Considerations for optimization and validation of an in vitro PBMC derived T cell assay for immunogenicity prediction of biotherapeutics. Clin Immunol. 2010;137:5–14.
26. López-Albaitero A, Mailliard R, Hackman T, Andrade Filho PA, Wang X, Gooding W, et al. Maturation pathways of dendritic cells determine TAP1 and TAP2 levels and cross-presenting function. J Immunother. NIH Public Access; 2009 Jun;32(5):465–73.
27. Dissanayake D, Murakami K, Tran MD, Elford AR, Millar DG, Ohashi PS. Peptide-Pulsed Dendritic Cells Have Superior Ability to Induce Immune-Mediated Tissue Destruction Compared to Peptide with Adjuvant. Labrecque N, editor. PLoS One. Public Library of Science; 2014 Mar 19;9(3):e92380.
28. Moser JM, Sassano ER, Leistritz DC, Eatrides JM, Phogat S, Koff W, et al. Optimization of a dendritic cell-based assay for the in vitro priming of naive human CD4+ T cells. J Immunol Methods. Elsevier B.V.; 2010;353(1–2):8–19.
29. Bullock TNJ, Colella TA, Engelhard VH. The Density of Peptides Displayed by Dendritic Cells Affects Immune Responses to Human Tyrosinase and gp100 in HLA-A2 Transgenic Mice. J Immunol. American Association of Immunologists; 2014 Jul 15;164(5):2354–61.
30. Pittet MJ, Valmori D, Dunbar PR, Speiser DE, Liénard D, Lejeune F, et al. High Frequencies of Naive Melan-a/Mart-1–Specific Cd8+ T Cells in a Large Proportion of Human Histocompatibility Leukocyte Antigen (Hla)-A2 Individuals. J Exp Med. Rockefeller University Press; 1999 Sep 6;190(5):705–16.
31. Dunbar PR, Smith CL, Chao D, Salio M, Shepherd D, Mirza F, et al. A Shift in the Phenotype of Melan-A-Specific CTL Identifies Melanoma Patients with an Active Tumor-Specific Immune Response. J Immunol. 2014;165(11):6644–52.
32. Aoun K, Bouratbine A. Cutaneous leishmaniasis in North Africa: a review. Parasite. 2014 Jan;21:14.
33. Chalghaf B, Chlif S, Mayala B, Ghawar W, Bettaieb J, Harrabi M, et al. Ecological Niche Modeling for the Prediction of the Geographic Distribution of Cutaneous Leishmaniasis in Tunisia. Am J Trop Med Hyg. 2016;94(4):844–51.

34. Frahm N, Yusim K, Suscovich TJ, Adams S, Sidney J, Hraber P, et al. Extensive HLA class I allele promiscuity among viral CTL epitopes. Eur J Immunol. John Wiley & Sons, Ltd; 2007 Sep 1;37(9):2419–33.
35. Axelsson-Robertson R, Weichold F, Sizemore D, Wulf M, Skeiky YAW, Sadoff J, et al. Extensive major histocompatibility complex class I binding promiscuity for Mycobacterium tuberculosis TB10.4 peptides and immune dominance of human leucocyte antigen (HLA)-B*0702 and HLA-B*0801 alleles in TB10.4 CD8+ T-cell responses. Immunology. John Wiley & Sons, Ltd (10.1111); 2010 Apr 1;129(4):496–505.

CHAPTER VI

GENERAL DISCUSSION AND FUTURE PERSPECTIVES

CHAPTER VI - GENERAL DISCUSSION AND FUTURE PERSPECTIVES

The development of new or more effective tools for leishmaniasis control is challenging and is often not a priority for standard product development pipelines, as the market does not allow the recovery of development costs, similarly to other NTDs. Furthermore, leishmaniasis is associated with poverty, which implies that costs are determinant for the feasibility of any tool or product (diagnosis, treatment, prevention). Also, when new products are available, these will likely need government and/or philanthropic subsidisation, and access in endemic areas must be facilitated.

Nevertheless, improving NTD control, including vaccine development against leishmaniasis, is part of the Millennium Development Goal 3, a United Nations initiative (target 3.3: by 2030, end the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases and combat hepatitis, water-borne diseases and other communicable diseases), and has received increasing focus and resources by pharmaceutical companies and other stake-holders.

Furthermore, peptide-based vaccines and therapeutics are a promising way of reducing production costs, rather than other expensive biologicals (purified recombinant proteins or antibodies). Additionally, peptide-based products do not need cold chain transport conditions, posing a great advantage particularly for low-resource settings.

1. A vaccine as the most promising tool for leishmaniasis control

The vaccine development pipeline proposed in this project helps mitigate some of the bottlenecks associated with product development for NTDs (Figure VI.1). Briefly, the exploitation of large proteomic datasets, comprising the most important pathogenic *Leishmania* species, allowed the selection of 52 relevant vaccine antigens from the *Leishmania* secretome, a known immunogenic antigen pool. Moreover, this screening approach was coupled to an HLA-based *in silico* epitope prediction and selection workflow leading to the synthesis of over 70 promising short peptide vaccine candidates. Then, these peptide candidates were tested exclusively in human samples, a key aspect that ensures peptide immunogenicity in human hosts and vaccine population coverage. Experimental validation with samples from both healed individuals from endemic areas and naive individuals is extremely important to effectively assess the vaccine candidate's feasibility (Figure VI.1).

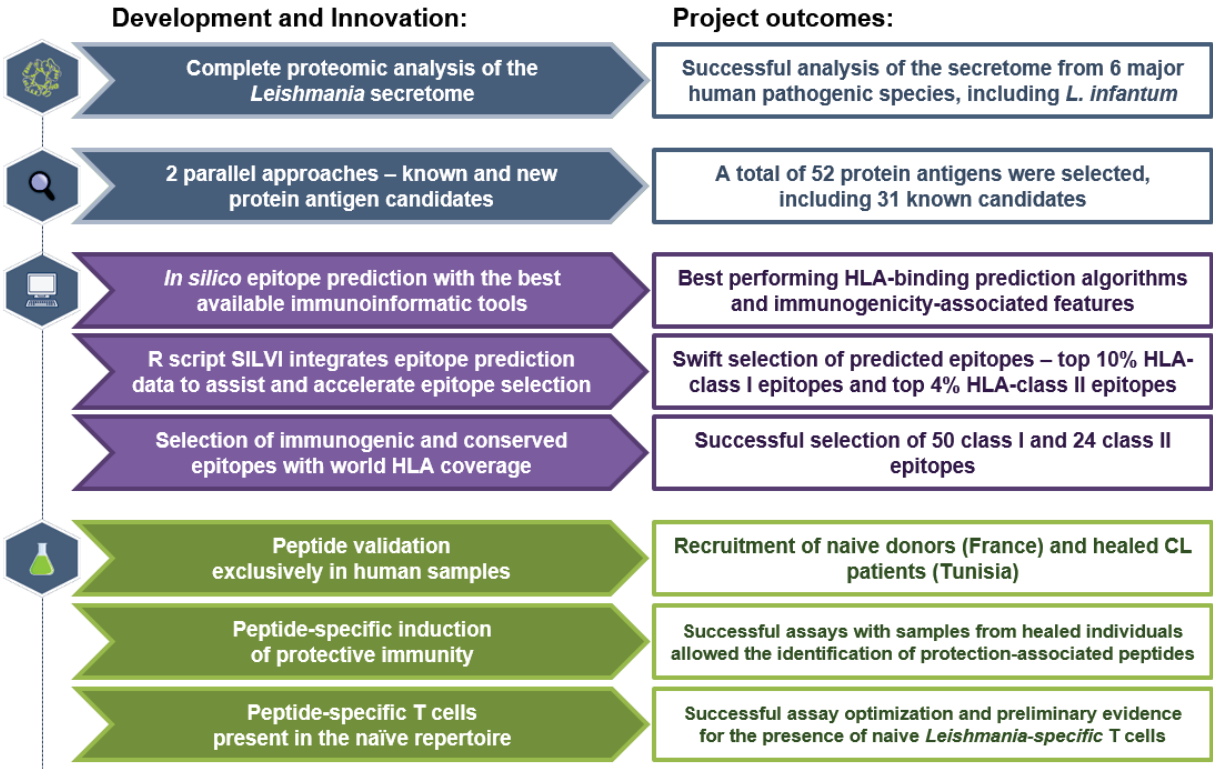


Figure VI. 1 Summary of major project outcomes.

Most studies assessing *Leishmania*-specific peptide immunogenicity performed so far (see Chapter I) compromise on either the number of antigens tested or the number of alleles considered, when HLA restriction is analysed. Our selected approach eliminates *a priori* assumptions, considering the full antigenic repertoire of the parasite's secretome and taking into account the vast HLA variability of the human host. The overall approach is decisive for the development of a pan-specific vaccine against multiple *Leishmania* species, and targeting multiple human populations.

In a more general view, vaccine development against *Leishmania* is like playing card without a complete deck. Information on correlates of protection is lacking, and many studies performed with animal models may skew our understanding of human-specific immunity mechanisms. The basic knowledge of *Leishmania*-specific immunity is quite limited, particularly if we compare with other protozoan parasites such as *Plasmodium*. Studies like the present that focus exclusively in human immunity and propose reproducible strategies for *Leishmania* vaccine development and peptide immunogenicity testing will greatly benefit the field.

Ultimately, the discovery of immunogenic epitopes in the *Leishmania* secretome associated with protective immune responses will fast-track the development of a peptide-based vaccine, which targets the main pathogenic *Leishmania* species and with worldwide HLA coverage (Figure VI.1).

2. Interest and limitations of using proteomics in a vaccine development pipeline

Proteomics is the most suitable tool for the analysis of *Leishmania* antigens as these parasites regulate gene expression mostly at the post-translational level, even if databases still need further improvement.

It is important to note that new protein annotations in the UniProtKB Database changed the proteomic analysis results due to a database update. The same raw data were used and the same bioinformatic analysis was performed, but the database was updated between June 2016 (analysis used for antigen selection, Chapter III) and March 2017 (analysis used for the proteomic analysis, Chapter II). The database annotation updates increased the number of species-specific proteins, as a higher number of unique proteins are identified. This affects protein identification and the number of common accessions among species. Briefly, if the Reverse Vaccinology approach for antigen discovery would be applied to the datasets generated with the new analysis (March 2017), only 326 accessions would be included, instead of 618. After BLASTp against the human proteome, only 33 common accessions are identified as non-homologous to human proteins. This corresponds to a loss of selected protein antigens (33 accessions/11 proteins) from the initial 24 protein antigens (76 accessions). Overall, 9 in 50 selected HLA-class I epitopes, and 7 in 24 selected HLA-class II epitopes are no longer identified. These peptides were already synthesised when the new analysis was performed, and they were included in subsequent experimental validation.

Nevertheless, database updates are frequent in -omics analysis, so these changes are fairly expected. In our case, the decrease in the total number of protein antigens analysed is actually undesirable for a comprehensive vaccine antigen selection. The core objective of the proposed peptide-based vaccine development pipeline was to screen the highest number possible of secreted proteins as antigen candidates, in order to maximise the chances of discovering immunogenic epitopes. Nevertheless, the epitope conservation filter included in the epitope selection step ensures that the peptides are identical among *Leishmania* target species even if the species-specific protein sequences are different. Finally, the database updates improved the molecular analysis of the secretome of the different *Leishmania* species (Chapter II), but did not interfere with the overarching goal of epitope-based vaccine development (Chapter III).

3. Lessons from other fields (cancer immunotherapy)

Importantly, although a prophylactic vaccine remains the gold-standard, a therapeutic vaccine (immunotherapeutic formulation) would also greatly advance leishmaniasis control. The proposed multiepitope peptides should also be tested for treatment, in combination and/or comparison with current chemotherapy. A peptide candidate that enhances treatment would be a major tool to help control leishmaniasis in endemic settings.

Currently, cancer immunotherapy is harnessing patients' immune system to recognize and kill cancer cells (1). Leishmaniasis chronic infection shares some immunoregulatory features with cancer, so it is possible to incorporate inhibitors of specific immune checkpoints into vaccine or immunotherapeutic formulations which can transiently reduce immunosuppression to allow the generation of robust vaccine-mediated anti-parasitic immunity. In this context, the CTLA-4/PD-1 (Cytotoxic T-lymphocyte-associated-protein 4/Programmed cell death-1) blockade could be explored to improve drug treatment and enhance vaccine efficacy against leishmaniasis. There is preliminary evidence that CTLA-4/PD-1 blockade promotes parasite clearance and treatment efficiency in animal models (2–4). Much like cancer blockade therapies, it is expected that specific types and combinations of immune checkpoint blockade will work differently for different types of leishmaniasis (VL/CL/MCL/PKDL/DCL). Further research into *Leishmania* immune responses and immunoregulation must be performed.

4. Considerations for the final peptide vaccine formulation

The peptide vaccine candidate against leishmaniases that will stem from this research will consist of synthetic peptides which include multiple *Leishmania*-specific immunogenic epitopes with worldwide HLA coverage. Evidently, the assembly design of the multi-epitope peptides will be carefully performed to maximise antigen uptake and presentation without affecting the induced immune response. Altogether, a successful vaccine must integrate the safety, efficacy and cost requirements to be deployed in the field.

Also, peptide dose and the administration regimen must be optimised, in order to ensure both large number of responding T cells and an appropriate quality of the induced response (5). Specific targeting of tissue-resident dendritic cells for the induction of protective responses and the presence of adjuvants to enhance memory generation should be included in vaccine design for improved *in situ* T cell memory generation (6). Dermal DC are the optimal target for the peptide vaccine formulations since they then migrate to the lymph nodes and provide the ideal antigen presentation to specific T cells inducing the desired immune response. Besides adequate adjuvantation, another important aspect to consider is the administration route, and how it affects

product dispersion in the dermis and the number of targeted dermal DC (7). The anatomic location for vaccine administration is essential for optimal efficacy (6). For example, the administration in the skin or nasal mucosa could boost the generation of T_{RM} cells *in vivo*. Early life vaccination is also preferential since memory T cells generated early in life have been shown to be maintained in peripheral blood, and there are fewer memory T cell clones to occupy pre-existent tissue niche environments where they were shown to persist until adulthood (6). Another key aspect of the successful assessment of vaccine immunogenicity and safety is the inclusion of sand fly challenges instead of intravenous infection in the design of clinical trials.

In this view, we believe the most well-suited final vaccine formulation against leishmaniases will include appropriate Th1-inducing and/or TLR agonist as adjuvant(s), and with optimised administration, possibly intradermally, with multiple or large injection sites, different prime-boost regimens or even novel administration devices, in order to maximise antigen uptake by DCs and ensure the ideal quality and quantity of the induced immune response.

5. Bibliography

1. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer. Nature Publishing Group; 2012 Apr 22;12(4):252–64.
2. Murphy ML, Cotterell SE, Gorak PM, Engwerda CR, Kaye PM. Blockade of CTLA-4 enhances host resistance to the intracellular pathogen, Leishmania donovani. J Immunol. 1998;161(8):4153–60.
3. Chiku VM, Silva KLO, de Almeida BFM, Venturin GL, Leal AAC, de Martini CC, et al. PD-1 function in apoptosis of T lymphocytes in canine visceral leishmaniasis. Immunobiology. Urban & Fischer; 2016 Aug 1;221(8):879–88.
4. Murray HW, Lu CM, Brooks EB, Fichtl RE, DeVecchio JL, Heinzel FP. Modulation of T-Cell Costimulation as Immunotherapy or Immunochemotherapy in Experimental Visceral Leishmaniasis. Infect Immun. 2003 Nov 1;71(11):6453–62.
5. Leggatt G. Peptide Dose and/or Structure in Vaccines as a Determinant of T Cell Responses. Vaccines. Multidisciplinary Digital Publishing Institute; 2014 Jul 2;2(3):537–48.
6. Farber DL, Yudanin NA, Restifo NP. Human memory T cells: generation, compartmentalization and homeostasis. Nat Rev Immunol. 2014 Jan;14(1):24–35.
7. Salerno EP, Shea SM, Olson WC, Petroni GR, Smolkin ME, McSkimming C, et al. Activation, dysfunction and retention of T cells in vaccine sites after injection of incomplete Freund's adjuvant, with or without peptide. Cancer Immunol Immunother. Springer-Verlag; 2013 Jul 9;62(7):1149–59.

Appendix II.1

Proteins identified in the secretome datasets ranked according to abundance (normalised iBAQ values) - 582 protein accessions with normalised iBAQ values over 1,10.

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
Q9GP00	**LEADING_SPECIES - DO NOT CHANGE!**	**Nucleoside diphosphate kinase**	**2552100000**	**9,4069**	**1,5186**
D1GJ50	LEADING_AUTO - DO NOT CHANGE!	Soluble promastigote surface antigen PSA-38S	1908300000	9,2806	1,4982
A4I8U5	Multi SPECIES	Tubulin beta chain	1456100000	9,1632	1,4792
A4HTR1	Multi SPECIES	Tubulin beta chain	1456100000	9,1632	1,4792
A4H727	LEADING_SPECIES - DO NOT CHANGE!	Tubulin alpha chain	1090500000	9,0376	1,4590
Q4QFP8	LEADING_AUTO - DO NOT CHANGE!	Putative small myristoylated protein-3	664170000	8,8223	1,4242
P69201	LEADING_AUTO - DO NOT CHANGE!	Ubiquitin-60S ribosomal protein L40	604610000	8,7815	1,4176
A4H3U0	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	582400000	8,7652	1,4150
Q25307	LEADING_AUTO - DO NOT CHANGE!	Infective insect stage-specific protein	579490000	8,7630	1,4146
Q07DU4	LEADING_AUTO - DO NOT CHANGE!	Peroxidoxin 2	578140000	8,7620	1,4145
D1GJ46	LEADING_AUTO - DO NOT CHANGE!	Soluble promastigote surface antigen PSA-34S (Fragment)	532290000	8,7261	1,4087
A4HPQ9	Multi SPECIES	Adenosylhomocysteinase	485670000	8,6863	1,4023
A4HPQ8	Multi SPECIES	Adenosylhomocysteinase	485670000	8,6863	1,4023
Q4Q7Y4	LEADING_AUTO - DO NOT CHANGE!	Putative heat-shock protein hsp70	482920000	8,6839	1,4019
E9ARD0	LEADING_AUTO - DO NOT CHANGE!	Elongation factor 1-alpha	481690000	8,6828	1,4017
O02614	LEADING_AUTO - DO NOT CHANGE!	Peptidyl-prolyl cis-trans isomerase	458460000	8,6613	1,3982
A4H3V1	LEADING_AUTO - DO NOT CHANGE!	Putative beta-fructofuranosidase	452080000	8,6552	1,3972
A4I058	LEADING_SPECIES - DO NOT CHANGE!	Putative endoribonuclease L-PSP (Pb5)	427180000	8,6306	1,3933
O43990	LEADING_SPECIES - DO NOT CHANGE!	Infective insect stage-specific protein	393100000	8,5945	1,3874
P15706	LEADING_AUTO - DO NOT CHANGE!	Leishmanolysin	366080000	8,5636	1,3824
A4I7Z7	LEADING_SPECIES - DO NOT CHANGE!	Superoxide dismutase	346670000	8,5399	1,3786
A4HYX4	LEADING_AUTO - DO NOT CHANGE!	Putative small myristoylated protein-1	345990000	8,5391	1,3785
Q4QJ67	LEADING_AUTO - DO NOT CHANGE!	Peptidyl-prolyl cis-trans isomerase	329800000	8,5183	1,3751
A4HW98	LEADING_SPECIES - DO NOT CHANGE!	Histone H4	327030000	8,5146	1,3745
Q27673	LEADING_AUTO - DO NOT CHANGE!	Leishmanolysin	318890000	8,5036	1,3728
Q4QDL6	Multi SPECIES	Histone H2B	307890000	8,4884	1,3703
Q9BJC7	Multi SPECIES	Histone H2B	307890000	8,4884	1,3703
Q4QHP1	Multi SPECIES	Histone H2B	307890000	8,4884	1,3703

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
A4HKT8	Multi SPECIES	Nucleoside diphosphate kinase	305540000	8,4851	1,3698
A4HKT7	Multi SPECIES	Nucleoside diphosphate kinase	305540000	8,4851	1,3698
A4HW62	LEADING_SPECIES - DO NOT CHANGE!	Enolase	294720000	8,4694	1,3672
A4I1P9	LEADING_AUTO - DO NOT CHANGE!	Putative 2,4-dihydroxyhept-2-ene-1,7-dioic acid aldolase	252150000	8,4017	1,3563
A4HSP4	LEADING_SPECIES - DO NOT CHANGE!	Histone H4	250580000	8,3989	1,3559
A4I2J4	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	237210000	8,3751	1,3520
E9AHM9	LEADING_SPECIES - DO NOT CHANGE!	Heat shock protein 83-1	218720000	8,3399	1,3463
A4IDG6	LEADING_SPECIES - DO NOT CHANGE!	Inosine-guanosine transporter	216660000	8,3358	1,3457
Q8MNZ1	LEADING_AUTO - DO NOT CHANGE!	Leishmanolysin	215660000	8,3338	1,3453
Q25225	LEADING_AUTO - DO NOT CHANGE!	Probable eukaryotic initiation factor 4A	215380000	8,3332	1,3452
Q4JI42	LEADING_AUTO - DO NOT CHANGE!	Promastigote surface antigen	204160000	8,3100	1,3415
E9BU45	LEADING_AUTO - DO NOT CHANGE!	14-3-3 protein-like protein	197980000	8,2966	1,3393
A4HQG6	LEADING_AUTO - DO NOT CHANGE!	Putative histidine secretory acid phosphatase	196730000	8,2939	1,3389
Q4QEN5	LEADING_AUTO - DO NOT CHANGE!	Cytochrome c	185550000	8,2685	1,3348
A4ICW8	LEADING_SPECIES - DO NOT CHANGE!	Elongation factor 2	180130000	8,2556	1,3327
Q4Q412	LEADING_AUTO - DO NOT CHANGE!	Macrophage migration inhibitory factor-like protein	156670000	8,1950	1,3229
A4I1D2	LEADING_AUTO - DO NOT CHANGE!	Peptidyl-prolyl cis-trans isomerase	146240000	8,1651	1,3181
A4I076	LEADING_SPECIES - DO NOT CHANGE!	Putative NADP-dependent alcohol dehydrogenase	144480000	8,1598	1,3173
E9B8C5	LEADING_AUTO - DO NOT CHANGE!	Trypanothione reductase	139940000	8,1459	1,3150
A4ICA5	LEADING_SPECIES - DO NOT CHANGE!	Putative histidine secretory acid phosphatase	139460000	8,1444	1,3148
A4I0Y8	LEADING_SPECIES - DO NOT CHANGE!	Putative IgE-dependent histamine-releasing factor	134210000	8,1278	1,3121
A1Y2D3	Multi SPECIES	Prostaglandin f2-alpha synthase (Fragment)	133050000	8,1240	1,3115
A4I6Z4	Multi SPECIES	Prostaglandin f2-alpha synthase/D-arabinose dehydrogenase	133050000	8,1240	1,3115
Q5SDH5	LEADING_AUTO - DO NOT CHANGE!	Putative calpain-like cysteine peptidase	130000000	8,1139	1,3099
A4HYX1	LEADING_SPECIES - DO NOT CHANGE!	Putative calpain-like cysteine peptidase	128160000	8,1078	1,3089
A4HU13	LEADING_SPECIES - DO NOT CHANGE!	Putative calmodulin	126090000	8,1007	1,3077
Q4QF80	Multi SPECIES	Tryparedoxin peroxidase	124200000	8,0941	1,3067
Q4QF76	Multi SPECIES	Tryparedoxin peroxidase	124200000	8,0941	1,3067
Q4QF68	Multi SPECIES	Thiol specific antioxidant	124200000	8,0941	1,3067
E9AGG5	LEADING_SPECIES - DO NOT CHANGE!	Putative surface antigen protein 2	120010000	8,0792	1,3042
A4ZZ66	LEADING_AUTO - DO NOT CHANGE!	Cytoplasmic tryparedoxin peroxidase	116510000	8,0664	1,3022
Q6RYT3	LEADING_SPECIES - DO NOT CHANGE!	Tryparedoxin	111070000	8,0456	1,2988

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
Q4QAP8	LEADING_AUTO - DO NOT CHANGE!	Triosephosphate isomerase	100230000	8,0010	1,2916
P40285	LEADING_AUTO - DO NOT CHANGE!	Histone H3	98197000	7,9921	1,2902
A4HE26	LEADING_AUTO - DO NOT CHANGE!	Peptidyl-prolyl cis-trans isomerase	97628000	7,9896	1,2898
A4HJ63	LEADING_AUTO - DO NOT CHANGE!	Putative amino acid permease	97306000	7,9881	1,2895
A4H7P0	Multi SPECIES	Putative calpain-like cysteine peptidase	96683000	7,9854	1,2891
A4H7N9	Multi SPECIES	Putative calpain-like cysteine peptidase	96683000	7,9854	1,2891
A4HVE5	LEADING_AUTO - DO NOT CHANGE!	Putative carboxypeptidase	96330000	7,9838	1,2888
Q4QGI0	LEADING_AUTO - DO NOT CHANGE!	Putative surface antigen protein	95339000	7,9793	1,2881
P48499	LEADING_AUTO - DO NOT CHANGE!	Triosephosphate isomerase	90700000	7,9576	1,2846
A4H6I8	LEADING_AUTO - DO NOT CHANGE!	Putative aminopeptidase	90500000	7,9566	1,2845
A4HS26	LEADING_AUTO - DO NOT CHANGE!	Putative beta-fructofuranosidase	89468000	7,9517	1,2837
A4I784	LEADING_SPECIES - DO NOT CHANGE!	Putative ADP-ribosylation factor	88068000	7,9448	1,2825
A4I8F6	LEADING_SPECIES - DO NOT CHANGE!	Dihydrolipoyl dehydrogenase	87918000	7,9441	1,2824
J9XRQ3	LEADING_SPECIES - DO NOT CHANGE!	Activated protein kinase C receptor	86331000	7,9362	1,2812
A4I4D6	LEADING_SPECIES - DO NOT CHANGE!	Cysteine peptidase C (CPC)	80992000	7,9084	1,2767
A4I2G1	LEADING_SPECIES - DO NOT CHANGE!	Putative 60S ribosomal protein L35	80572000	7,9062	1,2763
Q9U5N6	LEADING_AUTO - DO NOT CHANGE!	Fructose-bisphosphate aldolase	79363000	7,8996	1,2753
Q4Q740	LEADING_AUTO - DO NOT CHANGE!	Putative heat shock 70-related protein 1,mitochondrial	78529000	7,8950	1,2745
A4HWX3	LEADING_SPECIES - DO NOT CHANGE!	Transaldolase	77296000	7,8882	1,2734
A4IA34	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	76200000	7,8820	1,2724
Q4QDM4	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	75509000	7,8780	1,2718
A4IAQ1	LEADING_SPECIES - DO NOT CHANGE!	Pyruvate kinase	73487000	7,8662	1,2699
A4ICP0	Multi SPECIES	Putative 40S ribosomal protein S18	73320000	7,8652	1,2697
A4ICP1	Multi SPECIES	Putative 40S ribosomal protein S18	73320000	7,8652	1,2697
E9B376	LEADING_AUTO - DO NOT CHANGE!	Nucleoside diphosphate kinase	73229000	7,8647	1,2696
E9AN57	LEADING_AUTO - DO NOT CHANGE!	GP63, leishmanolysin	72288000	7,8591	1,2687
E9ALV8	LEADING_AUTO - DO NOT CHANGE!	Putative tryparedoxin	71414000	7,8538	1,2679
E9BCF2	LEADING_AUTO - DO NOT CHANGE!	Tryparedoxin peroxidase	70641000	7,8491	1,2671
A4H3T9	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	69862000	7,8442	1,2663
A4HTI0	LEADING_SPECIES - DO NOT CHANGE!	Superoxide dismutase	69739000	7,8435	1,2662
Q5UDS8	Multi SPECIES	Tryparedoxin peroxidase 2	69656000	7,8430	1,2661

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
A0A088RNK9	Multi SPECIES	Tryparedoxin peroxidase	69656000	7,8430	1,2661
Q71S90	LEADING_AUTO - DO NOT CHANGE!	Superoxide dismutase	68715000	7,8371	1,2652
E9AH84	LEADING_SPECIES - DO NOT CHANGE!	Inhibitor of cysteine peptidase	67939000	7,8321	1,2644
E9BPK5	LEADING_AUTO - DO NOT CHANGE!	Carboxypeptidase, putative	65230000	7,8144	1,2615
D1GJ51	LEADING_AUTO - DO NOT CHANGE!	Putative surface antigen protein 2	65151000	7,8139	1,2614
A4HZI9	LEADING_SPECIES - DO NOT CHANGE!	Proteasome subunit alpha type	65126000	7,8138	1,2614
A4HIL9	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	64451000	7,8092	1,2607
A4I5W4	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	62531000	7,7961	1,2585
A4HMB3	LEADING_AUTO - DO NOT CHANGE!	Putative aspartate aminotransferase	61566000	7,7893	1,2574
A4HUJ7	LEADING_SPECIES - DO NOT CHANGE!	Putative nuclear transport factor 2	58882000	7,7700	1,2543
A4HHC8	LEADING_AUTO - DO NOT CHANGE!	TXN1 protein	58595000	7,7679	1,2540
E9B701	LEADING_AUTO - DO NOT CHANGE!	Putative glycine cleavage system H protein	58571000	7,7677	1,2540
A4ID05	LEADING_SPECIES - DO NOT CHANGE!	Adenosylhomocysteinase	58281000	7,7655	1,2536
Q4QC13	LEADING_AUTO - DO NOT CHANGE!	Proteasome subunit alpha type	58068000	7,7639	1,2533
A4HVP9	LEADING_SPECIES - DO NOT CHANGE!	Nucleobase transporter	57780000	7,7618	1,2530
A4I0C0	LEADING_SPECIES - DO NOT CHANGE!	Putative 3-ketoacyl-CoA thiolase	56788000	7,7543	1,2518
A4IDG8	LEADING_SPECIES - DO NOT CHANGE!	Phosphomannomutase	56592000	7,7528	1,2515
A4I421	LEADING_SPECIES - DO NOT CHANGE!	Malate dehydrogenase	56273000	7,7503	1,2511
Q95U89	LEADING_AUTO - DO NOT CHANGE!	Peroxidoxin	55334000	7,7430	1,2500
A4I8S7	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	54538000	7,7367	1,2490
A4HUU6	LEADING_SPECIES - DO NOT CHANGE!	Putative 14-3-3 protein	54518000	7,7365	1,2489
E9ACG7	LEADING_AUTO - DO NOT CHANGE!	Putative delta-1-pyrroline-5-carboxylate dehydrogenase	54478000	7,7362	1,2489
C6KJE0	LEADING_SPECIES - DO NOT CHANGE!	Actin	53651000	7,7296	1,2478
A4I218	LEADING_SPECIES - DO NOT CHANGE!	Putative 40S ribosomal protein S16	52927000	7,7237	1,2468
A4IB24	LEADING_SPECIES - DO NOT CHANGE!	Putative ubiquitin-conjugating enzyme e2	52202000	7,7177	1,2459
A4HUX3	LEADING_AUTO - DO NOT CHANGE!	Putative aminopeptidase	50612000	7,7043	1,2437
Q4Q7X1	LEADING_AUTO - DO NOT CHANGE!	Glutamate dehydrogenase	49710000	7,6964	1,2425
E9B8I6	LEADING_AUTO - DO NOT CHANGE!	Dipeptidyl-peptidase III, putative	49125000	7,6913	1,2416
Q4Q1R5	LEADING_AUTO - DO NOT CHANGE!	Proteasome subunit alpha type	48011000	7,6813	1,2400
A4I7K4	LEADING_SPECIES - DO NOT CHANGE!	Putative ATP-dependent RNA helicase	47890000	7,6802	1,2398
A4ID39	LEADING_SPECIES - DO NOT CHANGE!	Cysteine synthase	47706000	7,6786	1,2396

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
E9ACZ6	LEADING_AUTO - DO NOT CHANGE!	Proteasome endopeptidase complex	47523000	7,6769	1,2393
A4HSK5	LEADING_SPECIES - DO NOT CHANGE!	S-methyl-5'-thioadenosine phosphorylase	47140000	7,6734	1,2387
A4H399	LEADING_AUTO - DO NOT CHANGE!	Histone H4	46531000	7,6677	1,2378
A4IA55	LEADING_SPECIES - DO NOT CHANGE!	Putative 40S ribosomal protein S19 protein	44623000	7,6496	1,2349
A4H3U1	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	44275000	7,6462	1,2343
Q4QJ78	LEADING_AUTO - DO NOT CHANGE!	Histone H4	44221000	7,6456	1,2342
E9BQW3	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	44193000	7,6454	1,2342
E9AWJ0	LEADING_AUTO - DO NOT CHANGE!	Putative uncharacterized protein	44191000	7,6453	1,2342
A4H638	LEADING_AUTO - DO NOT CHANGE!	GP63, leishmanolysin	43860000	7,6421	1,2337
A4IAX3	LEADING_SPECIES - DO NOT CHANGE!	Putative proteasome activator protein pa26	43473000	7,6382	1,2331
A4HMZ0	LEADING_SPECIES - DO NOT CHANGE!	Putative cystathione gamma lyase	43360000	7,6371	1,2329
Q4QEI9	LEADING_AUTO - DO NOT CHANGE!	Elongation factor 1-alpha	43095000	7,6344	1,2324
A4I0S4	LEADING_SPECIES - DO NOT CHANGE!	Triosephosphate isomerase	42952000	7,6330	1,2322
A4ICG5	LEADING_SPECIES - DO NOT CHANGE!	Putative membrane-bound acid phosphatase 2	41589000	7,6190	1,2299
A4HCN4	LEADING_AUTO - DO NOT CHANGE!	Putative endoribonuclease L-PSP (Pb5)	41346000	7,6164	1,2295
A4HUD9	Multi SPECIES	Isocitrate dehydrogenase [NADP]	40868000	7,6114	1,2287
A2CIA0	Multi SPECIES	Isocitrate dehydrogenase [NADP]	40868000	7,6114	1,2287
A4I120	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	40548000	7,6080	1,2282
A8I4U5	LEADING_AUTO - DO NOT CHANGE!	Cytosolic tryparedoxin	40238000	7,6046	1,2276
A4HXS6	LEADING_SPECIES - DO NOT CHANGE!	Aconitate hydratase	40189000	7,6041	1,2275
E9AGQ5	LEADING_AUTO - DO NOT CHANGE!	Putative heat shock protein	39876000	7,6007	1,2270
A4I4A3	LEADING_SPECIES - DO NOT CHANGE!	ADF/Cofilin	39746000	7,5993	1,2268
E9AGR8	LEADING_SPECIES - DO NOT CHANGE!	Putative aminopeptidase	39564000	7,5973	1,2264
A0A142BXV0	Multi SPECIES	Glucose-6-phosphate isomerase (Fragment)	39449000	7,5960	1,2262
A0A142BXU6	Multi SPECIES	Glucose-6-phosphate isomerase (Fragment)	39449000	7,5960	1,2262
Q3C165	Multi SPECIES	Glucose-6-phosphate isomerase	39449000	7,5960	1,2262
Q2PDE4	Multi SPECIES	Glucose-6-phosphate isomerase	39449000	7,5960	1,2262
Q2PDE2	Multi SPECIES	Glucose-6-phosphate isomerase	39449000	7,5960	1,2262
E9AE35	Multi SPECIES	Histone H2A	38536000	7,5859	1,2246
E9AE33	Multi SPECIES	Histone H2A	38536000	7,5859	1,2246
Q9N9V4	LEADING_AUTO - DO NOT CHANGE!	40S ribosomal protein S25	38350000	7,5838	1,2243
A4I650	Multi SPECIES	40S ribosomal protein S14	37676000	7,5761	1,2230
A4I3H3	Multi SPECIES	40S ribosomal protein S14	37676000	7,5761	1,2230

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
E9AF59	LEADING_AUTO - DO NOT CHANGE!	Putative aminopeptidase P	37283000	7,5715	1,2223
Q4QFX2	LEADING_AUTO - DO NOT CHANGE!	Putative inosine-guanine nucleoside hydrolase	37112000	7,5695	1,2220
E9AHQ2	LEADING_SPECIES - DO NOT CHANGE!	Succinyl-CoA:3-ketoacid-coenzyme A transferase	36958000	7,5677	1,2217
P42556	LEADING_AUTO - DO NOT CHANGE!	Pteridine reductase 1	36905000	7,5671	1,2216
E9AH20	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	36583000	7,5633	1,2210
A4HXU4	LEADING_AUTO - DO NOT CHANGE!	Probable citrate synthase, mitochondrial	36237000	7,5592	1,2203
A4HUK1	LEADING_SPECIES - DO NOT CHANGE!	Peptidylprolyl isomerase	35893000	7,5550	1,2196
E9AHH9	LEADING_SPECIES - DO NOT CHANGE!	Putative ribosomal protein S20	35723000	7,5529	1,2193
Q25332	Multi SPECIES	Secreted acid phosphatase 1 (SAP1)	35397000	7,5490	1,2186
Q25336	Multi SPECIES	Secreted acid phosphatase 2 (SAP2)	35397000	7,5490	1,2186
E9AU82	Multi SPECIES	Putative histidine secretory acid phosphatase	35397000	7,5490	1,2186
A4HW29	LEADING_SPECIES - DO NOT CHANGE!	Putative calpain-like cysteine peptidase	33911000	7,5303	1,2156
A4I212	LEADING_SPECIES - DO NOT CHANGE!	Glutathione peroxidase	33714000	7,5278	1,2152
E9AXG9	LEADING_AUTO - DO NOT CHANGE!	Peptidyl-prolyl cis-trans isomerase	33389000	7,5236	1,2146
A4HVS0	LEADING_SPECIES - DO NOT CHANGE!	Ubiquitin-conjugating enzyme-like protein	33161000	7,5206	1,2141
E9BRR9	LEADING_AUTO - DO NOT CHANGE!	ATP-dependent DEAD-box RNA helicase, putative	33156000	7,5206	1,2141
A4HC04	LEADING_SPECIES - DO NOT CHANGE!	Putative RNA helicase	33023000	7,5188	1,2138
Q9N856	Multi SPECIES	Uncharacterized protein	32890000	7,5171	1,2135
Q9N853	Multi SPECIES	Uncharacterized protein	32890000	7,5171	1,2135
Q9N852	Multi SPECIES	Uncharacterized protein	32890000	7,5171	1,2135
A4H7T6	LEADING_AUTO - DO NOT CHANGE!	ENOL protein	32754000	7,5153	1,2132
A4IB88	LEADING_SPECIES - DO NOT CHANGE!	Putative 60S ribosomal protein L5	32079000	7,5062	1,2117
Q95PT4	Multi SPECIES	Myo-inositol-1-phosphate synthase	31448000	7,4976	1,2104
E9APY4	Multi SPECIES	Myo-inositol-1-phosphate synthase	31448000	7,4976	1,2104
A4HUY6	Multi SPECIES	Putative 40S ribosomal protein S21	31231000	7,4946	1,2099
A4HUY7	Multi SPECIES	Putative 40S ribosomal protein S21	31231000	7,4946	1,2099
A4H7V6	LEADING_AUTO - DO NOT CHANGE!	INO1 protein	30521000	7,4846	1,2083
Q868G9	LEADING_AUTO - DO NOT CHANGE!	Inhibitor of cysteine peptidase	30102000	7,4786	1,2073
A0A0G3EHC9	LEADING_AUTO - DO NOT CHANGE!	Trypanothione reductase	29692000	7,4726	1,2063
E9BAX6	LEADING_SPECIES - DO NOT CHANGE!	Proteasome subunit beta type	29690000	7,4726	1,2063
P83851	LEADING_AUTO - DO NOT CHANGE!	Inosine-uridine preferring nucleoside hydrolase	29061000	7,4633	1,2048
A4I971	LEADING_SPECIES - DO NOT CHANGE!	Macrophage migration inhibitory factor-like protein	28576000	7,4560	1,2036
A4HT65	LEADING_SPECIES - DO NOT CHANGE!	Putative 3-hydroxyacyl-ACP dehydratase	28388000	7,4531	1,2032

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
E9AK13	Multi SPECIES	Putative beta-fructofuranosidase	28374000	7,4529	1,2031
A0A0D3RKQ2	Multi SPECIES	Secretory invertase	28374000	7,4529	1,2031
Q4QGJ9	LEADING_AUTO - DO NOT CHANGE!	Putative surface antigen protein 2	27751000	7,4433	1,2016
E9AGU0	LEADING_AUTO - DO NOT CHANGE!	Phosphoglycerate kinase	27664000	7,4419	1,2014
A4I307	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	27643000	7,4416	1,2013
A4I1I6	LEADING_SPECIES - DO NOT CHANGE!	GTP-binding nuclear protein	27585000	7,4407	1,2012
E9ADT5	LEADING_AUTO - DO NOT CHANGE!	Cysteine peptidase C (CPC)	27569000	7,4404	1,2011
Q3HL75	LEADING_AUTO - DO NOT CHANGE!	Enolase	27224000	7,4350	1,2002
Q25278	LEADING_AUTO - DO NOT CHANGE!	Probable 60S ribosomal protein L14	27119000	7,4333	1,2000
Q07DU5	LEADING_AUTO - DO NOT CHANGE!	Peroxidoxin 2	27078000	7,4326	1,1999
A4I9I3	LEADING_AUTO - DO NOT CHANGE!	Malate dehydrogenase	27077000	7,4326	1,1999
Q25298	LEADING_SPECIES - DO NOT CHANGE!	Kinetoplastid membrane protein 11C	26979000	7,4310	1,1996
Q4QI64	Multi SPECIES	Cathepsin L-like protease	26855000	7,4290	1,1993
K7PPA6	Multi SPECIES	Cysteine protease	26855000	7,4290	1,1993
K7PNP9	Multi SPECIES	Cysteine protease	26855000	7,4290	1,1993
K7PN53	Multi SPECIES	Cysteine protease	26855000	7,4290	1,1993
K7P5C3	Multi SPECIES	Cathepsin L-like protease	26855000	7,4290	1,1993
K7P522	Multi SPECIES	Cathepsin L-like protease	26855000	7,4290	1,1993
Q4QI61	Multi SPECIES	Cathepsin L-like protease	26855000	7,4290	1,1993
P90628	Multi SPECIES	Cathepsin L-like protease	26855000	7,4290	1,1993
Q4QBD1	LEADING_AUTO - DO NOT CHANGE!	Aldose 1-epimerase-like protein	26805000	7,4282	1,1992
E9AKP0	LEADING_AUTO - DO NOT CHANGE!	Peptidyl-prolyl cis-trans isomerase	26448000	7,4224	1,1982
O44012	LEADING_SPECIES - DO NOT CHANGE!	40S ribosomal protein S6	26281000	7,4196	1,1978
Q4QDZ7	LEADING_AUTO - DO NOT CHANGE!	Carboxypeptidase	25952000	7,4142	1,1969
Q4FX34	LEADING_AUTO - DO NOT CHANGE!	Aspartate aminotransferase	25771000	7,4111	1,1964
E9AHZ7	LEADING_SPECIES - DO NOT CHANGE!	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	25610000	7,4084	1,1960
E9ADX3	LEADING_AUTO - DO NOT CHANGE!	Tryparedoxin	25411000	7,4050	1,1954
E9BMG9	LEADING_AUTO - DO NOT CHANGE!	Cytoskeleton-associated protein CAP5.5, putative	25225000	7,4018	1,1949
Q4Q873	LEADING_AUTO - DO NOT CHANGE!	Putative zinc transporter	25212000	7,4016	1,1949
A4IDS4	LEADING_SPECIES - DO NOT CHANGE!	40S ribosomal protein SA	25134000	7,4003	1,1946
E9AGQ7	Multi SPECIES	Histone H2B	25050000	7,3988	1,1944
A4HU57	Multi SPECIES	Histone H2B	25050000	7,3988	1,1944
A4HY42	Multi SPECIES	Histone H2B	25050000	7,3988	1,1944

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
A0A088S198	LEADING_AUTO - DO NOT CHANGE!	Heat shock protein 83	24959000	7,3972	1,1941
A4H879	Multi SPECIES	Tryparedoxin peroxidase	24803000	7,3945	1,1937
A4H877	Multi SPECIES	TRYP1 protein	24803000	7,3945	1,1937
A4HWJ3	LEADING_SPECIES - DO NOT CHANGE!	Putative 40S ribosomal protein S3	24557000	7,3902	1,1930
A4HTZ8	LEADING_SPECIES - DO NOT CHANGE!	Oligopeptidase bwith=GeneDB:LmjF09.0770	24421000	7,3878	1,1926
A4H9T4	LEADING_SPECIES - DO NOT CHANGE!	Proteasome endopeptidase complex	24333000	7,3862	1,1924
Q6TDF7	LEADING_AUTO - DO NOT CHANGE!	70 kDa heat shock protein	23922000	7,3788	1,1912
A4GVE9	LEADING_SPECIES - DO NOT CHANGE!	Eukaryotic translation initiation factor 5A	23850000	7,3775	1,1910
A4I1G1	LEADING_SPECIES - DO NOT CHANGE!	ATP synthase subunit beta	23611000	7,3731	1,1903
A4HWS2	LEADING_SPECIES - DO NOT CHANGE!	Protein tyrosine phosphatase-like protein	23549000	7,3720	1,1901
E9AIM4	LEADING_AUTO - DO NOT CHANGE!	Contig, possible fusion of chromosomes 20 and 34	23089000	7,3634	1,1887
A4HSH2	LEADING_SPECIES - DO NOT CHANGE!	ATPase alpha subunit	22992000	7,3616	1,1884
Q4QDQ1	LEADING_AUTO - DO NOT CHANGE!	Putative pyruvate dehydrogenase E1 component alpha subunit	22768000	7,3573	1,1877
Q4QF35	LEADING_AUTO - DO NOT CHANGE!	Proliferating cell nuclear antigen	22599000	7,3541	1,1872
A4I115	LEADING_AUTO - DO NOT CHANGE!	Transketolase	22546000	7,3531	1,1870
E9ADX4	LEADING_AUTO - DO NOT CHANGE!	Tryparedoxin	22397000	7,3502	1,1866
A4HWV1	LEADING_SPECIES - DO NOT CHANGE!	Putative aspartate carbamoyltransferase	22299000	7,3483	1,1862
E9AHW0	Multi SPECIES	Putative 60S ribosomal protein L12	22190000	7,3462	1,1859
A4I130	Multi SPECIES	Putative 60S ribosomal protein L12	22190000	7,3462	1,1859
A4ICW2	LEADING_SPECIES - DO NOT CHANGE!	Eukaryotic translation initiation factor 3 subunit L	22141000	7,3452	1,1858
E9AFW0	LEADING_AUTO - DO NOT CHANGE!	Proteasome subunit alpha type	22054000	7,3435	1,1855
E9BBD0	LEADING_AUTO - DO NOT CHANGE!	Lectin, putative	21785000	7,3382	1,1846
E9AST7	LEADING_AUTO - DO NOT CHANGE!	Proteasome subunit beta type	21714000	7,3367	1,1844
A4HVQ0	Multi SPECIES	40S ribosomal protein S4	21692000	7,3363	1,1843
A4HVQ1	Multi SPECIES	Putative 40S ribosomal protein S4	21692000	7,3363	1,1843
A4ICY0	LEADING_SPECIES - DO NOT CHANGE!	Stress-inducible protein STI1 homolog	21384000	7,3301	1,1833
A4I7Q0	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	21333000	7,3291	1,1831
A4HZI8	LEADING_SPECIES - DO NOT CHANGE!	60S ribosomal protein L37a	21279000	7,3280	1,1830
P36400	Multi SPECIES	Cysteine proteinase B	21075000	7,3238	1,1823
Q05094	Multi SPECIES	Cysteine proteinase 2	21075000	7,3238	1,1823
A4HLC9	Multi SPECIES	Tubulin beta chain	20900000	7,3201	1,1817
A4HLC8	Multi SPECIES	Tubulin beta chain	20900000	7,3201	1,1817
A4HT77	Multi SPECIES	Putative 60S ribosomal protein L7a	20811000	7,3183	1,1814

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
A4HT78	Multi SPECIES	Putative 60S ribosomal protein L7a	20811000	7,3183	1,1814
A4I7N3	LEADING_SPECIES - DO NOT CHANGE!	Putative ras-related rab-4	20793000	7,3179	1,1813
E9BIG7	LEADING_AUTO - DO NOT CHANGE!	Thimet oligopeptidase, putative	20741000	7,3168	1,1812
E9AN53	LEADING_AUTO - DO NOT CHANGE!	GP63, leishmanolysin	20634000	7,3146	1,1808
P27891	LEADING_AUTO - DO NOT CHANGE!	Histone H2A.1	20575000	7,3133	1,1806
A4HLA1	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	20569000	7,3132	1,1806
A4H5S5	LEADING_AUTO - DO NOT CHANGE!	Elongation factor-1 gamma	20520000	7,3122	1,1804
E9AKI9	LEADING_AUTO - DO NOT CHANGE!	S-methyl-5'-thioadenosine phosphorylase	20503000	7,3118	1,1804
A4ID21	LEADING_SPECIES - DO NOT CHANGE!	Putative 40S ribosomal protein S27-1	20320000	7,3079	1,1797
A4I3W2	LEADING_SPECIES - DO NOT CHANGE!	Eukaryotic translation initiation factor 3 subunit E	20250000	7,3064	1,1795
A4HS64	LEADING_SPECIES - DO NOT CHANGE!	Putative ubiquitin-conjugating enzyme e2	20070000	7,3025	1,1789
A4I1K7	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	19840000	7,2975	1,1781
A4ID83	LEADING_SPECIES - DO NOT CHANGE!	Succinate-CoA ligase subunit beta	19824000	7,2972	1,1780
A4ID08	LEADING_AUTO - DO NOT CHANGE!	Eukaryotic translation initiation factor 3 subunit I	19788000	7,2964	1,1779
A4HY22	LEADING_AUTO - DO NOT CHANGE!	Plasma membrane ATPase	19619000	7,2927	1,1773
Q711P7	LEADING_AUTO - DO NOT CHANGE!	Putative trypanothione synthetase	19570000	7,2916	1,1771
A4IBY7	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	19434000	7,2886	1,1766
Q4Q0A9	LEADING_AUTO - DO NOT CHANGE!	Putative histidine secretory acid phosphatase	19316000	7,2859	1,1762
Q4QDF8	LEADING_AUTO - DO NOT CHANGE!	Putative histone H3 variant	19279000	7,2851	1,1760
Q8WT31	Multi SPECIES	Cysteine proteinase	19202000	7,2833	1,1758
A4HYH2	Multi SPECIES	Cysteine peptidase A (CPA)	19202000	7,2833	1,1758
A4HFQ7	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	19146000	7,2821	1,1756
A4HZS1	LEADING_SPECIES - DO NOT CHANGE!	Putative 40S ribosomal protein S15	19074000	7,2804	1,1753
A4I2Y7	LEADING_AUTO - DO NOT CHANGE!	Glycosomal phosphoenolpyruvate carboxykinase, putative	19028000	7,2794	1,1751
A4HYI1	LEADING_SPECIES - DO NOT CHANGE!	Peptidylprolyl isomerase	18939000	7,2774	1,1748
A4I1L9	LEADING_SPECIES - DO NOT CHANGE!	Putative pyruvate dehydrogenase E1 beta subunit	18842000	7,2751	1,1744
E9AG92	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	18819000	7,2746	1,1744
E9B6R2	LEADING_AUTO - DO NOT CHANGE!	Proteasome subunit beta type	18648000	7,2706	1,1737
A4HJ18	LEADING_AUTO - DO NOT CHANGE!	TDR1 protein	18510000	7,2674	1,1732
A4HV26	LEADING_SPECIES - DO NOT CHANGE!	Putative 40S ribosomal protein S15A	18472000	7,2665	1,1730
A4HP21	LEADING_SPECIES - DO NOT CHANGE!	Putative universal minicircle sequence binding protein	18180000	7,2596	1,1719

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
A4I841	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	17935000	7,2537	1,1710
A4ID74	LEADING_SPECIES - DO NOT CHANGE!	40S ribosomal protein S24	17933000	7,2537	1,1710
Q4Q8S3	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	17791000	7,2502	1,1704
A4IE56	LEADING_SPECIES - DO NOT CHANGE!	Putative oxidoreductase	17719000	7,2484	1,1701
P23223	LEADING_AUTO - DO NOT CHANGE!	Leishmanolysin	17636000	7,2464	1,1698
A4IBR8	LEADING_SPECIES - DO NOT CHANGE!	Putative 60S ribosomal protein L23	17432000	7,2413	1,1690
A4HSZ7	LEADING_SPECIES - DO NOT CHANGE!	Protein disulfide isomerase	17418000	7,2410	1,1689
Q4Q6Z5	Multi SPECIES	Glyceraldehyde-3-phosphate dehydrogenase	17309000	7,2383	1,1685
Q4Q6Z4	Multi SPECIES	Glyceraldehyde-3-phosphate dehydrogenase	17309000	7,2383	1,1685
E9ARK6	LEADING_AUTO - DO NOT CHANGE!	Citrate synthase	17114000	7,2334	1,1677
E9B3L2	LEADING_AUTO - DO NOT CHANGE!	Heat shock protein 83-1	17075000	7,2324	1,1675
E9B8M5	LEADING_SPECIES - DO NOT CHANGE!	Proteasome subunit beta type	16932000	7,2287	1,1669
A4I7P2	LEADING_SPECIES - DO NOT CHANGE!	Putative RNA binding protein	16759000	7,2242	1,1662
Q4QHH2	Multi SPECIES	GP63, leishmanolysin	16224000	7,2102	1,1640
Q4QHH1	Multi SPECIES	GP63, leishmanolysin	16224000	7,2102	1,1640
Q4QHG9	Multi SPECIES	GP63, leishmanolysin	16224000	7,2102	1,1640
B8YDG1	Multi SPECIES	GP63	16224000	7,2102	1,1640
Q9N9V3	Multi SPECIES	Putative ribosomal protein L10	16206000	7,2097	1,1639
A4HS71	Multi SPECIES	Putative 60S ribosomal protein L10	16206000	7,2097	1,1639
P62884	LEADING_SPECIES - DO NOT CHANGE!	Guanine nucleotide-binding protein subunit beta-like protein	16001000	7,2041	1,1630
A4I426	LEADING_SPECIES - DO NOT CHANGE!	Glutamate dehydrogenase	15985000	7,2037	1,1629
A4I2C4	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	15795000	7,1985	1,1621
P14700	LEADING_AUTO - DO NOT CHANGE!	Membrane antigen containing repeating peptides (Fragment)	15610000	7,1934	1,1612
A4HFK0	LEADING_SPECIES - DO NOT CHANGE!	Putative small GTP-binding protein Rab1	15572000	7,1923	1,1611
A4IAD2	LEADING_SPECIES - DO NOT CHANGE!	Putative 60S ribosomal protein L21	15564000	7,1921	1,1610
A4IB25	LEADING_SPECIES - DO NOT CHANGE!	Histone H4	15499000	7,1903	1,1607
Q5EXB3	LEADING_AUTO - DO NOT CHANGE!	Hemoglobin receptor	15466000	7,1894	1,1606
A4I935	LEADING_SPECIES - DO NOT CHANGE!	Peptidyl-prolyl cis-trans isomerase	15412000	7,1879	1,1604
E9BHI2	LEADING_AUTO - DO NOT CHANGE!	Protein phosphatase, putative	15357000	7,1863	1,1601
E9B099	LEADING_AUTO - DO NOT CHANGE!	Putative heat-shock protein hsp70	15261000	7,1836	1,1597
A4HGY1	LEADING_AUTO - DO NOT CHANGE!	Putative heat-shock protein hsp70	15244000	7,1831	1,1596
A4H9Q9	LEADING_SPECIES - DO NOT CHANGE!	Nonspecific nucleoside hydrolasewith=GeneDB:LmjF18.1580	15134000	7,1800	1,1591

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
A4I574	LEADING_SPECIES - DO NOT CHANGE!	Putative aspartyl-tRNA synthetase	15031000	7,1770	1,1586
A4IBL4	LEADING_SPECIES - DO NOT CHANGE!	Putative cystathione gamma lyase	15003000	7,1762	1,1585
A0A1E1J358	Multi SPECIES	ADP-ribosylation factor, putative	14981000	7,1755	1,1584
A0A088RXX3	Multi SPECIES	ADP-ribosylation factor, putative	14981000	7,1755	1,1584
A4IDB2	LEADING_SPECIES - DO NOT CHANGE!	Putative translation elongation factor 1-beta	14958000	7,1749	1,1583
A0A1E1IN58	Multi SPECIES	Putative small GTP binding protein rab6-like protein	14863000	7,1721	1,1578
A0A088RH69	Multi SPECIES	Small GTP binding protein rab6-like protein	14863000	7,1721	1,1578
A4I0C2	Multi SPECIES	Acetyl-coenzyme A synthetase	14842000	7,1715	1,1577
A4I093	Multi SPECIES	Acetyl-coenzyme A synthetase	14842000	7,1715	1,1577
E9AE96	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	14748000	7,1687	1,1573
A4HWZ0	LEADING_SPECIES - DO NOT CHANGE!	Sucrose-phosphate synthase-like protein	14734000	7,1683	1,1572
Q8T6M2	LEADING_AUTO - DO NOT CHANGE!	Guanosine permease	14696000	7,1672	1,1570
Q9U4E3	LEADING_AUTO - DO NOT CHANGE!	ADP-ribosylation factor-like 3A	14634000	7,1654	1,1567
A4IAU1	Multi SPECIES	40S ribosomal protein S3a-2	14547000	7,1628	1,1563
A4IAU0	Multi SPECIES	40S ribosomal protein S3a-1	14547000	7,1628	1,1563
Q4QJG7	LEADING_AUTO - DO NOT CHANGE!	Trypanothione reductase	14514000	7,1618	1,1561
E9ACW0	LEADING_AUTO - DO NOT CHANGE!	Putative heat shock protein DNAJ	14428000	7,1592	1,1557
A4IAS5	LEADING_SPECIES - DO NOT CHANGE!	60S ribosomal protein L30	14388000	7,1580	1,1555
A9LNR9	LEADING_AUTO - DO NOT CHANGE!	Tryparedoxin peroxidase	14334000	7,1564	1,1553
A4H8V4	LEADING_SPECIES - DO NOT CHANGE!	Elongation factor 1-alpha	14289000	7,1550	1,1550
A4ID12	LEADING_SPECIES - DO NOT CHANGE!	Putative glycyl tRNA synthetase	14283000	7,1548	1,1550
E9BTH9	LEADING_SPECIES - DO NOT CHANGE!	Short chain 3-hydroxyacyl-CoA dehydrogenase, putative	14209000	7,1526	1,1547
A4HS12	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	14148000	7,1507	1,1544
A4HKH0	Multi SPECIES	Superoxide dismutase	13808000	7,1401	1,1526
A4HKH1	Multi SPECIES	Superoxide dismutase	13808000	7,1401	1,1526
A4ICS4	Multi SPECIES	Putative ubiquitin/ribosomal protein S27a	13696000	7,1366	1,1521
A4HW73	Multi SPECIES	Putative ubiquitin/ribosomal protein S27a	13696000	7,1366	1,1521
E9AK40	LEADING_AUTO - DO NOT CHANGE!	Putative spermidine synthase	13618000	7,1341	1,1517
A4H6Z3	LEADING_SPECIES - DO NOT CHANGE!	Putative NADH:flavin oxidoreductase/NADH oxidase	13527000	7,1312	1,1512
E9AUR8	LEADING_AUTO - DO NOT CHANGE!	Putative calpain-like cysteine peptidase	13524000	7,1311	1,1512
E9ARW7	LEADING_AUTO - DO NOT CHANGE!	Putative aminopeptidase	13469000	7,1293	1,1509
A4IAZ8	LEADING_SPECIES - DO NOT CHANGE!	Putative casein kinase	13413000	7,1275	1,1506

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
A0A142BXX3	Multi SPECIES	6-phosphogluconate dehydrogenase, decarboxylating (Fragment)	13367000	7,1260	1,1504
A0A142BXX2	Multi SPECIES	6-phosphogluconate dehydrogenase, decarboxylating (Fragment)	13367000	7,1260	1,1504
I3VJT5	Multi SPECIES	6-phosphogluconate dehydrogenase, decarboxylating (Fragment)	13367000	7,1260	1,1504
Q18L52	Multi SPECIES	6-phosphogluconate dehydrogenase, decarboxylating	13367000	7,1260	1,1504
Q18L04	Multi SPECIES	6-phosphogluconate dehydrogenase, decarboxylating	13367000	7,1260	1,1504
Q5SDH3	LEADING_AUTO - DO NOT CHANGE!	Putative small myristoylated protein 4	13217000	7,1211	1,1496
Q8I496	LEADING_SPECIES - DO NOT CHANGE!	Hs1vu complex proteolytic subunit-like	13207000	7,1208	1,1495
A4HDR8	LEADING_SPECIES - DO NOT CHANGE!	Transketolase	13203000	7,1207	1,1495
A4I8P2	LEADING_SPECIES - DO NOT CHANGE!	Thiol-dependent reductase 1	13162000	7,1193	1,1493
A4IB89	Multi SPECIES	Putative 60S Ribosomal protein L36	13160000	7,1193	1,1493
A4HZ73	Multi SPECIES	Putative 60S Ribosomal protein L36	13160000	7,1193	1,1493
A4HWN5	LEADING_SPECIES - DO NOT CHANGE!	Putative ribosomal protein S6	13075000	7,1164	1,1488
E9BET4	LEADING_AUTO - DO NOT CHANGE!	Calpain-like cysteine peptidase, putative	13065000	7,1161	1,1488
A4HHL6	LEADING_SPECIES - DO NOT CHANGE!	Rab GDP dissociation inhibitor	12988000	7,1135	1,1484
Q4QDX9	LEADING_AUTO - DO NOT CHANGE!	Putative 60S ribosomal protein L10a	12928000	7,1115	1,1480
Q4QIE0	LEADING_AUTO - DO NOT CHANGE!	Superoxide dismutase	12908000	7,1109	1,1479
A4I7Q4	LEADING_SPECIES - DO NOT CHANGE!	60S ribosomal protein L18a	12906000	7,1108	1,1479
A4I1Q0	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	12794000	7,1070	1,1473
A4I996	LEADING_SPECIES - DO NOT CHANGE!	Putative aminopeptidase	12740000	7,1052	1,1470
Q4QD68	LEADING_AUTO - DO NOT CHANGE!	Cysteine peptidase A (CPA)	12736000	7,1050	1,1470
E9B1Z9	LEADING_AUTO - DO NOT CHANGE!	GP63-like protein, leishmanolysin-like protein	12688000	7,1034	1,1467
A4H445	Multi SPECIES	Surface antigen-like protein	12624000	7,1012	1,1464
A4H446	Multi SPECIES	Surface antigen-like protein	12624000	7,1012	1,1464
A4H440	Multi SPECIES	Surface antigen-like protein	12624000	7,1012	1,1464
A4H3Y2	Multi SPECIES	Surface antigen-like protein	12624000	7,1012	1,1464
E9AP91	LEADING_AUTO - DO NOT CHANGE!	Putative lectin	12620000	7,1011	1,1463
E9BUX4	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	12590000	7,1000	1,1462
E9ADY9	LEADING_AUTO - DO NOT CHANGE!	Putative carnitine/choline acetyltransferase	12582000	7,0997	1,1461
A4ICV5	LEADING_SPECIES - DO NOT CHANGE!	Putative proteasome beta 2 subunit	12464000	7,0957	1,1455
E9AHL7	LEADING_SPECIES - DO NOT CHANGE!	Putative ADP-ribosylation factor	12447000	7,0951	1,1454
A4H4D9	LEADING_SPECIES - DO NOT CHANGE!	Putative dipeptidyl-peptidase III	12376000	7,0926	1,1450
E9AKY4	LEADING_AUTO - DO NOT CHANGE!	Protein disulfide isomerase	12352000	7,0917	1,1448

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
E9AZH2	LEADING_AUTO - DO NOT CHANGE!	Proteasome subunit beta type	12337000	7,0912	1,1447
A4I191	Multi SPECIES	RNA-binding protein, putative, UPB2	12286000	7,0894	1,1445
E9AH94	Multi SPECIES	RNA-binding protein, putative, UPB1	12286000	7,0894	1,1445
A4HYZ5	LEADING_SPECIES - DO NOT CHANGE!	Putative 40S ribosomal protein S11	12261000	7,0885	1,1443
O00912	LEADING_SPECIES - DO NOT CHANGE!	Histone H4	12223000	7,0872	1,1441
E9AGX4	LEADING_SPECIES - DO NOT CHANGE!	Putative 40S ribosomal protein S23	12165000	7,0851	1,1438
A4HQG9	LEADING_AUTO - DO NOT CHANGE!	Putative histidine secretory acid phosphatase	12145000	7,0844	1,1436
E9AWS6	LEADING_AUTO - DO NOT CHANGE!	Triosephosphate isomerase	12118000	7,0834	1,1435
A4I1V1	LEADING_SPECIES - DO NOT CHANGE!	Adenine phosphoribosyltransferase	12044000	7,0808	1,1431
E9BBL6	LEADING_AUTO - DO NOT CHANGE!	Pyrroline-5-carboxylate reductase	11965000	7,0779	1,1426
A0A1E1IRP4	Multi SPECIES	Uncharacterized protein	11957000	7,0776	1,1426
A0A088RML4	Multi SPECIES	Uncharacterized protein	11957000	7,0776	1,1426
E8NHG6	LEADING_AUTO - DO NOT CHANGE!	Putative uncharacterized protein LmxM_30_0440_1	11954000	7,0775	1,1425
E9B2V9	LEADING_AUTO - DO NOT CHANGE!	Superoxide dismutase	11930000	7,0766	1,1424
E9AHK3	LEADING_SPECIES - DO NOT CHANGE!	S-adenosylmethionine synthase	11867000	7,0743	1,1420
A4HVX3	LEADING_AUTO - DO NOT CHANGE!	Proteasome subunit alpha type	11839000	7,0733	1,1419
A4HRR9	LEADING_SPECIES - DO NOT CHANGE!	Dipeptylcarboxypeptidase	11757000	7,0703	1,1414
A4HVW1	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	11671000	7,0671	1,1409
E9AHJ2	Multi SPECIES	Putative paraflagellar rod protein 1D	11634000	7,0657	1,1406
Q66V59	Multi SPECIES	Paraflagellar rod protein 1	11634000	7,0657	1,1406
A4I4N5	Multi SPECIES	Putative paraflagellar rod protein 1D	11634000	7,0657	1,1406
O18699	LEADING_SPECIES - DO NOT CHANGE!	Cysteine protease	11603000	7,0646	1,1404
A4HW72	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	11565000	7,0631	1,1402
A4ICB5	LEADING_SPECIES - DO NOT CHANGE!	Putative prolyl oligopeptidase	11466000	7,0594	1,1396
E9B3P3	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	11466000	7,0594	1,1396
A4I4E1	LEADING_SPECIES - DO NOT CHANGE!	Putative guanine deaminase	11359000	7,0553	1,1390
A4HX65	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	11357000	7,0553	1,1389
A4HV05	LEADING_SPECIES - DO NOT CHANGE!	40S ribosomal protein S5	11343000	7,0547	1,1389
E9BCQ4	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	11332000	7,0543	1,1388
Q4Q504	LEADING_SPECIES - DO NOT CHANGE!	Putative ribosomal protein L27	11260000	7,0515	1,1383
A4HTJ5	LEADING_SPECIES - DO NOT CHANGE!	Putative small ubiquitin protein	11221000	7,0500	1,1381
A4ICK8	LEADING_SPECIES - DO NOT CHANGE!	Fructose-bisphosphate aldolase	11178000	7,0484	1,1378

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
E8NHC1	LEADING_AUTO - DO NOT CHANGE!	Proteasome subunit beta type	11054000	7,0435	1,1371
A4HLW4	LEADING_AUTO - DO NOT CHANGE!	Putative carboxypeptidase	10882000	7,0367	1,1360
A4I885	LEADING_SPECIES - DO NOT CHANGE!	Putative cysteine desulfhydrase	10870000	7,0362	1,1359
Q9BHZ6	Multi SPECIES	Elongation factor-1 gamma	10862000	7,0359	1,1358
A4HU19	Multi SPECIES	Elongation factor-1 gamma	10862000	7,0359	1,1358
A4HU18	Multi SPECIES	Elongation factor-1 gamma	10862000	7,0359	1,1358
A4I3Z9	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	10842000	7,0351	1,1357
E9AW99	LEADING_AUTO - DO NOT CHANGE!	Putative uncharacterized protein	10740000	7,0310	1,1350
Q4QF83	LEADING_AUTO - DO NOT CHANGE!	NAD-specific glutamate dehydrogenase	10736000	7,0308	1,1350
O43992	LEADING_AUTO - DO NOT CHANGE!	40S ribosomal protein S2	10688000	7,0289	1,1347
E9ACC9	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	10528000	7,0223	1,1336
A0A088RLI0	LEADING_AUTO - DO NOT CHANGE!	Histone H2B	10499000	7,0211	1,1334
A4HVR4	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	10498000	7,0211	1,1334
E9AYP9	LEADING_AUTO - DO NOT CHANGE!	Putative uncharacterized protein	10449000	7,0191	1,1331
A0A1E1J8P0	LEADING_AUTO - DO NOT CHANGE!	Putative S-adenosylhomocysteine hydrolase	10445000	7,0189	1,1331
A4HQL6	LEADING_AUTO - DO NOT CHANGE!	Protein disulfide-isomerase	10442000	7,0188	1,1331
E9AHB0	Multi SPECIES	Putative 60S ribosomal protein L7	10367000	7,0157	1,1326
A4I1V4	Multi SPECIES	Putative 60S ribosomal protein L7	10367000	7,0157	1,1326
Q4QBL1	LEADING_AUTO - DO NOT CHANGE!	Farnesyl pyrophosphate synthase	10321000	7,0137	1,1322
A0A1E1J0K7	Multi SPECIES	40S ribosomal protein S17, putative	10315000	7,0135	1,1322
A0A088RXB2	Multi SPECIES	40S ribosomal protein S17, putative	10315000	7,0135	1,1322
A4HMQ9	LEADING_AUTO - DO NOT CHANGE!	Putative aminopeptidase P	10241000	7,0103	1,1317
A4I5C0	LEADING_SPECIES - DO NOT CHANGE!	Putative adenosine kinase	10235000	7,0101	1,1317
A4I5Z0	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	10164000	7,0071	1,1312
A4HY61	LEADING_SPECIES - DO NOT CHANGE!	Putative 40S ribosomal protein S13	10143000	7,0062	1,1310
Q4QDX3	LEADING_AUTO - DO NOT CHANGE!	Probable citrate synthase, mitochondrial	10117000	7,0051	1,1308
A4HI30	LEADING_SPECIES - DO NOT CHANGE!	Putative 4-methyl-5(Beta-hydroxyethyl)-thiazole monophosphate synthesis protein	10093000	7,0040	1,1307
A4I890	LEADING_SPECIES - DO NOT CHANGE!	Putative ribosomal protein L27	10091000	7,0039	1,1307
A4HU23	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	10064000	7,0028	1,1305
A4H746	LEADING_AUTO - DO NOT CHANGE!	40S ribosomal protein S12	10051000	7,0022	1,1304
E9AII6	LEADING_SPECIES - DO NOT CHANGE!	Contig, possible fusion of chromosomes 20 and 34	10038000	7,0016	1,1303
A4I067	LEADING_SPECIES - DO NOT CHANGE!	Pteridine reductase 1	9968900	6,9986	1,1298

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
Q4Q6F6	LEADING_AUTO - DO NOT CHANGE!	Biotin/lipoate protein ligase-like protein	9954600	6,9980	1,1297
A4HX73	Multi SPECIES	Elongation factor 1-alpha	9950300	6,9978	1,1297
E9AGP5	Multi SPECIES	Elongation factor 1-alpha	9950300	6,9978	1,1297
A4HT92	LEADING_SPECIES - DO NOT CHANGE!	Putative 40S ribosomal protein S9	9923600	6,9967	1,1295
A4HV19	LEADING_SPECIES - DO NOT CHANGE!	Putative 60S ribosomal protein L28	9922600	6,9966	1,1295
A4HDN8	LEADING_AUTO - DO NOT CHANGE!	Inhibitor of cysteine peptidase	9877400	6,9946	1,1292
A4HCH8	LEADING_AUTO - DO NOT CHANGE!	FPPS protein	9716400	6,9875	1,1280
E9AEF4	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	9692000	6,9864	1,1278
E9ACQ0	LEADING_AUTO - DO NOT CHANGE!	Surface antigen-like protein	9659700	6,9850	1,1276
A4HXT8	LEADING_AUTO - DO NOT CHANGE!	Putative 60S ribosomal protein L10a	9620100	6,9832	1,1273
A4HE56	LEADING_SPECIES - DO NOT CHANGE!	Aldehyde dehydrogenase, mitochondrial	9561600	6,9805	1,1269
Q27686	LEADING_AUTO - DO NOT CHANGE!	Pyruvate kinase	9543900	6,9797	1,1268
A4I2F5	LEADING_AUTO - DO NOT CHANGE!	Putative nitrilase	9474400	6,9766	1,1262
Q4QIX1	LEADING_AUTO - DO NOT CHANGE!	Protein disulfide isomerase	9325200	6,9697	1,1251
F8QV42	LEADING_AUTO - DO NOT CHANGE!	Tubulin beta chain	9321600	6,9695	1,1251
A4H942	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	9313200	6,9691	1,1250
A4I1Z8	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	9266800	6,9669	1,1247
A4I9G0	LEADING_SPECIES - DO NOT CHANGE!	Small nuclear ribonucleoprotein SmD2	9223300	6,9649	1,1244
B8Y658	LEADING_SPECIES - DO NOT CHANGE!	Methylene-tetrahydrofolate dehydrogenase/cyclohydrolase	9209300	6,9642	1,1242
Q70GE8	LEADING_AUTO - DO NOT CHANGE!	Thiol-dependent reductase 1	9171100	6,9624	1,1240
A4I849	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	9114400	6,9597	1,1235
E9AGX7	LEADING_SPECIES - DO NOT CHANGE!	T-complex protein 1 subunit delta	9067600	6,9575	1,1232
E9BQ78	LEADING_SPECIES - DO NOT CHANGE!	Mitogen activated protein kinase, putative	8964200	6,9525	1,1224
A4HB47	LEADING_SPECIES - DO NOT CHANGE!	Contig, possible fusion of chromosomes 20 and 34	8918500	6,9503	1,1220
A4I412	Multi SPECIES	Putative heat-shock protein hsp70	8894100	6,9491	1,1218
E9AHH1	Multi SPECIES	Putative heat-shock protein hsp70	8894100	6,9491	1,1218
Q4QGD8	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	8874200	6,9481	1,1217
A4HWC9	LEADING_SPECIES - DO NOT CHANGE!	Putative ribonucleoprotein p18, mitochondrial	8809100	6,9449	1,1211
A4I2T4	LEADING_SPECIES - DO NOT CHANGE!	Putative arginyl-tRNA synthetase	8701500	6,9396	1,1203
E9AHP6	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	8670600	6,9380	1,1200
A4HSP6	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	8649100	6,9370	1,1198

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
A4HFU7	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	8636800	6,9364	1,1197
Q6QMI0	Multi SPECIES	Nonspecific nucleoside hydrolase (Fragment)	8601200	6,9346	1,1195
Q6QMJ1	Multi SPECIES	Nonspecific nucleoside hydrolase (Fragment)	8601200	6,9346	1,1195
Q6QMI7	Multi SPECIES	Nonspecific nucleoside hydrolase (Fragment)	8601200	6,9346	1,1195
Q2PDC6	Multi SPECIES	Inosine-uridine preferring nucleoside hydrolase	8601200	6,9346	1,1195
Q2PDC5	Multi SPECIES	Inosine-uridine preferring nucleoside hydrolase	8601200	6,9346	1,1195
Q2PDB9	Multi SPECIES	Inosine-uridine preferring nucleoside hydrolase	8601200	6,9346	1,1195
Q2PD97	Multi SPECIES	Inosine-uridine preferring nucleoside hydrolase	8601200	6,9346	1,1195
Q2PD96	Multi SPECIES	Inosine-uridine preferring nucleoside hydrolase	8601200	6,9346	1,1195
A4I4C9	LEADING_SPECIES - DO NOT CHANGE!	Putative lipophosphoglycan biosynthetic protein	8589400	6,9340	1,1194
C7EX18	LEADING_AUTO - DO NOT CHANGE!	Gp63	8559700	6,9325	1,1191
A4ICD5	LEADING_SPECIES - DO NOT CHANGE!	Protein disulfide-isomerase	8539900	6,9315	1,1190
A4H8B2	LEADING_SPECIES - DO NOT CHANGE!	Proliferating cell nuclear antigen	8539700	6,9314	1,1190
A4I1F4	LEADING_SPECIES - DO NOT CHANGE!	Aldehyde dehydrogenase, mitochondrial	8522100	6,9305	1,1188
Q4QFJ8	LEADING_AUTO - DO NOT CHANGE!	Inositol-3-phosphate synthase	8477500	6,9283	1,1184
E9AWR7	LEADING_AUTO - DO NOT CHANGE!	Malic enzyme	8476600	6,9282	1,1184
A4HXD8	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	8463000	6,9275	1,1183
A0A1E1ISN0	Multi SPECIES	Putative cystathionine beta-synthase	8460200	6,9274	1,1183
A0A088RP82	Multi SPECIES	Cystathionine beta-synthase	8460200	6,9274	1,1183
A4ID19	LEADING_SPECIES - DO NOT CHANGE!	Nascent polypeptide-associated complex subunit beta	8399300	6,9242	1,1178
A4I7N0	LEADING_SPECIES - DO NOT CHANGE!	Profilin	8323000	6,9203	1,1172
A4I082	LEADING_SPECIES - DO NOT CHANGE!	Aldose 1-epimerase-like protein	8314600	6,9198	1,1171
A0A0S1M2A5	Multi SPECIES	Putative spermidine synthase 1 (Fragment)	8296200	6,9189	1,1169
A0A0S1M298	Multi SPECIES	Putative spermidine synthase 1 (Fragment)	8296200	6,9189	1,1169
A0A0S1M285	Multi SPECIES	Putative spermidine synthase 1 (Fragment)	8296200	6,9189	1,1169
A4H3X8	Multi SPECIES	Putative spermidine synthase	8296200	6,9189	1,1169
A4IDL0	LEADING_SPECIES - DO NOT CHANGE!	Tyrosine aminotransferase	8254100	6,9167	1,1166
A4I3Q8	LEADING_SPECIES - DO NOT CHANGE!	Putative replication factor A, 51kDa subunit	8238700	6,9159	1,1164
E9BB38	LEADING_AUTO - DO NOT CHANGE!	Alanine aminotransferase	8238400	6,9158	1,1164
A4HRT6	LEADING_SPECIES - DO NOT CHANGE!	Putative ribosomal protein L38	8214700	6,9146	1,1162
A4H635	LEADING_AUTO - DO NOT CHANGE!	GP63-4 protein	8210400	6,9144	1,1162

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
A4I253	LEADING_AUTO - DO NOT CHANGE!	Heat shock protein 70-related protein	8198100	6,9137	1,1161
A9YYK9	Multi SPECIES	p4 nuclease (Fragment)	8185800	6,9131	1,1160
A4I5I0	Multi SPECIES	p1/s1 nuclease	8185800	6,9131	1,1160
E9JUH3	Multi SPECIES	Trypanothione reductase (Fragment)	8182100	6,9129	1,1160
E9JUG5	Multi SPECIES	Trypanothione reductase (Fragment)	8182100	6,9129	1,1160
E9BB84	LEADING_AUTO - DO NOT CHANGE!	Carboxypeptidase, putative	8178000	6,9126	1,1159
A4IAG9	LEADING_SPECIES - DO NOT CHANGE!	Nucleolar protein family a member-like protein	8167600	6,9121	1,1158
Q8I8E1	Multi SPECIES	Protein disulfide-isomerase	8159700	6,9117	1,1158
Q4Q059	Multi SPECIES	Protein disulfide-isomerase	8159700	6,9117	1,1158
E9BG32	LEADING_AUTO - DO NOT CHANGE!	GDP-mannose pyrophosphorylase	8137900	6,9105	1,1156
A4HWJ8	LEADING_AUTO - DO NOT CHANGE!	60S ribosomal protein L6	8124700	6,9098	1,1155
I3VJK5	Multi SPECIES	Fumarate hydratase (Fragment)	8066700	6,9067	1,1150
A2CIQ7	Multi SPECIES	Fumarate hydratase	8066700	6,9067	1,1150
A2CIQ1	Multi SPECIES	Fumarate hydratase	8066700	6,9067	1,1150
A2CIP4	Multi SPECIES	Fumarate hydratase	8066700	6,9067	1,1150
A2CIN2	Multi SPECIES	Fumarate hydratase	8066700	6,9067	1,1150
A2CIN0	Multi SPECIES	Fumarate hydratase	8066700	6,9067	1,1150
A4H9H8	LEADING_AUTO - DO NOT CHANGE!	Probable citrate synthase, mitochondrial	8021700	6,9043	1,1146
A4IBC5	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	7993600	6,9027	1,1143
A2CIM6	Multi SPECIES	Glucose-6-phosphate 1-dehydrogenase	7927800	6,8992	1,1137
A2CIL1	Multi SPECIES	Glucose-6-phosphate 1-dehydrogenase	7927800	6,8992	1,1137
A2CIK1	Multi SPECIES	Glucose-6-phosphate 1-dehydrogenase	7927800	6,8992	1,1137
A2CIJ8	Multi SPECIES	Glucose-6-phosphate 1-dehydrogenase	7927800	6,8992	1,1137
A4HVI6	LEADING_SPECIES - DO NOT CHANGE!	40S ribosomal protein S12	7922800	6,8989	1,1137
P42865	LEADING_AUTO - DO NOT CHANGE!	Probable quinone oxidoreductase	7871000	6,8960	1,1132
A7UFI6	Multi SPECIES	Malate dehydrogenase	7838600	6,8942	1,1130
A4HAC0	Multi SPECIES	Malate dehydrogenase	7838600	6,8942	1,1130
Q1A5Y1	LEADING_AUTO - DO NOT CHANGE!	Tubulin beta chain (Fragment)	7770400	6,8904	1,1123
Q9N9V2	Multi SPECIES	Nascent polypeptide associated complex homologue, alpha chain	7764500	6,8901	1,1123
A4HS73	Multi SPECIES	Nascent polypeptide associated complex subunit-like protein, copy 1	7764500	6,8901	1,1123
A4HS72	Multi SPECIES	Nascent polypeptide associated complex subunit-like protein, copy 2	7764500	6,8901	1,1123
P25204	LEADING_AUTO - DO NOT CHANGE!	40S ribosomal protein S8	7762500	6,8900	1,1123
E9B3B0	LEADING_AUTO - DO NOT CHANGE!	Putative chaperonin alpha subunit	7704400	6,8867	1,1117
Q4QGJ6	LEADING_AUTO - DO NOT CHANGE!	Putative surface antigen protein 2	7645300	6,8834	1,1112

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
Q2HZY7	Multi SPECIES	Elongation factor 2	7632600	6,8827	1,1111
A4HNM7	Multi SPECIES	Elongation factor 2	7632600	6,8827	1,1111
A4I478	LEADING_SPECIES - DO NOT CHANGE!	Putative thymine-7-hydroxylase	7508000	6,8755	1,1099
Q4Q276	LEADING_AUTO - DO NOT CHANGE!	Uncharacterized protein	7437800	6,8714	1,1093
E9ARS1	LEADING_AUTO - DO NOT CHANGE!	Putative heat shock protein	7388800	6,8686	1,1088
A4HCV5	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	7387700	6,8685	1,1088
A4HXQ3	LEADING_SPECIES - DO NOT CHANGE!	Glycogen synthase kinase 3	7374100	6,8677	1,1087
Q9N9V5	LEADING_SPECIES - DO NOT CHANGE!	Putative 40S ribosomal protein S33	7367100	6,8673	1,1086
A4HKX6	LEADING_SPECIES - DO NOT CHANGE!	Dihydrolipoyl dehydrogenase	7358900	6,8668	1,1085
A4HCZ3	LEADING_SPECIES - DO NOT CHANGE!	T-complex protein 1 subunit gamma	7346600	6,8661	1,1084
Q2PD92	LEADING_SPECIES - DO NOT CHANGE!	Aspartate aminotransferase	7265000	6,8612	1,1076
A0A1E1J180	Multi SPECIES	Aspartyl aminopeptidase,putative,metallo-peptidase, Clan MH, Family M20	7237900	6,8596	1,1074
A0A088RY37	Multi SPECIES	Aspartyl aminopeptidase, putative	7237900	6,8596	1,1074
A4ICM4	LEADING_SPECIES - DO NOT CHANGE!	Putative ribosomal protein L24	7176300	6,8559	1,1068
Q4Q8S1	LEADING_AUTO - DO NOT CHANGE!	6-phosphogluconolactonase	7162800	6,8551	1,1066
A4HSS8	LEADING_SPECIES - DO NOT CHANGE!	Putative glutamine synthetase	7136800	6,8535	1,1064
A4HFS1	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	7075200	6,8497	1,1058
A4HI66	LEADING_SPECIES - DO NOT CHANGE!	Importin subunit alpha	7069400	6,8494	1,1057
A4HUB4	LEADING_SPECIES - DO NOT CHANGE!	Putative ribosomal protein l35a	7035600	6,8473	1,1054
E9BEP5	LEADING_AUTO - DO NOT CHANGE!	1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	7032900	6,8471	1,1053
A4IA81	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	6980300	6,8439	1,1048
P42922	LEADING_SPECIES - DO NOT CHANGE!	60S ribosomal protein L11	6869200	6,8369	1,1037
E9AKK2	LEADING_AUTO - DO NOT CHANGE!	Putative dipeptidyl-peptidase III (Metallo-peptidase, clan m-, family m49)	6803900	6,8328	1,1030
E9AYH6	LEADING_AUTO - DO NOT CHANGE!	Putative uncharacterized protein	6794400	6,8322	1,1029
A4HUL2	LEADING_SPECIES - DO NOT CHANGE!	Nucleoside phosphorylase-like protein	6661600	6,8236	1,1015
E9AGM7	Multi SPECIES	Putative nucleoside transporter 1	6651400	6,8229	1,1014
E9AGM6	Multi SPECIES	Putative nucleoside transporter 1	6651400	6,8229	1,1014
E9AGM5	Multi SPECIES	Putative nucleoside transporter 1	6651400	6,8229	1,1014
A4HWK9	Multi SPECIES	Putative nucleoside transporter 1	6651400	6,8229	1,1014
A4HU05	LEADING_SPECIES - DO NOT CHANGE!	Uncharacterized protein	6633800	6,8218	1,1013
A4HHW2	LEADING_SPECIES - DO NOT CHANGE!	Putative 2-hydroxy-3-oxopropionate reductase	6632500	6,8217	1,1012

Appendix II.1

LEADING_CHECK	CASE_CHECK	DESC_CHECK	iBAQ	LOG10	NORMALISED TO LOG10 AVERAGE
E9AQI0	LEADING_AUTO - DO NOT CHANGE!	Protein tyrosine phosphatase-like protein	6618900	6,8208	1,1011
E9AU19	LEADING_AUTO - DO NOT CHANGE!	Putative kinetoplast-associated protein	6607400	6,8200	1,1010
Q4QAI0	LEADING_SPECIES - DO NOT CHANGE!	Putative IgE-dependent histamine-releasing factor	6601000	6,8196	1,1009
A4ID58	Multi SPECIES	Putative ribosomal protein L29	6597900	6,8194	1,1009
A4ID57	Multi SPECIES	Putative ribosomal protein L29	6597900	6,8194	1,1009
A4HNP0	LEADING_SPECIES - DO NOT CHANGE!	Putative proteasome beta 2 subunit	6574500	6,8179	1,1006
E9ATD8	LEADING_AUTO - DO NOT CHANGE!	Putative cysteine synthase	6562000	6,8170	1,1005
Q4QBF5	LEADING_AUTO - DO NOT CHANGE!	Putative endoribonuclease L-PSP (Pb5)	6543800	6,8158	1,1003
E9B6B5	LEADING_AUTO - DO NOT CHANGE!	Putative aminopeptidase P	6530900	6,8150	1,1002
A4HSB6	LEADING_SPECIES - DO NOT CHANGE!	Putative RNA-binding protein	6519300	6,8142	1,1000

Appendix V.1

Data was collected from the Allele Frequency Net Database for the following world regions: South and Central America (Table 1), North Africa (Table 2), Western Asia (Table 3), South Asia (Table 4).

Data were retrieved from AFND through a 'HLA allele freq (Classical)' search with the following filters: i) region name; ii) level of resolution=2-digits; iii) Population standard=Gold and Silver (8); iv) show frequencies=only positives

Table 1 Population studies from AFND included in the HLA allele frequency analysis of the South and Central America World region.

World region: SOUTH AND CENTRAL AMERICA				
Population	**Ethnicity**	**Study**	**Sample Size**	**HLA Loci**
Argentina Buenos Aires	Caucasoid	HLA	466	A, B, C, DPB1, DQA1, DQB1, DRB1
Argentina Buenos Aires pop 2	Caucasoid	HLA	1216	A, B, DRB1
Argentina Chiriguano	Amerindian	HLA	54	A, B, DPB1, DQA1, DQB1, DRB1
Argentina Chubut Tehuelche	Amerindian	HLA	23	DRB1
Argentina Corrientes	Mixed	HLA	155	A, B, DRB1
Argentina Cuyo Region	Caucasoid	HLA	420	A, B, DRB1
Argentina Gran Chaco Eastern Toba	Amerindian	HLA	135	A, B, DPB1, DQA1, DQB1, DRB1
Argentina Gran Chaco Mataco Wichi	Amerindian	HLA	49	A, B, DPB1, DQA1, DQB1, DRB1
Argentina Gran Chaco Western Toba Pilaga	Amerindian	HLA	19	A, B, DPB1, DQA1, DQB1, DRB1
Argentina Kolla	Amerindian	HLA	61	DQB1, DRB1
Argentina La Plata	Caucasoid	HLA	100	A, B, C, DRB1
Argentina Mapuche	Amerindian	HLA	48	DPB1, DQA1, DQB1, DRB1
Argentina Rio Negro Mapuche	Amerindian	HLA	34	DRB1
Argentina Rosario Toba	Amerindian	HLA	86	A, B, DQA1, DQB1, DRB1
Argentina Salta Wichi pop 2	Amerindian	HLA	19	A, B, DPB1, DQA1, DQB1, DRB1
Bolivia Aymara	Amerindian	HLA	102	A, B, DQB1, DRB1
Bolivia Quechua		HLA	80	DQB1, DRB1, A, B
Brazil Belem Mixed	Mixed	HLA	100	B
Brazil Belo Horizonte Caucasian	Caucasoid	HLA	95	A, B
Brazil Central Plateau Xavante	Amerindian	HLA	74	DPB1, DQA1, DQB1, DRB1
Brazil Curitiba-Parana Mixed	Mixed	HLA	264	A, B, C, DRB1
Brazil Guarani Kaiowa	Amerindian	HLA	155	DQA1, DQB1, DRB1
Brazil Guarani M bya	Amerindian	HLA	93	DQA1, DQB1, DRB1

Appendix V.1

Table 1 (continued)

Population	Ethnicity	Study	Sample Size	HLA Loci
Brazil Guarani Nandeva	Amerindian	HLA	86	DQA1, DQB1, DRB1
Brazil Kaingang	Amerindian	HLA	235	DQA1, DQB1, DRB1
Brazil Mato Grosso do Sul	Mixed	HLA	203	DRB1
Brazil Minas Gerais State Mixed	Mixed	HLA	1	A, B, DRB1
Brazil North East Mixed	Mixed	HLA	205	DQA1, DQB1, DRB1
Brazil Parana Afro Brazilian	Black	HLA	77	A, B, DRB1
Brazil Parana Cafuzo	Mixed	HLA	319	A, B, DRB1
Brazil Parana Caucasian	Caucasoid	HLA	2775	A, B, DRB1
Brazil Parana Japanese		HLA	192	A B C
Brazil Parana Mulatto	Mulatto	HLA	186	A, B, DRB1
Brazil Parana Oriental	Oriental	HLA	33	A, B, DRB1
Brazil Pernambuco Mixed	Mixed	HLA	101	A, B, C
Brazil Piaui Mixed	Mixed	HLA	21943	A, B, DRB1
Brazil REDOME Acre	Mixed	HLA	859	A, B, DRB1
Brazil REDOME Alagoas	Mixed	HLA	25349	A, B, DRB1
Brazil REDOME Amapa	Mixed	HLA	17,864	A, B, DRB1
Brazil REDOME Amazonas	Mixed	HLA	24,129	A, B, DRB1
Brazil REDOME Bahia	Mixed	HLA	47399	A, B, DRB1
Brazil REDOME Ceara	Mixed	HLA	101217	A, B, DRB1
Brazil REDOME Espirito Santo	Mixed	HLA	88,485	A, B, DRB1
Brazil REDOME Federal District	Mixed	HLA	29549	A, B, DRB1
Brazil REDOME Goias	Mixed	HLA	88,574	A, B, DRB1
Brazil REDOME Maranhao	Mixed	HLA	10180	A, B, DRB1
Brazil REDOME Mato Grosso	Mixed	HLA	34649	A, B, DRB1
Brazil REDOME Mato Grosso do Sul	Mixed	HLA	95667	A, B, DRB1
Brazil REDOME Minas Gerais	Mixed	HLA	211275	A, B, DRB1
Brazil REDOME Para	Mixed	HLA	72637	A, B, DRB1
Brazil REDOME Paraiba	Mixed	HLA	43868	A, B, DRB1
Brazil REDOME Parana	Mixed	HLA	341639	A, B, DRB1
Brazil REDOME Pernambuco	Mixed	HLA	92332	A, B, DRB1
Brazil REDOME Piaui	Mixed	HLA	46140	A, B, DRB1
Brazil REDOME Rio de Janeiro	Mixed	HLA	139322	A, B, DRB1
Brazil REDOME Rio Grande do Norte	Mixed	HLA	46603	A, B, DRB1
Brazil REDOME Rio Grande do Sul	Mixed	HLA	241329	A, B, DRB1
Brazil REDOME Rondonia	Mixed	HLA	54396	A, B, DRB1
Brazil REDOME Roraima	Mixed	HLA	4140	A, B, DRB1
Brazil REDOME Santa Catarina	Mixed	HLA	106673	A, B, DRB1
Brazil REDOME Sao Paulo	Mixed	HLA	800809	A, B, DRB1

Appendix V.1

Table 1 (continued)

Population	Ethnicity	Study	Sample Size	HLA Loci
Brazil REDOME Sergipe	Mixed	HLA	7321	A, B, DRB1
Brazil REDOME Tocantins	Mixed	HLA	20692	A, B, DRB1
Brazil Rio Grande do Norte Mestizo	Mestizo	HLA	12973	A, B, DRB1
Brazil Rio Grande do Sul Black	Black	HLA	248	A, B, DRB1
Brazil Rio Grande do Sul Caucasoid	Caucasoid	HLA	4428	A, B, DRB1
Brazil Rio Grande do Sul Mixed	Mixed	HLA	324	A, B, DRB1
Brazil Sao Paulo Bauru Mixed	Mixed	HLA	3542	A, B, DRB1
Brazil Sao Paulo Mixed	Mixed	HLA	239	A, B, DQA1, DQB1, DRB1
Brazil South East Cord Bllod	Unknown	HLA	11409	A B DRB1
Brazil South Ribeirao Preto	Mixed	HLA	184	A, B, DRB1
Brazil Southeast Campinas Mixed	Mixed	HLA	99	DRB1
Chile Easter Island	Amerindian	HLA	21	A, B, C, DRB1
Chile Huilliche	Amerindian	HLA	40	DRB1
Chile Santiago		HLA	920	A, B, DRB1
Chile Santiago Mixed	Mixed	HLA	70	A, B, C, DQB1, DRB1
Colombia Antioquia Paisa	Amerindian	HLA	100	DQB1, DRB1
Colombia Barranquilla		HLA	188	A, B, DRB1
Colombia Bogota and Medellin Mestizo	Mestizo	HLA	65	DQB1, DRB1
Colombia Guajira Peninsula Wayuu	Amerindian	HLA	88	DQA1, DQB1, DRB1
Colombia Jaidukama		HLA	39	A, B, DQB1, DRB1
Colombia Northwest Tule	Amerindian	HLA	29	DPB1, DQA1, DQB1, DRB1
Colombia San Basilio de Palenque		HLA	42	A B DQB1 DRB1
Colombia Sierra Nevada de Santa Marta Arhuaco	Amerindian	HLA	107	DQA1, DQB1, DRB1
Colombia Sierra Nevada de Santa Marta Arsario	Amerindian	HLA	18	DQA1, DQB1, DRB1
Colombia Sierra Nevada de Santa Marta Kogi	Amerindian	HLA	42	DQA1, DQB1, DRB1
Colombia Wayu from Guajira Peninsula	Wayu	HLA	48	A B DQB1 DRB1
Costa Rica Central Valley Mestizo	Mestizo	HLA	130	A, B, C, DQB1, DRB1, DRB3 DRB4 DRB5
Cuba Caucasian	Caucasoid	HLA	70	A, B
Cuba Mixed	Mixed	HLA	78	A, B, DQA1, DQB1, DRB1
Cuba Mixed pop 2	Mixed	HLA	189	A, B, C, DRB1
Cuba Mulatto	Mulatto	HLA	42	A, B
Ecuador African	Black	HLA	58	DPA1, DPB1, DQA1, DQB1
Ecuador Amazonia Mixed Ancestry		HLA	39	A, B, DQB1, DRB1
Ecuador Andes Mixed Ancestry		HLA	824	A, B, DQB1, DRB1
Ecuador Cayapa	Amerindian	HLA	183	A, DPB1, DQA1, DQB1, DRB1
Ecuador Coast Mixed Ancestry		HLA	238	A, B, DQB1, DRB1
Ecuador Mixed Ancestry		HLA	1173	A, B, DQB1, DRB1
Guatemala Mayan	Amerindian	HLA	132	A, B, DQB1, DRB1
Jamaica	Black	HLA	132	DQA1, DQB1, DRB1

Appendix V.1

Table 2 Population studies from AFND included in the HLA allele frequency analysis for the North Africa World region.

World region: NORTH AFRICA				
Population	**Ethnicity**	**Study**	**Sample Size**	**HLA Loci**
Algeria pop 2	Arab	HLA	106	A, B, C, DQA1, DQB1, DRB1
Ethiopia Amhara	Black	HLA	98	DQA1, DQB1, DRB1
Ethiopia Oromo	Black	HLA	83	DQA1, DQB1, DRB1
Morocco	Arab	HLA	96	A, B, DQA1, DQB1, DRB1
Morocco Casablanca	Arab	HLA	100	A, B
Morocco pop 2	Mixed	HLA	110	A, B, DQB1, DRB1
Morocco Souss Region	Arab	HLA	98	DQA1, DQB1, DRB1
Sudan Central Shaigiya Mixed	Mixed	HLA	36	A, B, C, DRB1
Sudan East Rashaida	Arab	HLA	27	A, B, C, DRB1
Sudan Mixed	Mixed	HLA	200	A, B, C, DPB1, DQB1, DRB1
Sudan South Nuba	Black	HLA	46	A, B, C, DRB1
Tunisia	Arab	HLA	100	A, B, C, DPB1, DQA1, DQB1, DRB1
Tunisia Ghannouch	Arab	HLA	82	A, B, DQB1, DRB1
Tunisia pop 3	Arab	HLA	104	A, B, DQB1, DRB1

Table 3 Population studies from AFND included in the HLA allele frequency analysis for the Western Asia World region.

World region: WESTERN ASIA				
Population	**Ethnicity**	**Study**	**Sample Size**	**HLA Loci**
Armenia combined Regions	Arab	HLA	100	A, B, DRB1
Armenia Ghegharkunik	Arab	HLA	242	A, B, DRB1
Armenia living in Iran	Arab	HLA	85	A, B, DRB1
Armenia living in Karabakh	Arab	HLA	445	A, B, DRB1
Armenia living in Lebanon	Arab	HLA	368	A, B, DRB1
Armenia living in USA	Arab	HLA	233	A, B, DRB1
Armenia Lori	Arab	HLA	102	A, B, DRB1
Armenia Shirak	Arab	HLA	76	A, B, DRB1
Armenia Syunik	Arab	HLA	117	A, B, DRB1
Armenia Yerevan	Arab	HLA	445	A, B, DRB1
Gaza Palestinians	Arab	HLA	165	A, B, DQB1, DRB1
Georgia Kurds	Kurds	HLA	30	A, B, DQB1, DRB1
Georgia Svaneti Region Svan	Caucasoid	HLA	80	A, B, DPB1, DQA1, DQB1, DRB1
Iran	Persian	HLA	58	DQA1, DQB1, DRB1
Iran Baloch	Persian	HLA	100	A, B, C, DQA1, DQB1, DRB1
Iran pop 1	Arab	HLA	64	A, B
Iran pop 3	Arab	HLA	100	DQB1, DRB1

Appendix V.1

Table 3 (continued)

Population	Ethnicity	Study	Sample Size	HLA Loci
Iran Royan Cord Blood Bank	Persian	HLA	15600	A, B, DRB1
Iran Tehran	Persian	HLA	120	C
Iraq Erbil		HLA	372	A B C DQB1 DRB1
Israel Arab	Arab	HLA	109	DQA1, DQB1, DRB1
Israel Arab		HLA	12301	DRB1, A, B
Israel Ashkenazi Jews	Jew	HLA	80	DQA1, DQB1, DRB1
Israel Ashkenazi Jews		HLA	4625	DRB1, A, B
Israel Ashkenazi Jews pop 2	Jew	HLA	132	A, B, DQA1, DQB1, DRB1
Israel Ethiopian Jews	Jew	HLA	122	DQA1, DQB1, DRB1
Israel Iranian Jews	Jew	HLA	101	DQA1, DQB1, DRB1
Israel Jews pop 2	Jew	HLA	23000	A, B, DRB1
Israel Libyan Jews	Jew	HLA	119	DQA1, DQB1, DRB1
Israel Moroccan Jews	Jew	HLA	113	A, B, DQA1, DQB1, DRB1
Israel Non Ashkenazi Jews	Jew	HLA	80	DQA1, DQB1, DRB1
Israel Yemenite Jews	Jew	HLA	76	DQA1, DQB1, DRB1
Jordan	Arab	HLA	15141	A, B, DRB1
Jordan Amman	Arab	HLA	146	A, B, C, DQA1, DQB1, DRB1
Iraq Kurdistan Region	Kurd	HLA	209	A, B, C, DQB1, DRB1
Lebanon Kafar Zubian	Arab	HLA	94	DPB1, DQB1, DRB1
Lebanon Niha el Shouff	Arab	HLA	61	DPB1, DQB1, DRB1
Lebanon Yuhmur	Arab	HLA	82	DPB1, DQB1, DRB1
Oman	Arab	HLA	118	A, B
Saudi Arabia	Arab	HLA	18	A, B, DPB1, DQB1, DRB1
Saudi Arabia pop 2	Arab	HLA	383	A, B, DQB1, DRB1
Saudi Arabia pop 4	Arab	HLA	499	A, B, C, DQB1, DRB1
Turkey Ankara	Caucasoid	HLA	50	DQA1, DQB1, DRB1
Turkey Istanbul	Caucasoid	HLA	250	DRB1
Turkey pop 1	Caucasoid	HLA	250	DQA1, DQB1, DRB1
Turkey pop 2	Caucasoid	HLA	228	A, B, DQB1, DRB1
Turkey pop 3	Caucasoid	HLA	50	DQA1, DQB1
Turkey pop 5	Caucasoid	HLA	142	A, B, C
United Arab Emirates pop 2	Arab	HLA	373	A, B, DRB1

Appendix V.1

Table 4 Population studies from AFND included in the HLA allele frequency analysis for the South Asia World region.

World Region: SOUTH ASIA				
Population	**Ethnicity**	**Study**	**Sample Size**	**HLA Loci**
Bangladesh Dhaka Bangalee	Asian	HLA	141	A, B, DRB1
India Andhra Pradesh Brahmin	Asian	HLA	98	DRB1
India Andhra Pradesh Sunni	Asian	HLA	100	DRB1
India Bombay	Asian	HLA	59	DPA1, DPB1, DQA1, DQB1
India Delhi pop 2	Asian	HLA	90	A, B, C
India Jalpaiguri Toto	Asian	HLA	40	A, B
India Kerala Adiya	Asian	HLA	21	B, C
India Kerala Hindu Ezhava	Asian	HLA	24	A, B, C
India Kerala Hindu Nair	Asian	HLA	41	A, B, C
India Kerala Hindu Namboothiri	Asian	HLA	40	A, B, C
India Kerala Hindu Pulaya	Asian	HLA	16	A, B, C
India Kerala Kanikkar	Asian	HLA	22	B, C
India Kerala Kattunaikka	Asian	HLA	17	B, C
India Kerala Kurichiya	Asian	HLA	10	B, C
India Kerala Kuruma	Asian	HLA	15	B, C
India Kerala Malabar Muslim	Asian	HLA	34	A, B, C
India Kerala Malapandaram	Asian	HLA	10	B, C
India Kerala Paniya	Asian	HLA	10	B, C
India Kerala Syrian Christian	Asian	HLA	31	A, B, C
India Lucknow	Asian	HLA	123	DPB1, DQA1, DQB1, DRB1
India New Delhi pop 2	Asian	HLA	102	DQB1, DRB1
India North	Asian	HLA	85	DRB1
India North Gujarat	Asian	HLA	338	A, B, C, DQB1, DRB1
India North pop 2	Asian	HLA	72	A, B, C, DQA1, DQB1, DRB1
India North pop 3	Asian	HLA	587	A, B, DRB1
India Northeast Kayastha	Asian	HLA	190	DQA1, DQB1, DRB1
India Northeast Lachung	Asian	HLA	58	DQA1, DQB1, DRB1
India Northeast Mathur	Asian	HLA	155	DQA1, DQB1, DRB1
India Northeast Mech	Asian	HLA	63	DQA1, DQB1, DRB1
India Northeast Rajbanshi	Asian	HLA	98	DQA1, DQB1, DRB1
India Northeast Rastogi	Asian	HLA	196	DQA1, DQB1, DRB1
India Northeast Shia	Asian	HLA	190	DQA1, DQB1, DRB1
India Northeast Sunni	Asian	HLA	188	DQA1, DQB1, DRB1
India Northeast Vaish	Asian	HLA	198	DQA1, DQB1, DRB1
India Tamil Nadu Chennai	Asian	HLA	137	DRB1
India Tamil Nadu Dravidian	Asian	HLA	156	DRB1
India Uttar Pradesh	Asian	HLA	202	DQA1, DQB1, DRB1

Table 4 (continued)

Population	Ethnicity	Study	Sample Size	HLA Loci
Pakistan Baloch	Asian	HLA	66	A, B, C, DQB1, DRB1
Pakistan Brahui	Asian	HLA	104	A, B, C, DQB1, DRB1
Pakistan Burusho	Asian	HLA	92	A, B, C, DQB1, DRB1
Pakistan Kalash	Asian	HLA	69	A, B, C, DQB1, DRB1
Pakistan Karachi Parsi	Asian	HLA	91	A, B, C, DQB1, DRB1
Pakistan Mixed Pathan	Mixed	HLA	100	A, B, C, DQB1, DRB1
Pakistan Mixed Sindhi	Mixed	HLA	101	A, B, C, DQB1, DRB1
Sri Lanka Colombo Sinhalese	Asian	HLA	101	A, B, C, DQB1, DRB1

Appendix V.2

Peptide screening in samples from healed individuals
- Total PBMC stimulation IFN-γ ELISpot results -

1st healed donor series (20/09/2017)

Whole blood samples from 10 healed individuals (Gafsa): TUN1 to 10

1 blood sample was hemolysed (TUN2) so PBMC were not purified

---

HLA-CLASS I peptides:

- 5 µM each peptide;
- All solubilized in sterile water (or sodium bicarbonate);
- 14 peptide pools with 7 peptides per pool (except I_pool7 and G with 6 peptides);
- pool ALL_I (49 peptides), I_pool1-7 and I_poolA-G.

	1	2	3	4	5	6	7
A	pi1	pi2	pi3	pi4	pi5	pi6	pi7
B	pi8	pi9	pi10	pi11	pi12	pi13	pi14
C	pi15	pi16	pi17	pi18	pi19	pi20	pi21
D	pi22	pi23	pi24	pi25	pi26	pi27	pi28
E	pi29	pi30	pi31	pi32	pi33	pi34	pi35
F	pi36	pi37	pi38	pi39	pi40	pi41	pi42
G	pi43	pi44	pi45	pi46	pi47	pi48	pi49

HLA-CLASS II peptides:

- 5 µM each peptide;
- All solubilized in 10% DMSO + sterile water;
- 10 peptide pools with 5 peptides per pool (except II_pool5 and E with 4 peptides);
- pool ALL_II (24 peptides), II_pool1-5 and II_poolA-E

	1	2	3	4	5
A	pii1	pii2	pii3	pii4	pii5
B	pii6	pii7	pii8	pii9	pii10
C	pii11	pii12	pii13	pii14	pii15
D	pii16	pii17	pii18	pii19	pii20
E	pii21	pii22	pii23	pii24	

Controls:

PHA 10 µg/mL

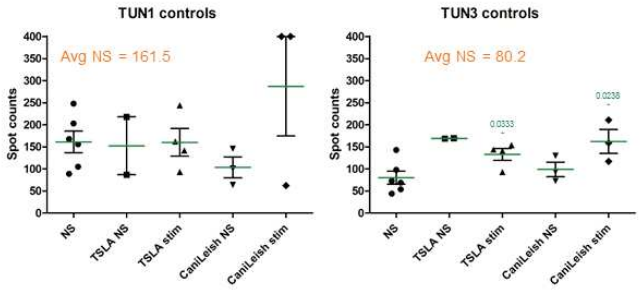
CaniLeish® 10 µg/mL

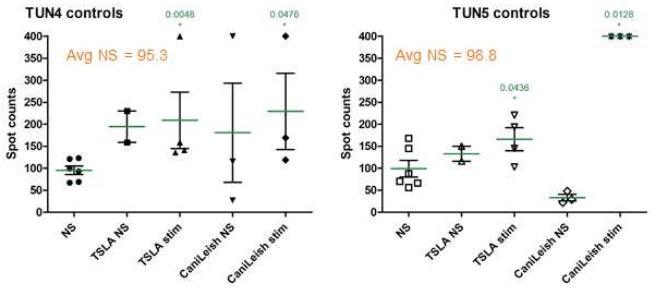
TSLA 10 µg/mL

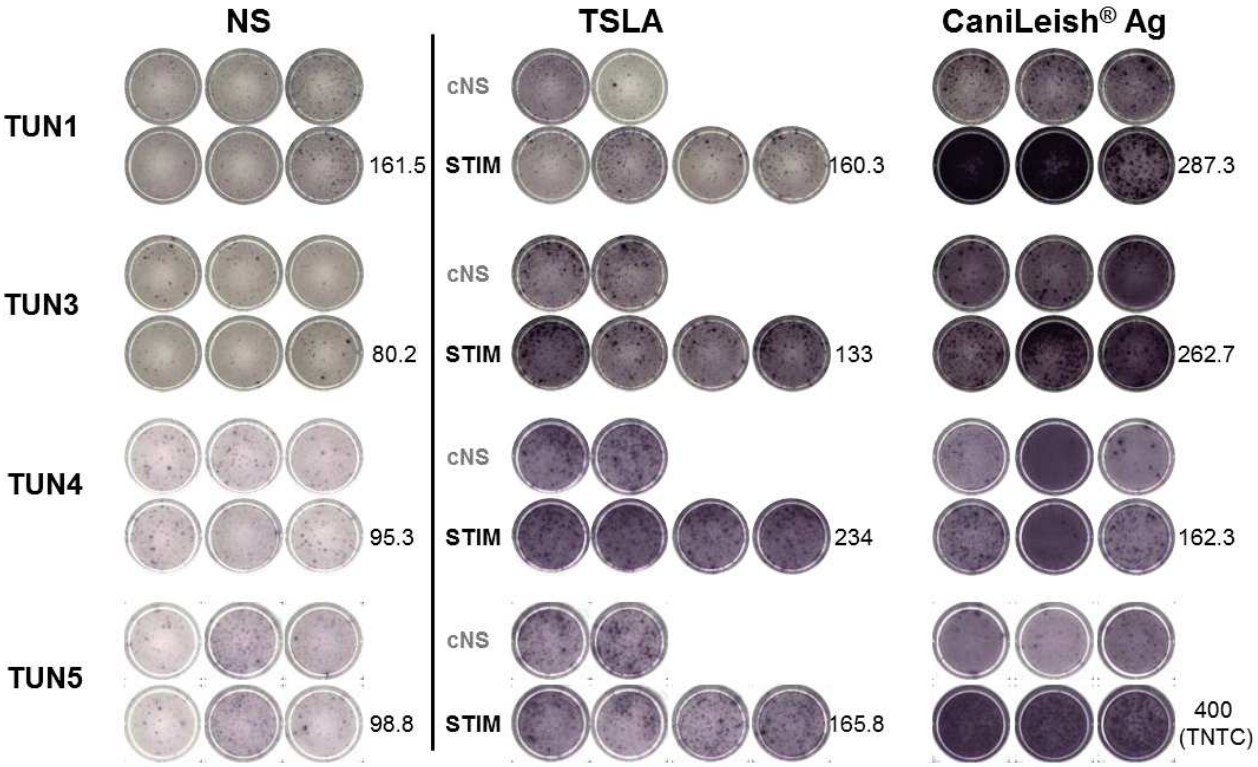
Appendix V.2

OVERALL RESULTS - CONTROLS

Non-parametric significance test (Mann-Whitney)
* p-value < 0.05
TNTC = spot count 400



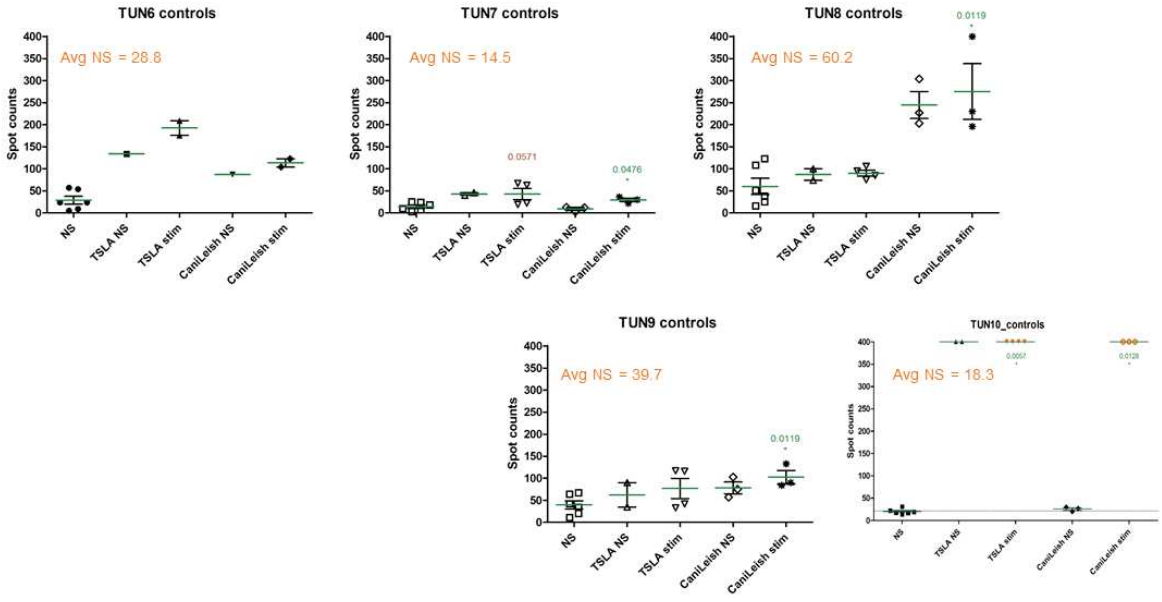


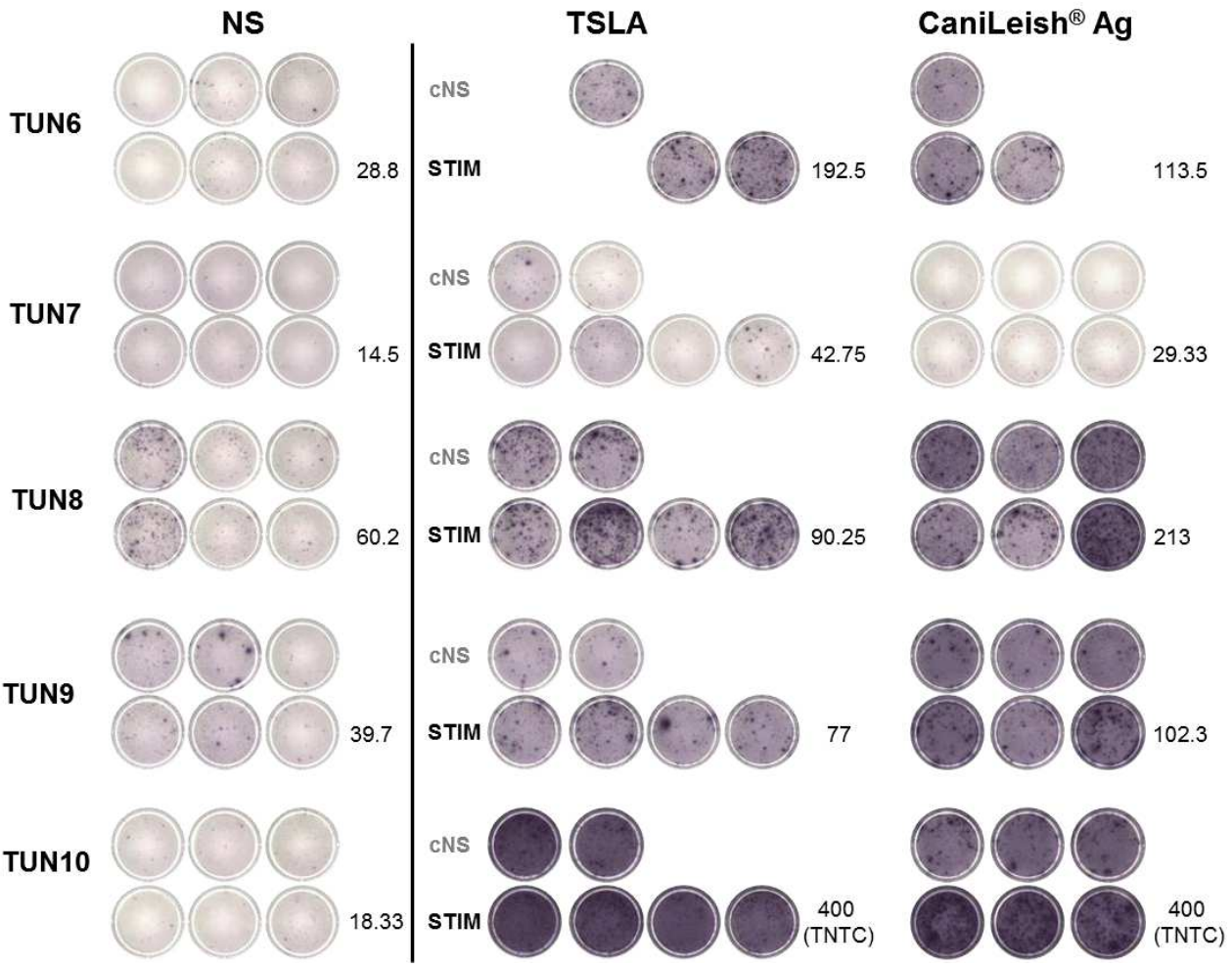


Appendix V.2

OVERALL RESULTS - CONTROLS

Non-parametric significance test (Mann-Whitney)
* p-value < 0.05
TNTC = spot count 400





Appendix V.2

OVERALL RESULTS - CONTROLS

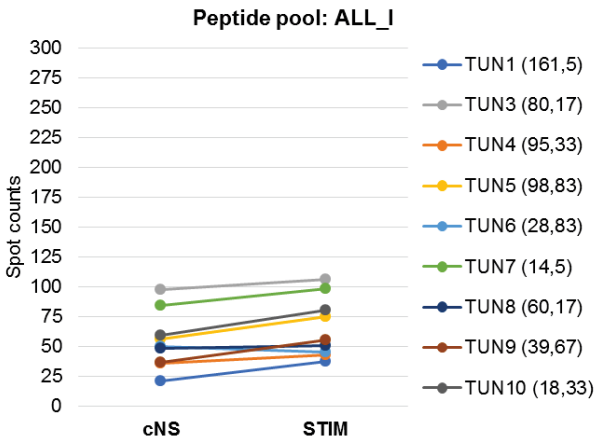
Non-parametric significance test (Mann-Whitney)
* p-value < 0.05
TNTC = spot count 400

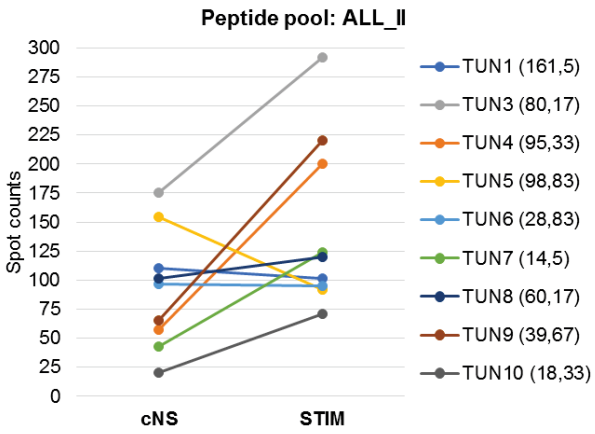
Donor	NS	TSLA stim	CaniLeish® stim
TUN1	161.5	160,3	287.3
TUN3	80.2	**133***	**162.3***
TUN4	95.3	**209***	**229.3***
TUN5	98.8	**165,75***	**400 (TNTC)***
TUN6	28.8	192.5	113.5
TUN7	14.5	42.75 (0.0571)	**29.33***
TUN8	60.2	90.25	**213***
TUN9	39.7	77	**102.3***
TUN10	18.333	**400 (TNTC)***	**400 (TNTC)***

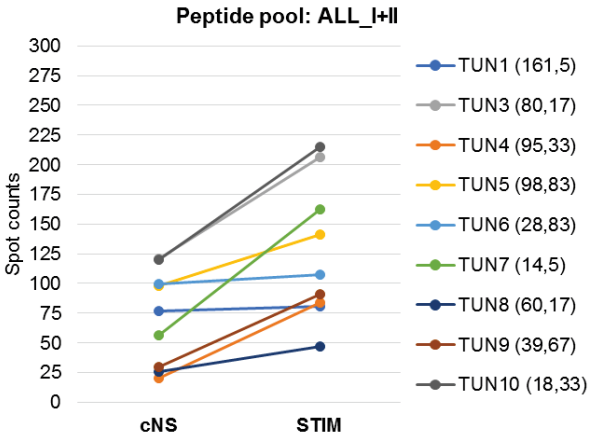
Appendix V.2

OVERALL RESULTS - Pools 'ALL'

cNS = culture well not stimulated in ELISPOT
NeverStimulated average in the legend



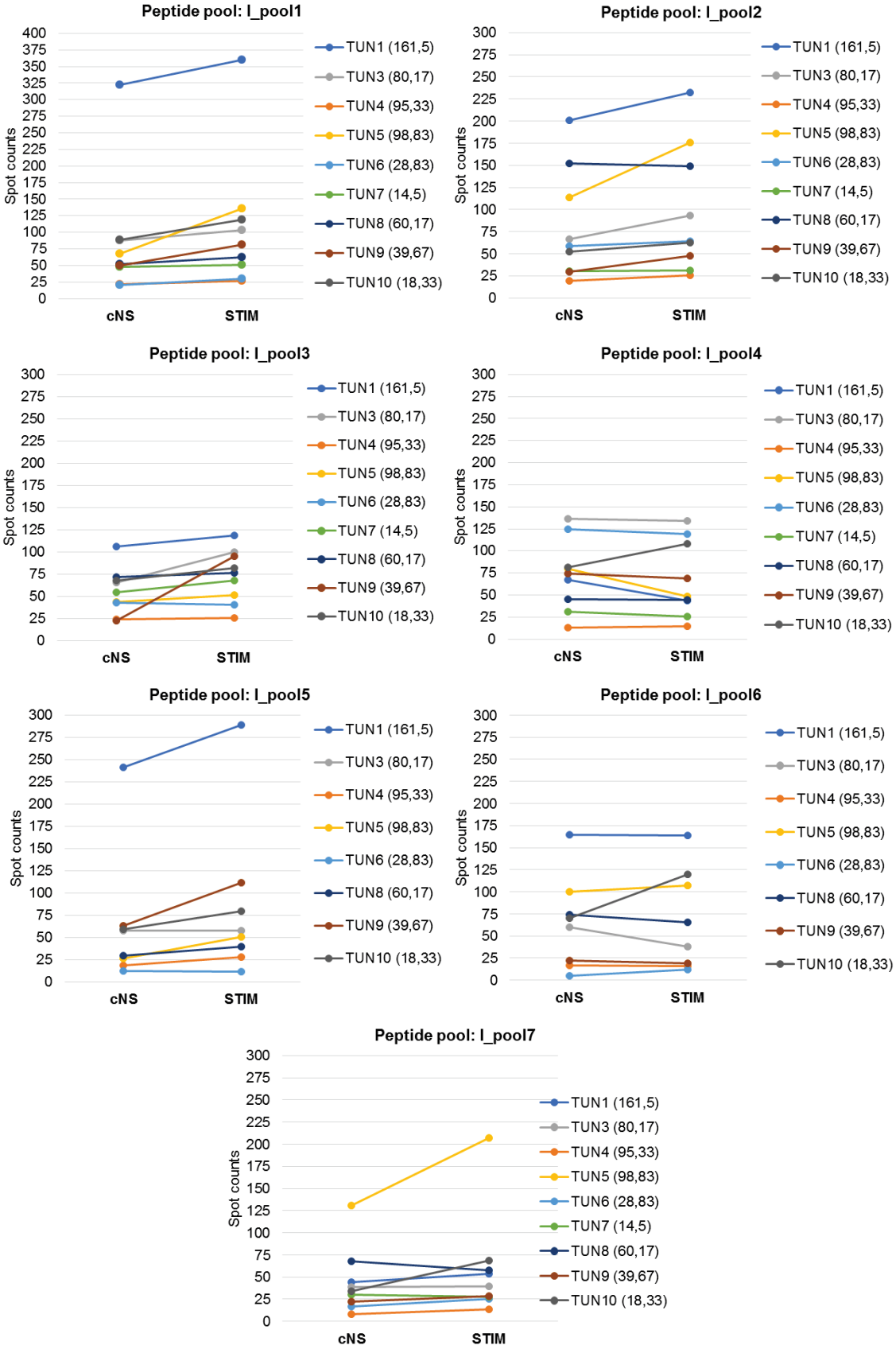




Appendix V.2

OVERALL RESULTS – HLA-CLASS I Pools 1-7

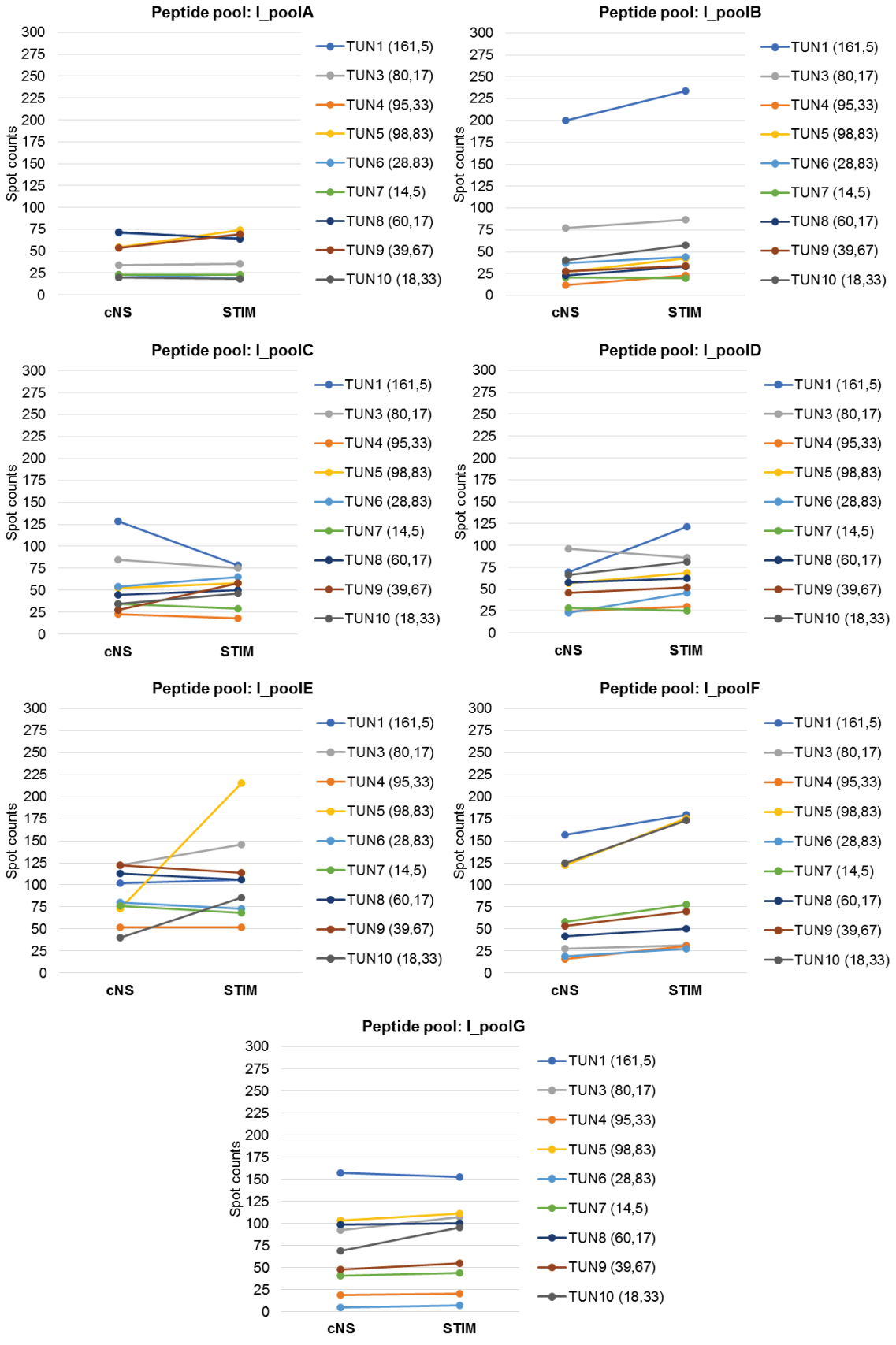
cNS = culture well not stimulated in ELISPOT
NeverStimulated average in the legend



Appendix V.2

OVERALL RESULTS – HLA-CLASS I Pools A-B

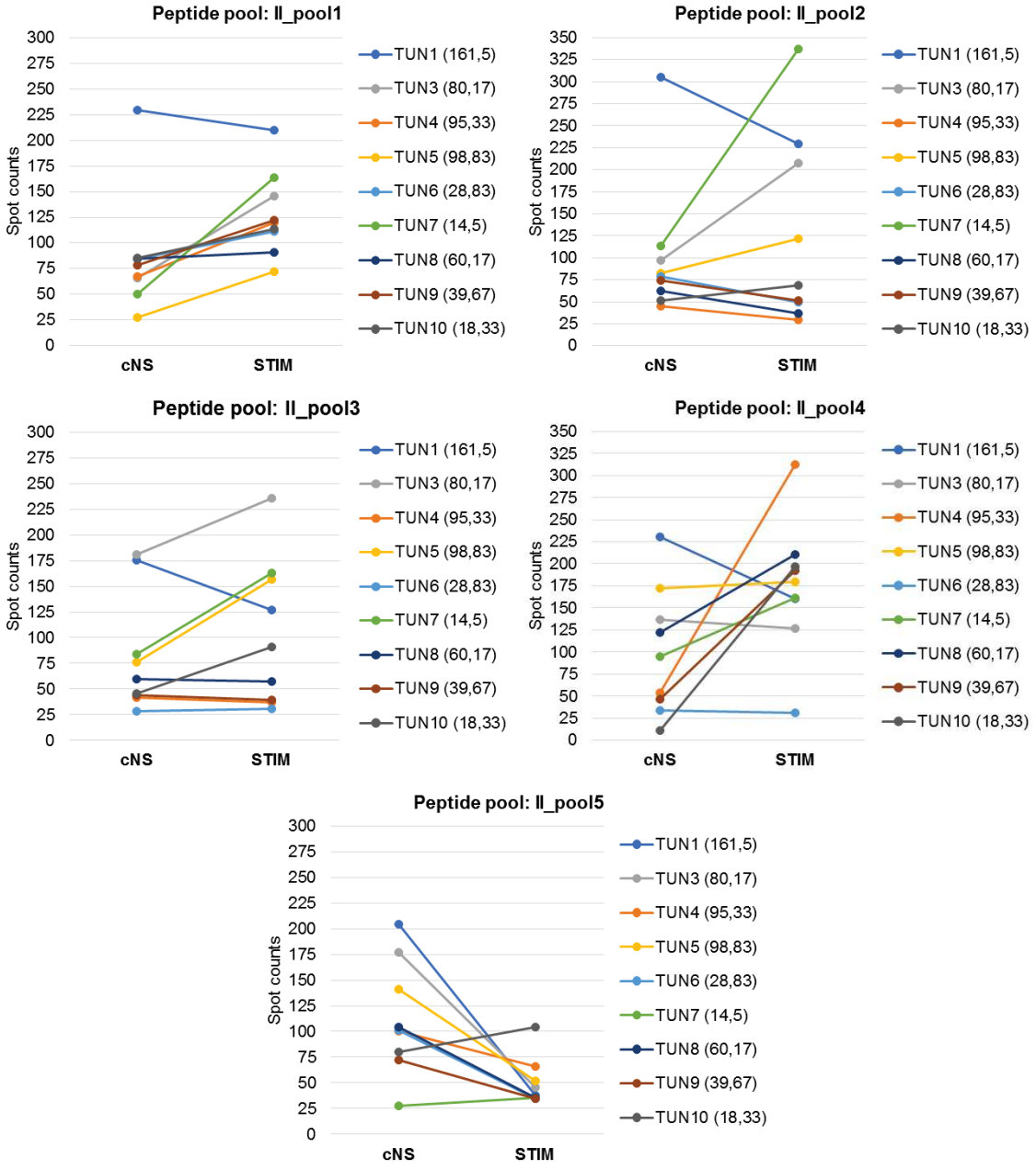
cNS = culture well not stimulated in ELISPOT
NeverStimulated average in the legend



Appendix V.2

OVERALL RESULTS – HLA-CLASS II Pools 1-5

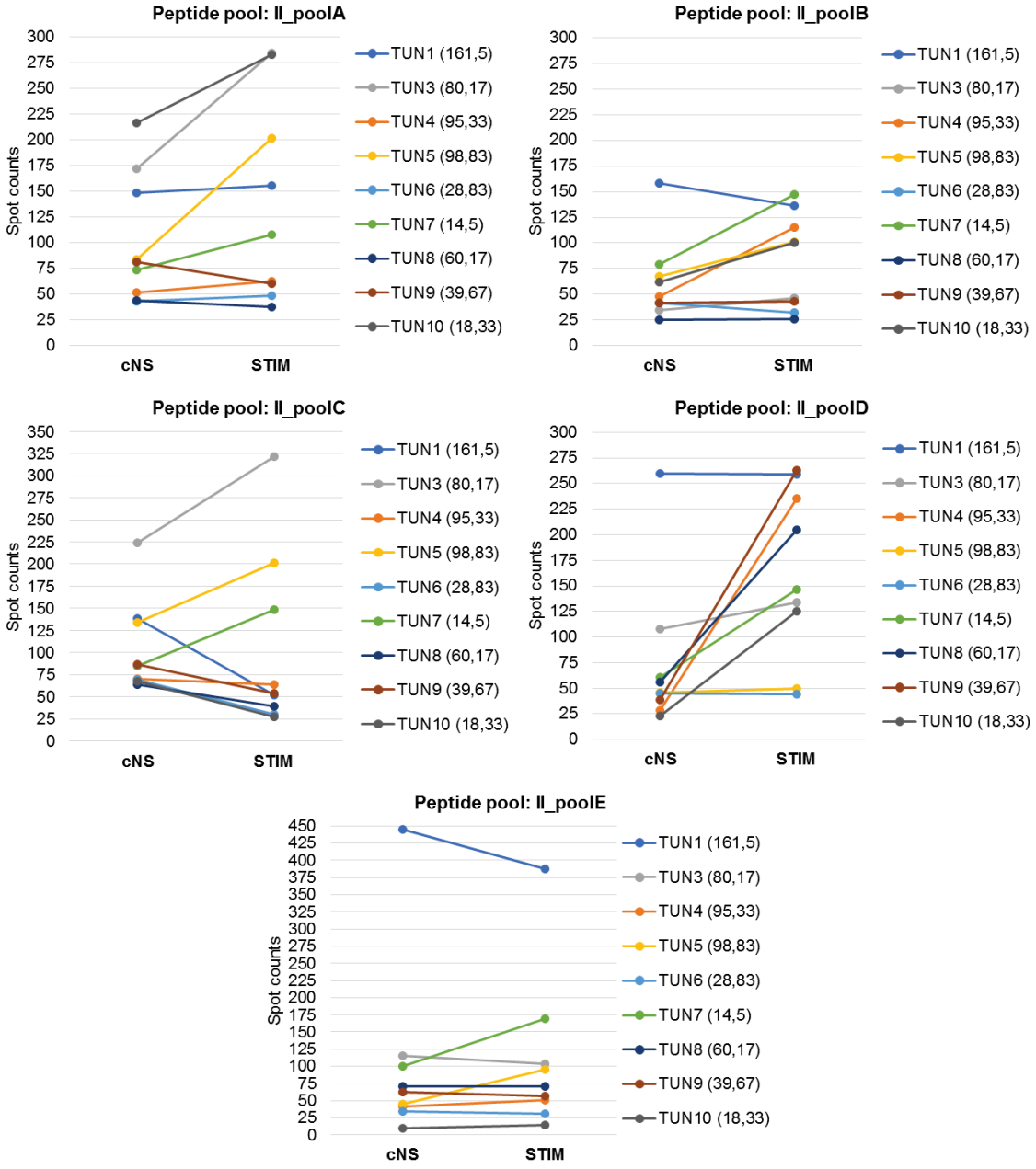
cNS = culture well not stimulated in ELISPOT
NeverStimulated average in the legend



Appendix V.2

OVERALL RESULTS – HLA-CLASS II Pools A-E

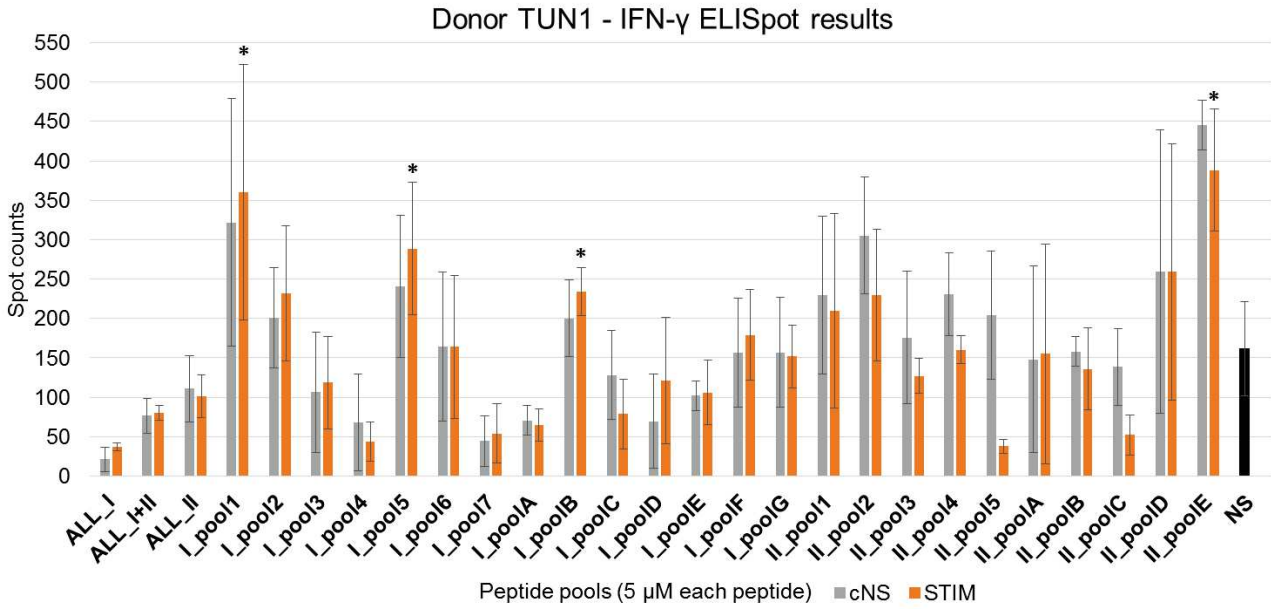
cNS = culture well not stimulated in ELISPOT
NeverStimulated average in the legend



Appendix V.2

TUN1

M / Age 33
1 scar (leg)



Donor ID	HLA-A1	HLA-A1 supertype	HLA-A2	HLA-A2 supertype	HLA-B1	HLA-B1 supertype	HLA-B2	HLA-B2 supertype	HLA-C1	HLA-C2	HLA-DRB1	HLA-DRB2	HLA-DQB1	HLA-DQB2	HLA-DPB1	HLA-DPB2
TUN1	30:01:01G	A01 A03	02:05:01G	A02	50:01:01G	B44	42:01:01	B07	06:02:01G	17:01:01G	03:02:01	15:02:01G	04:02	06:01	01:AGXDR	04:AHPJH

Class I

	1	2	3	4	5	6	7
A	pi1	pi2	pi3	pi4	pi5	pi6	pi7
B	**pi8**	**pi9**	**pi10**	**pi11**	**pi12**	**pi13**	**pi14**
C	pi15	pi16	pi17	pi18	pi19	pi20	pi21
D	pi22	pi23	pi24	pi25	pi26	pi27	pi28
E	pi29	pi30	pi31	pi32	pi33	pi34	pi35
F	pi36	pi37	pi38	pi39	pi40	pi41	pi42
G	pi43	pi44	pi45	pi46	pi47	pi48	pi49

pi8	**B27_1**
pi12	**B58_4**

Class II

	1	2	3	4	5
A	pii1	pii2	pii3	pii4	pii5
B	pii6	pii7	pii8	pii9	pii10
C	pii11	pii12	pii13	pii14	pii15
D	pii16	pii17	pii18	pii19	pii20
E	pii21	pii22	pii23	pii24	

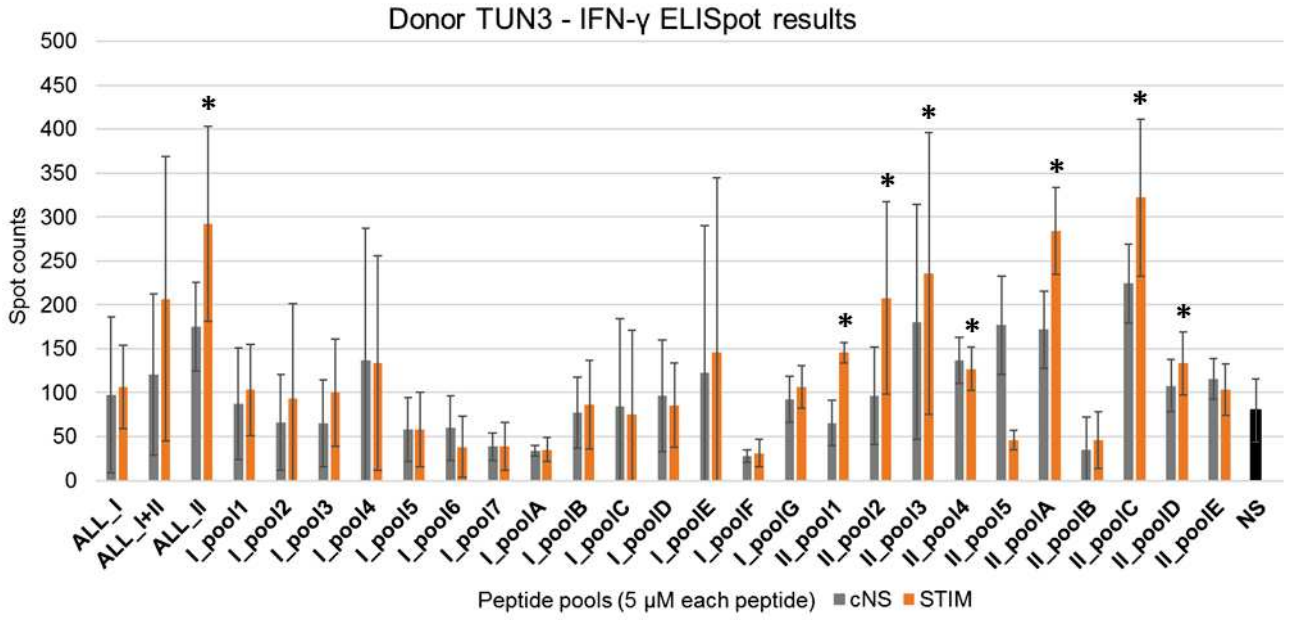
Appendix V.2

TUN1
M / Age 33 1 scar (leg)

Appendix V.2

TUN3

M / Age 33
1 scar (leg)



Donor ID	HLA-A1	HLA-A1 supertype	HLA-A2	HLA-A2 supertype	HLA-B1	HLA-B1 supertype	HLA-B2	HLA-B2 supertype	HLA-C1	HLA-C2	HLA-DRB1	HLA-DRB2	HLA-DQB1	HLA-DQB2	HLA-DPB1	HLA-DPB2
TUN3	01:01:01G	A01	03:01:01G	A03	15:03:01G	B27	44:02:01G	B44	02:10:01G	16:04:01G	11:04:01	10:01:01	05:01	03:01:01	17:01:01	23:01:01

No class I peptide pools

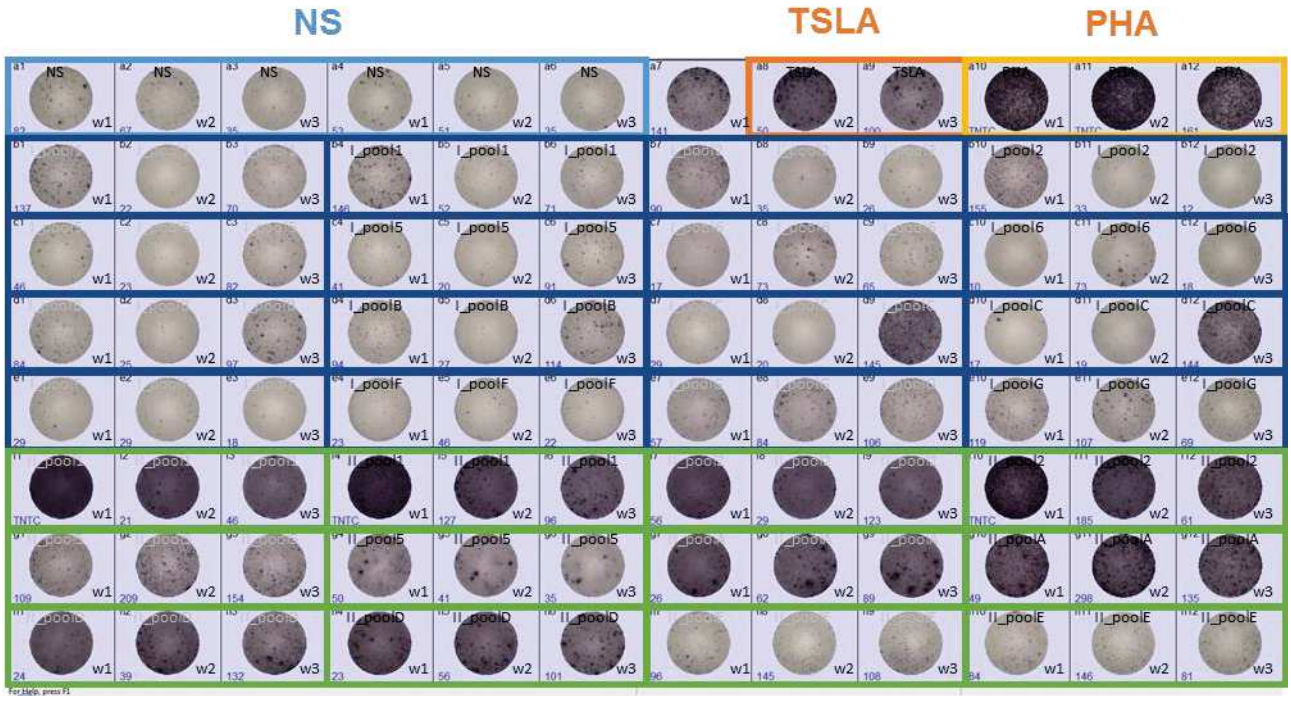
Class II

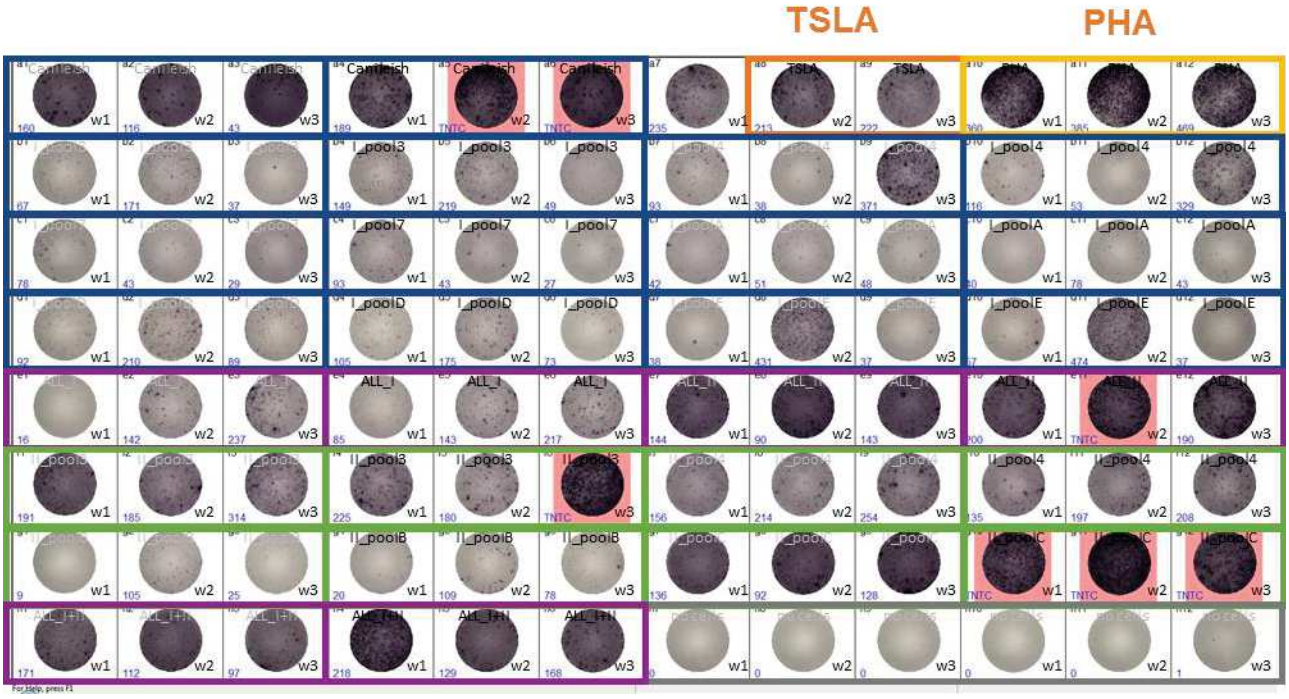
	1	2	3	4	5
A	**pii1**	**pii2**	**pii3**	**pii4**	pii5
B	pii6	pii7	pii8	pii9	pii10
C	**pii11**	**pii12**	**pii13**	**pii14**	pii15
D	**pii16**	**pii17**	**pii18**	**pii19**	pii20
E	pii21	pii22	pii23	pii24	

pii1	14_b12
pii2	16_b13
pii3	3_a11
pii4	22_b4
pii11	7_a14
pii12	12_a32
pii13	21_b4
pii14	17_b15
pii16	4_a11
pii17	5_a12
pii18	9_a25
pii19	13_b12

Appendix V.2

TUN3
M / Age 33 1 scar (leg)



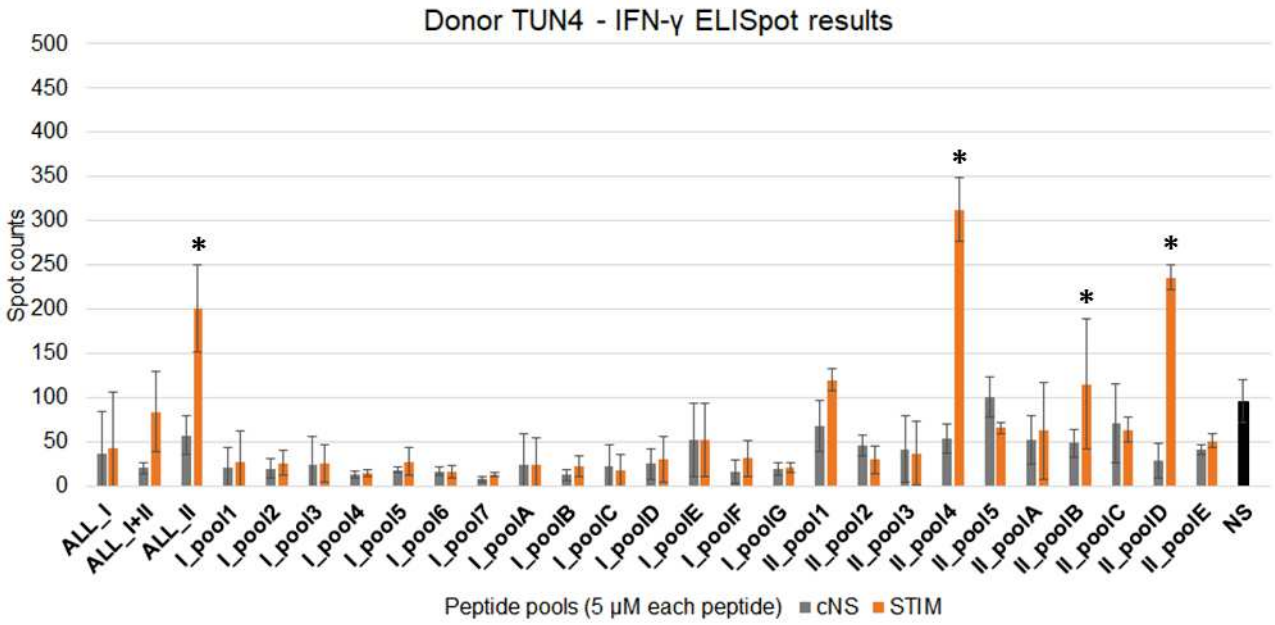


Appendix V.2

TUN4

F / Age 36

2 scars (nose, wrist)



Donor ID	HLA-A1	HLA-A1 supertype	HLA-A2	HLA-A2 supertype	HLA-B1	HLA-B1 supertype	HLA-B2	HLA-B2 supertype	HLA-C1	HLA-C2	HLA-DRB1	HLA-DRB2	HLA-DQB1	HLA-DQB2	HLA-DPB1	HLA-DPB2
TUN4	23:01:01G	A24	02:22:01G	A02	58:01:01G	B58	44:03:01G	B44	03:02:01G	04:01:01G	07:DFRJ	07:DFRJ	03:03	02:GKDU	01:01:02G	17:JECX

No class I peptide pools

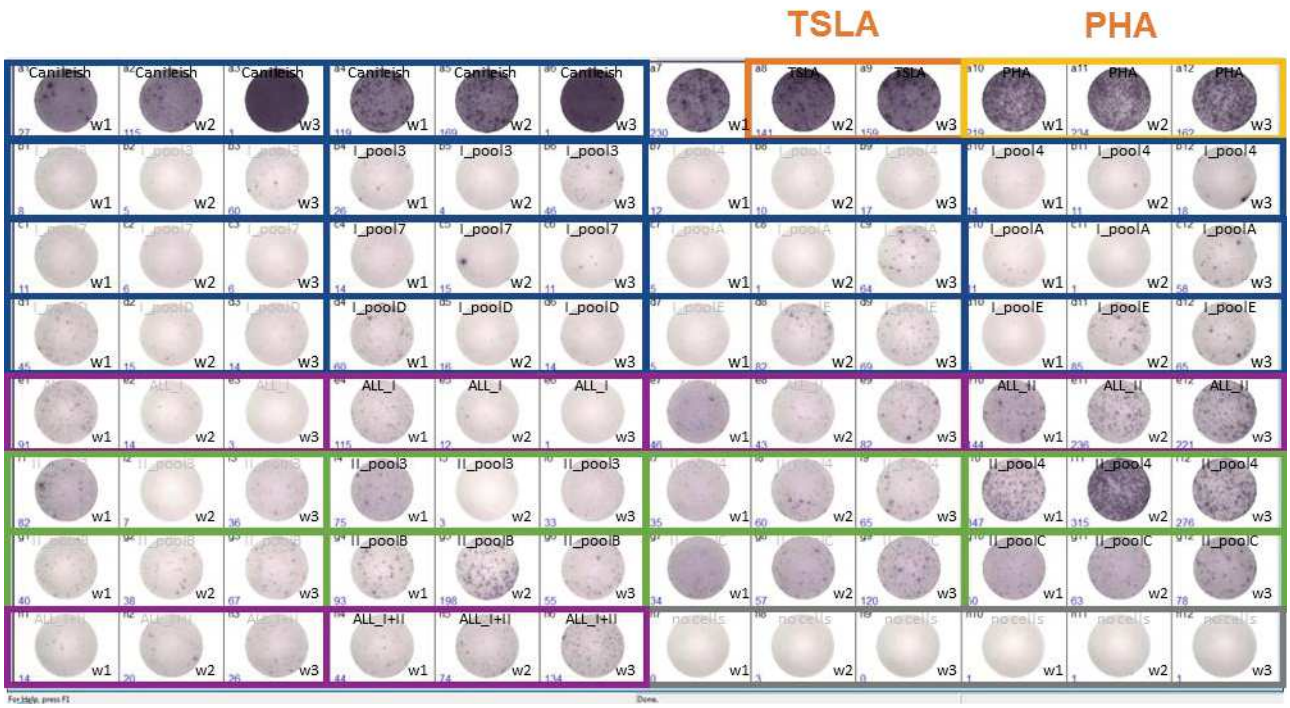
Class II

	1	2	3	4	5
A	pii1	pii2	pii3	pii4	pii5
B	pii6	pii7	pii8	pii9	pii10
C	pii11	pii12	pii13	pii14	pii15
D	pii16	pii17	pii18	**pii19**	pii20
E	pii21	pii22	pii23	pii24	

pii19	13_b12

Appendix V.2

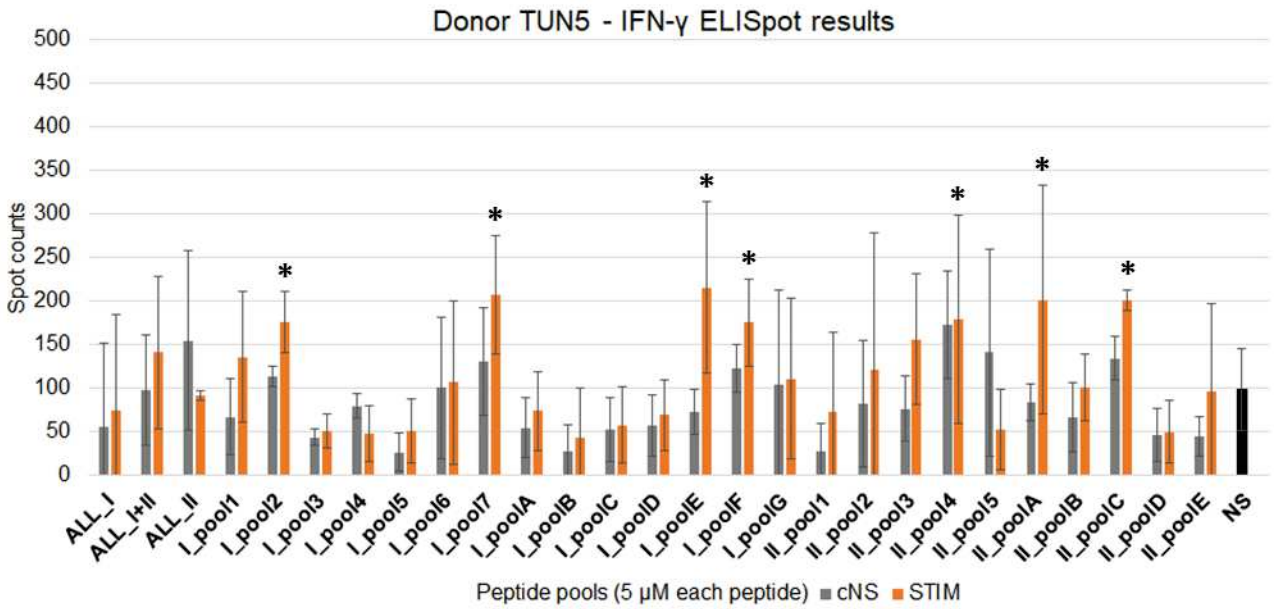
TUN4
F / Age 36 2 scars (nose, wrist)



Appendix V.2

TUN5

M / Age 49
1 scar (leg)



Donor ID	HLA-A1	HLA-A1 supertype	HLA-A2	HLA-A2 supertype	HLA-B1	HLA-B1 supertype	HLA-B2	HLA-B2 supertype	HLA-C1	HLA-C2	HLA-DRB1	HLA-DRB2	HLA-DQB1	HLA-DQB2	HLA-DPB1	HLA-DPB2
TUN5	26:01:01G	A26	34:02:01	A03	07:02:01G	B07	08:01:01G	B08	07:01:01G	07:02:01G	03:01:01	15:01:01	06:02:01	02:01:01	01:AETTG	04:ADCGE

Class I

	1	2	3	4	5	6	7
A	pi1	pi2	pi3	pi4	pi5	pi6	**pi7**
B	pi8	pi9	pi10	pi11	pi12	pi13	**pi14**
C	pi15	pi16	pi17	pi18	pi19	pi20	**pi21**
D	pi22	pi23	pi24	pi25	pi26	pi27	**pi28**
E	pi29	**pi30**	pi31	pi32	pi33	pi34	**pi35**
F	pi36	**pi37**	pi38	pi39	pi40	pi41	**pi42**
G	pi43	pi44	pi45	pi46	pi47	pi48	**pi49**

pi30	**B62_2**
pi35	**B44_1**
pi37	B8_4
pi42	**B44_4**

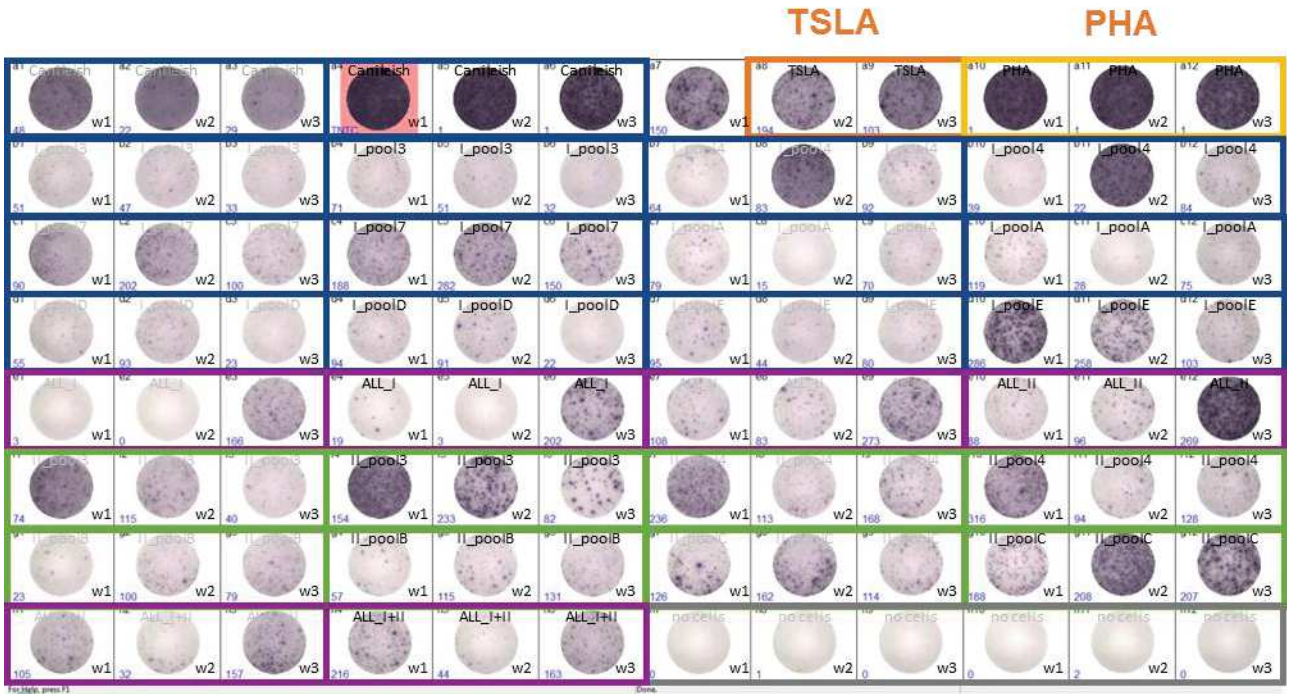
Class II

	1	2	3	4	5
A	pii1	pii2	pii3	**pii4**	pii5
B	pii6	pii7	pii8	pii9	pii10
C	pii11	pii12	pii13	**pii14**	pii15
D	pii16	pii17	pii18	pii19	pii20
E	pii21	pii22	pii23	pii24	

pii4	22_b4
pii14	17_b15

Appendix V.2

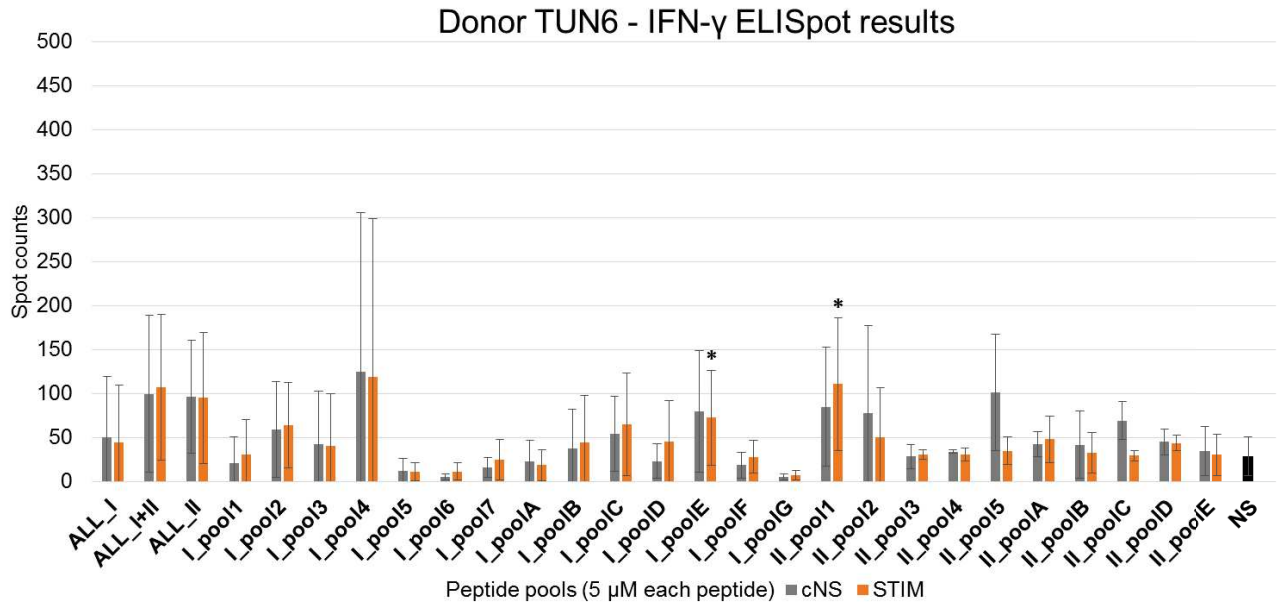
TUN5
M / Age 49 1 scar (leg)



Appendix V.2

TUN6

M / Age 26
1 scar (leg)



Donor ID	HLA-A1	HLA-A1 supertype	HLA-A2	HLA-A2 supertype	HLA-B1	HLA-B1 supertype	HLA-B2	HLA-B2 supertype	HLA-C1	HLA-C2	HLA-DRB1	HLA-DRB2	HLA-DQB1	HLA-DQB2	HLA-DPB1	HLA-DPB2
TUN6	01:01:01G	A01	29:02:01G	A01 A24	15:03:01G	B27	08:01:01G	B08	02:10:01G	07:01:01G	03:01:01	07:DFRJ	02:01:01	02:GKDU	11:AWXGR	04:01

Class I

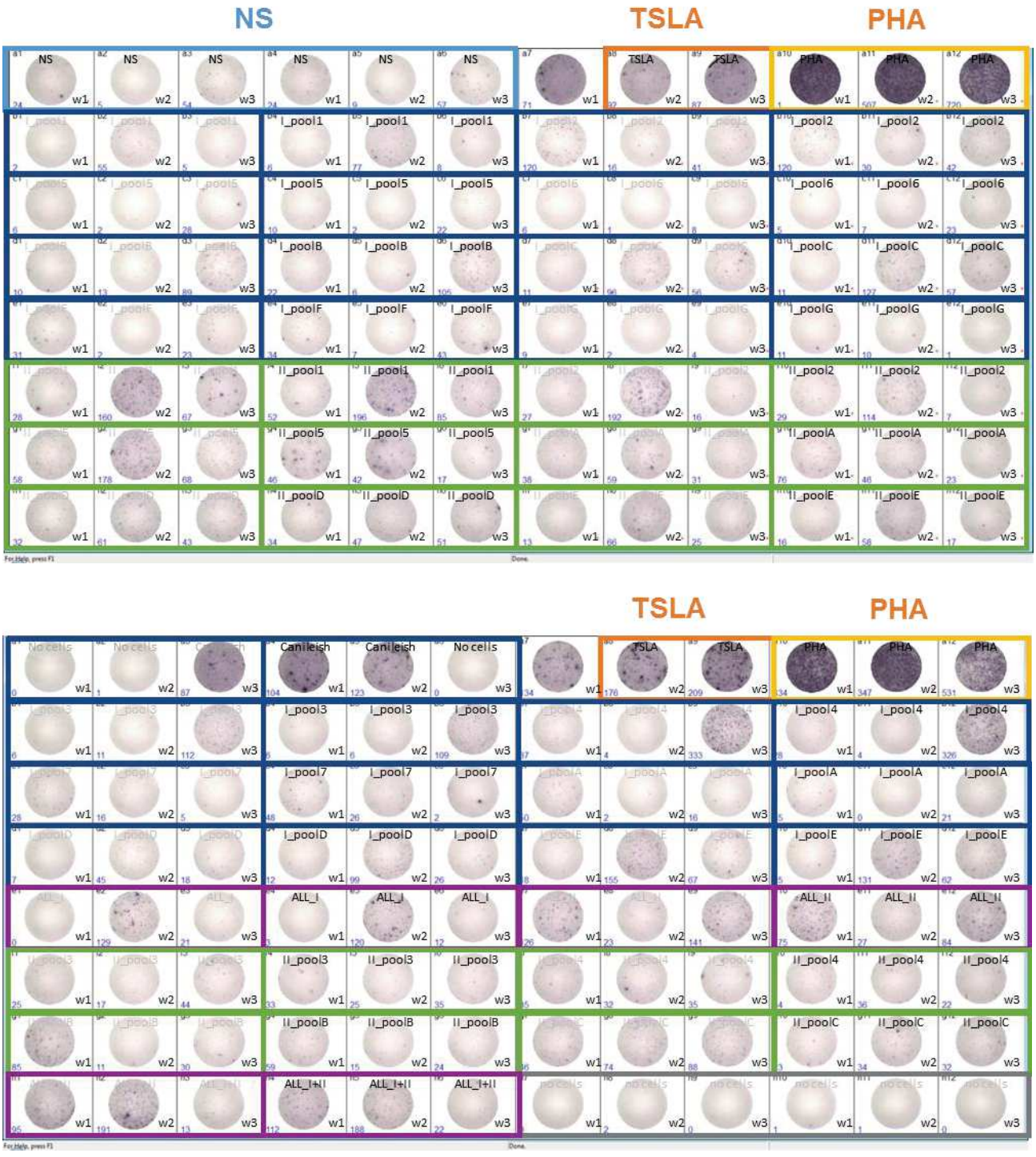
	1	2	3	4	5	6	7
A	pi1	pi2	pi3	pi4	pi5	pi6	pi7
B	pi8	pi9	pi10	pi11	pi12	pi13	pi14
C	pi15	pi16	pi17	pi18	pi19	pi20	pi21
D	pi22	pi23	pi24	pi25	pi26	pi27	pi28
E	pi29	pi30	pi31	pi32	pi33	pi34	pi35
F	pi36	pi37	pi38	pi39	pi40	pi41	pi42
G	pi43	pi44	pi45	pi46	pi47	pi48	pi49

Class II

	1	2	3	4	5
A	pii1	pii2	pii3	pii4	pii5
B	pii6	pii7	pii8	pii9	pii10
C	pii11	pii12	pii13	pii14	pii15
D	pii16	pii17	pii18	pii19	pii20
E	pii21	pii22	pii23	pii24	

Appendix V.2

TUN6
M / Age 26 1 scar (leg)

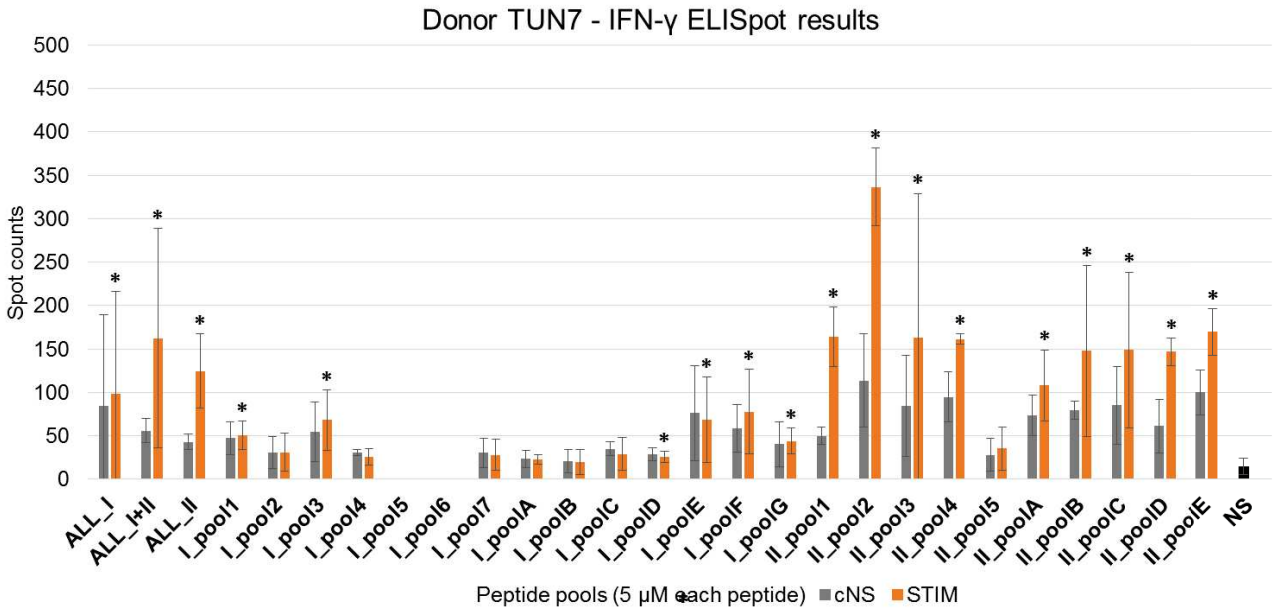


Appendix V.2

TUN7

M / Age 51

5 scars (arm, leg)



Donor ID	HLA-A1	HLA-A1 supertype	HLA-A2	HLA-A2 supertype	HLA-B1	HLA-B1 supertype	HLA-B2	HLA-B2 supertype	HLA-C1	HLA-C2	HLA-DRB1	HLA-DRB2	HLA-DQB1	HLA-DQB2	HLA-DPB1	HLA-DPB2
TUN7	29:02:01G	A01 A24	32:01:01G	A01	35:03:01G	B07	44:03:01G	B44	16:01:01G	04:01:01G	11:01:01	15:01:01	06:02:01	03:01:01	04:FNVS	04:HJMR

Class I

	1	2	3	4	5	6	7
A	pi1	pi2	pi3	pi4	pi5	pi6	pi7
B	pi8	pi9	pi10	pi11	pi12	pi13	pi14
C	pi15	pi16	pi17	pi18	pi19	pi20	pi21
D	pi22	pi23	pi24	pi25	pi26	pi27	pi28
E	pi29	pi30	pi31	pi32	pi33	pi34	pi35
F	pi36	pi37	pi38	pi39	pi40	pi41	pi42
G	pi43	pi44	pi45	pi46	pi47	pi48	pi49

pi22	A1_1
pi24	A3_3
pi29	A24_1
pi31	A3_1
pi36	A2_4
pi38	B7_2
pi43	B58_5
pi45	B8_1

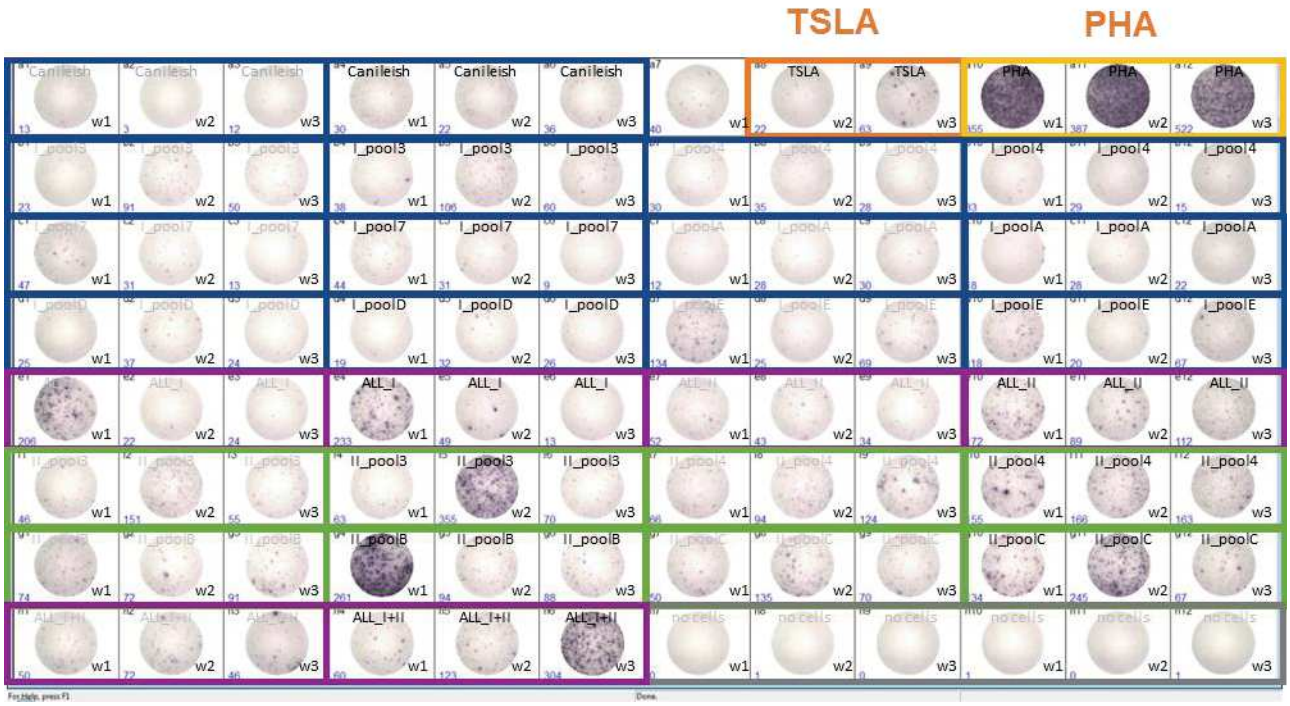
Class II

	1	2	3	4	5
A	pii1	pii2	pii3	pii4	pii5
B	pii6	pii7	pii8	pii9	pii10
C	pii11	pii12	pii13	pii14	pii15
D	pii16	pii17	pii18	pii19	pii20
E	pii21	pii22	pii23	pii24	

All except pii5, 10, 15 and 20

Appendix V.2

TUN7
M / Age 51 5 scars (arm, leg)

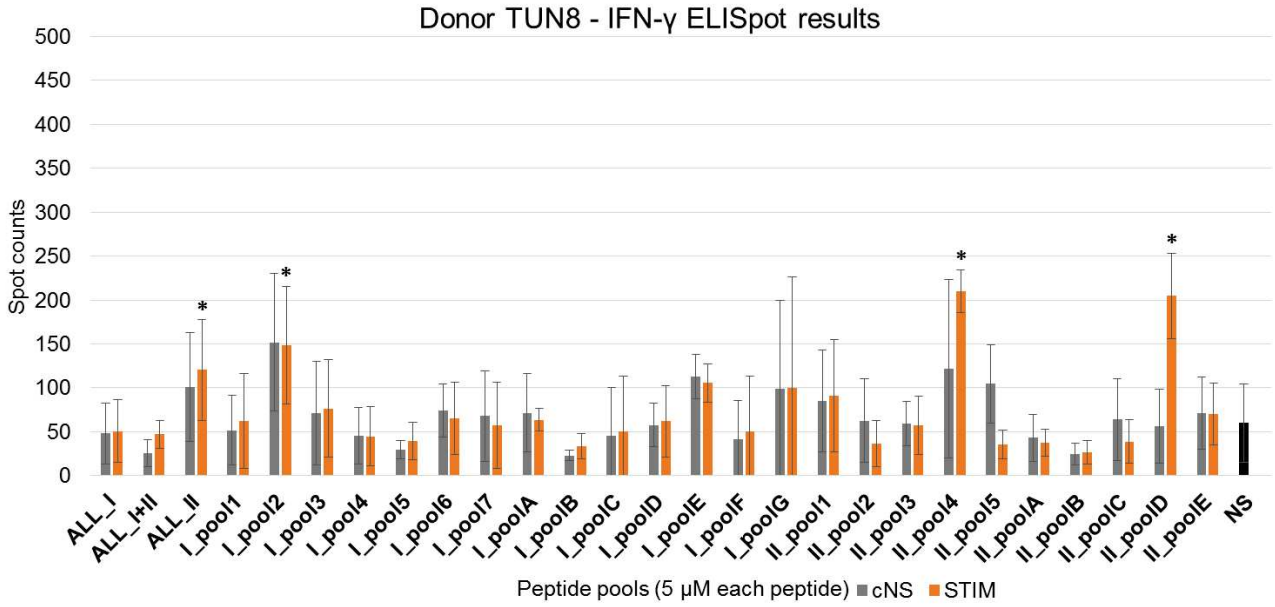


Appendix V.2

TUN8

M / Age 45

2 scars (abdomen, foot)



Donor ID	HLA-A1	HLA-A1 supertype	HLA-A2	HLA-A2 supertype	HLA-B1	HLA-B1 supertype	HLA-B2	HLA-B2 supertype	HLA-C1	HLA-C2	HLA-DRB1	HLA-DRB2	HLA-DQB1	HLA-DQB2	HLA-DPB1	HLA-DPB2
TUN8	01:01:01G	A01	02:01:01G	A02	52:01:01G	B62	47:01:01G	unclassified	12:02:01G	06:02:01G	15:02:01G	07:DFRJ	03:03:02	06:01	09:01:01	04:01

Class I

	1	2	3	4	5	6	7
A	pi1	pi2	pi3	pi4	pi5	pi6	pi7
B	pi8	pi9	pi10	pi11	pi12	pi13	pi14
C	pi15	pi16	pi17	pi18	pi19	pi20	pi21
D	pi22	pi23	pi24	pi25	pi26	pi27	pi28
E	pi29	pi30	pi31	pi32	pi33	pi34	pi35
F	pi36	pi37	pi38	pi39	pi40	pi41	pi42
G	pi43	pi44	pi45	pi46	pi47	pi48	pi49

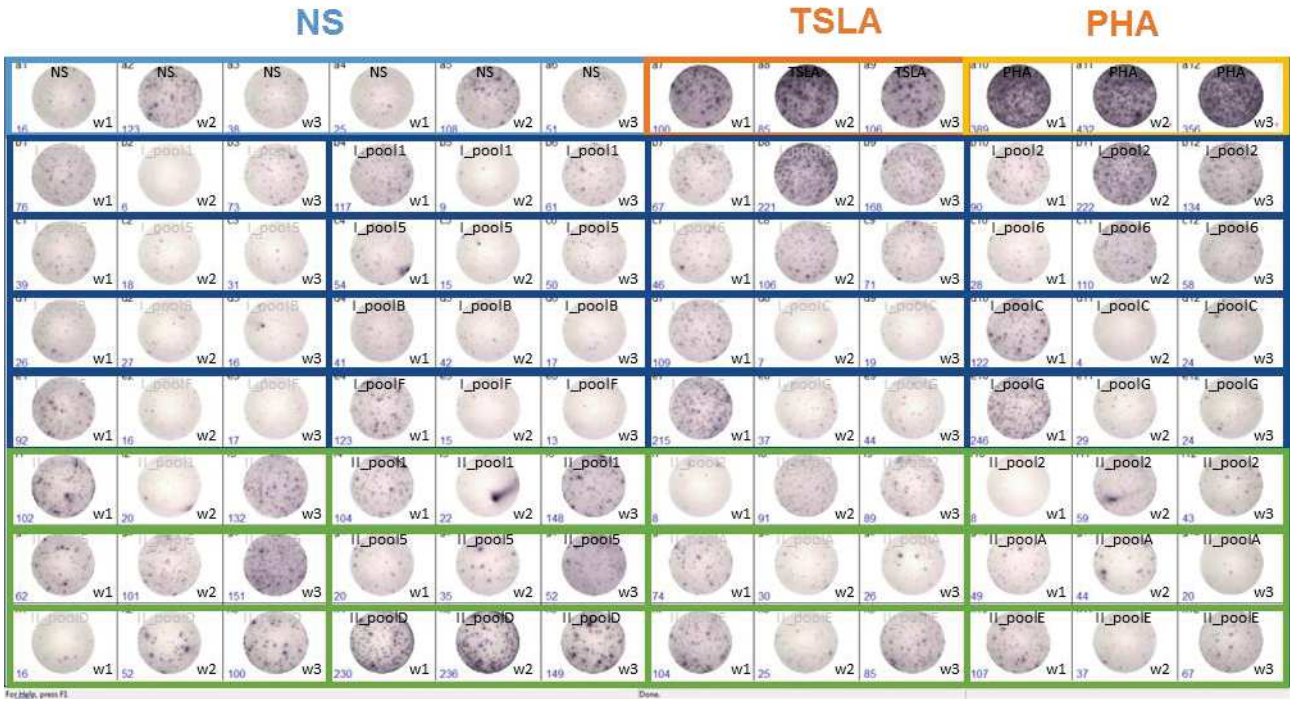
Class II

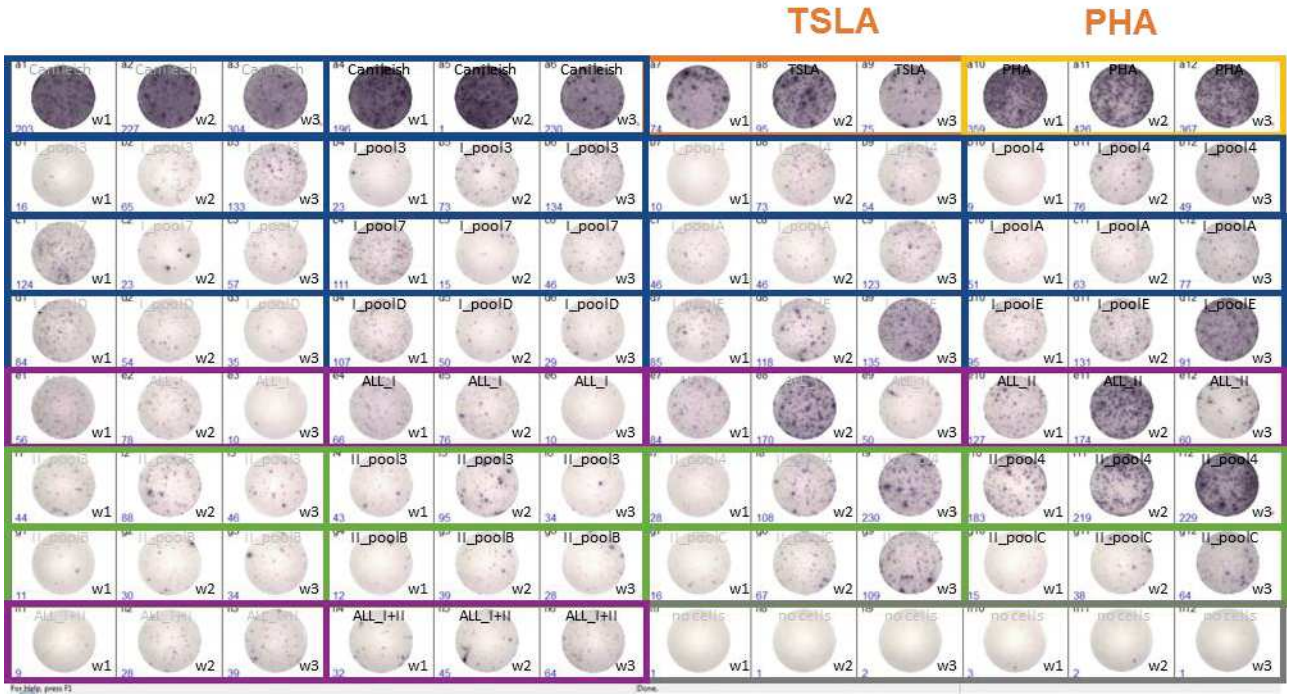
	1	2	3	4	5
A	pii1	pii2	pii3	pii4	pii5
B	pii6	pii7	pii8	pii9	pii10
C	pii11	pii12	pii13	pii14	pii15
D	pii16	pii17	pii18	**pii19**	pii20
E	pii21	pii22	pii23	pii24	

pii19	13_b12

Appendix V.2

TUN8
M / Age 45 2 scars (abdomen, foot)



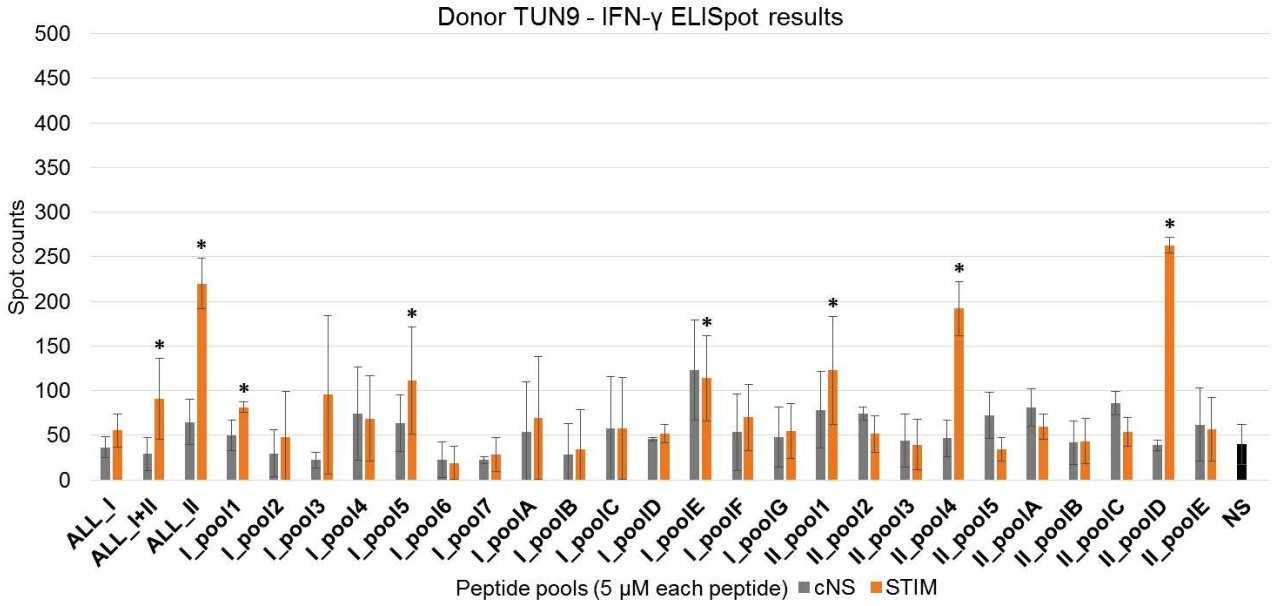


Appendix V.2

TUN9

M / Age 64

1 scars (leg)



Donor ID	HLA-A1	HLA-A1 supertype	HLA-A2	HLA-A2 supertype	HLA-B1	HLA-B1 supertype	HLA-B2	HLA-B2 supertype	HLA-C1	HLA-C2	HLA-DRB1	HLA-DRB2	HLA-DQB1	HLA-DQB2	HLA-DPB1	HLA-DPB2
TUN9	02:02:01G	A02	32:01:01G	A01	53:01:01G	B07	35:03:01G	B07	04:01:01G	04:01:01G	11:01:02	11:01:01	03:19:01	03:01:01	13:AKXDF	04:02

Class I

	1	2	3	4	5	6	7
A	pi1	pi2	pi3	pi4	pi5	pi6	pi7
B	pi8	pi9	pi10	pi11	pi12	pi13	pi14
C	pi15	pi16	pi17	pi18	pi19	pi20	pi21
D	pi22	pi23	pi24	pi25	pi26	pi27	pi28
E	**pi29**	pi30	pi31	pi32	**pi33**	pi34	pi35
F	pi36	pi37	pi38	pi39	pi40	pi41	pi42
G	pi43	pi44	pi45	pi46	pi47	pi48	pi49

pi29	**A24_1**
pi33	**B27_5**

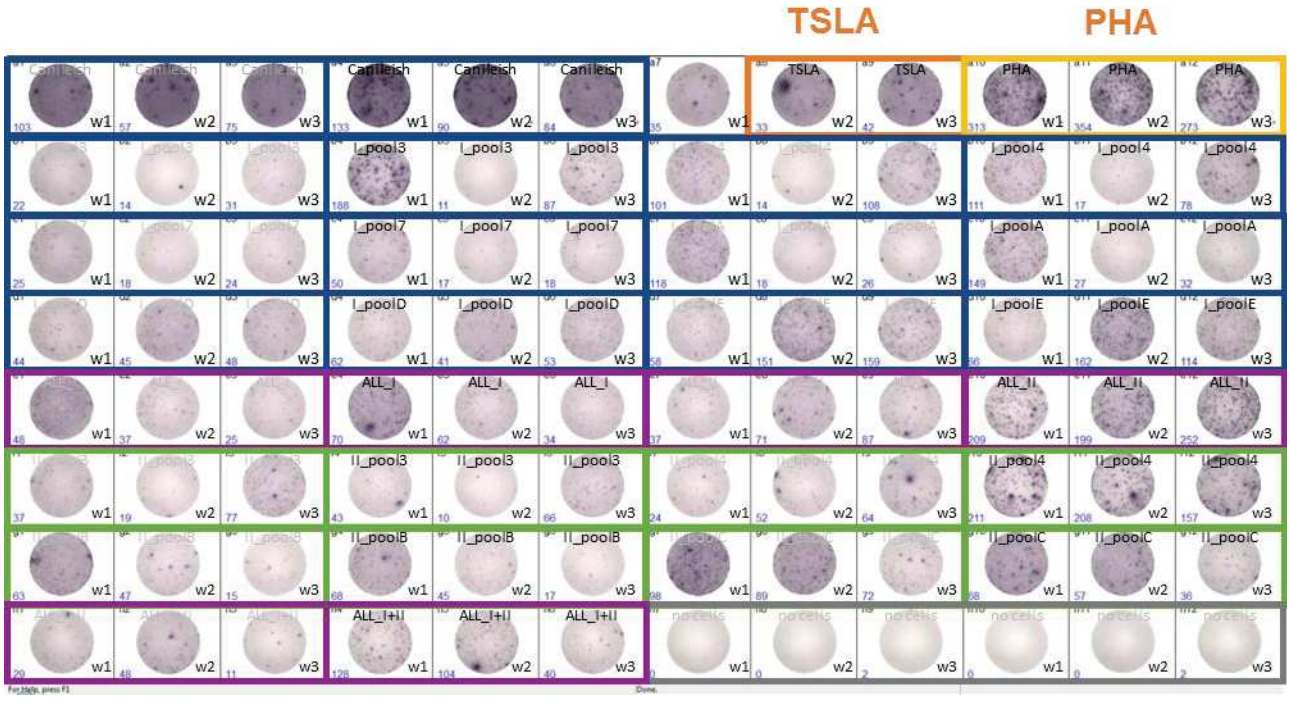
Class II

	1	2	3	4	5
A	pii1	pii2	pii3	pii4	pii5
B	pii6	pii7	pii8	pii9	pii10
C	pii11	pii12	pii13	pii14	pii15
D	pii16	pii17	pii18	**pii19**	pii20
E	pii21	pii22	pii23	pii24	

pii16	4_a11
pii19	13_b12

Appendix V.2

TUN9
M / Age 64 1 scars (leg)

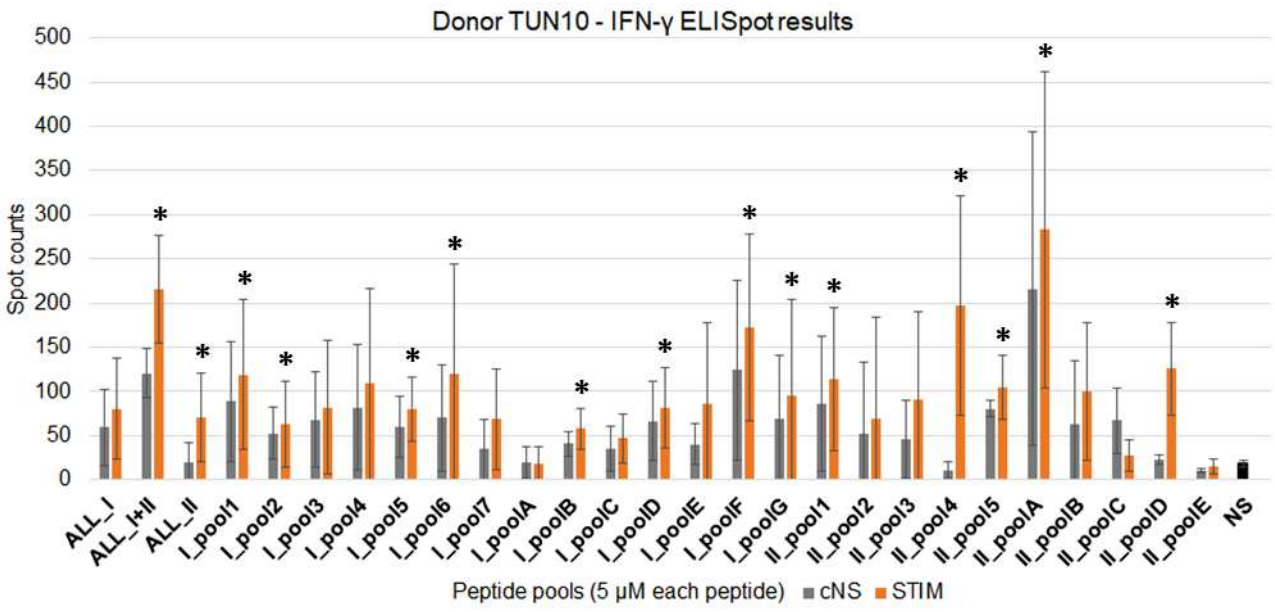


Appendix V.2

TUN10

M / Age 64

1 scars (leg)



Donor ID	HLA-A1	HLA-A1 supertype	HLA-A2	HLA-A2 supertype	HLA-B1	HLA-B1 supertype	HLA-B2	HLA-B2 supertype	HLA-C1	HLA-C2	HLA-DRB1	HLA-DRB2	HLA-DQB1	HLA-DQB2	HLA-DPB1	HLA-DPB2
TUN10	02:05:01G	A02	02:02:01G	A02	35:08:01G	B07	44:03:02G	B44	07:01:01G	04:01:01G	04:03:01	11:01:02	06:02:01	03:02:01	17:JECX	04:ADCGE

Class I

	1	2	3	4	5	6	7
A	pi1	pi2	pi3	pi4	pi5	pi6	pi7
B	**pi8**	**pi9**	pi10	pi11	**pi12**	**pi13**	pi14
C	pi15	pi16	pi17	pi18	pi19	pi20	pi21
D	**pi22**	**pi23**	pi24	pi25	**pi26**	**pi27**	pi28
E	pi29	pi30	pi31	pi32	pi33	pi34	pi35
F	**pi36**	**pi37**	pi38	pi39	**pi40**	**pi41**	pi42
G	**pi43**	**pi44**	pi45	pi46	**pi47**	**pi48**	pi49

pi8	B27_1
pi9	**B7_3**
pi12	B58_4
pi13	A24_3
pi22	A1_1
pi23	B8_3
pi26	**B44_5**
pi27	B58_3
pi36	**A2_4**
pi37	B8_4
pi40	**B7_5**
pi41	B27_3
pi43	B58_5
pi44	A3_2
pi47	A1_4
pi48	**B7_4**

Class II

	1	2	3	4	5
A	**pii1**	pii2	pii3	**pii4**	**pii5**
B	pii6	pii7	pii8	pii9	pii10
C	pii11	pii12	pii13	pii14	pii15
D	**pii16**	pii17	pii18	**pii19**	**pii20**
E	pii21	pii22	pii23	pii24	

pii1	14_b12
pii4	22_b4
pii5	18_b15
pii16	4_a11
pii19	13_b12
pii20	10_a25

Appendix V.2

TUN10
M / Age 64 1 scars (leg)