Tracking genome engineering outcome at individual DNA breakpoints

Michael T Certo[1,2], Byoung Y Ryu[2], James E Annis[3], Mikhail Garibov[2], Jordan Jarjour[2,4,6], David J Rawlings[2,4,5] & Andrew M Scharenberg[2,4,5]

Site-specific genome engineering technologies are increasingly important tools in the postgenomic era, where biotechnological objectives often require organisms with precisely modified genomes. Rare-cutting endonucleases, through their capacity to create a targeted DNA strand break, are one of the most promising of these technologies. However, realizing the full potential of nuclease-induced genome engineering requires a detailed understanding of the variables that influence resolution of nuclease-induced DNA breaks. Here we present a genome engineering reporter system, designated 'traffic light', that supports rapid flow-cytometric analysis of repair pathway choice at individual DNA breaks, quantitative tracking of nuclease expression and donor template delivery, and high-throughput screens for factors that bias the engineering outcome. We applied the traffic light system to evaluate the efficiency and outcome of nuclease-induced genome engineering in human cell lines and identified strategies to facilitate isolation of cells in which a desired engineering outcome has occurred.

© 2011 Nature America, Inc. All rights reserved.

The explosive accumulation of genomic sequence data is driving demand for technologies to site-specifically engineer genomes[1]. One promising approach for genome engineering is the use of rare-cutting endonucleases to exploit endogenous DNA repair pathways[2–4]. However, nuclease-induced DNA breaks may engage any one of several DNA repair pathways that can produce distinct genetic outcomes[5,6]. Therefore, an important technological goal is to understand how experimental variables influence the choice of DNA-repair pathway and to develop methods to bias break resolution toward a desired outcome. Although several factors are known to influence what repair pathway is used after a DNA break, including cell-cycle status[7], DNA repair protein expression and post-translational modification[8], availability of donor templates[6,9] and usage of single- versus double-strand breaks[10], their application to bias outcome in a genome engineering context have not been systematically explored. An important limitation in developing such applications has been the lack of a method to rapidly assess different types of repair outcomes occurring at an individual DNA breakpoint. Although various nuclease-induced double-strand break repair reporters have been developed[9,11–17], none afford the ability to directly measure the efficiency and competition between DNA repair pathways that resolve a DNA break.

We constructed a reporter, traffic light, that generates a flow-cytometric readout of homology-directed repair (HDR)-mediated gene targeting and mutagenic nonhomologous end-joining (mutNHEJ)-mediated gene disruption occurring at an individual DNA breakpoint. We integrated the reporter into a system that provides for quantitative single-cell tracking of nuclease and donor template delivery, and supports efficient siRNA-mediated manipulation of endogenous DNA repair pathways. Using this system, we demonstrated that high concentration of donor template can promote gene targeting while suppressing mutNHEJ, that single-strand breaks can induce gene targeting without eliciting mutNHEJ and that limiting the classical nonhomologous end-joining pathway through DNA-PKcs silencing increases the efficiency of gene targeting.

RESULTS

Fluorescent reporter for HDR and mutNHEJ

We designed a construct in which a double-strand break is produced at an embedded nuclease cleavage site (in this case, an I-SceI site), and repair of the break generates distinct fluorescent signals upon resolution either through HDR with an exogenous donor template or through mutNHEJ (**Fig. 1a**,**b**): in the former case, a functional enhanced GFP (*eGFP*) open reading frame is restored by the exogenously provided donor template to signal gene targeting[9]; in the latter case a frameshift places a monomeric (m)Cherry coding sequence in-frame to signal gene disruption. By design, the eGFP coding sequence contains an alternative +3 reading frame (**Supplementary Fig. 1**), and the T2A 'dis-linker' enables the downstream-encoded mCherry to escape degradation of the misfolded protein encoded in this +3 reading frame of *eGFP* (**Supplementary Fig. 2** and **Supplementary Note 1**). We designated our construct the traffic light reporter (TLR).

[1]Program in Molecular and Cellular Biology, University of Washington, Seattle, Washington, USA. [2]Center of Immunity and Immunotherapies, Seattle Children's Research Institute, Seattle, Washington, USA. [3]Quellos High Throughput Core, Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, Washington, USA. [4]Department of Immunology, University of Washington, Seattle, Washington, USA. [5]Department of Pediatrics, University of Washington, Seattle, Washington, USA. [6]Present address: Precision Genome Engineering Inc., Seattle, Washington, USA. Correspondence should be addressed to A.M.S. (andrewms@u.washington.edu) and D.J.R. (drawling@u.washington.edu).

RECEIVED 6 APRIL; ACCEPTED 7 JUNE; PUBLISHED ONLINE 10 JULY 2011; DOI:10.1038/NMETH.1648

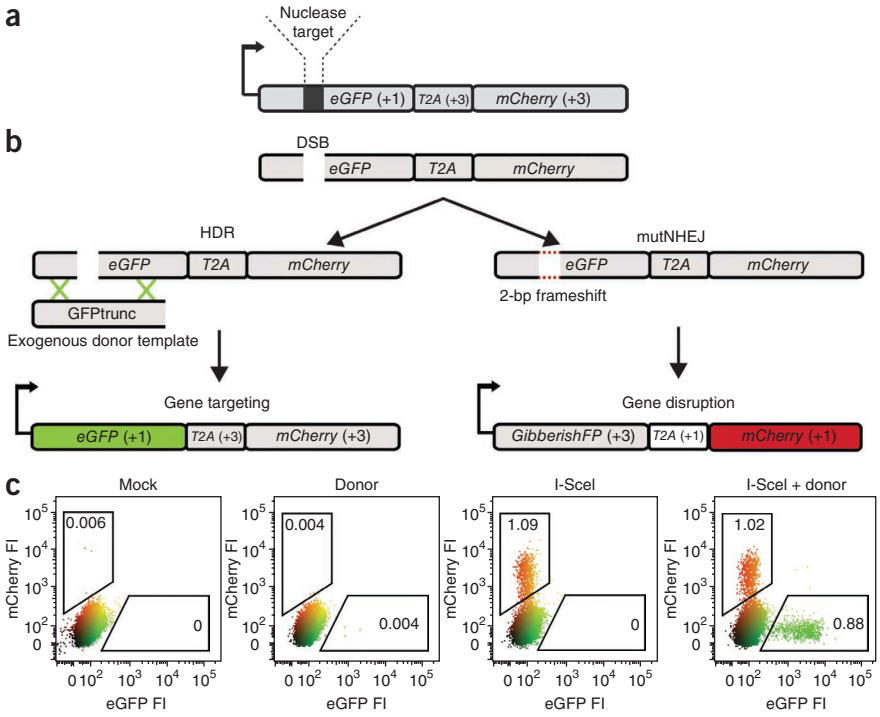

Figure 1 | The traffic light reporter. (**a**) Diagram of the TLR. Arrow represents promoter and initial eGFP start codon. Reading frames relative to the initial eGFP start codon are indicated in parentheses. (**b**) Schematic depicting different engineering outcomes after the induction of a site specific double-strand break (DSB). If the break is resolved through the HDR pathway, the full *eGFP* sequence will be reconstituted, and cells will fluoresce green; if the break undergoes mutNHEJ, *eGFP* will be translated out of frame (*gibberishFP*, +3 reading frame) and the *T2A* and *mCherry* sequences are rendered in frame to produce red fluorescent cells. (**c**) Flow cytometric analysis of HEK293T *TLR-Sce* cells 72 h after transduction with the indicated lentiviral constructs. Numbers shown inside plots indicate percentages of live cells. FI, relative fluorescence intensity reported in arbitrary units.

We derived a polyclonal population of human HEK293T cells containing a single copy of the TLR containing an I-SceI target site integrated into the genome (HEK293T *TLR-Sce*). We then expressed I-SceI with or without donor template and analyzed the cells by flow cytometry (**Fig. 1c**). As expected, cells transduced with I-SceI lentivirus alone resulted in a proportion of cells expressing mCherry (mCherry$^+$), indicative of mutNHEJ at the reporter locus. Cells transduced with both I-SceI and donor template lentivirus yielded either mCherry$^+$ or eGFP$^+$ cells. Sequence analysis confirmed that mCherry$^+$ cells were generated by a variety of mutagenic events at the I-SceI target site that resulted in frameshifts to the +3 reading frame, and that these frameshifts represented about one-third of all the mutagenic events (**Supplementary Fig. 3**). We performed similar experiments in several single-cell clones (**Supplementary Fig. 4**) and observed that all clones were capable of break resolution through either HDR or mutNHEJ, although some exhibited relative preferences for a specific pathway, suggesting a potential influence of genomic location or local chromatin environment on double-strand-break repair[6].

Effect of nuclease and donor delivery on TLR readout

We evaluated the effect of an increasing dose of I-SceI and donor template, both delivered with a single lentivirus (**Fig. 2a**). At low viral dose, we observed relatively few fluorescent cells, and nearly all were mCherry$^+$, indicating that the repair events occurred via mutNHEJ pathways. As we increased the amount of virus, we observed the expected dose-dependent increase in the total number of repair events, with an increasing fraction of events accounted for by the HDR pathway (**Fig. 2b**,**c**).

We reasoned that the bias toward HDR at high viral doses could be explained by increasing copies of donor template, as this has been previously shown to increase gene targeting rates at the population level[9]. To test this hypothesis, we used an integration-deficient lentivirus[18] to provide increasing amounts of episomal donor template while holding the dose of I-SceI–encoding lentivirus constant. As expected, gene-targeting events increased in correlation with the amount of template transduced, but we also observed a concomitant decrease in mutNHEJ events (**Fig. 2d–f**). The loss of mutNHEJ events with increasing donor template is an important observation, as it suggests that inadequate delivery of a suitable donor can lead to failed homology searches that default to mutNHEJ repair. This is consistent with recent work suggesting that mutagenic alternative nonhomologous end-joining pathways may have a resection step in common with HDR[19]. Therefore, if HDR-based genome modification is desired, high-level donor delivery is a key variable for concurrently promoting gene targeting and suppressing undesirable mutNHEJ.

Tracking endonuclease and donor template delivery

As genome engineering efficiency and outcome are both dependent on nuclease and donor template delivery, we tagged the nuclease and donor template with genes encoding unique fluorescent markers, infrared fluorescent protein 1.4 (IFP) and mTagBFP (BFP), respectively, to allow their simultaneous readout with the TLR (**Supplementary Fig. 5**). We used viral transduction to deliver *I-SceI-T2A-IFP* and *Donor-T2A-BFP* into HEK293T *TLR-Sce* cells and evaluated the four fluorescence parameters (eGFP, mCherry, IFP and BFP signals) by flow cytometry (**Fig. 3**). Mock-transduced cells were nonfluorescent, whereas cells transduced with *Donor-T2A-BFP* integration-deficient lentivirus were blue, with no detectable repair events (**Fig. 3a**). Cells transduced with *I-SceI-T2A-IFP* lentivirus fluoresced in the infrared channel and resulted in mCherry$^+$ mutNHEJ events. Upon transduction with both *I-SceI-T2A-IFP* and *Donor-T2A-BFP*, we observed a spectrum of cells with a variety of mean fluorescence intensity values for nuclease and donor, and both mutNHEJ and gene-targeting repair events. The aggregate mutNHEJ and gene targeting repair event frequency correlated well with experiments using untagged expression and donor template vectors, indicating that the inclusion of the fluorescent markers was not adversely influencing the TLR readout.

Because mean fluorescence intensity value has been shown to correlate with both protein amounts and vector copy number[20], we hypothesized that cells with unique mean fluorescence intensity values for nuclease and donor template after co-transduction

© 2011 Nature America, Inc. All rights reserved.

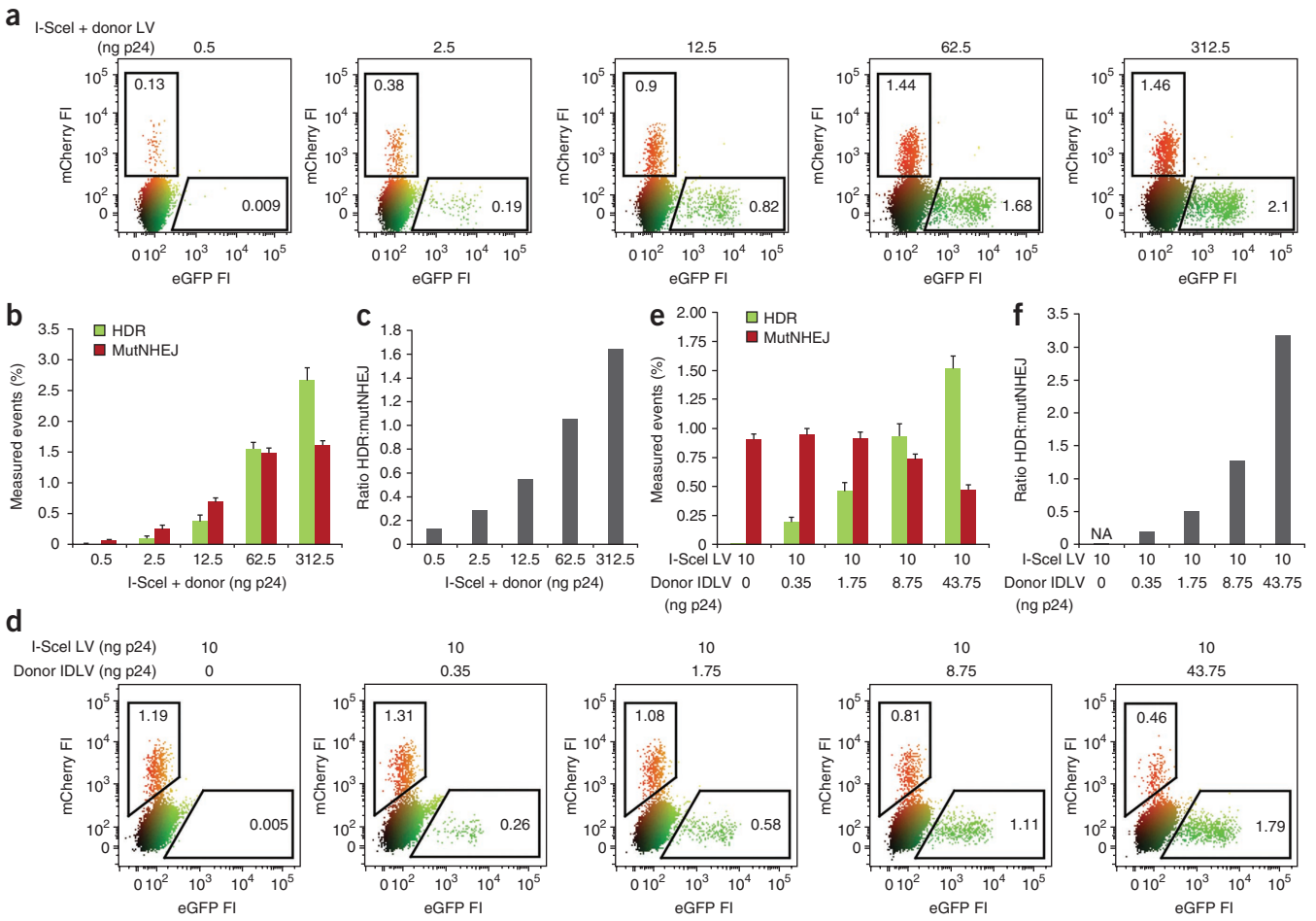

Figure 2 | Titration of nuclease and donor template. (**a**) Representative flow plot after transduction of HEK293T *TLR-Sce* cells with indicated amounts of *I-SceI* plus donor lentivirus (LV). p24 values indicate the amount of lentiviral capsid protein added to cells. Numbers inside plots indicate percentages of live cells. FI, relative fluorescence intensity reported in arbitrary units. (**b**) Quantification of data from **a**. Bar graphs represent a minimum of three independent experiments performed in duplicate, with s.e. shown. (**c**) Ratio of HDR to mutNHEJ based on data in **b**. (**d**) Representative flow plot after titration of donor integrase-deficient lentivirus in fivefold increments with constant I-SceI lentiviral vector dose in HEK293T *TLR-Sce* cells. (**e**) Quantification of data from **d**. Bars represent a minimum of three independent experiments performed in duplicate, with s.e. shown. (**f**) Ratio of HDR to mutNHEJ in **e**. NA, not applicable.

© 2011 Nature America, Inc. All rights reserved.

would have unique repair profiles. We examined the TLR readout in a population of co-transduced cells as a function of nuclease and donor template amounts by gating on cells with unique mean fluorescence intensity values for each parameter. Control gates agreed with previous experiments, such that cells expressing I-SceI alone (IFP^+, BFP^-) predominately generated mutNHEJ events, and cells with high amounts of both I-SceI and donor template (IFP^+ and BFP^+) exhibited both mutNHEJ and gene-targeting signals (**Fig. 3b**). In gates where donor template delivery was constant and nuclease expression increased, the absolute number of observed engineering events increased, whereas the ratio of HDR to mutNHEJ remained nearly constant (**Fig. 3c**). Conversely, in gates where nuclease expression was constant and donor template delivery increased, the ratio of gene targeting to mutNHEJ changed drastically, with cells containing large amounts of donor template increasingly trending toward gene targeting (**Fig. 3d**). We also observed a shift from mutNHEJ to HDR as a function of donor delivery for a zinc-finger nuclease–mediated DNA break (**Supplementary Fig. 6**).

Together, these results confirmed that the efficiency and outcome of genome engineering vary substantially at the population level, with increased nuclease delivery being a strong predictor of the likelihood of an engineering 'event' and increased donor delivery being a strong predictor of break resolution via gene targeting. They also illustrate how quantitative tracking of nuclease and donor delivery in single cells may be applied to isolate cell populations with high frequencies of desired engineering outcomes.

Single- versus double-strand break–induced engineering

As single-strand breaks are a potential alternative approach to catalyzing site-specific HDR while minimizing mutagenic outcomes[10,21,22], 'nickases' may be useful for genome engineering when safety or fidelity is necessary. We compared single-strand and double-strand break–induced genome engineering using the TLR (**Fig. 4**). We derived a TLR cell line containing an I-AniI nuclease target site (HEK293T *TLR-Ani*) and initiated readout with either a double-strand break–inducing nuclease, the optimized I-AniI Y2 variant (ref. 23) (cleavase) or a variant that creates a single-strand break at the identical target site, I-AniI Y2 K227M[21] (nickase) (**Fig. 4a**). To control for nuclease delivery, we added a T2A-linked BFP tracking fluorophore to both enzymes. Upon transduction of the HEK293T *TLR-Ani* cells with *I-AniIY2-T2A-BFP* or *I-AniIY2K227M-T2A-BFP* lentivirus and donor integration-deficient lentivirus, we observed a 100-fold reduction in

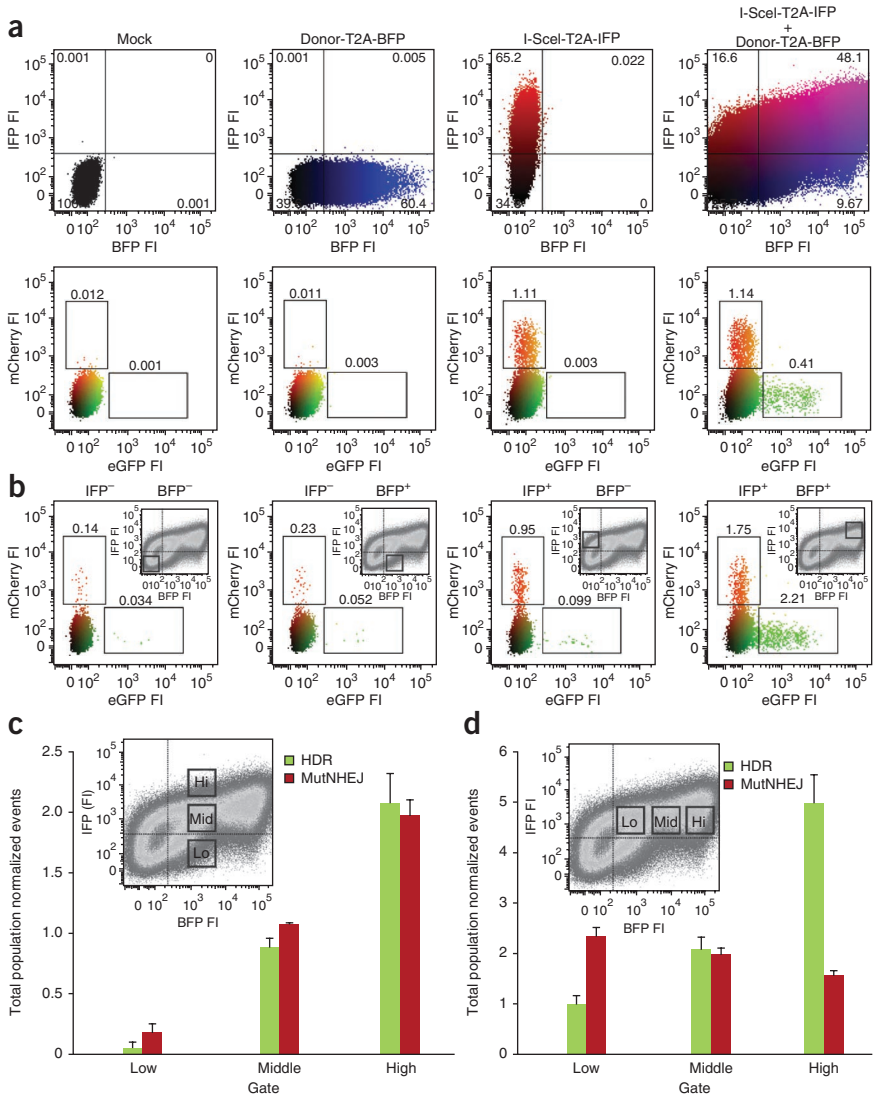

Figure 3 | Four-color system to track nuclease and donor template delivery simultaneously with the TLR. (**a**) Representative flow plot 72 h after transduction of HEK293T *TLR-Sce* cells with *I-SceI-T2A-IFP* lentivirus and *Donor-T2A-BFP* integration-deficient lentivirus. Shown are nuclease and donor tracking (top), and engineering outcome (bottom). Numbers inside plots indicate percentages of live cells. (**b**) Control gating analysis of HEK293T *TLR-Sce* cells transduced with both *I-SceI-T2A-IFP* lentivirus and *Donor-T2A-BFP* integration-deficient lentivirus. Inset, flow plots showing nuclease and donor template expression levels as indicated by mean fluorescence intensity (FI). Main plots show readout from the TLR as a function of the gate shown in the inset. (**c**) Quantification of TLR readout when applying a nuclease titration gating analysis in cells transduced with both *I-SceI-T2A-IFP* and *Donor-T2A-BFP*. Bars represent the amount of gene targeting and mutNHEJ present in the indicated inset gates (low, middle and high mean fluorescence intensity values) normalized to the HDR and mutNHEJ values for the total ungated population. Average data of three independent experiments are shown with s.e. (**d**) Quantification of TLR readout when applying donor template titration gating analysis as indicated above. Error bars, s.e. (n = 3).

total TLR for the nickase at the population level (**Fig. 4b**). Of the detected repair events, the nickase exhibited a strong bias toward gene targeting over mutNHEJ (**Fig. 4c**). We also consistently observed a small portion of red fluorescent cells presumed to result from mutNHEJ events after nickase expression, suggesting that a fraction of events were processed through a double-strand break intermediate. Using the tracking BFP, we conducted a gating analysis to determine the effect of nickase expression on TLR readout (**Fig. 4d**). This revealed a dose-dependent increase in gene-targeting events, approaching 0.25% in the highest-expressing cells, without an associated increase in mutNHEJ.

These results strongly support the concept that substituting a nickase in place of a cleavase is an effective means to induce HDR while minimizing mutNHEJ events. They also highlight a key utility of the TLR system: by providing concurrent quantification of gene targeting and mutNHEJ events, we could rapidly analyze rare repair events occurring at a nuclease-induced DNA single-strand break.

siRNA kinome screen for genome engineering enhancers

In many contexts, efficiency of genomic modification is likely to be the highest priority of a genome engineer. Therefore, we evaluated the TLR system as a high-throughput screening platform for DNA-repair modulators that might improve engineering efficiency. As the DNA damage response is highly regulated by phosphorylation[24], we evaluated a human kinome siRNA library in an initial screen. As seen in the *Z*-score scatter plot (**Fig. 5a** and **Supplementary Fig. 7**), we observed a variety of HDR and mutNHEJ phenotypes. Several kinases known to have a role in DNA repair appeared at or near the top of ranked lists (**Supplementary Table 1**). These included DNA-PKcs, TLK-1 and several kinases in the higher-order inositol phosphate pathway (IHPK3, IMPK and PIK3C2B), which have recently been shown to be involved in classical nonhomologous end joining[25,26]. In addition, knockdown of several genes encoding kinases not previously implicated in DNA repair also resulted in positive *Z* scores.

As we had observed suppression of mutNHEJ in cells with abundant donor template (**Figs. 2e** and **3c**), we hypothesized that coupling knockdown of a repair modulator with gating for cells with high enzyme and donor levels would allow for high rates of gene targeting while minimizing the associated mutNHEJ. To test this, we conducted a gating analysis for various transduction levels of HEK293T *TLR-Sce* cells pretreated with siRNAs to *PRKDC*, which encodes DNA-PKcs and also transduced with both *I-SceI-T2A-IFP* lentivirus and donor integration-deficient lentivirus virus (**Fig. 5b**). As transduction levels increased, we observed a tenfold increase in gene targeting events with only a modest associated increase in mutNHEJ events, achieving absolute gene targeting rates of nearly 10% in highly transduced, *PRKDC* knockdown cells, compared to <1% in ungated, control siRNA–treated cells.

© 2011 Nature America, Inc. All rights reserved.

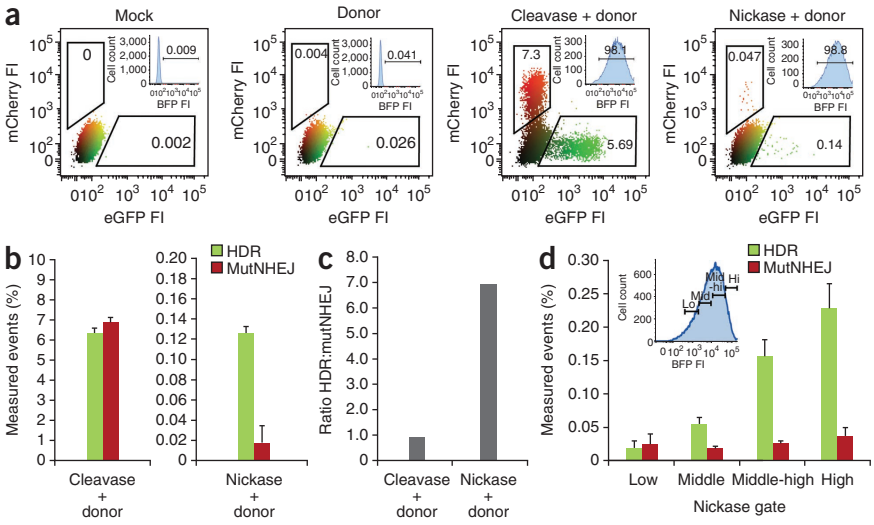

Figure 4 | Effect of single versus double-strand DNA breaks on engineering outcome. (**a**) Representative flow plots showing TLR readout of HEK293T *TLR-Ani* cells transduced with either *I-AniY2-T2A-BFP* lentivirus (cleavase) or *I-AniIK227M-T2A-BFP* lentivirus (nickase). Insets, gating for nuclease expression to control for transduction levels. (**b**) Quantification of data shown in **a** from three independent experiments in duplicate. Percentage measured events have had the background rates from cells transduced with donor alone subtracted to control for the low numbers. Error bars, s.e. (**c**) Comparison of the ratio of HDR to mutNHEJ between cleavase- and nickase-induced engineering. (**d**) Gating analysis showing TLR readout across nickase expression levels. Error bars, s.e. (three independent experiments performed in duplicate). Inset, BFP histogram gated for relative nickase expression.

Although we consistently observed increases in gene targeting after *PRKDC* knockdown in multiple TLR clones (data not shown) and several different reporter systems using the I-SceI nuclease, the magnitude of the effect varied with the enzyme used to initiate the break (**Supplementary Fig. 8**).

DISCUSSION

The TLR system has two key advantages over previously described systems (for example, refs. 9,11–17). (i) It provides a rapid positive signal readout of both mutNHEJ and gene targeting occurring at a single break site, allowing direct assessment of the competition between repair mechanisms. (ii) We combined it with the capacity to track and control for both nuclease and donor template delivery. Previously described reporters have all focused on either nonhomologous end joining or HDR. Assessing repair profiles in cell lines with multiple reporters is complicated by epigenetic effects at different reporter loci, a requirement for multiple breakpoints and a lack of direct competition for resolution[27]. Although it would be possible to obtain information on competition between repair processes at a single DNA break site by using one of the previously described HDR reporters, doing so would require high-throughput DNA sequencing of PCR-amplified loci to obtain statistically meaningful information, an approach incompatible with even medium-throughput screening for genome engineering outcome modulators.

Our siRNA kinome screen for modulators of genome engineering outcome identified several genes whose silencing resulted in enhanced rates of both mutNHEJ and gene targeting. Prominent among these was *PRKDC*, which encodes a component of the classical nonhomologous end-joining complex that has previously been shown to improve homologous recombination rates[28,29]. This observation is consistent with a growing literature indicating that nuclease-induced breaks are often subject to precise rejoining of ends[15,19], preventing them from engaging 'engineering-productive' pathways. Whereas we consistently observed the effect of *PRKDC* knockdown in several DNA-repair contexts, it varied among nucleases with varying affinities and activities toward the same target site, suggesting that the biochemical and biophysical properties of the nuclease may influence processing of the break by the host DNA repair machinery. This is an observation with important implications for nuclease-induced genome engineering, as certain enzymes may have a propensity to engage particular DNA repair pathways, or be more or less responsive to repair pathway manipulations. It also raises important considerations for investigators studying DNA repair using nuclease-induced breaks, who have largely focused on breaks induced by a single enzyme, I-SceI.

The TLR should be applicable in diverse organisms for identifying new proteins involved in DNA repair or pathway choice, evaluating new approaches to inducing targeted breaks, screening for small-molecule modulators of specific repair pathways, and rapidly and comprehensively vetting 'third party' manipulations aimed at increasing endonuclease-induced engineering efficiency.

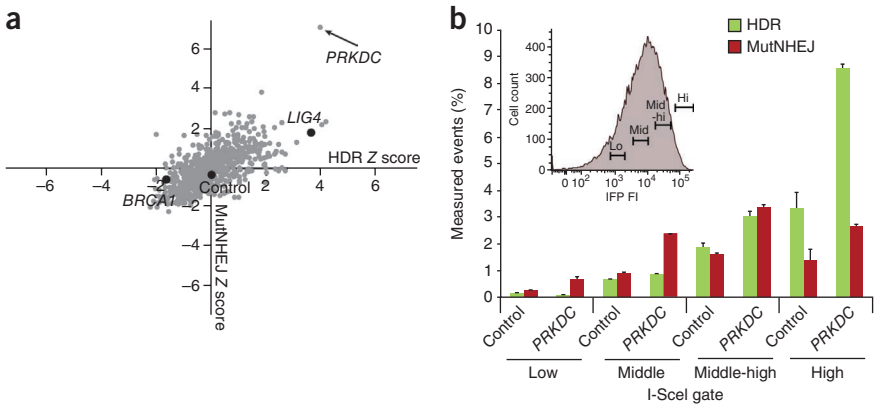

Figure 5 | High-throughput siRNA kinome screen to identify modifiers of engineering outcome. (**a**) Scatter plot of gene targeting and mutNHEJ *Z* scores obtained from the siRNA screen. Library data are in gray. Control siRNA values are an average for at least three independent transfections. (**b**) Gating analysis comparing TLR readout from control and *PRKDC* siRNA treatment as a function of nuclease expression 72 h after transduction of HEK293T *TLR-Sce* cells transduced with *I-SceI-T2A-IFP* lentivirus and donor integration-deficient lentivirus. Inset, nuclease expression gates. Data are derived from three independent experiments performed in duplicate. Error bars, s.e.

© 2011 Nature America, Inc. All rights reserved.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

Accession codes. Addgene: 31475 (pCVL SFFV d14GFP donor), 31476 (pCVL SFFV d14GFP EF1s HA.NLS.Sce(opt)), 31477 (pCVL SFFV HA.NLS.I-AniIY2(reo).T2A.TagBFP), 31478 (pCVL SFFV HA.NLS.I-AniIY2K227M(reo).T2A.TagBFP), 31479 (pCVL SFFV-EF1s HA.NLS.Sce(opt)), 31480 (pCVL traffic light reporter 1.1 (Ani target) Ef1a Puro), 31481 (pCVL traffic light reporter 1.1 (Sce target) Ef1a BFP), 31482 (pCVL traffic light reporter 1.1 (Sce target) Ef1a Puro), 31483 (pCVL traffic light reporter 2.1 (VF2468 ZFN target) Ef1a), 31484 (pRRL sEF1a HA.NLS.Sce(opt).T2A.IFP) and 31485 (pRRL SFFV d20GFP.T2A.mTagBFP donor).

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

M.T.C. was supported in part by Public Health Service, National Research Service Award, T32 GM07270, from the US National Institute of General Medical Sciences. Additional funding was from US National Institutes of Health (RL1CA133832, UL1DE019582, R01-HL075453, PL1-HL092557 and RL1-HL092553) and Seattle Children's Center for Immunity and Immunotherapies. We thank C. Ramirez and K. Joung for zinc-finger nuclease constructs (Harvard University, Massachusetts General Hospital), and all members of the Northwest Genome Engineering Consortium for their many insightful discussions.

AUTHOR CONTRIBUTIONS

M.T.C. designed and performed experiments, analyzed data and wrote the paper; B.Y.R., J.E.A. and M.G. performed experiments; J.J. and D.J.R. designed experiments; and A.M.S. designed experiments and wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturemethods/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Carr, P.A. & Church, G.M. Genome engineering. *Nat. Biotechnol.* **27**, 1151–1162 (2009).
2. Pâques, F. & Duchateau, P. Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. *Curr. Gene Ther.* **7**, 49–66 (2007).
3. Durai, S. *et al.* Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Res.* **33**, 5978–5990 (2005).
4. Porteus, M.H. & Carroll, D. Gene targeting using zinc finger nucleases. *Nat. Biotechnol.* **23**, 967–973 (2005).
5. Caldecott, K.W. Single-strand break repair and genetic disease. *Nat. Rev. Genet.* **9**, 619–631 (2008).
6. Shrivastav, M., De Haro, L.P. & Nickoloff, J.A. Regulation of DNA double-strand break repair pathway choice. *Cell Res.* **18**, 134–147 (2008).
7. Branzei, D. & Foiani, M. Regulation of DNA repair throughout the cell cycle. *Nat. Rev. Mol. Cell Biol.* **9**, 297–308 (2008).
8. Cann, K.L. & Hicks, G.G. Regulation of the cellular DNA double-strand break response. *Biochem. Cell Biol.* **85**, 663–674 (2007).
9. Porteus, M.H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**, 763 (2003).
10. Metzger, M.J., McConnell-Smith, A., Stoddard, B.L. & Miller, A.D. Single-strand nicks induce homologous recombination with less toxicity than double-strand breaks using an AAV vector template. *Nucleic Acids Res.* **39**, 926–935 (2011).
11. Bennardo, N., Cheng, A., Huang, N. & Stark, J.M. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet.* **4**, e1000110 (2008).
12. Stark, J.M., Pierce, A.J., Oh, J., Pastink, A. & Jasin, M. Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Mol. Cell. Biol.* **24**, 9305–9316 (2004).
13. Nagaraju, G., Hartlerode, A., Kwok, A., Chandramouly, G. & Scully, R. XRCC2 and XRCC3 regulate the balance between short- and long-tract gene conversions between sister chromatids. *Mol. Cell. Biol.* **29**, 4283–4294 (2009).
14. Brenneman, M.A., Wagener, B.M., Miller, C.A., Allen, C. & Nickoloff, J.A. XRCC3 controls the fidelity of homologous recombination: roles for XRCC3 in late stages of recombination. *Mol. Cell* **10**, 387–395 (2002).
15. Guirouilh-Barbat, J., Rass, E., Plo, I., Bertrand, P. & Lopez, B.S. Defects in XRCC4 and KU80 differentially affect the joining of distal nonhomologous ends. *Proc. Natl. Acad. Sci. USA* **104**, 20902–20907 (2007).
16. Pierce, A.J., Johnson, R.D., Thompson, L.H. & Jasin, M. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev.* **13**, 2633–2638 (1999).
17. Aubert, M. *et al.* Successful targeting and disruption of an integrated reporter lentivirus using the engineered homing endonuclease Y2 I-AniI. *PLoS ONE* **6**, e16825 (2011).
18. Sarkis, C., Philippe, S., Mallet, J. & Serguera, C. Non-integrating lentiviral vectors. *Curr. Gene Ther.* **8**, 430–437 (2008).
19. Bennardo, N., Gunn, A., Cheng, A., Hasty, P. & Stark, J.M. Limiting the persistence of a chromosome break diminishes its mutagenic potential. *PLoS Genet.* **5**, e1000683 (2009).
20. Kustikova, O.S. *et al.* Dose finding with retroviral vectors: correlation of retroviral vector copy numbers in single cells with gene transfer efficiency in a cell population. *Blood* **102**, 3934–3937 (2003).
21. McConnell Smith, A. *et al.* Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-AniI LAGLIDADG homing endonuclease. *Proc. Natl. Acad. Sci. USA* **106**, 5099–5104 (2009).
22. Lee, G.S., Neiditch, M.B., Salus, S.S. & Roth, D.B. RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. *Cell* **117**, 171–184 (2004).
23. Takeuchi, R., Certo, M., Caprara, M.G., Scharenberg, A.M. & Stoddard, B.L. Optimization of in vivo activity of a bifunctional homing endonuclease and maturase reverses evolutionary degradation. *Nucleic Acids Res.* **37**, 877–890 (2009).
24. Matsuoka, S. *et al.* ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160–1166 (2007).
25. Ma, Y. & Lieber, M.R. Binding of inositol hexakisphosphate (IP6) to Ku but not to DNA-PKcs. *J. Biol. Chem.* **277**, 10756–10759 (2002).
26. Kumar, A., Fernandez-Capetillo, O. & Carrera, A.C. Nuclear phosphoinositide 3-kinase β controls double-strand break DNA repair. *Proc. Natl. Acad. Sci. USA* **107**, 7491–7496 (2010).
27. Guirouilh-Barbat, J. *et al.* Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells. *Mol. Cell* **14**, 611–623 (2004).
28. Słabicki, M. *et al.* A genome-scale DNA repair RNAi screen identifies SPG48 as a novel gene associated with hereditary spastic paraplegia. *PLoS Biol.* **8**, e1000408 (2010).
29. Pierce, A.J., Hu, P., Han, M., Ellis, N. & Jasin, M. Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. *Genes Dev.* **15**, 3237–3242 (2001).

© 2011 Nature America, Inc. All rights reserved.

ONLINE METHODS

Construct assembly. All constructs were cloned into the 'RRL' (Addgene 12252) or 'CVL' lentiviral backbones using standard molecular biology techniques and have been deposited with Addgene. Sequences and maps are available in **Supplementary Note 2**. The *mCherry* sequence was mutated by standard site-directed mutagenesis to remove internal start codons that contributed to background fluorescence M9S and M16L) (primers are listed in **Supplementary Table 2**). *mTagBFP* was purchased from Evrogen. *IFP1.4* (ref. 30) was gene-synthesized (Genscript).

Lentivirus generation. Lentivirus was produced by transient co-transfection of HEK293T cells in 10-cm dishes in 10 ml of medium using PEI transfection reagent (Polysciences) with 6 μg RRL or CVL backbone plamids, 1.5 μg pMD2G envelope plasmid (VSV-G) and 3 μg psPAX2 for integrating lentivirus and psPAX2 D64V for integration-deficient lentivirus, per plate. Viral supernatants were concentrated 100× by overnight centrifugation at 8,000*g*, followed by aspiration of supernatant and resuspension in 1/100th the original volume. We aliquoted 100× stocks and stored them at −80 °C. Virus was titered using Lenti-x p24 rapid titer ELISA kit (Clontech) according to the manufacturer's protocol. Multiplicity of infection (MOI) was calculated for fluorescent protein–tagged virus experiments by dividing the amount of infectious units added by the number of cells plated. Infectious units were calculated by transducing 0.2×10^6 HEK293T cells with various amounts of viral stocks and analyzing them on a flow cytometer 72 h after transduction and applying the following formula for volumes of virus that yielded between 5–20% fluorescent cells (($0.2 \times 10^6 \times$ percentage fluorescent cells)/100) / volume viral stock added).

Cell-line generation. Cell lines containing the TLR were generated by transducing 0.2×10^6 HEK293T cells with 1 μl of unconcentrated reporter lentivirus, typically yielding ~5% transduction based on fluorescent lentivirus prepared in parallel. Three days after transduction, cells with integrated reporters were selected for in 1 μg ml^{-1} puromycin for 5 d. Puromycin-resistant cells were then sorted to remove the ~0.1% of cells exhibiting red (mCherry) fluorescence, thought to be due to integration errors.

Transduction. We seeded 0.1×10^6 HEK293T cells in a 24-well plate 24 h before transduction. Cells were transduced with the following amounts of lentivirus: for experiments shown in **Figure 1**, 25 ng for each lentivirus; in **Figure 2**, as indicated in the text; in **Figure 3**, 25 ng each virus, corresponding to an MOI of 8; in **Figure 4**, 125 ng p24 for *I-AniIY2-T2A-BFP* and *I-AniIY2K227M-T2A-BFP* lentivirus, corresponding to an MOI of 13, and 25 ng donor integration-deficient lentivirus; in **Figure 5**, 3.0 ng p24 *I-Sce* plus donor lentivirus per well for high-throughput screen, for *PRKDC* knockdown experiments, *I-SceI-T2A-IFP* lentivirus was added at an MOI of 13 along with 25 ng p24 donor integration-deficient lentivirus. All transductions were done in the presence of 4 μg of polybrene. Twenty-four hours after transduction, medium was changed, and cells were split and analyzed 72 h after transduction.

Flow cytometry. Cells were collected 72 h after transduction and analyzed on a BD LSRII or BD Facs ARIA. eGFP fluorescence was measured using a 488-nm laser for excitation and a 530/30 filter for detection. mCherry fluorescence was measured by using a 561-nm laser for excitation and a 610/20 filter for detection. mTagBFP fluorescence was measured using a 405-nm laser for excitation and a 450/50 filter for detection. IFP1.4 signal was measured using a 640 nm laser and acquired with a 710/50 filter for detection. Biliverdin was not used for detecting IFP1.4, as it was determined unnecessary for detection at high transduction levels (data not shown). Data were analyzed using FloJo software. For sorting experiments, cells were sorted on BD Facs ARIA.

High-throughput siRNA kinome screen. RNA inhibition was carried out using a library to the human kinome (Sigma) containing three pooled siRNAs per gene in 384-well format containing positional constant controls on every plate. Using predetermined optimized conditions (over 80% transfection efficiency), six replicate transfections per gene were performed. We aliquoted 3.0 ng p24 lentivirus containing both I-SceI and donor template to each well 24 h after siRNA tranfection. Seventy-two hours after infection the cells were lifted via the addition of 10 mM EDTA, and the six replicates were pooled into a single well for flow-cytometry analysis. All automation was performed using a CyBio Vario and a Thermo Fisher Wellmate. Cytometry data were analyzed using FloJo software. Wells with less than 2,000 total cell counts were excluded from the analysis. HDR and mutNHEJ gates were initially applied to all samples universally, then inspected manually and shifted as necessary. Annotated data from flow cytometry analysis underwent *Z*-score analysis for each metric obtained and were used to generate ranks. Screening hits beyond these thresholds were followed up on via the determinants of biological interest and validation of previously reported trends. Hit validation was initially performed with three pooled siRNAs for each target from Qiagen in triplicate.

siRNA knockdown. For knockdown experiments, 0.05×10^6 HEK293T TLR cells were plated 24 h before transfection in a 24-well plate with 0.5 ml of medum. A total of 10 nM of siRNA pool was used per well and transfected using RNAi-max reagent (Invitrogen) according to the manufacturer's protocol. Twenty-four hours after transfection, medium was changed, and cells were transduced as indicated above and analyzed 72 h after transduction. A minimum of two experimentally validated siRNAs (shown by Qiagen to knock down greater than 70% mRNA) per indicated target were obtained from Qiagen and pooled. Control siRNA was a 'universal' control from Qiagen.

Statistical analysis. Error bars in bar graphs represent s.e., as calculated by dividing the s.d. of replicates by the square root of the number of total replicates. For normalized data, the percent measured value for HDR and mutNHEJ after flow-cytometric analysis of a given treatment was divided by the percent measured value of HDR and mutNHEJ for the indicated parent population (control siRNA or ungated population). *Z* score was calculated by subtracting the average of a metric (HDR or mutNHEJ) from the raw score of a given siRNA and dividing by the s.d. of the same metric. Metric averages and s.d. were calculated from the library values where more than 2,000 cells were acquired during the screen.

30. Shu, X. *et al.* Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science* **324**, 804–807 (2009).

© 2011 Nature America, Inc. All rights reserved.