

# Comparative Omics Analysis of Historic and Recent Isolates of *Bordetella pertussis* and Effects of Genome Rearrangements on Evolution

Ana Dienstbier, Fabian Amman, Denisa Petráčková, Daniel Štipl, Jan Čapek, Jana Zavadilová, Kateřina Fabiánová, Jakub Držmíšek, Dilip Kumar, Mark Wildung, Derek Pouchnik, Branislav Večerek

Despite high vaccination coverage, pertussis is increasing in many industrialized countries, including the Czech Republic. To better understand *Bordetella pertussis* resurgence, we analyzed historic strains and recent clinical isolates by using a comparative omics approach. Whole-genome sequencing showed that historic and recent isolates of *B. pertussis* have substantial variation in genome organization and form separate phylogenetic clusters. Subsequent RNA sequence analysis and liquid chromatography with mass tandem spectrometry analyses showed that these variations translated into discretely separated transcriptomic and proteomic profiles. When compared with historic strains, recent isolates showed increased expression of flagellar genes and genes involved in lipopolysaccharide biosynthesis and decreased expression of polysaccharide capsule genes. Compared with reference strain Tohama I, all strains had increased expression and production of the type III secretion system apparatus. We detected the potential link between observed effects and insertion sequence element-induced changes in gene context only for a few genes.

*Bordetella pertussis* is a gram-negative, strictly human pathogen of the respiratory tract and the major causative agent of whooping cough. This highly contagious disease is especially severe in infants and

remains a major cause of infant illness and death worldwide, predominantly in industrialized countries (1). Although pertussis is a vaccine-preventable disease, increased incidence is being observed in some countries that have highly vaccinated populations, including the Czech Republic (2–4). Although several factors are contributing to pertussis resurgence in these countries (5–7), the 2 prominent factors are incomplete and short-lived immunity induced by current acellular vaccines (8–10) and genetic variation, leading to escape from immunity by antigenic variation (11–13).

*B. pertussis* has an efficient mechanism of genome structure diversification because it contains >200 copies of insertion sequence 481 (IS481) in its genome (14). IS element-mediated homologous recombination results in excision or insertion of flanking genome regions and leads to genome reduction and decay (14–16), as well as genome rearrangements (17,18) and large duplications (19). Furthermore, a previous study indicated that gene order rearrangements associated with IS elements can alter gene expression profiles in *B. pertussis* (20). Recently, we have shown that, besides their effect on genome structure and stability, ISs can affect expression profiles of neighboring genes by IS element-specific promoters (21).

On the basis of these observations, we hypothesized that strains with different genomic organization should display altered global transcriptomic and, consequently, proteomic profiles, and thereby genome rearrangements might contribute to strain variation and adaptation. To validate this assumption, we have performed genomic, transcriptomic, and proteomic analyses of recent clinical isolates from the Czech

Author affiliations: Institute of Microbiology of the Czech Academy of Sciences, Prague, Czech Republic (A. Dienstbier, D. Petráčková, D. Štipl, J. Čapek, J. Držmíšek, D. Kumar, B. Večerek); Institute for Theoretical Chemistry of the University of Vienna, Vienna, Austria (F. Amman); National Institute of Public Health, Prague (J. Zavadilová, K. Fabiánová); Washington State University, Pullman, Washington, USA (M. Wildung, D. Pouchnik)

DOI: <https://doi.org/10.3201/eid2701.191541>

Republic obtained during 2008–2015, previously characterized vaccine strains isolated during 1954–1965 (22) (hereafter referred to as historic strains), and the reference strain Tohama I.

## Materials and Methods

### Bacterial Strains and Growth Conditions

Recent isolates of *B. pertussis* from the Czech Republic were obtained from the National Institute of Public Health in Prague (Table 1). Historic strains from the Czech Republic (22) and reference strain Tohama I (23) have been described. All strains were cultivated on Bordet-Gengou agar plates supplemented with 15% sheep blood for 3–4 days at 37°C. For liquid cultures, bacteria were grown in Stainer-Scholte medium (24) supplemented with 0.1% cyclodextrin and 0.5% casamino acids (Difco, <https://www.fishersci.com>) at 37°C. To harvest samples for DNA, RNA, and protein isolation, *B. pertussis* cells were grown overnight in Stainer-Scholte medium to mid-exponential phase of growth (optical density ≈1.0). Three independent cultivations were performed to collect 3 biologic replicates of each of the strains for RNA and protein isolation.

### Genomic Analyses

For the genome organization analysis, genomic sequences were aligned by using the progressive Mauve algorithm (25) and clustered on the basis of their genome organization similarity by using the maximum-likelihood for the gene order pipeline (26). For single-nucleotide polymorphism (SNP) analysis, IS elements within the genomes were masked with Ns, and resulting sequences were aligned by using Mugsy software (27). SNPs were extracted by using custom scripts (<https://genohub.com>). Maximum-parsimony phylogenetic analysis was performed on sequences with masked IS elements by using the kSNP3 program with a k number of 23 (28). The unrooted phylogenetic tree was visualized by using iTOL (29).

### RNA Isolation, Sequencing, and Data Analysis

We provide information on RNA isolation, sequencing, and data analysis (Appendix 1, <https://wwwnc.cdc.gov/EID/article/27/1/19-1541-App1.pdf>). RNA sequencing data from sequencing runs were deposited in the European Nucleotide Archive under project accession no. PRJEB34096. We defined significance as a *q* value <0.05 (*p* value adjusted for multiple testing correction [Appendix 1]).

### Protein Sample Preparation and Proteomic Analysis

We compiled information on protein sample preparation and label-free proteomic analysis, which used liquid chromatography with mass tandem spectrometry analyses (Appendix 1). Proteomics data were deposited in the ProteomeXchange Consortium by using the PRIDE partner repository with the dataset identifier PXD015184.

## Results

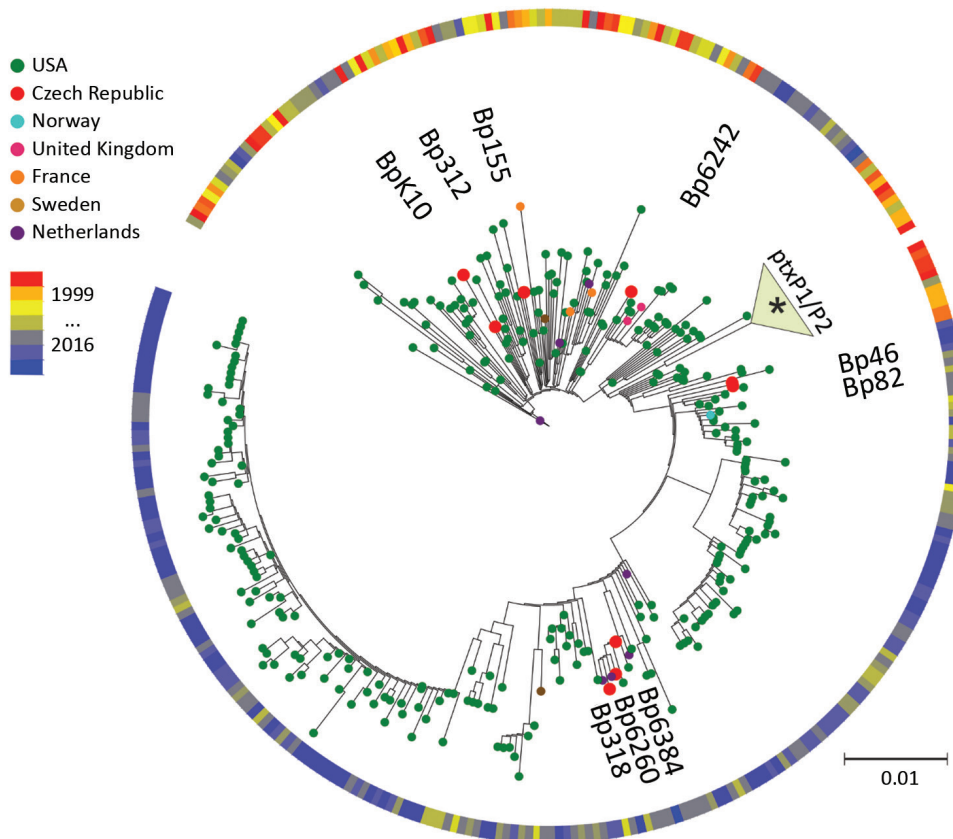
### Genome Organization and Content of Recent Isolates

We determined complete *de novo* genome assemblies of 9 recent isolates of *B. pertussis* strains collected in the Czech Republic during 2008–2015 from patients representing different age groups and vaccination status (Table 1). Genotyping of recent strains showed that they belonged to *ptxP3* lineage. SNP-based phylogenetic analysis of these strains and >350 complete *B. pertussis* genome sequences currently deposited in GenBank (Appendix 2 Table 1, <https://wwwnc.cdc.gov/EID/article/27/1/19-1541-App2.xlsx>) showed that recent *B. pertussis* isolates cluster with *ptxP3* isolates from other countries, demonstrating worldwide spread and lack of geographic signature (Figure 1). The genome alignment of recent isolates and previously characterized historic strains belonging to the *ptxP1* lineage (22) showed that all genomes contain large-scale structural rearrangements (Figure 2, panel A). According to their genome organization, sequenced strains could be classified into 8 groups. None of the historic strains clustered with any of the

**Table 1.** Characteristics of recent isolates of *Bordetella pertussis* and 9 infected patients, Czech Republic\*

Year	Strain information		Patient information	
	Name	Genotype	Age, y/sex	Vaccination status
2008	Bp155	<i>ptxP3, fim2-1, fim3B, prn2</i>	<1/M	Not vaccinated
2008	Bp312	<i>ptxP3, fim2-1, fim3B, prn2</i>	45/F	wP
2012	Bp6260	<i>ptxP3, fim2-1, fim3A, prn2</i>	<1/F	Not vaccinated
2012	Bp6242	<i>ptxP3, fim2-1, fim3B, prn2</i>	67/F	Not vaccinated
2012	Bp6384	<i>ptxP3, fim2-1, fim3A, prn2</i>	69/M	Not vaccinated
2012	K10	<i>ptxP3, fim2-1, fim3B, prn3</i>	8/F	aP
2014	Bp82	<i>ptxP3, fim2-1, fim3A, prn2</i>	14/F	wP plus aP
2014	Bp46	<i>ptxP3, fim2-1, fim3A, prn2</i>	15/M	wP plus aP
2015	Bp318	<i>ptxP3, fim2-1, fim3A, prn2</i>	7/F	aP

\*aP, acellular vaccine; wP, whole-cell vaccine.



**Figure 1.** Maximum-parsimony, unrooted phylogenetic tree based on single-nucleotide polymorphism analysis of available genome sequences of *Bordetella pertussis*. Red dots indicate recent isolates from the Czech Republic. Year and country of isolation are color-coded. The 3 black dots indicate time span between 1999 and 2016. Asterisk (\*) indicates association of historic strains from the Czech Republic with the *ptxP1/ptxP2* clade. Scale bar indicates nucleotide substitutions per site.

recent isolates. The separation of these 2 groups was verified by using a maximum-likelihood phylogenetic tree, which was constructed on the basis of the genome organization of all sequenced strains from the Czech Republic (Figure 2, panel B).

To check whether there are also sequence signatures differentiating these 2 groups of strains, we performed SNP analysis, which yielded 35 SNPs (15 synonymous, 13 nonsynonymous, and 7 intergenic) (Appendix 2 Table 2) and distinguished historic and recent isolates. Variants found in historic strains were also present in Tohama I. One of the new SNPs specific for recent isolates was identified in the promoter region of the *bteA* gene, which encodes the type III secretion system (T3SS) effector. Approximately one third of the SNPs have been reported to be specific for the *ptxP3* lineage isolates from other countries (30). When compared with those of historic strains, the genome size of recent isolates was substantially reduced, thereby confirming ongoing gene loss within the global population of *B. pertussis* (Appendix 2 Table 3). Analysis of genome alignments showed 2 regions of difference (RD) between the 2 groups of strains from the Czech Republic, which corresponded to previously

identified regions RD3 and RD10 (20). Although RD3 (28.7 kb, spans genes *BP0910A–BP0937*) is absent in all recent isolates, the RD10 (25.1 kb, spans genes *BP1948–BP1968*) is absent in all recent isolates and historic strain V67.

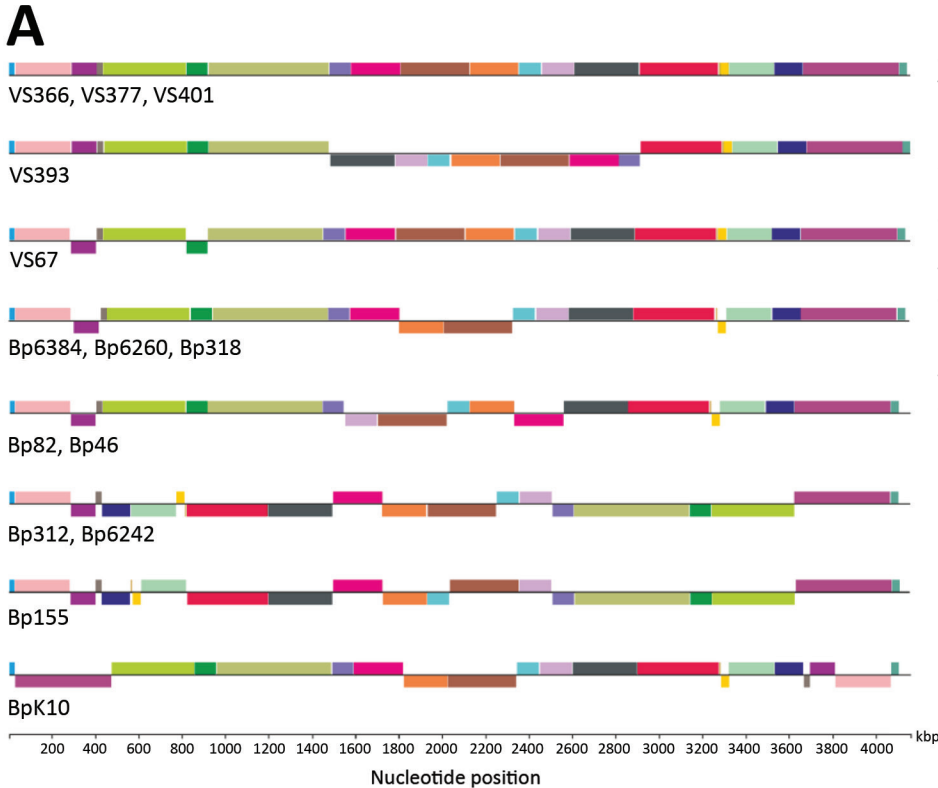
#### Transcriptomic Profiles and Genomic Structure Alterations

Total RNA was isolated from biologic triplicates of *B. pertussis* Tohama I strain; historic strains VS393, VS67, and VS401; and recent isolates Bp318, Bp155, Bp46, Bp6242, and BpK10 and analyzed by using RNA sequencing. These strains were selected on the basis of genome organization and phylogenetic distances to encompass the highest variability among the studied strains (Figure 2, panel B). Hierarchical clustering of RNA sequence data showed that samples from both groups of strains from the Czech Republic clustered separately from each other and from Tohama I (Figure 3). Consistent with phylogenetic analysis (Figure 2, panel B), we found that samples of strain VS67 formed a separate cluster. These analyses suggested that among historic strains, the VS67 strain displays closest distance to recent strains, which is consistent with our

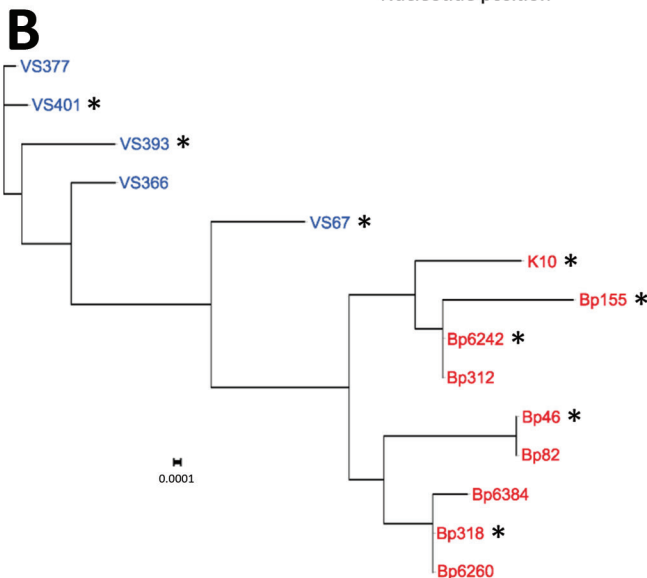
previous observation that the VS67 strain clusters together with a recent U.S. isolate (22).

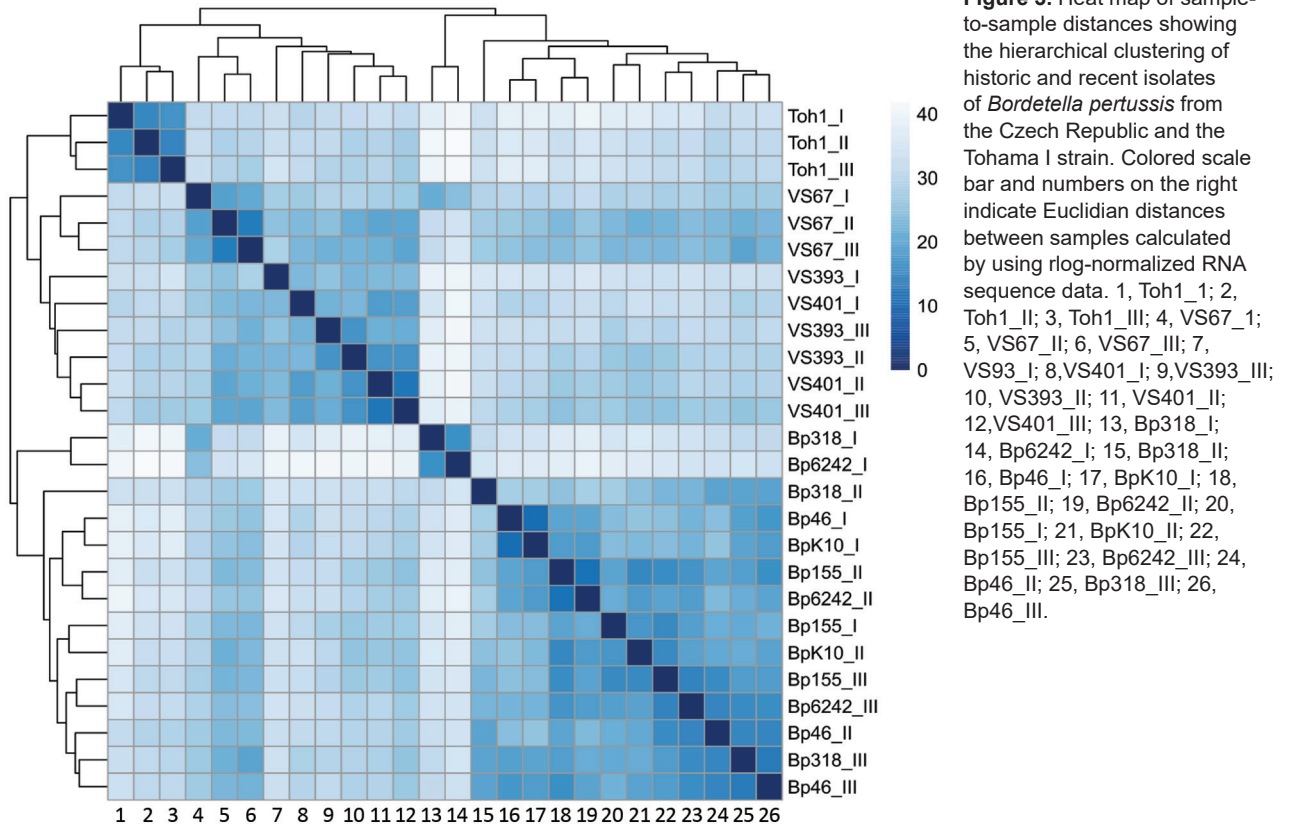
Differential expression (DE) analysis identified 78, 124, and 115 significantly ( $q < 0.05$  for all comparisons) modulated *B. pertussis* genes ( $-1 \geq \log_2 FC \geq 1$ ) between recent isolates and historic strains, recent isolates and Tohama I, and historic strains and Tohama I, respectively (Appendix 2 Table 4). Among the DE genes, 30 were up-regulated in recent isolates compared with historic strains, including those encoding

the flagella apparatus (*flgB-J*), LuxR (*BP1969*), and ArsR (*BP2946*) families of transcriptional regulators, phosphoglucomutase (*pgm*), phosphoglucose isomerase (*pgi*), and nicotinate-nucleotide diphosphorylase (*nadC*). Conversely, among the 48 DE genes down-regulated in recent isolates were genes encoding the polysaccharide capsule proteins (*kpsEMT*), several ABC transporters, and central metabolism enzymes, including those involved in tryptophan synthesis (*trpDEG*). Expression of several virulence



**Figure 2.** Genomic analyses of *Bordetella pertussis* isolates from the Czech Republic. A) Genome alignment of historic and recent isolates showing large-scale genome rearrangements. Homologous gene blocks are denoted by the same color. B) Maximum-likelihood phylogenetic tree based on genomic organization of historic (blue) and recent (red) isolates. Asterisk (\*) indicates strains selected for transcriptomic and proteomic analyses. Scale bar indicates nucleotide substitutions per site. kbp, kilobasepairs.





factors, including pertactin, tracheal colonization factor, filamentous hemagglutinin, and pertussis toxin subunit S3, was significantly up-regulated in recent isolates. However, the increase did not reach the 2-fold threshold. A recent isolate-specific SNP, which was identified in the promoter region of the *bteA* gene, did not result in a significant change of gene expression (Appendix 2 Table 4). Among the DE genes that showed increased expression in both groups of strains from the Czech Republic compared with Tohama I, we identified numerous genes within the T3SS *bcs/btr* locus, including *bsp22*, *bopN*, *bopB*, and *bopD* and several genes involved in sulfate metabolism (*cysADITW*).

Gene ontology enrichment (Figure 4, panel A) showed that within the set of genes, which were significantly modulated between recent and historic isolates, categories such as bacterial type flagellum-dependent cell motility, polysaccharide biosynthesis, and tryptophan biosynthesis were highly enriched. Conversely, when we compared both groups of isolates from the Czech Republic to Tohama I (Figure 4, panel C), genes associated with sulfate transmembrane transport, pathogenesis, and protein secretion by the type III secretion system terms were enriched among the DE genes.

### Clustering of Proteomic Profiles of Recent Isolates, Historic Strains, and Tohama I Strain

The cell-associated (bacterial pellets) and cell-free (culture supernatants) fractions of selected *B. pertussis* strain cultures were analyzed by using liquid chromatography with mass tandem mass spectrometry. First, hierarchical clustering of the cell-associated protein profiles showed that consistent with RNA sequencing data, strains from the Czech Republic cluster separately from Tohama I and despite high variability among biologic replicates, separation of historic and recent strains was still apparent (Figure 5, panel A). Similarly, hierarchical clustering of secreted proteins indicated that recent isolates cluster apart from historic strains and Tohama I (Figure 5, panel B).

Label-free quantification of cell-associated protein intensities identified 33, 132, and 87 proteins showing significantly changed abundance between recent and historic strains, recent strains and Tohama I, and historic strains and Tohama I, respectively (Appendix 2 Table 5). In good correlation with transcriptomic data, we found that protein levels of hydroxymethylglutaryl-CoA lyase (BP3695), ArsR family transcriptional factor BP2946, small lipoprotein BP2782, and nicotinate-nucleotide diphosphorylase NadC were increased, but levels of several central metabolism

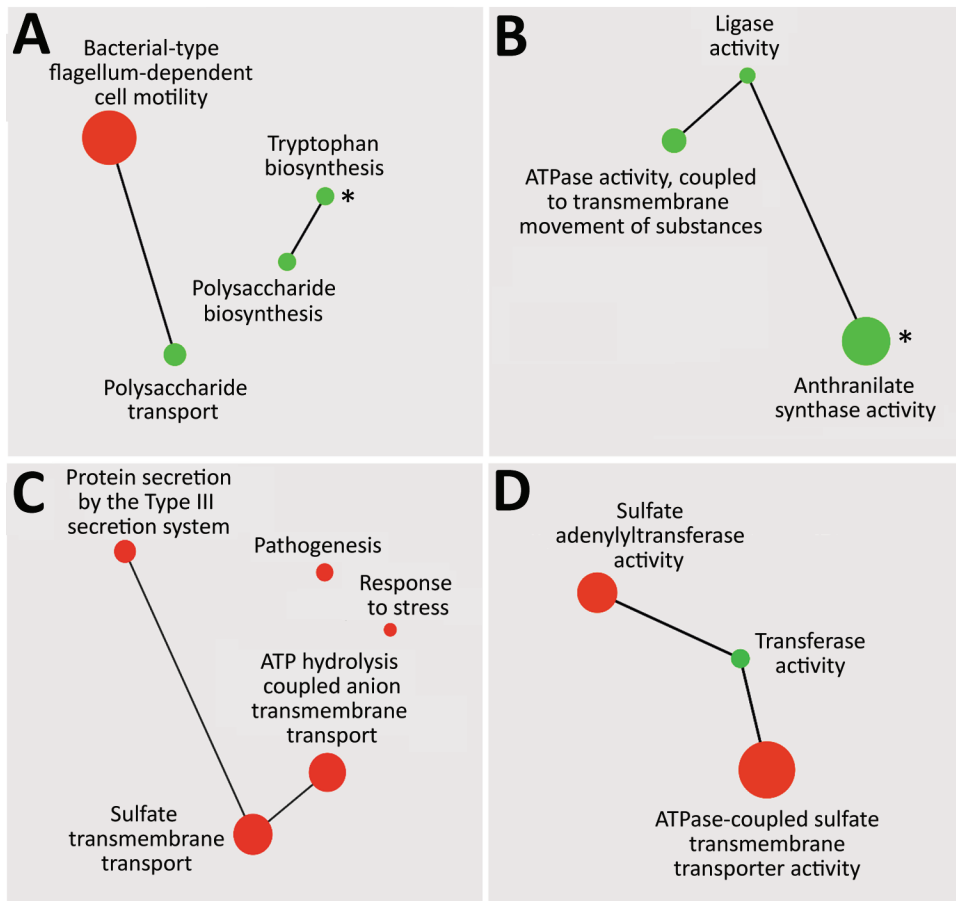
enzymes (BP0624–BP0629), tryptophan synthesis genes, and polysaccharide capsule proteins were decreased in recent isolates compared with historic strains (Table 2). Also in support of RNA sequencing data, we determined that various components of the T3SS apparatus and several proteins involved in metabolism of sulfate showed increased abundance compared with Tohama I (Table 3). All strains from the Czech Republic produced pertactin, but flagellar proteins were not detected in any of the studied strains.

Label-free quantification analysis of secreted proteins showed that 121, 130, and 43 proteins displayed significant changes in abundance between recent and historic strains, recent strains and Tohama I, and historic strains and Tohama I, respectively (Appendix 2 Table 6). Levels of several secreted proteins were in good agreement with transcriptomic data (e.g., strains from the Czech Republic and in particular recent isolates secreted increased amounts of several T3SS proteins compared with Tohama I) (Table 3). Abundance of all pertussis toxin subunits and associated transport protein PtlE was higher in recent isolates than in historic strains,

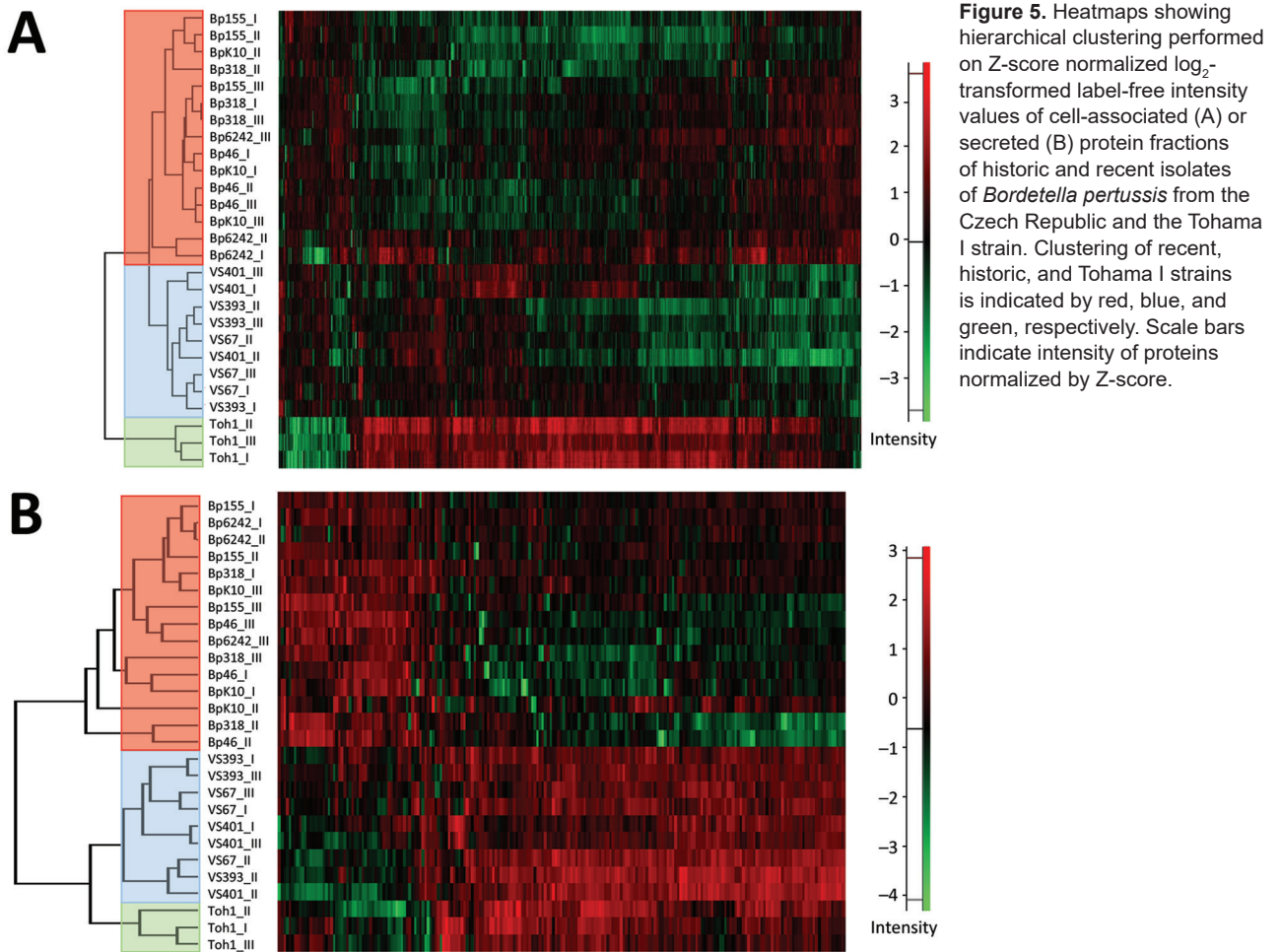
which suggests that some of the differences between the *ptxP1* and *ptxp3* strains are also manifested at the level of protein secretion.

#### Changes in Genome Structure and Alterations in Gene Expression

Considering the observed differences between historic and recent strains, we attempted to track back the modulated gene expression profiles to alterations in the genome sequence and structure. Because our SNP analysis (Table 2) suggested that there were no SNPs that could explain the altered expression of the DE genes, we have additionally inspected the upstream regions of all DE genes for larger sequence variations. We identified such variations in 4 genes. The gene *toh\_02779* (*BP2827*) was preceded by an *IS481* element in Bp155, Bp6242, and BpK10, but not in other strains. In addition, 3 DE genes with an upstream *IS481* element had varying gene context further upstream of the transposase in the studied strains (Table 4). Apparently, the observed differences in expression of these genes could be potentially linked to the upstream IS elements.



**Figure 4.** Gene ontology enrichment analysis of genes down-regulated or up-regulated between recent and historic strains of *Bordetella pertussis* from the Czech Republic (A, B) or between both groups of strains from the Czech Republic and Tohama I (C, D). Enriched terms from the domains' biological process (A, C) and molecular function (B, D) and their catenations, shown as green circles (down-regulated genes) and red circles (up-regulated genes), were summarized by using Revigo (<http://revigo.irb.hr>) and visualized by using Cytoscape (<https://cytoscape.org>) as interactive scatter plots. Circle size indicates level of enrichment. Asterisks (\*) in panels A and B indicate gene ontology terms that were enriched also for genes down-regulated in recent isolates compared with Tohama I.



**Figure 5.** Heatmaps showing hierarchical clustering performed on Z-score normalized  $\log_2$ -transformed label-free intensity values of cell-associated (A) or secreted (B) protein fractions of historic and recent isolates of *Bordetella pertussis* from the Czech Republic and the Tohama I strain. Clustering of recent, historic, and Tohama I strains is indicated by red, blue, and green, respectively. Scale bars indicate intensity of proteins normalized by Z-score.

We then tested the possible effect of genome rearrangements on the distance of the DE genes from the origin of replication (*oriC*), a parameter that can greatly affect gene expression (31). We determined the distance from *oriC* to all the genes significantly deregulated between recent and historic strains, and although expression of some of the genes inversely correlated with the distance from *oriC*, the differences were not significant.

## Discussion

We conducted a comparative study analyzing the link between genomic organization, gene expression profiles, and protein production/secretion in historic and recent strains of *B. pertussis*. Our results indicate that global changes in genomic structures observed between historic and recent isolates of *B. pertussis* from the Czech Republic translated into different gene expression and protein production profiles. Similarly to other countries, the IS element-driven recombination led to large changes in genomic structures and to

considerable gene loss in the isolates from the Czech Republic over the past 50–60 years. Results of our integrated omics analysis support our assumption that genomic rearrangements might affect global expression profiles and phenotypic diversity in *B. pertussis*. Hierarchical clustering of our omics data indicates that strains, which cluster apart at genomic structure level, also have distinct transcriptomic and proteomic profiles.

Given the extent of genome structural variability among both groups of strains, the number of differentially expressed genes was rather low ( $\approx 2\%$  of all coding genes). Earlier DNA microarray studies suggested that gene expression profiles between *ptxP1* strains and recent resurgence-associated *ptxP3* lineage differ only subtly (32,33). Although we have identified an increased number of significantly modulated genes, our data on historic (*ptxP1*) and recent (*ptxP3*) isolates are consistent with these reports. None of the gene expression alterations could be shown to result from nucleotide polymorphism, and only a few could be

**Table 2.** Genes expressing consistently changed RNA and protein levels between recent and historic isolates of *Bordetella pertussis*, Czech Republic\*

Gene in Tohama I	Gene group	Gene name	Transcriptome	Proteome	Annotation
toh_00606	Group_2206	BP0624	-2.2	-2.5	Substrate-CoA ligase
toh_00607	Group_2604	BP0625	-2.3	-3.3	Acyl-CoA dehydrogenase
toh_00609	Group_2725	BP0627	-2.0	-3.3	Enoyl-CoA hydratase/isomerase
toh_00610	Group_895	BP0628	-2.4	-2.8	Pyruvate dehydrogenase component
toh_00611	<i>pdhA</i>	BP0629	-2.3	-2.8	Pyruvate dehydrogenase component
toh_01576	Group_23	WP_003811211.1	-6.8	-3.3	Capsular biosynthesis protein
toh_01584	<i>wza</i>	BP1628	-6.8	-2.2	Capsular polysaccharide export protein
toh_02732	Group_1068	BP2782	-3.9	-3.0	Lipoprotein
toh_02896	Group_128	BP2946	3.6	2.3	<i>ArsR</i> family transcriptional regulator
toh_03214	<i>trpD</i>	BP3262	-2.1	<b>-1.4</b>	Anthranilate phosphoribosyltransferase
toh_03215	<i>trpG</i>	BP3263	-2.4	<b>-1.5</b>	Anthranilate synthase component II
toh_03216	<i>trpE</i>	BP3264	-2.4	<b>-1.4</b>	Anthranilate synthase component I
toh_03637	Group_1064	BP3695	3.3	5.5	Hydroxymethylglutaryl-CoA lyase
toh_03667	<i>nadC</i>	BP3725	2.5	11.8	Nicotinate-nucleotide diphosphorylase

\*Fold change values for recent isolates/historic strains comparison are shown for RNA sequencing and proteomic analyses. Values that did not indicate statistical significance ( $FC \geq 2$ ; adjusted  $p < 0.05$ ) are shown in bold. toh, Tohama

linked to IS element-induced changes in the local gene context. Upstream of the *IS481* element adjacent to the *BP1492* gene, we identified the *BP120* tRNA gene in most historic strains and Tohama I, which is, however, missing in recent isolates and historic strain VS67. Thus, it is possible that the activity of the strong tRNA promoter is responsible for increased expression of the *BP1492* gene in historic strains. Also, the presence of an *IS481* element in front of the *BP2827* gene in Bp155, Bp6242, and BpK10 strains might explain the increased expression of this gene in recent isolates. In support of this possibility, when compared with all other strains lacking this IS element, these 3 recent isolates showed highly increased expression of this gene. It is also possible that the observed differences in gene expression between historic and recent strains result from changes in genome organization or gene loss. Bacterial chromosome organization appears to favor a conserved gene order (34), and changes in genome architecture and topology can affect gene expression (35,36). Therefore, it is conceivable that genome rearrangements, resulting in changes in gene order and orientation or in large deletions, might affect transcriptomic profiles in *B. pertussis*.

We have identified 2 previously characterized regions of difference between historic and recent strains, which might offer alternative explanation for the observed differences in gene expression (20,37). Consistent with these reports, our reports found that

genes within RD3 and RD10 are missing in all recent isolates. RD3 contains 2 putative transcriptional regulators (*BP0924* and *BP0928*) of unknown function. Thus, it is probable that absence of these regulators in recent isolates might be accountable for some of the identified alterations in gene expression.

Furthermore, RNA sequencing analysis identified 2 transcriptional regulator genes that are expressed at higher levels in recent isolates, and suggested that some of the observed differences between historic and recent strains might also result from altered expression of regulatory genes. *BP1969*, which encodes a LuxR family transcriptional factor, lies upstream of the *BP1970* and *BP1971* genes, which encode phosphoglucomutase Pgm and phosphoglucose isomerase Pgi. Similarly to *BP1969*, *pgm* and *pgi* genes were significantly up-regulated in recent strains. Therefore, we assume that the *BP1969* gene probably represents a cognate regulator for these glycolytic genes. Besides its role in glycolysis, Pgm catalyzes the generation of sugar nucleotides needed for biosynthesis of lipopolysaccharide and cell wall and was shown to be required for virulence of *B. bronchiseptica* (38) and several other pathogens (39,40). Strains lacking the *pgm* gene showed increased susceptibility to antimicrobial peptides and were attenuated in in vivo models of infection (38–40). Pgi catalyzes the second step in glycolysis and was shown to be required for virulence of *Xanthomonas campestris* (41). Thus, we presume that



**Table 3.** Genes expressing consistently increased RNA and protein levels in historic and recent isolates of *Bordetella pertussis*, Czech Republic, compared with Tohama I\*

Gene in Tohama I	Gene group	Gene name	Transcriptome		Proteome		Secretome		Annotation
			Cz/Toh	VS/Toh	Cz/Toh	VS/Toh	Cz/Toh	VS/Toh	
toh_00485	Group_1902	BP0500	2.1	2.8	<b>1.5</b>	<b>1.3</b>	6.0	5.1	T3SS effector BopC
toh_00936	<i>Sbp</i>	BP0966	14.7	13.0	5.6	4.6	ND	ND	Sulfate-binding protein
toh_02200	<i>bscI</i>	BP2249	2.1	3.1	<b>1.2</b>	ND	2.8	6.2	T3SS protein BscI
toh_02203	<i>bopB</i>	BP2252	2.1	3.0	3.1	2.8	2.1	1.8	T3SS protein BopB
toh_02204	<i>bopD</i>	BP2253	<b>1.8</b>	2.5	2.0	1.5	3.7	2.3	T3SS protein BopD
toh_02205	<i>bcrH1</i>	BP2254	2.3	4.0	2.5	2.1	<b>1.6</b>	<b>3.1</b>	T3SS protein
toh_02206	Group_1710	BP2255	2.4	3.8	3.9	4.0	4.8	<b>1.7</b>	Hypothetical protein
toh_02207	<i>bsp22</i>	BP2256	2.4	3.6	<b>1.6</b>	<b>1.5</b>	3.6	<b>1.8</b>	T3SS protein Bsp22
toh_02208	<i>bopN</i>	BP2257	2.5	2.9	<b>1.3</b>	<b>1.0</b>	3.1	<b>1.5</b>	T3SS protein BopN
toh_02210	Group_2630	BP2259	2.0	2.4	<b>1.4</b>	<b>1.9</b>	5.1	<b>2.4</b>	Putative T3SS protein
toh_02214	<i>bscE</i>	BP2263	2.3	<b>1.5</b>	5.2	4.5	2.0	<b>2.3</b>	T3SS protein BscE
toh_03375	Group_1130	BP3434	6.1	4.4	3.3	1.9	2.6	<b>1.8</b>	Exported protein

\*Fold change values resulting from comparison of either recent isolates (Cz) or historic strains (VS) with Tohama I strain (Toh) are shown for RNA sequencing and proteomic analyses. Values that did not indicate statistical significance (FC≥2; adjusted p<0.05) are shown in bold. ND, not determined; toh, Tohama.

increased expression and production of both enzymes might contribute to increased virulence and fitness of the *ptxP3* lineage.

Among other modulated genes, expression of numerous genes within the operon encoding the flagellar apparatus was significantly increased in recent isolates. However, we could not corroborate this finding because we did not detect any flagellar proteins in our samples. Recent observations suggest that *B. pertussis* is motile under modulatory Bvg-conditions (42) and that motility genes are up-regulated during adaptation to the mouse respiratory tract (43). Apparently, in vivo conditions, prevailing during *B. pertussis* infections in mice, cannot be completely reproduced under standard laboratory growth conditions, as documented (43,44), and further experiments are required to determine whether the increased expression translates into higher motility of recent isolates

and contributes to improved ability of *ptxP3* strains to colonize the respiratory tract (33).

Conversely, expression of an almost complete operon that encodes genes involved in polysaccharide capsule synthesis was substantially down-regulated in recent isolates. This observation is consistent with that of a previous report (45) and demonstrates that capsule proteins are produced by *B. pertussis*. This finding also involves the protein responsible for polysaccharide biosynthesis TviD (*BP1618*), which has been reported to be encoded by a pseudogene (14).

Data on the role of the capsule in the virulence and physiologic fitness of *B. pertussis* are contradictory. Hoo et al. (46) showed that the capsule proteins are expressed during the infection and are required for an efficient colonization of mouse lungs. In contrast, in vitro assays showed that the capsule did

**Table 4.** Proteins encoded in regions upstream of an *IS481* element adjacent to differentially expressed genes in recent and historic strains of *Bordetella pertussis*, Czech Republic\*

Strain	Protein		
	toh_01451 (BP1492)	toh_01915 (BP1969)	toh_02005 (BP2055)
Tohama I	tRNA	Partial phosphonate monoester hydrolase	Partial cyclopropane-fatty-acyl-phospholipid synthase
VS393	tRNA	Partial phosphonate monoester hydrolase	Partial cyclopropane-fatty-acyl-phospholipid synthase
VS401	tRNA	Partial phosphonate monoester hydrolase	IS481 element
VS67	Partial FUSC family protein	Partial phosphonate monoester hydrolase	Partial cyclopropane-fatty-acyl-phospholipid synthase
Bp318	Partial FUSC family protein	MarR family transcriptional regulator	Partial cyclopropane-fatty-acyl-phospholipid synthase
Bp155	Partial FUSC family protein	MarR family transcriptional regulator	Partial cyclopropane-fatty-acyl-phospholipid synthase
Bp46	Partial FUSC family protein	MarR family transcriptional regulator	Partial cyclopropane-fatty-acyl-phospholipid synthase
Bp6242	Partial FUSC family protein	MarR family transcriptional regulator	Partial cyclopropane-fatty-acyl-phospholipid synthase
BpK10	Partial FUSC family protein	MarR family transcriptional regulator	Partial cyclopropane-fatty-acyl-phospholipid synthase

\*IS481, insertion sequence 481.

not protect *B. pertussis* cells from phagocytosis and serum killing (45) and that the capsule locus was not expressed during infection of mouse respiratory tract (43). Therefore, it is difficult to assess whether reduced production of capsule proteins provides recent strains with any selective advantage. Nevertheless, the capsular polysaccharides of several gram-negative bacteria are highly immunogenic and were used to formulate carbohydrate-protein conjugate vaccines (47). Therefore, it is possible that in circulating isolates of *B. pertussis*, reduced production of the capsule synthesis apparatus contributes to evasion from the host immune response.

Our omics data manifest that, in spite of being isolated at the similar period of time, historic strains are substantially distinct from the reference strain Tohama I. Previous genomic analyses documented that several different clusters of *B. pertussis* circulated in Europe and the United States already in prevaccine and early vaccine eras and that their genomes were different from Tohama I (20,37). Our results with strains from the Czech Republic are consistent with these observations and also confirm this distinction at the transcriptomic and proteomic levels. For example, expression and production of various sulfate metabolism factors (*sbp*, *cysT*, *cysA*) were strongly reduced in Tohama I compared with strains from the Czech Republic. Likewise, we demonstrated that recent and historic strains had significantly increased expression, production, and secretion of several T3SS components. This observation is consistent with previous reports (48,49) and confirms that not only recent isolates but also low-passage historic strains of *B. pertussis* are T3SS proficient (48,49). We conclude that, in agreement with previous reports (37,50), the Tohama I strain is not a good representative of the circulating *B. pertussis* population.

Collectively, our data suggest that, besides shaping the evolution of *B. pertussis* on a genomic scale, the genome rearrangement and genome reduction processes also affect global transcriptomic and proteomic profiles. In agreement with results of a previous report (20), we assume that these mechanisms counterbalance the low level of genetic variability observed in this pathogen and strongly contribute to adaptation of the global population of *B. pertussis*.

### Acknowledgments

We thank Richard Neuboeck helping to setup software for some of the analyses, Michael Weigand for providing the sequence of the *oriC* region, and Karel Harant and Pavel Talacko for performing the liquid chromatography with mass tandem spectrometry analyses runs.

This study was supported by Czech Health Research Council (<http://www.azvcr.cz>) grant 16-30782A to B.V., Czech Research Foundation (<http://www.qacr.cz>) grant 19-12338S to B.V., institutional funding RVO61388971, and a mobility grant from the Czech Academy of Sciences (MSM200201702) to A.D. This study was also supported by the project “BIOCEV – Biotechnology and Biomedicine Centre of the Academy of Sciences and Charles University” (CZ.1.05/1.1.00/02.0109) from the European Regional Development Fund.

### About the Author

Dr. Dienstbier is a postdoctoral fellow in the Laboratory of Post-transcriptional Control of Gene Expression, Institute of Microbiology, Czech Academy of Sciences, Prague, Czech Republic. Her research interests are genomics and transcriptomics of *B. pertussis* strains and molecular mechanisms underlying the pathogenesis of pertussis.

### References

- Centers for Disease Control and Prevention. Vaccine preventable deaths and the Global Immunization Vision and Strategy, 2006–2015. *MMWR Morb Mortal Wkly Rep*. 2006;55:511–5.
- Raguckas SE, VandenBussche HL, Jacobs C, Klepser ME. Pertussis resurgence: diagnosis, treatment, prevention, and beyond. *Pharmacotherapy*. 2007;27:41–52. <https://doi.org/10.1592/phco.27.1.41>
- van Gent M, Heuvelman CJ, van der Heide HG, Hallander HO, Advani A, Guiso N, et al. Analysis of *Bordetella pertussis* clinical isolates circulating in European countries during the period 1998–2012. *Eur J Clin Microbiol Infect Dis*. 2015;34:821–30. <https://doi.org/10.1007/s10096-014-2297-2>
- Fabiánová K, Benes C, Kriz B. A steady rise in incidence of pertussis since nineties in the Czech Republic. *Epidemiol Mikrobiol Imunol*. 2010;59:25–33.
- Sealey KL, Belcher T, Preston A. *Bordetella pertussis* epidemiology and evolution in the light of pertussis resurgence. *Infect Genet Evol*. 2016;40:136–43. <https://doi.org/10.1016/j.meegid.2016.02.032>
- Espósito S, Stefanelli P, Fry NK, Fedele G, He Q, Paterson P, et al.; World Association of Infectious Diseases and Immunological Disorders (WAidid) and the Vaccine Study Group of the European Society of Clinical Microbiology and Infectious Diseases (EVASG). Pertussis prevention: reasons for resurgence, and differences in the current acellular pertussis vaccines. *Front Immunol*. 2019;10:1344. <https://doi.org/10.3389/fimmu.2019.01344>
- Cherry JD. Epidemic pertussis and acellular pertussis vaccine failure in the 21st century. *Pediatrics*. 2015;135:1130–2. <https://doi.org/10.1542/peds.2014-4118>
- Klein NP, Bartlett J, Rowhani-Rahbar A, Fireman B, Baxter R. Waning protection after fifth dose of acellular pertussis vaccine in children. *N Engl J Med*. 2012;367:1012–9. <https://doi.org/10.1056/NEJMoa1200850>
- Liko J, Robison SG, Cieslak PR. Priming with whole-cell versus acellular pertussis vaccine. *N Engl J Med*. 2013;368:581–2. <https://doi.org/10.1056/NEJMc1212006>
- Burdin N, Handy LK, Plotkin SA. What is wrong with pertussis vaccine immunity? The problem of waning effectiveness of

- pertussis vaccines. Cold Spring Harb Perspect Biol. 2017;9:a029454. <https://doi.org/10.1101/cshperspect.a029454>
11. Bart MJ, Harris SR, Advani A, Arakawa Y, Bottero D, Bouchez V, et al. Global population structure and evolution of *Bordetella pertussis* and their relationship with vaccination. MBiol. 2014;5:e01074. <https://doi.org/10.1128/mBio.01074-14>
  12. Mooi FR, Van Der Maas NA, De Melker HE. Pertussis resurgence: waning immunity and pathogen adaptation: two sides of the same coin. Epidemiol Infect. 2014;142:685–94. <https://doi.org/10.1017/S0950268813000071>
  13. Bouchez V, Hegerle N, Strati F, Njamkepo E, Guiso N. New data on vaccine antigen deficient *Bordetella pertussis* isolates. Vaccines (Basel). 2015;3:751–70. <https://doi.org/10.3390/vaccines3030751>
  14. Parkhill J, Sebahia M, Preston A, Murphy LD, Thomson N, Harris DE, et al. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. Nat Genet. 2003;35:32–40. <https://doi.org/10.1038/ng1227>
  15. Cummings CA, Brinig MM, Lepp PW, van de Pas S, Relman DA. *Bordetella* species are distinguished by patterns of substantial gene loss and host adaptation. J Bacteriol. 2004;186:1484–92. <https://doi.org/10.1128/JB.186.5.1484-1492.2004>
  16. Linz B, Ivanov YV, Preston A, Brinkac L, Parkhill J, Kim M, et al. Acquisition and loss of virulence-associated factors during genome evolution and speciation in three clades of *Bordetella* species. BMC Genomics. 2016;17:767. <https://doi.org/10.1186/s12864-016-3112-5>
  17. Bowden KE, Weigand MR, Peng Y, Cassiday PK, Sammons S, Knipe K, et al. Genome structural diversity among 31 *Bordetella pertussis* isolates from two recent U.S. whooping cough statewide epidemics. MSphere. 2016; 1:e00036–16. <https://doi.org/10.1128/mSphere.00036-16>
  18. Weigand MR, Peng Y, Loparev V, Batra D, Bowden KE, Burroughs M, et al. The history of *Bordetella pertussis* genome evolution includes structural rearrangement. J Bacteriol. 2017;199:e00806–16. <https://doi.org/10.1128/JB.00806-16>
  19. Ring N, Abrahams JS, Jain M, Olsen H, Preston A, Bagby S. Resolving the complex *Bordetella pertussis* genome using barcoded nanopore sequencing. Microb Genom. 2018;4. <https://doi.org/10.1099/mgen.0.000234>
  20. Brinig MM, Cummings CA, Sanden GN, Stefanelli P, Lawrence A, Relman DA. Significant gene order and expression differences in *Bordetella pertussis* despite limited gene content variation. J Bacteriol. 2006;188:2375–82. <https://doi.org/10.1128/JB.188.7.2375-2382.2006>
  21. Amman F, D'Halluin A, Antoine R, Huot L, Bibova I, Keidel K, et al. Primary transcriptome analysis reveals importance of IS elements for the shaping of the transcriptional landscape of *Bordetella pertussis*. RNA Biol. 2018; 15:967–75. <https://doi.org/10.1080/15476286.2018.1462655>
  22. Dienstbier A, Pouchnik D, Wildung M, Amman F, Hofacker IL, Parkhill J, et al. Comparative genomics of Czech vaccine strains of *Bordetella pertussis*. Pathog Dis. 2018;76. <https://doi.org/10.1093/femspd/fty071>
  23. Kasuga T, Nakase Y, Ukishima K, Takatsu K. Studies on *Haemophilus pertussis*. III. Some properties of each phase of *H. pertussis*. Kitasato Arch Exp Med. 1954;27:37–47.
  24. Stainer DW, Scholte MJ. A simple chemically defined medium for the production of phase I *Bordetella pertussis*. J Gen Microbiol. 1970;63:211–20. <https://doi.org/10.1099/00221287-63-2-211>
  25. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One. 2010;5:e11147. <https://doi.org/10.1371/journal.pone.0011147>
  26. Hu F, Lin Y, Tang J. MLGO: phylogeny reconstruction and ancestral inference from gene-order data. BMC Bioinformatics. 2014;15:354. <https://doi.org/10.1186/s12859-014-0354-6>
  27. Angiuoli SV, Salzberg SL. Mugsy: fast multiple alignment of closely related whole genomes. Bioinformatics. 2011;27:334–42. <https://doi.org/10.1093/bioinformatics/btq665>
  28. Gardner SN, Slezak T, Hall BG. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. Bioinformatics. 2015; 31:2877–8. <https://doi.org/10.1093/bioinformatics/btv271>
  29. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44(W1):W242–5. <https://doi.org/10.1093/nar/gkw290>
  30. Sealey KL, Harris SR, Fry NK, Hurst LD, Gorringer AR, Parkhill J, et al. Genomic analysis of isolates from the United Kingdom 2012 pertussis outbreak reveals that vaccine antigen genes are unusually fast evolving. J Infect Dis. 2015;212:294–301. <https://doi.org/10.1093/infdis/jiu665>
  31. Block DH, Hussein R, Liang LW, Lim HN. Regulatory consequences of gene translocation in bacteria. Nucleic Acids Res. 2012;40:8979–92. <https://doi.org/10.1093/nar/gks694>
  32. de Gouw D, Hermans PW, Bootsma HJ, Zomer A, Heuvelman K, Diavatopoulos DA, et al. Differentially expressed genes in *Bordetella pertussis* strains belonging to a lineage which recently spread globally. PLoS One. 2014; 9:e84523. <https://doi.org/10.1371/journal.pone.0084523>
  33. King AJ, van der Lee S, Mohangoo A, van Gent M, van der Ark A, van de Waterbeemd B. Genome-wide gene expression analysis of *Bordetella pertussis* isolates associated with a resurgence in pertussis: elucidation of factors involved in the increased fitness of epidemic strains. PLoS One. 2013;8:e66150. <https://doi.org/10.1371/journal.pone.0066150>
  34. Kang Y, Gu C, Yuan L, Wang Y, Zhu Y, Li X, et al. Flexibility and symmetry of prokaryotic genome rearrangement reveal lineage-associated core-gene-defined genome organizational frameworks. MBiol. 2014;5:e01867. <https://doi.org/10.1128/mBio.01867-14>
  35. Dorman CJ. Genome architecture and global gene regulation in bacteria: making progress towards a unified model? Nat Rev Microbiol. 2013;11:349–55. <https://doi.org/10.1038/nrmicro3007>
  36. Bryant JA, Sellars LE, Busby SJ, Lee DJ. Chromosome position effects on gene expression in *Escherichia coli* K-12. Nucleic Acids Res. 2014;42:11383–92. <https://doi.org/10.1093/nar/gku828>
  37. Kallonen T, Gröndahl-Yli-Hannuksela K, Elomaa A, Lutyńska A, Fry NK, Mertsola J, et al. Differences in the genomic content of *Bordetella pertussis* isolates before and after introduction of pertussis vaccines in four European countries. Infect Genet Evol. 2011;11:2034–42. <https://doi.org/10.1016/j.meegid.2011.09.012>
  38. West NP, Jungnitz H, Fitter JT, McArthur JD, Guzmán CA, Walker MJ. Role of phosphoglucomutase of *Bordetella bronchiseptica* in lipopolysaccharide biosynthesis and virulence. Infect Immun. 2000;68:4673–80. <https://doi.org/10.1128/IAI.68.8.4673-4680.2000>
  39. Ugalde JE, Czibener C, Feldman MF, Ugalde RA. Identification and characterization of the *Brucella abortus* phosphoglucomutase gene: role of lipopolysaccharide in virulence and intracellular multiplication. Infect Immun. 2000;68:5716–23. <https://doi.org/10.1128/IAI.68.10.5716-5723.2000>

40. Hardy GG, Magee AD, Ventura CL, Caimano MJ, Yother J. Essential role for cellular phosphoglucosyltransferase in virulence of type 3 *Streptococcus pneumoniae*. *Infect Immun*. 2001;69:2309–17. <https://doi.org/10.1128/IAI.69.4.2309-2317.2001>
41. Tung SY, Kuo TT. Requirement for phosphoglucose isomerase of *Xanthomonas campestris* in pathogenesis of citrus canker. *Appl Environ Microbiol*. 1999;65:5564–70. <https://doi.org/10.1128/AEM.65.12.5564-5573.1999>
42. Hoffman CL, Gonyar LA, Zacca F, Sisti F, Fernandez J, Wong T, et al. *Bordetella pertussis* can be motile and express flagellum-like structures. *MBiol*. 2019;10:e00787–19. <https://doi.org/10.1128/mBio.00787-19>
43. van Beek LF, de Gouw D, Eleveld MJ, Bootsma HJ, de Jonge MI, Mooi FR, et al. Adaptation of *Bordetella pertussis* to the respiratory tract. *J Infect Dis*. 2018;217:1987–96. <https://doi.org/10.1093/infdis/jiy125>
44. Wong TY, Hall JM, Nowak ES, Boehm DT, Gonyar LA, Hewlett EL, et al. Analysis of the *in vivo* transcriptome of *Bordetella pertussis* during infection of mice. *MSphere*. 2019;4:e00154–19. <https://doi.org/10.1128/mSphereDirect.00154-19>
45. Neo Y, Li R, Howe J, Hoo R, Pant A, Ho S, et al. Evidence for an intact polysaccharide capsule in *Bordetella pertussis*. *Microbes Infect*. 2010;12:238–45. <https://doi.org/10.1016/j.micinf.2009.12.002>
46. Hoo R, Lam JH, Huot L, Pant A, Li R, Hot D, et al. Evidence for a role of the polysaccharide capsule transport proteins in pertussis pathogenesis. *PLoS One*. 2014;9:e115243. <https://doi.org/10.1371/journal.pone.0115243>
47. Ada G, Isaacs D. Carbohydrate-protein conjugate vaccines. *Clin Microbiol Infect*. 2003;9:79–85. <https://doi.org/10.1046/j.1469-0691.2003.00530.x>
48. Fennelly NK, Sisti F, Higgins SC, Ross PJ, van der Heide H, Mooi FR, et al. *Bordetella pertussis* expresses a functional type III secretion system that subverts protective innate and adaptive immune responses. *Infect Immun*. 2008;76:1257–66. <https://doi.org/10.1128/IAI.00836-07>
49. Gaillard ME, Bottero D, Castuma CE, Basile LA, Hozbor D. Laboratory adaptation of *Bordetella pertussis* is associated with the loss of type three secretion system functionality. *Infect Immun*. 2011;79:3677–82. <https://doi.org/10.1128/IAI.00136-11>
50. Caro V, Bouchez V, Guiso N. Is the sequenced *Bordetella pertussis* strain Tohama I representative of the species? *J Clin Microbiol*. 2008;46:2125–8. <https://doi.org/10.1128/JCM.02484-07>

Address for correspondence: Branislav Vecerek, Institute of Microbiology, Czech Academy of Sciences, Videnska 1083, 14220 Prague 4, Czech Republic; email: vecerek@biomed.cas.cz

## EID Podcast

### Tickborne Ehrlichia in North Carolina

While caring for patients in North Carolina, Dr. Ross Boyce began to suspect that tickborne *Ehrlichia* was being underdiagnosed. His study showed that *Ehrlichia*, despite being relatively common, was only tested for in about a third of patients thought to have a tickborne illness.

In this EID podcast, Dr. Ross Boyce, an infectious disease physician at the University of North Carolina at Chapel Hill, examines the prevalence and diagnosis of *Ehrlichia* in North Carolina.

Visit our website to listen: <https://go.usa.gov/xy6UH> EMERGING INFECTIOUS DISEASES®

# Comparative Omics Analysis of Historic and Recent Isolates of *Bordetella pertussis* and Effects of Genome Rearrangements on Evolution

## Appendix 1

### Materials and Methods

#### Whole-Genome Sequencing, Assembly, and Annotation

Genomic DNA was isolated from bacterial cells by using QIAamp UCP Pathogen Mini Kit (QIAGEN, <https://www.qiagen.com>). Genomes were sequenced on the RSII Platform (PacBio, <https://www.pacb.com>) and assembled by using the HGAP SMRT Portal protocol (PacBio). Sanger sequencing was used to polish the assemblies. Completed genome sequences were annotated by using Prokka software (1) and deposited in the European Nucleotide Archive under project PRJEB26966. For comparison of sequenced strains with the reference strain the genome sequence of Tohama I (strain E476; Appendix Table 1) was reannotated by using our annotation pipeline.

#### RNA Isolation, Sequencing, and Data Analysis

Total RNA was isolated by using TRI Reagent (Sigma, <https://www.sigmaaldrich.com>) according to manufacturer's protocol. Removal of DNA was achieved by treatment of samples with the TURBO DNA-Free Kit (Thermo Fisher Scientific, <https://www.thermofisher.com>). RNA quality and quantity was determined by agarose gel electrophoresis and using the Nanodrop 2000 Machine (Thermo Fisher Scientific). RNA quality was assessed at sequencing facility by using the 2100 Bioanalyzer Device (Agilent, <https://www.agilent.com>). All samples displayed RNA integrity numbers >9. Ribosomal RNA was removed by using the Ribo-Zero rRNA Removal Kit for Bacteria (Illumina, <https://www.agilent.com>). Libraries were prepared by using the Next Ultra II DNA Library Prep Kit (New England Biolabs, <https://www.neb.com>), and sequenced on an Illumina HiSeq 2500 platform using HiSeqV4 chemistry with single-end 50-bp

reads at the Vienna Biocenter Core Facilities Next Generation Sequencing Unit (Vienna, Austria). After quality control, the reads were demultiplexed and quality trimming and adaptor removal from the reads was performed by using Trimmomatic (2).

Reads were mapped to combined transcriptome made of all strains by using the Salmon algorithm (3). Combined transcriptome was built from all annotated transcripts in all strains, and homologous transcripts were conflated into multistrain-representing genes. Gene homology between the strains was determined by using the Roary pipeline (4). Differential gene expression analysis was performed by using DESeq2 (5). Genes with a  $\log_2$ -fold change  $\leq -1$  or  $\geq 1$  and a  $q$  value  $< 0.05$  ( $p$  value adjusted for multiple testing correction by the method of Benjamini and Hochberg [6]) were considered significantly deregulated.

To see the overall gene expression differences between the groups of recent clinical isolates and vaccine strains, all isolates within each group were treated as replicates of the same sample. RNA sequencing data from the sequencing runs were deposited in the European Nucleotide Archive under project accession no. PRJEB34096.

#### **Protein Sample Preparation and Label-Free Proteomic Analysis by Using Liquid Chromatography with Mass Tandem Spectrometry Analyses**

Cultures of *B. pertussis* were pelleted by centrifugation at  $10,000\times g$  at  $4^\circ\text{C}$  for 10 min to separate cell pellets and culture supernatants. Cells were resuspended in digestion buffer (100 mmol/L triethylammonium bicarbonate, pH 8.5, 2% sodium deoxycholate) and lysed by sonication. For analysis of supernatant fractions, supernatants were filtered through 0.22- $\mu\text{m}$  filters and precipitated with 10% (wt/vol) trichloroacetic acid (Sigma) overnight at  $4^\circ\text{C}$ . Precipitated proteins were collected by centrifugation at  $14,000\times g$  at  $4^\circ\text{C}$ , for 20 min, washed with 80% (wt/vol) acetone, and dissolved in 100 mmol/L triethylammonium bicarbonate, pH 8.5, 2% sodium deoxycholate digestion buffer. Protein concentrations were determined by using the BCA Protein Assay Kit (Thermo Fischer Scientific), and 20  $\mu\text{g}$  of protein/sample were used for protein analysis. Cysteines were reduced with 5 mol/L Tris(2-carboxyethyl)phosphine (at  $60^\circ\text{C}$  for 60 min) and blocked with 10 mmol/L methyl methanethiosulfonate (at room temperature for 10 min). Samples were digested with trypsin (trypsin:protein ratio 1:20) at  $37^\circ\text{C}$  overnight. Digestion of samples was stopped by addition of trifluoroacetic acid (Sigma) to a final concentration of 1% (vol/vol). Sodium deoxycholate was removed by extraction with ethyl

acetate, and peptides were desalted on a C18 column (Michrom Bio, <https://www.bioprocessonline.com>).

A Nano Reversed Phase Column (EASY-Spray Column, 50 cm × 75 μm internal diameter, PepMap C18, 2-μm particles, 100 Å pore size; Thermo Fisher Scientific) was used for liquid chromatography–mass spectrometry analysis. Mobile phase buffer A was composed of water and 0.1% formic acid. Mobile phase B was composed of acetonitrile and 0.1% formic acid. Samples were loaded onto the trap column (Acclaim PepMap300, C18, 5 μm, 300 Å wide pore, 300 μm × 5 mm; Thermo Fisher Scientific) at a flow rate of 15 μL/min. Loading buffer was composed of water, 2% acetonitrile, and 0.1% trifluoroacetic acid. Peptides were eluted with a gradient of phase B ranging from 4% to 35% over 60 min at a flow rate of 300 nL/min. Eluting peptide cations were converted to gas-phase ions by electrospray ionization and analyzed by Orbitrap Fusion (Thermo Fisher Scientific). Survey scans of peptide precursors from 350 m/z to 1,400 m/z were performed at 120 K resolution (at 200 m/z) with a  $5 \times 10^5$  ion count target.

Tandem mass spectrometry (MS2) was performed by isolation within a 1.5-Th window with the quadrupole, higher-energy collisional dissociation fragmentation with normalized collision energy of 30, and rapid scan mass spectrometry analysis in the ion trap. The MS2 ion count target value was set to 104 and the maximal injection time was 35 ms. The precursors with charge state of 2–6 were sampled for MS2. The dynamic exclusion duration was set to 45 s with a 10 ppm tolerance around the selected precursor and its isotopes. Monoisotopic precursor selection was turned on. The instrument was run at top speed mode with cycles of 2 s (7).

Raw data were imported into MaxQuant software version 1.5.3.8 (8) for identification and label-free quantification of proteins. The false discovery rate was set to 1% for peptides and minimum specific length of 7 amino acids. The Andromeda search engine (9) was used for the MS/MS spectra search against the Uniprot *Bordetella pertussis* Database (<https://www.uniprot.org>) containing 3,258 entries.

Enzyme specificity was set as C-terminal to Arg and Lys, also enabling cleavage at proline bonds and a maximum of 2 missed cleavages. Dithiomethylation of cysteine was selected as a fixed modification, and N-terminal protein acetylation and methionine oxidation as variable modifications. The match between runs feature of MaxQuant was used to transfer identifications to other liquid chromatography; mass tandem spectrometry runs were based on their masses and

retention time (maximum deviation 0.7 min), and this feature was also used in quantification experiments. Protein abundance was calculated from obtained label-free protein intensities by using the MaxLFQ algorithm (10). Proteins with <4 mass tandem spectrometry spectral counts were removed from the analysis. Statistics and data interpretation were performed by using Perseus version 1.6.2.3 software (11). Normalized label-free intensities were compared pairwise between recent clinical isolates, vaccine strains, and Tohama I. Similarly to transcriptomic analysis, all isolates within each group of isolates from the Czech Republic were treated as replicates of the same sample, and each abundance ratio was tested for significance by using a 2-group *t*-test ( $p < 0.05$ ). The *p* values were further adjusted for multiple testing correction to control the false-discovery rate at a cutoff value of 0.05 by using the permutation test (number of randomizations = 250). Proteins with corrected *p* values (*q* values)  $< 0.05$  were considered as significantly modulated. For downstream analyses (e.g., gene ontology [GO] term enrichment), only proteins that were detected by  $\geq 2$  unique peptides in  $\geq 2$  of the 3 biologic replicates were considered. Proteins for which label-free intensities were not obtained in any of the replicates of either the vaccine strains, clinical isolates, or the Tohama I strain were considered as strongly modulated and defined as ON/OFF. Proteomics data were deposited in the ProteomeXchange Consortium by using the PRIDE (12) partner repository with the dataset identifier PXD015184.

The hierarchical clustering analysis was generated by using Perseus 1.6.2.3 software (11). In brief, intensities of label-free quantified proteins were log<sub>2</sub>-transformed to reduce the effect of outliers. For analysis, the greatly modulated proteins were separated by using a multiple-sample test with the false discovery rate at a cutoff value of 0.05 by using the permutation test (250 randomizations). Hierarchical clustering was performed on Z-score normalized log<sub>2</sub>-transformed label-free quantified intensities of greatly modulated proteins within either cell-associated or cell-free fractions.

### **GO Term Enrichment Analysis**

To gain a comprehensive functional annotation of the reference genome, GO terms per gene were deduced by using blast2go (13). For the GO term enrichment analysis, deregulated genes were split into up-regulated and down-regulated genes, and each gene set was analyzed separately. Each GO term, which was associated with  $> 1$  gene in the gene set, was tested for enrichment in comparison to the whole transcriptome by using the Fisher exact test. Afterwards, determined *p* values were corrected for multiple testing by using the method of Benjamini and



Hochberg (6), summarized by using the Revigo tool (14), and visualized by using Cytoscape (<https://cytoscape.org>).

## References

1. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068–9. [PubMed](#) <https://doi.org/10.1093/bioinformatics/btu153>
2. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20. [PubMed](#) <https://doi.org/10.1093/bioinformatics/btu170>
3. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017;14:417–9. [PubMed](#) <https://doi.org/10.1038/nmeth.4197>
4. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*. 2015;31:3691–3. [PubMed](#) <https://doi.org/10.1093/bioinformatics/btv421>
5. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550. [PubMed](#) <https://doi.org/10.1186/s13059-014-0550-8>
6. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B*. 1995;57:289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
7. Hebert AS, Richards AL, Bailey DJ, Ulbrich A, Coughlin EE, Westphall MS, et al. The one hour yeast proteome. *Mol Cell Proteomics*. 2014;13:339–47. [PubMed](#) <https://doi.org/10.1074/mcp.M113.034769>
8. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008;26:1367–72. [PubMed](#) <https://doi.org/10.1038/nbt.1511>
9. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res*. 2011;10:1794–805. [PubMed](#) <https://doi.org/10.1021/pr101065j>
10. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics*. 2014;13:2513–26. [PubMed](#) <https://doi.org/10.1074/mcp.M113.031591>

11. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods*. 2016;13:731–40. [PubMed https://doi.org/10.1038/nmeth.3901](https://doi.org/10.1038/nmeth.3901)
12. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res*. 2019;47(D1):D442–50. [PubMed https://doi.org/10.1093/nar/gky1106](https://doi.org/10.1093/nar/gky1106)
13. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;21:3674–6. [PubMed https://doi.org/10.1093/bioinformatics/bti610](https://doi.org/10.1093/bioinformatics/bti610)
14. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*. 2011;6:e21800. [PubMed https://doi.org/10.1371/journal.pone.0021800](https://doi.org/10.1371/journal.pone.0021800)