

Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9

Dominik Paquet^{1*}, Dylan Kwart^{1*}, Antonia Chen¹, Andrew Sproul^{2†}, Samson Jacob², Shaun Teo¹, Kimberly Moore Olsen¹, Andrew Gregg^{1,3}, Scott Noggle² & Marc Tessier-Lavigne¹

The bacterial CRISPR/Cas9 system allows sequence-specific gene editing in many organisms and holds promise as a tool to generate models of human diseases, for example, in human pluripotent stem cells^{1,2}. CRISPR/Cas9 introduces targeted double-stranded breaks (DSBs) with high efficiency, which are typically repaired by non-homologous end-joining (NHEJ) resulting in nonspecific insertions, deletions or other mutations (indels)². DSBs may also be repaired by homology-directed repair (HDR)^{1,2} using a DNA repair template, such as an introduced single-stranded oligo DNA nucleotide (ssODN), allowing knock-in of specific mutations³. Although CRISPR/Cas9 is used extensively to engineer gene knockouts through NHEJ, editing by HDR remains inefficient^{3–8} and can be corrupted by additional indels⁹, preventing its widespread use for modelling genetic disorders through introducing disease-associated mutations. Furthermore, targeted mutational knock-in at single alleles to model diseases caused by heterozygous mutations has not been reported. Here we describe a CRISPR/Cas9-based genome-editing framework that allows selective introduction of mono- and bi-allelic sequence changes with high efficiency and accuracy. We show that HDR accuracy is increased dramatically by incorporating silent CRISPR/Cas-blocking mutations along with pathogenic mutations, and establish a method termed ‘CORRECT’ for scarless genome editing. By characterizing and exploiting a stereotyped inverse relationship between a mutation’s incorporation rate and its distance to the DSB, we achieve predictable control of zygosity. Homozygous introduction requires a guide RNA targeting close to the intended mutation, whereas heterozygous introduction can be accomplished by distance-dependent suboptimal mutation incorporation or by use of mixed repair templates. Using this approach, we generated human induced pluripotent stem cells with heterozygous and homozygous dominant early onset Alzheimer’s disease-causing mutations in amyloid precursor protein (APP^{Swe})¹⁰ and presenilin 1 (PSEN1^{M146V})¹¹ and derived cortical neurons, which displayed genotype-dependent disease-associated phenotypes. Our findings enable efficient introduction of specific sequence changes with CRISPR/Cas9, facilitating study of human disease.

While attempting to knock-in early onset Alzheimer’s disease mutations into iPS cells using CRISPR/Cas9, we detected HDR by presence of an intended mutation provided via the cognate ssODNs, however most HDR events also contained unwanted indels (Fig. 1a). This is presumably due to the high nuclease activity of CRISPR/Cas9 (refs 3, 4, 6, 8), which may continuously re-cut edited loci until sufficient modification by NHEJ prevents further targeting. If so, this re-editing may be blocked by simultaneously mutating the NGG protospacer adjacent motif (PAM) or guide RNA binding sequence, which CRISPR/Cas9 requires for targeting², as shown in

prokaryotes¹². As the efficacy of potential blocking mutations has not been systematically studied in eukaryotic cells, we tested their effect on HDR accuracy in wild-type human induced pluripotent stem cells (iPS cells) (Extended Data Fig. 1) and, for comparison, human embryonic kidney (HEK293) cells. We introduced Cas9–eGFP and single guide RNA (sgRNA) plasmids together with five pooled repair ssODN templates, which in addition to the APP^{Swe} or PSEN1^{M146V} pathogenic mutation also contained a putative silent CRISPR/Cas-blocking mutation in the PAM or guide RNA target sequence, or a control non-blocking mutation outside those regions (Fig. 1b, c and Supplementary Tables 1 and 2).

We analysed genomic loci of Cas9–eGFP-expressing cells by next-generation sequencing and determined the fraction of HDR reads that were ‘accurate’, that is, without undesirable indel modifications (Fig. 1d, e; Extended Data Table 1 lists overall HDR rates for all experiments). Without blocking mutations, only 6 to 35% of reads that incorporated pathogenic mutations had accurate HDR, but presence of a CRISPR/Cas-blocking PAM mutation increased HDR accuracy in both iPS cells and HEK293 cells two- to tenfold, depending on locus and cell type, which may increase the probability of accurately editing both alleles in a cell up to 100-fold (assuming independent allele editing). The remaining ‘inaccurate’ HDR events were presumably generated by prior or concomitant NHEJ. Blocking mutations targeting the guide RNA sequence increased HDR accuracy to a similar extent for APP, but much less for PSEN1 (Fig. 1d, e). Therefore, whereas PAM-site mutations seem broadly effective, guide RNA target mutations may have variable effects at different loci. Similar results were obtained for ssODNs transfected individually rather than pooled (Extended Data Fig. 2a, b). Indel frequency, position, and size had expected distributions^{1,13} (Extended Data Fig. 3a). Interestingly, in experiments with pooled ssODNs, up to 11% of HDR reads contained multiple blocking or control mutations (Extended Data Fig. 2c, d), showing that cells used multiple oligonucleotides in multiple rounds of repair, and highlighting the propensity of CRISPR/Cas9 for re-editing. Thus, CRISPR/Cas-blocking mutations, preferably in the PAM, minimize undesirable re-editing during derivation of knock-in mutant clones.

Introducing silent blocking mutations in coding regions is often possible, though in some cases silent mutations may be precluded by the PAM reading frame or prove ineffective in the guide RNA target. Furthermore, in non-coding regions, blocking mutations may have unwanted consequences. Intended mutations may occasionally double as blocking mutations, but this is not always the case. We therefore developed a method to remove blocking mutations when desired, termed CORRECT (consecutive re-guide or re-Cas steps to erase CRISPR/Cas-blocked targets), with two variants: re-guide and re-Cas. In both, blocking mutations are first introduced together with

¹Laboratory of Brain Development and Repair, The Rockefeller University, 1230 York Avenue, New York, New York 10065, USA. ²The New York Stem Cell Foundation Research Institute, New York, New York 10032, USA. ³Weill Cornell Graduate School of Medical Sciences, The Rockefeller University and Sloan-Kettering Institute Tri-institutional MD-PhD Program, 1300 York Avenue, New York, New York 10065, USA. †Present address: Department of Pathology and Cell Biology and the Taub Institute for Research on Alzheimer’s Disease and the Aging Brain, Columbia University Medical Center, 630 West 168th Street, New York, New York 10032, USA.

*These authors contributed equally to this work.

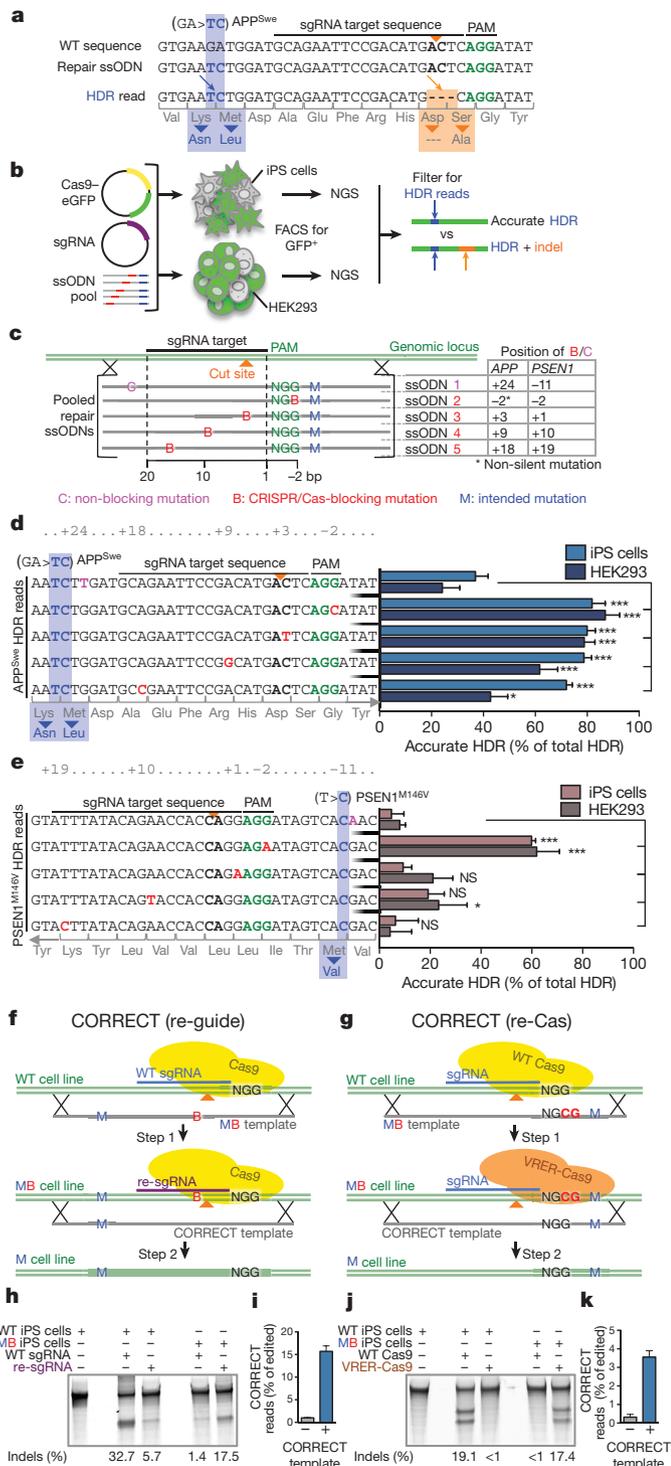


Figure 1 | CRISPR/Cas9 blocking mutations increase HDR accuracy by preventing re-editing and can be used for scarless CORRECT editing.
a, APP sequencing alignment showing concomitant HDR (blue arrow) and indels (orange arrow) after editing. **b**, Experimental setup for gene editing analysis by next-generation sequencing (NGS). **c**, Pooled ssODNs used to test effects of CRISPR/Cas9-blocking mutations. **d**, **e**, Percentages of accurate HDR for blocking or control mutations at *APP* (**d**) and *PSEN1* (**e**) loci in iPS cells and HEK293 cells. Values represent mean \pm s.e.m. ($n = 3$). *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, one-way ANOVA. **f**, **g**, Two-step workflow for CORRECT variants re-guide (**f**) and re-Cas (**g**): re-guide uses a blocking mutation B in the guide RNA target sequence, whereas for re-Cas the PAM is mutated to a sequence detected by a Cas9 variant. Blocking mutations are removed in step 2 using re-sgRNA/WT-Cas9 or WT-sgRNA/VRER-Cas9, while pathogenic mutations M are retained. **h**, **j**, Surveyor mismatch cleavage assay detecting CRISPR/Cas9 activity shows specificity of WT-Cas9/WT-sgRNAs for wild-type targets, and WT-Cas9/re-sgRNA (**h**) or VRER-Cas9/WT-sgRNA (**j**) for mutated loci. **i**, **k**, Next-generation sequencing quantification of genomes with sequence inserted by HDR with CORRECT templates in pooled iPS cells ($n = 2$). WT, wild-type.

Cas9/re-sgRNA (for APP^{Swe}) or VRER-Cas9/sgRNA (for APP^{A673T}) (Fig. 1h, j). The expected editing events were detected with high efficiency by next-generation sequencing (Fig. 1i, k). Thus, CORRECT enables efficient scarless introduction of just an intended mutation. We next examined mutational status of the two alleles in individual iPS cell clones. We could readily isolate clones with homozygous early onset Alzheimer’s disease mutations, but, interestingly, in clones heterozygous for early onset Alzheimer’s disease mutations, the ‘non-HDR’ allele almost always contained indels (Extended Data Figs 3b and 4a, b). This is possibly due to the high efficiency of Cas9 (refs 3, 4, 6, 8), which results mostly in bi-allelic modifications^{4–7}, and raises the question of how to isolate heterozygous clones.

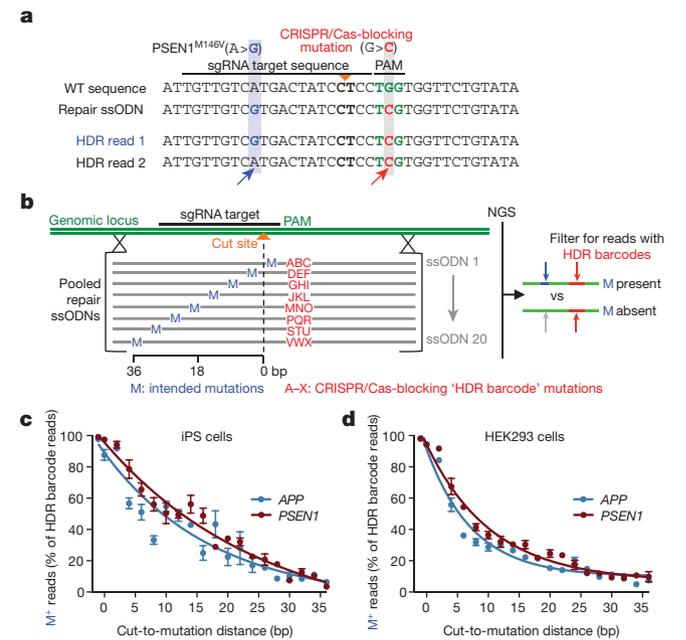


Figure 2 | A monotonic inverse relationship between mutation incorporation and distance from the CRISPR/Cas9 cleavage site.
a, *PSEN1* sequencing alignment showing introduction of a CRISPR/Cas9-blocking mutation (red arrow) with or without the pathogenic mutation (blue arrow) during HDR. **b**, Pooled ssODNs used to scan mutation incorporation rates based on cut-to-mutation distance. Barcode mutations (red) identify HDR-reads and mutation M position during next-generation sequencing (NGS) analysis. **c**, **d**, A monotonic relationship governs rate of mutation M incorporation and cut-to-mutation distance during HDR in both iPS cells (**c**) and HEK293 cells (**d**) ($n = 4$ for iPS, $n = 3$ for HEK293); goodness of fit: R^2 (*APP*) = 0.75 (iPS) / 0.96 (HEK293), R^2 (*PSEN1*) = 0.94 (iPS) / 0.97 (HEK293); curves for *APP* and *PSEN1* are not significantly different, two-tailed t -test: $P = 0.31$ (iPS) / 0.06 (HEK293).

intended mutations by HDR. Whereas the re-guide blocking mutation interferes with guide RNA targeting, the re-Cas mutation blocks PAM detection by mutating the NGG to the target sequence of a specificity-modified Cas9 (in our experiments, NGCG, target of the recently described VRER-Cas9 (ref. 14)). The blocking mutation is then removed with modified reagents: for re-guide, a re-sgRNA targeting the modified sequence is used with wild-type Cas9 (Fig. 1f). For re-Cas, the modified PAM is targeted with the Cas9 variant (Fig. 1g). We tested the feasibility of CORRECT by re-guide using an APP^{Swe} iPS cell line containing a guide RNA target mutation (see Fig. 1); for re-Cas, we generated an APP^{A673T} mutant iPS cell line with a NGCG PAM mutation (Extended Data Fig. 2e, f). We then removed the blocking mutations from both lines with CORRECT templates and wild-type

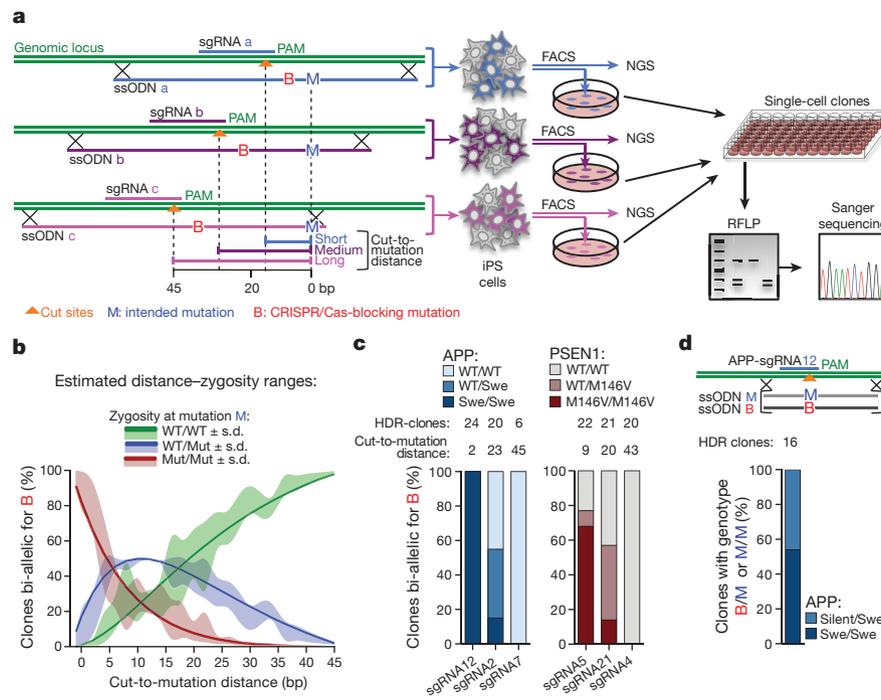


Figure 3 | Introduction of heterozygous or homozygous mutations into iPSC cells by manipulating the cut-to-mutation distance or using mixed HDR templates. **a**, Experimental setup with three sgRNA/ssODN pairs per locus with increasing cut-to-mutation distance. Edited iPSC cells were analysed by next-generation sequencing (NGS; see Extended Data Fig. 5) or grown for clonal analysis. **b**, Predicted distance ranges for desired zygosity, calculated based on oligonucleotide scan data (see Fig. 2c and

Methods). **c**, Frequency of different APP and PSEN1 mutation genotypes in single-cell clones with bi-allelic HDR of blocking mutations. Indicated zygosity fits to predicted values. **d**, Introduction of heterozygous mutations with mixed repair ssODNs. One ssODN contained the pathogenic APP^{Swe} mutation, which blocks sgRNA12 (M), the other a silent blocking mutation (B) (see alternative approach with two silent blocking mutations in Extended Data Fig. 5d).

A first approach came from our observation that many alleles that incorporated a silent CRISPR/Cas-blocking mutation did not contain the intended pathogenic mutation, particularly if it was distant from the CRISPR/Cas9 cleavage site (Fig. 2a). This was similar to reports of distance dependence for editing with CRISPR/Cas9 (refs 9, 15) or other systems^{13,16–19}. We reasoned that a predictable relationship between distance and mutation incorporation could be exploited to control allelic mutation incorporation. We therefore characterized distance dependence at the APP and PSEN1 loci by scanning mutation incorporation rates with 20 different pooled ssODNs, each containing a unique CRISPR/Cas-blocking three-base-pair barcode sequence, as well as single point mutations at increasing distances from the cleavage site (Fig. 2b and Supplementary Table 2). Notably, we found a clear monotonic inverse relationship between rate of mutation incorporation and distance from cleavage site that did not differ significantly for APP and PSEN1 in either iPSC cells or HEK293 cells (Fig. 2c, d). The relationship was also similar for longer ssDNA or dsDNA HDR repair templates (Extended Data Fig. 4d), and for three distinct sgRNA/ssODN pairs (Supplementary Tables 1 and 2) targeting DSBs at short, intermediate and long cut-to-mutation distances (Fig. 3a and Extended Data Fig. 5a, b). Thus, a general and predictable ‘distance effect’ may govern mutation incorporation by HDR during gene editing in these human cells.

from iPSC cells electroporated with the abovementioned sgRNA/ssODN pairs (Fig. 3a), and selected those with bi-allelic incorporation of silent CRISPR/Cas-blocking mutations (Fig. 3c). The rate of homozygosity and heterozygosity for the pathogenic mutation correlated with our predictions, indicating that cut-to-mutation distance can be exploited to control zygosity using Fig. 3b to select distance.

Our data imply that cut-to-mutation distance needs to be minimized for efficient homozygous mutation incorporation and, conversely, that frequencies of mono-allelic alterations should increase at greater distances, as mutation incorporation probability drops. We determined overall probability of mutation incorporation for iPSC cells by combining APP^{Swe} and PSEN1^{M146V} oligonucleotide scan data (from Fig. 2c) and calculated expected distance ranges favouring homozygous, heterozygous and wild-type genotypes by multiplying single allele probabilities (assuming independent editing at both alleles) (Fig. 3b). To test these predictions, we derived single-cell clones

At certain loci, only guide RNAs targeting close to the intended mutation may be available, which could preclude isolation of heterozygous clones using the distance effect. As an alternative, we considered equimolar mixing of two ssODNs that both possess a blocking mutation, but only one of which contains the pathogenic mutation (Fig. 3d, Extended Data Fig. 5c; alternative approach in Extended Data Fig. 5d). We validated this approach using the closely targeting APP-sgRNA12, which previously only yielded clones homozygous for the APP^{Swe} mutation, and detected many with mono-allelic incorporation. We also verified this strategy for PSEN1^{M146V} (Extended Data Fig. 5e), suggesting it is widely applicable.

Many genetic disorders have been studied by deriving iPSC cells from patients with a disease, but this approach takes several months, and is limited by availability of patient cells and variable genetic backgrounds. These problems can be circumvented by knock-in of disease mutations in a reference ‘wild-type’ cell line, which only takes a few weeks and provides isogenic controls. Alzheimer’s disease has mostly been studied in animal models relying on non-physiological mutant gene overexpression²⁰. Human iPSC cells derived from patients with early onset Alzheimer’s disease mutations were recently established^{21–25}, but only TALEN-mediated gene editing has been used to knock-in an early onset Alzheimer’s disease mutation²⁶. Using distance to control zygosity, we generated allelic series with knock-in mutations in APP^{Swe} or PSEN1^{M146V} (Extended Data Fig. 6a, b and Extended Data Table 2). We differentiated iPSC cells into cortical neurons (Extended Data Fig. 6c–j), and examined whether APP^{Swe} and PSEN1^{M146V} mutations increase total amyloid-β (Aβ) generation or the ratio of the 42-residue

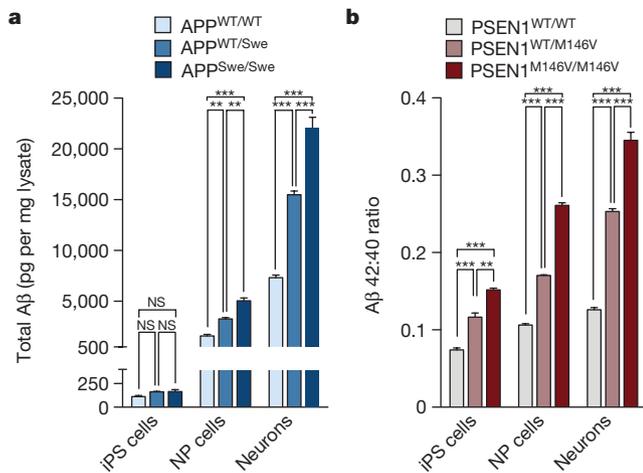


Figure 4 | APP^{Swe} and PSEN1^{M146V} knock-in lines display genotype-dependent disease-associated changes in Aβ secretion. **a**, Mutation load dependent changes in total Aβ in APP^{Swe} mutant iPS cells, neural precursor and cortical neurons. NP, neural precursor. **b**, Mutation load dependent changes in Aβ_{42:40} ratios in PSEN1^{M146V} mutant cells. Values represent mean ($n = 3$ biological replicates) \pm s.e.m. ****** $P < 0.05$ and ******* $P < 0.001$, one-way ANOVA.

versus the 40-residue Aβ peptide (Aβ_{42:40}), respectively, as predicted from patient data and model systems^{10,11}. We found more than threefold higher Aβ levels in homozygous, and twofold higher Aβ levels in heterozygous APP^{Swe} mutant cells, and up to threefold increase in secreted Aβ_{42:40} ratio in homozygous and twofold increase in heterozygous PSEN1^{M146V} mutant cells, compared to isogenic controls (Fig. 4a, b). Changes in Aβ levels and Aβ_{42:40} ratios correlated with neuronal identity and maturity (Fig. 4a, b). Thus Alzheimer's disease related phenotypes can be faithfully modelled in human neurons by introducing early onset Alzheimer's disease associated mutations, and these phenotypes correlate with mutation load.

Widespread application of CRISPR/Cas9 to induce specific genomic changes depends on strategies to improve HDR. Manipulations of cell cycle and small molecules inhibiting NHEJ have recently been reported to increase HDR rates^{27–30}, but these approaches do not directly aim to improve HDR accuracy, achieved here using CRISPR/Cas-blocking mutations. This allowed us to isolate one accurately edited line by picking just 20 to 40 clones on average (Extended Data Table 1), a rate compatible with manual picking, which might be further improved by combination with small molecule inhibitors of NHEJ. Titrating down Cas9 or guide RNA levels may also improve accuracy, but in our experiments this greatly reduced HDR rates such that manual single-clone picking became impractical (data not shown). Methods improving rate and accuracy of HDR can also be combined with CORRECT, enabling efficient scarless editing in dividing cells.

To enable control of zygosity during CRISPR/Cas9 editing, we extended previous studies^{9,13,15–19} by characterizing in two human cell types the stereotyped inverse relationship between incorporation rate of a base by HDR and its distance from CRISPR/Cas9 cleavage site. The length of gene conversion tracts we observed for CRISPR/Cas9 editing (~30–35 bp) was similar to that for TALENs in human cells^{13,18}, but differed markedly for zinc finger nucleases in *Drosophila* (over 3,000 bp)¹⁷ and restriction enzymes in rodent cells (80–200 bp)^{16,19}, potentially reflecting experimental or species differences (for example, in activities of repair pathways). Controlling zygosity by exploiting the distance effect may work best in systems with short gene conversion tracts. Our alternative approach of oligonucleotide mixing is more universally applicable.

The distance relationship did not change with altered editing conditions including HDR template types and therefore probably reflects intrinsic features of the repair mechanism. Distance dependence

may reflect the distribution of different size deletions after CRISPR/Cas9-mediated DSBs, which require only the part of the ssODN overlapping the deletion for repair¹⁶ (see model in Extended Data Fig. 7). Regardless of mechanism, the observation of a stereotyped distance effect implies that HDR is most efficiently achieved by selecting guide RNAs targeting close to the intended sequence change, and allows definition of optimal distance ranges for improved guide RNA selection to generate mono- or bi-allelic modifications.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 May 2015; accepted 11 March 2016.

Published online 27 April 2016.

- Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–827 (2013).
- Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
- Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
- Dow, L. E. *et al.* Inducible *in vivo* genome editing with CRISPR-Cas9. *Nature Biotechnol.* **33**, 390–394 (2015).
- Platt, R. J. *et al.* CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* **159**, 440–455 (2014).
- Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918 (2013).
- Canver, M. C. *et al.* Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *J. Biol. Chem.* **289**, 21312–21324 (2014).
- Yang, H. *et al.* One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**, 1370–1379 (2013).
- Inui, M. *et al.* Rapid generation of mouse models with defined point mutations by the CRISPR/Cas9 system. *Sci. Rep.* **4**, 5396 (2014).
- Haass, C. *et al.* The Swedish mutation causes early-onset Alzheimer's disease by beta-secretase cleavage within the secretory pathway. *Nature Med.* **1**, 1291–1296 (1995).
- Alzheimer's Disease Collaborative Group. The structure of the presenilin 1 (S182) gene and identification of six novel mutations in early onset AD families. *Nature Genet.* **11**, 219–222 (1995).
- Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnol.* **31**, 233–239 (2013).
- Yang, L. *et al.* Optimization of scarless human stem cell genome editing. *Nucleic Acids Res.* **41**, 9049–9061 (2013).
- Kleinstiver, B. P. *et al.* Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015).
- Bialk, P., Rivera-Torres, N., Strouse, B. & Kmiec, E. B. Regulation of gene editing activity directed by single-stranded oligonucleotides and CRISPR/Cas9 systems. *PLoS ONE* **10**, e0129308 (2015).
- Elliott, B., Richardson, C., Winderbaum, J., Nickoloff, J. A. & Jasin, M. Gene conversion tracts from double-strand break repair in mammalian cells. *Mol. Cell. Biol.* **18**, 93–101 (1998).
- Beumer, K. J., Trautman, J. K., Mukherjee, K. & Carroll, D. Donor DNA utilization during gene targeting with zinc-finger nucleases. *Genes Genomes Genetics* (G3) **3**, 657–664 (2013).
- Rivera-Torres, N., Strouse, B., Bialk, P., Niamat, R. A. & Kmiec, E. B. The position of DNA cleavage by TALENs and cell synchronization influences the frequency of gene editing directed by single-stranded oligonucleotides. *PLoS ONE* **9**, e96483 (2014).
- Taghian, D. G. & Nickoloff, J. A. Chromosomal double-strand breaks induce gene conversion at high frequency in mammalian cells. *Mol. Cell. Biol.* **17**, 6386–6393 (1997).
- Götz, J. & Ittner, L. M. Animal models of Alzheimer's disease and frontotemporal dementia. *Nature Rev. Neurosci.* **9**, 532–544 (2008).
- Yagi, T. *et al.* Modeling familial Alzheimer's disease with induced pluripotent stem cells. *Hum. Mol. Genet.* **20**, 4530–4539 (2011).
- Israel, M. A. *et al.* Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells. *Nature* **482**, 216–220 (2012).
- Kondo, T. *et al.* Modeling Alzheimer's disease with iPSCs reveals stress phenotypes associated with intracellular Aβ and differential drug responsiveness. *Cell Stem Cell* **12**, 487–496 (2013).
- Muratore, C. R. *et al.* The familial Alzheimer's disease APPV717I mutation alters APP processing and Tau expression in iPSC-derived neurons. *Hum. Mol. Genet.* **23**, 3523–3536 (2014).
- Sproul, A. A. *et al.* Characterization and molecular profiling of PSEN1 familial Alzheimer's disease iPSC-derived neural progenitors. *PLoS ONE* **9**, e84547 (2014).
- Woodruff, G. *et al.* The presenilin-1 ΔE9 mutation results in reduced γ-secretase activity, but not total loss of PS1 function, in isogenic human stem cells. *Cell Rep.* **5**, 974–985 (2013).

27. Lin, S., Staahl, B. T., Alla, R. K. & Doudna, J. A. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* **3**, e04766 (2015).
28. Yu, C. *et al.* Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* **16**, 142–147 (2015).
29. Maruyama, T. *et al.* Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nature Biotechnol.* **33**, 538–542 (2015).
30. Chu, V. T. *et al.* Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nature Biotechnol.* **33**, 543–548 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements This research was supported by The Rockefeller University, The New York Stem Cell Foundation, The Ellison Foundation, Cure Alzheimer's Fund, the Empire State Stem Cell fund through New York State Department of Health contract number C023046, and CTSA, RUCCTS grant number 8 UL1 TR000043 from the National Center for Advancing Translational Sciences (NCATS, NIH). D.P. is a New York Stem Cell Foundation Druckenmiller Fellow and received a fellowship from the German Academy of Sciences Leopoldina. D.K. is a Howard Hughes Medical Institute International Student Research Fellow and received a fellowship from the National Sciences and Engineering Research Council of Canada. S.T. is supported by the Agency for Science, Technology and Research of Singapore. A.G. is supported by a Medical Scientist Training Program grant from the National Institute of

General Medical Sciences of the National Institutes of Health under award number T32GM007739 to the Weill Cornell/Rockefeller/Sloan-Kettering Tri-institutional MD-PhD program. We thank members of the Tessier-Lavigne laboratory and L. Marraffini for discussions. Our thanks also go to S. Mazel and the team at the Rockefeller University Flow Cytometry Resource Center, J. Gonzalez and the team at the Rockefeller University Translational Technology Core Laboratory, C. Zhao and the team at the Rockefeller University Genomics Resource Center, and D. Paull and M. Duffield for technical help. Opinions expressed here are solely those of the authors and do not necessarily reflect those of the Empire State Stem Cell Fund, the New York State Department of Health, or the State of New York. The content of this study is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions D.P., D.K. and M.T.-L. conceived and designed the study. D.P. and D.K. performed and analysed the experiments. A.C. and A.G. helped perform the experiments. S.T. helped analyse next-generation sequencing data. A.S., S.J. and S.N. generated and characterized the iPS cells. K.M.O. performed and analysed the electrophysiology assays. D.P., D.K., and M.T.-L. wrote the manuscript with input from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.T.-L. (marctl@rockefeller.edu).

METHODS

sgRNA and Cas9-VRER plasmid design and construction. sgRNAs were designed using the Zhang laboratory CRISPR design tool (<http://crispr.mit.edu>). sgRNA sequences targeting *APP* or *PSEN1* (Supplementary Table 1) were cloned into plasmid MLM3636 (a gift from K. Joung, Addgene number 43860) as previously described³¹. To generate the Cas9-VRER variant¹⁴ with human codon usage, we introduced the 4 mutations into pCas9_GFP (a gift from K. Musunuru, Addgene plasmid number 44719). Briefly, we amplified fragments around the intended mutation sites by PCR with mutated primers (Supplementary Table 1), digested the plasmid with BamHI/BsrGI and fused all fragments by Gibson assembly.

Design of ssODN repair templates. The 100-nt ssODN repair templates (PAGE-purified, IDT) were designed with homologous genomic flanking sequence centred around the predicted CRISPR/Cas9 cleavage site and containing pathogenic and/or CRISPR/Cas9-blocking mutations (Supplementary Table 2). CRISPR/Cas9-blocking silent (that is, that do not alter the amino acid sequence) mutations were selected based on codon-usage of the edited gene by changing the codon to another codon already used in the same mRNA for the respective amino acid.

Generation of long ssDNA and dsDNA repair templates. To generate 200 bp and 400 bp ssDNA and dsDNA repair templates, 1,000 bp of *PSEN1* sequence around the edited locus was first PCR-amplified and TOPO-cloned. Then, a library of 20 ssODN oligonucleotides or gBlocks (IDT) containing the required mutations was integrated into the TOPO-vector by Gibson assembly (NEB), resulting in a library of 20 plasmid templates, each containing CRISPR/Cas9-blocking barcode mutations and an intended mutation at varying cut-to-mutation distances (as described in Fig. 2b). From each plasmid template, 200 bp and 400 bp dsDNA PCR amplicons were generated (primers in Supplementary Table 1), and mixed in equal amounts to generate pools of either size PCR template amplicons. Template pools were then gel extracted to remove residual plasmid. These were then re-amplified by PCR and concentrated before transfection. To generate ssDNA templates, dsDNA amplicons were generated as described above with 5' phosphorylated forward primers. Re-amplified dsDNA amplicons were then digested with lambda exonuclease (NEB) to generate ssDNA. Reactions were column purified before transfection (see Extended Data Fig. 4c).

Immunocytochemistry and microscopy. Cells were fixed in 4% paraformaldehyde, permeabilized in PBS/0.1% Triton X-100 and stained with primary and secondary antibodies (see later). Stained cells were imaged on a Nikon Eclipse Ti inverted microscope and acquired using NIS Elements imaging software (Nikon). Fiji (<http://www.fiji.sc>) and Adobe Photoshop were used to pseudo-colour images, adjust contrast and add scale bars.

Antibodies. The following antibodies were used: Oct4 (1:500, Stemgent S090023), Tra160 (1:500, Millipore MAB4360), SSEA4 (1:500, Abcam ab16287), Nanog (1:500, Cell Signaling 4903), MAP2 (1:2000, Abcam 5392), Pax6 (1:300, Covance PRB-278P), Tuj1 (mouse 1:1,000, Covance MMS-435P / rabbit 1:1,000, Covance MRB-435P), Otx2 (1:100, Millipore AB9566), Nestin (1:200, Millipore 2C13B9), FoxG1 (1:300, Abcam ab18259), CTIP2 (1:300, Abcam ab18465), Tbr1 (1:500, Millipore AB2261), SatB2 (1:100, Abcam ab51502), MAGUK (1:100, NeuroMab K28_86), Synapsin (1:200, Cell Signaling Technologies 5297), anti-mouse/rabbit/rat/chicken Alexa Fluor 488/568/647 (Invitrogen 1:500).

iPS cell lines. iPS cells were reprogrammed from human skin fibroblasts (Coriell Institute, catalog ID: AG07889) of a 18-year-old male individual using the Cytotune-iPS Sendai Reprogramming Kit (Life Technologies) according to the manufacturer's instructions, following Rockefeller University Institutional Review Board approval. Informed consent was obtained from all subjects upon sample submission to Coriell Institute. Fibroblasts were confirmed to be wild-type for all studied loci by genotyping. Multiple clones were selected based on characteristic morphology. Genetic fingerprinting confirmed iPS cells were derived from corresponding fibroblast lines. Clone 7889SA possessed a normal karyotype (Cell Line Genetics), and was characterized for typical iPS cell properties and absence of mycoplasma contamination.

Expression of pluripotency genes was analysed by NanoString nCounter gene expression system using a pre-designed codeset³². Data was normalized to the geometric mean of three housekeeping genes (ACTB, POLR2A, ALAS1) using the nSolver Analysis Software v1.0 (NanoString). 100 ng of total RNA from line 7889SA was compared to RNA extracted from the human embryonic stem cell lines HUES9 (ref. 33). Gene expression for 7 pluripotency markers and the four Yamanaka factors (Oct4, Sox2, Klf4, c-Myc) introduced as Sendai transgenes (s-t) was compared. Note that the s-tSox2 probe detects some expression of endogenous Sox2, leading to larger values for both lines.

Expression of pluripotency markers Oct4, Tra160, SSEA4 and Nanog was confirmed by immunofluorescence. *In vivo* pluripotency was confirmed by teratoma analysis as described^{25,32}. Briefly, undifferentiated iPS cells were embedded

into Matrigel and subcutaneously injected into the dorsal flank of two immune-compromised three-month-old male or female mice (NOD.Cg-Prkdc^{scid}Il2rg^{tm1Wjl}/SzJ, stock no. 005557, The Jackson Laboratory). Paraffin sections of the teratomas were subjected to haematoxylin and eosin (H&E) staining and structures characteristic for the three germ layers (ectoderm, mesoderm and endoderm) were identified by microscopy. Animal work was approved by the Columbia Institutional Animal Care and Use Committee and no randomization or blinding was used for analysis.

To generate homozygous and heterozygous APP^{Swe} iPS cell lines, cells were electroporated with the sgRNA2/ssODN and sgRNA12/ssODN combinations described in Fig. 3c and Supplementary Table 2. To study heterozygous and homozygous PSEN1^{M146V} mutations, cells were electroporated with the sgRNA5/ssODN combination described in Fig. 3c and Supplementary Table 2. Electroporated cells were isolated by FACS, followed by single-cell clone generation, RFLP and sequencing analysis as described below. One iPS cell line per genotype was isolated and characterized. The newly established gene-edited lines displayed normal karyotypes and expressed pluripotency markers Oct4, Tra160, SSEA4, Nanog and alkaline phosphatase (data not shown).

Cell culture and transfection. iPS cells were maintained on irradiated MEFs (Globalstem) plated on cell culture plates coated with 0.1% gelatin and grown in HUESM (Knockout Dulbecco's modified Eagle's Medium (KO-DMEM), 20% knockout serum, 0.1 mM non-essential amino acids, 2 mM Glutamax, 100 U per ml penicillin, 0.1 mg per ml streptomycin (all Life Technologies), 0.1 mM 2-mercaptoethanol (Sigma-Aldrich), 10 ng ml⁻¹ FGF2 (Stemgent), at 37 °C with 5% CO₂). Prior to transfection, iPS cells were transferred to Geltrex-coated (Life Technologies) cell culture plates and grown in MEF-conditioned HUESM containing 10 μM ROCK inhibitor (Stemgent).

iPS cells were transfected with Cas9- and sgRNA-expressing plasmids, and ssODNs by electroporation. Two million cells were resuspended in 100 μl cold BTXpress electroporation buffer (Harvard Apparatus) with 20 μg pCas9_GFP, 5 μg sgRNA plasmid, and 30 μg ssODN (100 bp ssODN, PAGE-purified, IDT). Cells were electroporated at 65 mV for 20 ms in a 1 mm cuvette (Harvard Apparatus). After electroporation, cells were transferred to Geltrex-coated cell culture plates and grown in MEF-conditioned HUESM containing ROCK inhibitor for 2 days. In all transfections, 7889SA-derived iPS cells wild-type at genome-edited loci were used.

HEK293T cells (Life Technologies) were maintained in DMEM with 10% FBS, 2 mM Glutamax and 100 U per ml penicillin and 0.1 mg per ml streptomycin (all Life Technologies) at 37 °C with 5% CO₂. HEK293 cells were seeded on 12-well plates at 250,000 cells per ml. When approximately 70% confluent, HEK293 cells were transfected with 800 ng Cas9 plasmid, 400 ng sgRNA plasmid and 1 μg ssODN Cells using X-tremeGENE 9 (Roche).

Fluorescence-activated cell sorting. All GFP-positive cells, regardless of expression levels, were collected in the Rockefeller University Flow Cytometry Resource Center using a FACSAria II flow cytometer (BD Biosciences). Then 48 h following transfection, cells were resuspended in PBS with 0.5% BSA fraction V solution, 10 mM HEPES, 100 U per ml penicillin, 0.1 mg per ml streptomycin (all from Life Technologies), 0.5 M EDTA, 20 mM glucose, 10 ng per ml DAPI in the presence of ROCK inhibitor for iPS cell sorts. For pooled cell next-generation sequencing analysis, 150,000 to 250,000 cells were collected and immediately frozen in liquid N₂ for further study. For single-cell derived iPS cell clonal analysis 30,000 GFP⁺ cells were immediately plated on a 10 cm plate of MEFs in HUESM and ROCK inhibitor following cell sorting.

Next-generation sequencing analysis of HDR-mediated mutation incorporation. Genomic DNA was extracted from sorted cells and the genomic region around the CRISPR/Cas9 target site for *APP* and *PSEN1* genes was amplified by PCR with primers positioned outside of the HDR repair template sequence to avoid template amplification for 25 cycles using Q5 polymerase (NEB) according to the manufacturer's protocol (PCR primers listed in Supplementary Table 1). Primers contained sample-specific barcodes. 25 cycles were previously determined to be optimal for exponential amplification of the template as well as visibility for gel extraction (data not shown). To eliminate PCR byproducts and genomic DNA, PCR products were gel purified. 25–100 ng of pooled barcoded PCR products were submitted to the Rockefeller University Genomics Resource Center for targeted MiSeq (Illumina) 300 bp paired-end next-generation sequencing with library preparation using the v3 reagent kit (Illumina).

Data analysis was performed using Galaxy^{34,35} (<http://usegalaxy.org>) or Unix-based software tools listed below (summarized in Extended Data Fig. 8). First, quality of paired-end sequencing reads (R1 and R2 fastq files) was assessed using FastQC³⁶. Raw paired-end reads were combined using paired end read merger (PEAR)³⁷ to generate single merged high-quality full-length reads. Reads with sample-specific forward and reverse barcodes were de-multiplexed using the

FASTX-Toolkit³⁸ barcode splitter. The barcodes were then trimmed using seqtk (<https://github.com/lh3/seqtk>). Reads were then filtered by quality (using Filter FASTQ³⁹) removing reads with a mean PHRED quality score under 30 and a minimum per base score under 24. Only reads shorter than or equal to the length of the PCR amplicons plus 40 bp (to account for insertions) were considered for analysis.

For the accurate HDR and indel analysis in Fig. 1, reads were filtered to assess the presence of HDR or NHEJ-induced indels. To isolate sequences with HDR, reads were first filtered to remove unedited wild-type reads. Next, HDR reads containing *APP* or *PSEN1* mutations were isolated by matching a 6-nt HDR motif around the pathogenic mutation. HDR reads were then analysed for incorporation of CRISPR/Cas-blocking mutations by matching 6-nt to 8-nt HDR motifs around each mutation and categorized into unique groups of reads containing all possible combinations (32) of CRISPR/Cas-blocking mutations to account for measurable HDR after re-editing (Extended Data Fig. 8b). Each group of reads was then aligned to a corresponding reference sequence using *bwa mem*⁴⁰ (which has been successfully used for this purpose by others^{41,42}) with option -M to determine the rate of accurate HDR and indel or substitution mutations (Extended Data Fig. 8c). Reads with multiple blocking mutations were analysed separately. Accurate HDR reads were calculated in each group as the percentage of HDR reads without indels. To determine indel frequency, size and distribution, all edited reads from each experimental replicate were combined and aligned, as described above. Indels were then marked at each base using *bam-readcount* (<https://github.com/genome/bam-readcount>), quantified in R⁴³ and plotted using GraphPad Prism.

In all other experiments (all figures except Fig. 1a–e), reads were first filtered for experiment-specific barcode and quality as described earlier (Extended Data Fig. 8a). Next, reads were considered to have HDR if they matched the repair ssODN template plus an additional 3-nt genomic sequence on each side to ensure proper genomic context during HDR and contained the pathogenic mutation and/or CRISPR/Cas-blocking silent mutation (Extended Data Fig. 8d). For all next-generation sequencing experiments, HDR rates were calculated and listed in Extended Data Table 1. *n* values represent independent biological replicates.

To exclude a significant contribution of oligonucleotide synthesis and sequencing errors to our analysis, we sequenced *PSEN1* PCR amplicons from *APP*-edited iPSC cells, and *APP/PSEN1* repair ssODNs annealed to a complementary ssODN. Errors introduced by sequencing were 2.7% ± 0.1% per 100 bp, and 2.3% ± 1.7% of the 100 bp ssODN sequences contained errors.

Calculation of optimal distance ranges for homozygous or heterozygous genotypes. Mutation scan data for *APP* and *PSEN1* loci determined by next-generation sequencing for iPSC cells from Fig. 2c were combined to determine single allelic mutation incorporation probabilities p_a as a function of cut-to-mutation distance (p_a^{mut}). The probability of wild-type incorporation (p_a^{wt}) was determined as ($p_a^{wt} = 1 - p_a^{mut}$). Assuming gene editing and HDR at each allele in a single cell are independent events, we calculated the zygosity probabilities (p_z) for each allele combination given two alleles per cell. Specifically, probability of a homozygous, wild-type, and heterozygous zygosity was calculated as $p_z^{mut/mut} = p_a^{mut} \times p_a^{mut}$, $p_z^{wt/wt} = p_a^{wt} \times p_a^{wt}$ and $p_z^{wt/mut} = 2 \times (p_a^{wt} \times p_a^{mut})$, respectively. These calculations were made using the entire range of data derived from Fig. 2c, extrapolated for distance values above 36 and plotted in Fig. 3b as fit curve ± s.d. of raw values.

RFLP analysis and Sanger sequencing for genotyping of single-cell clones. To facilitate single-cell clone genotyping, the ssODN HDR templates used for gene editing were designed to introduce a restriction endonuclease motif with the blocking or pathogenic mutation. Genome edited single-cell-derived iPSC cell clones grown on MEF-containing 10-cm plates (in HUESM + ROCK inhibitor) were manually picked into a single well of a U-bottom 96-well tissue culture plate in 100 µl HUESM + ROCK inhibitor. Cells were pelleted by centrifugation, and plates were immediately frozen in liquid N₂ and stored at -80 °C. Genomic DNA was extracted as previously described⁴⁴. Briefly, cells were resuspended in 25 µl lysis buffer (0.75 µl 10 mg ml⁻¹ proteinase K (Ambion), 2.5 µl 10× PCR buffer (Sigma-Aldrich), transferred to 96-well PCR plates and incubated at 55 °C for 4 h. Proteinase K was inactivated by incubating plates at 96 °C for 10 min.

To identify clones with HDR events, the genomic region surrounding the APP^{Swe} or PSEN1^{M146V} loci were amplified by Taq polymerase (Roche) and digested with restriction enzymes to screen for a novel restriction site introduced by the blocking or pathogenic mutation (primers, repair ssODNs and restriction enzymes used are listed in Supplementary Table 1 and 2). Digested DNA was analysed by agarose gel electrophoresis. The zygosity of the pathogenic mutation in clones that had undergone incorporation of the silent CRISPR/Cas-blocking mutations was determined by Sanger sequencing (Genewiz). Bi-allelic HDR rates for single-cell clones were calculated and listed in Extended Data Table 1.

To determine the frequency and distribution of indels in mono-allelic HDR single-cell clones with NHEJ at the other allele, Sanger sequencing reads were

separated into single reads for HDR and indel-containing alleles using PolyPeak Parser⁴⁵. Indel-containing reads were then combined into a single FASTA file and analysed for indel distribution by aligning to the reference sequence as described earlier.

CORRECT. Re-guide and re-Cas use a two-step gene editing workflow: two million iPSC cells were electroporated with sgRNA and Cas9 plasmids. In addition, during the first step, a ssODN containing the intended mutation (M) and a CRISPR/Cas blocking mutation (B) was introduced (MB template). Cas9-eGFP expressing cells were FACS sorted and single-cell iPSC cell clones were derived. The presence of B and M mutations was detected by RFLP. A single clone containing homozygous B and M mutations was then expanded for use in the second step of CORRECT. These 'MB iPSC cells' were then electroporated with re-sgRNA and wild-type Cas9 plasmids (for re-guide) or wild-type sgRNA and mutant VRER Cas9 plasmids (for re-Cas). In addition, at this step the CORRECT template was provided to remove blocking mutation B. The efficacy of CRISPR/Cas blocking mutation removal was determined by next-generation sequencing. Alternatively, after the second CORRECT step, cells can be plated to derive single-cell scarless 'M iPSC cell' clones.

Off-target analysis. Gene edited homozygous and heterozygous APP^{Swe} and PSEN1^{M146V} iPSC cell lines were tested for off-target editing events predicted for each sgRNA by the Zhang laboratory CRISPR design tool (<http://crispr.mit.edu>) and the COSMID⁴⁶ tool (<http://crispr.bme.gatech.edu>), which also considers insertions or deletions in the guide RNA target sequence. The top five non-overlapping predicted off-target sites for each sgRNA from each tool were used. The region surrounding each off-target site was PCR-amplified, Sanger sequenced (Genewiz) and compared to the unedited cell line.

Cortical neuron differentiation. iPSC-cell-derived cortical neurons were generated as previously described⁴⁷ with modifications. Specifically, to generate neural precursor cells (NP cells), iPSC cells were plated on 12-well tissue culture plates coated with Geltrex (Life Technologies) in MEF-conditioned HUESM with ROCK inhibitor. When cells were 100% confluent, medium was replaced with neural induction (NI) medium (day *in vitro* 0 (DIV0)) and maintained for 8 days. On DIV8 cells were dissociated using Accutase (Life Technologies) and resuspended in NI medium with ROCK inhibitor at 30 million cells per ml. Cells were plated on dried poly-L-ornithine (Sigma-Aldrich) and laminin-coated (Life Technologies) 6-well plates in 10-µl spots. Cells were left to adhere for ~45 min and NI medium with ROCK inhibitor was added. On DIV10 NI was replaced with neural maintenance (NM) medium. Upon the appearance of neural rosettes, 20 ng ml⁻¹ FGF2 was added for 2 days. When neurons started to emerge from rosettes, those were isolated manually after treatment with STEMdiff Neural Rosette Selection Reagent (STEMCELL Technologies) for 1 h. Rosettes were washed and plated on poly-L-ornithine/laminin-coated 6-well plates. Between DIV30 and DIV 36 NPCs were frozen in NM supplemented with 10% DMSO and 20 ng ml⁻¹ FGF2.

For cortical neuron maturation, ~200,000–500,000 NPCs were plated on 24-well poly-L-ornithine/laminin-coated plates and maintained in Neurobasal medium supplemented with B-27 serum-free supplement, 2 mM Glutamax and 100 U per ml penicillin and 0.1 mg per ml streptomycin (all Life Technologies). During the first 7 days after plating, cells were treated with 10 µM DAPT (Sigma-Aldrich) to augment neuronal maturation.

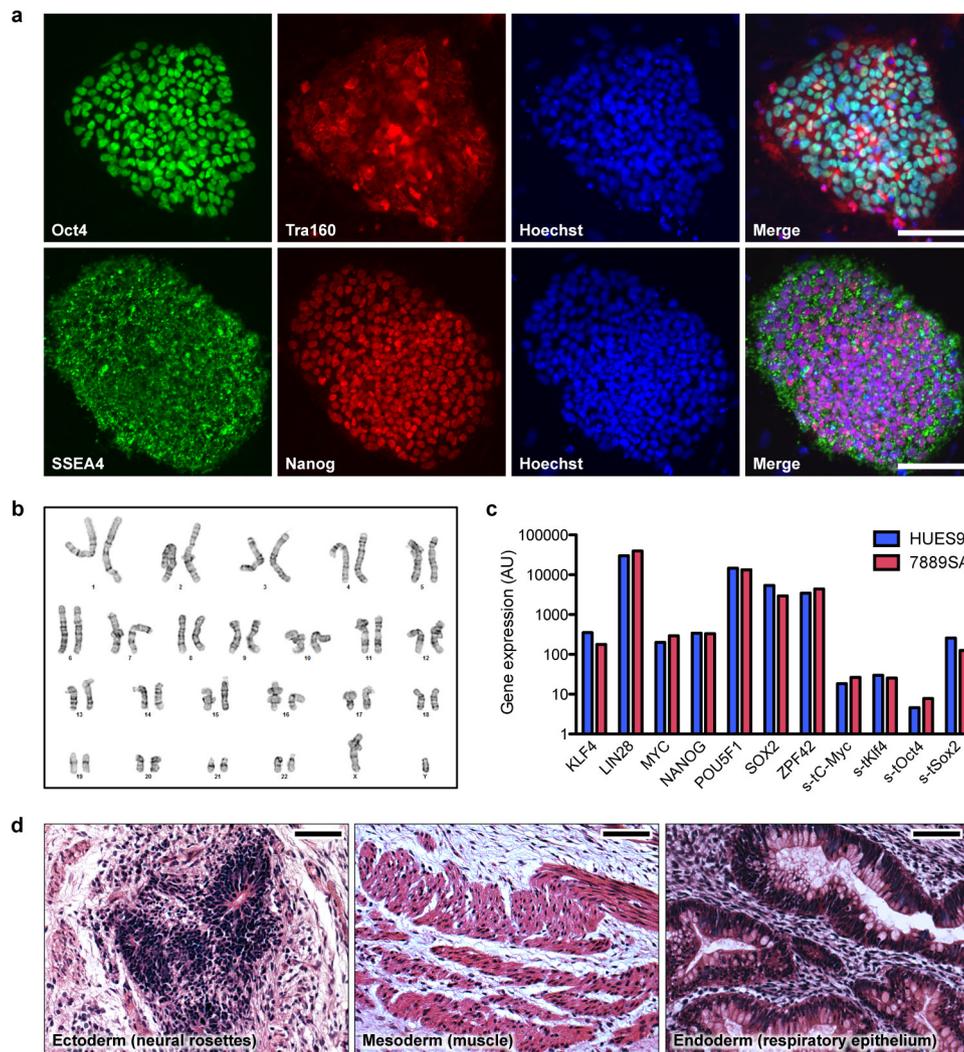
Cortical neuron characterization. Canonical neural precursor cell markers (Nestin, Pax6, FoxG1, Otx2) and mature cortical neuronal markers (Tbr1, CTIP2, Satb2) were analysed by immunofluorescence staining at DIV10 and DIV65, respectively. Electrophysiological properties of iPSC-cell-derived cortical neurons were assessed between DIV71 and 85 using a submerged recording chamber mounted on an Olympus BX51 microscope equipped for infrared-DIC microscopy. Neurons were perfused with 95% O₂/5% CO₂ equilibrated ACSF (in mM): 119 NaCl, 2.5 KCl, 1.3 MgSO₄, 2.5 CaCl₂, 1 NaH₂PO₄, 26 NaHCO₃ and 11 glucose. Whole-cell patch clamp pipettes (5 MΩ) were filled with (in mM): 123 K-gluconate, 10 HEPES, 0.2 EGTA, 8 NaCl, 2 Na₂ATP, 0.3 Na₃GTP. Action potentials were elicited by step current injections and recorded in current-clamp mode (-65 mV). Properties (threshold, overshoot) of the largest action potential elicited in each cell were measured. Spontaneous synaptic activity was recorded in voltage-clamp mode (-70 mV). Data was digitized at 10 kHz and recorded using a Multiclamp 700B amplifier and Clampex 10.3.0.2 software (Molecular Devices).

Amyloid-β measurements. Aβ was measured in cell supernatant conditioned for 2 days (iPSC cells), 3 days (DIV34 neural precursors), or 4 days (DIV72 cortical neurons). Experiments were performed in 3 biological replicates. Supernatants from experiments collected at different time points were frozen at -80 °C. Secreted Aβ₁₋₃₈, Aβ₁₋₄₀ and Aβ₁₋₄₂ were measured with MSD Human (6E10) Aβ V-PLEX kits (Meso Scale Discovery) according to the manufacturer's directions. iPSC cell and neuronal total Aβ levels were normalized to total protein levels from cell lysate determined by BCA assay (Pierce).

Surveyor assays. Genomic DNA was extracted from gene-edited iPS cells as described above. 300–500 bp around the gene-edited locus were amplified by PCR using Herculase II (Agilent) and column purified. PCR amplicons were rehybridized and treated with Surveyor nuclease according to the manufacturer's directions (IDT). Digested DNA was separated on a 4–20% TBE polyacrylamide gel (BioRad) and imaged using SYBR Gold (Life Technologies). Densitometry was performed using Fiji. Per cent indel quantification was based on relative band intensities using the formula $100 \times (1 - (1 - (b + c)/(a + b + c))^{1/2})$, where a is the undigested PCR product intensity and b and c are the intensities of each cleavage product⁴⁸.

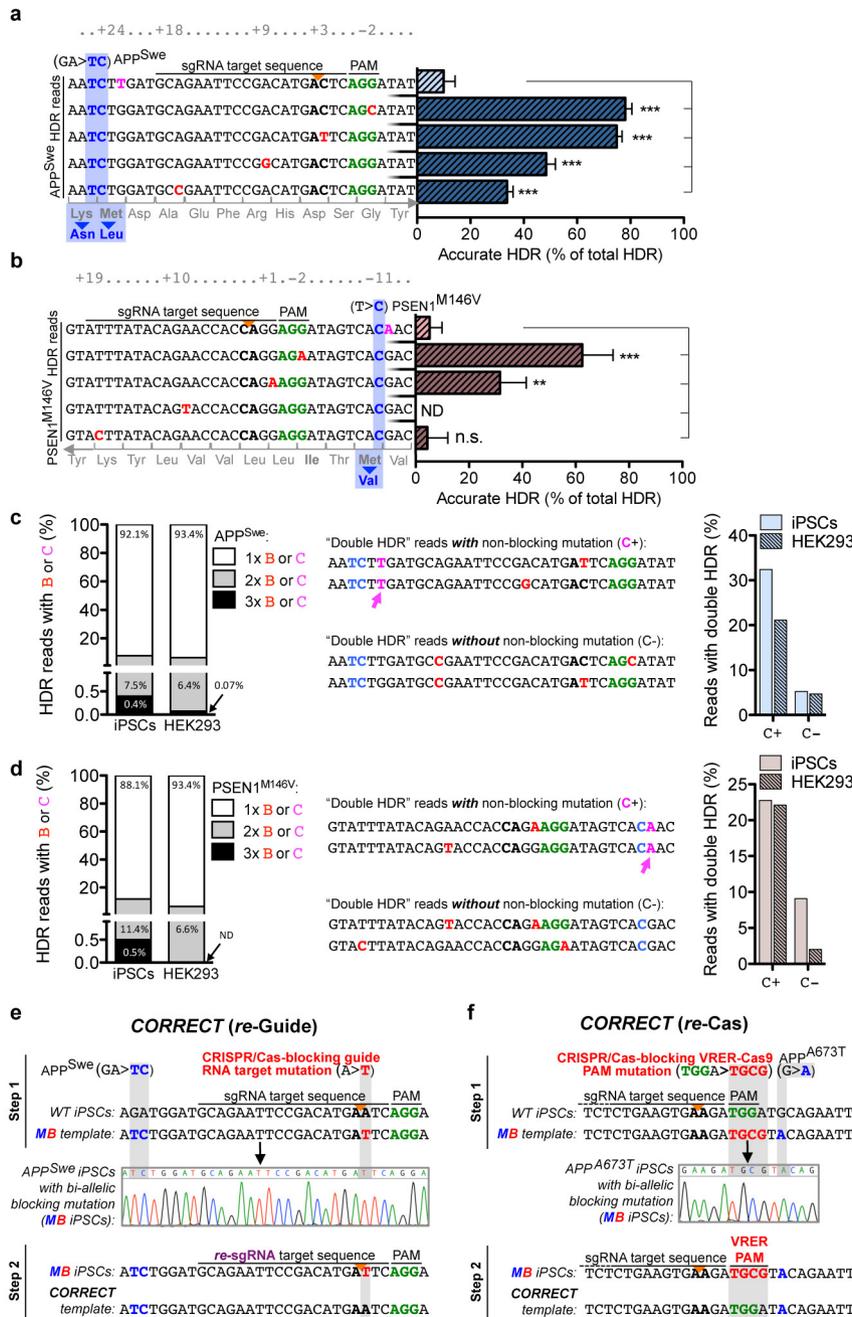
Statistical analysis. No statistical methods were used to predetermine sample size and the experiments were not randomized. Experimental data was analysed for significance using GraphPad Prism 6. $P < 0.05$ was considered statistically significant. All experiments except the oligonucleotide scan were analysed by one-way ANOVA followed by post-testing with either Tukey's test, if multiple values were compared to each other, or Dunnett's method, if alterations were compared to controls. Similarity of variance was confirmed with Bartlett's test where appropriate. For the oligonucleotide scan, nonlinear regression analysis was performed to fit exponential decay equation model curves to experimental values; R square values were determined to test goodness of fit. To analyse if distance-incorporation relationships were significantly different for genomic loci, the rate constant k was determined for each individual data set and the k values of the two loci were compared using the unpaired t -test. The analysis approaches have been justified as appropriate by previous biological studies, and all data met the criteria of the tests. The investigators were not blinded to allocation during experiments and outcome assessment.

31. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnol.* **31**, 822–826 (2013).
32. Kahler, D. J. *et al.* Improved methods for reprogramming human dermal fibroblasts using fluorescence activated cell sorting. *PLoS ONE* **8**, e59867 (2013).
33. Cowan, C. A. *et al.* Derivation of embryonic stem-cell lines from human blastocysts. *N. Engl. J. Med.* **350**, 1353–1356 (2004).
34. Goecks, J., Nekrutenko, A. & Taylor, J. & The Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* **11**, R86 (2010).
35. Blankenberg, D. *et al.* Galaxy: a web-based genome analysis tool for experimentalists. *Curr. Protoc. Mol. Biol.* **Chapter 19**, Unit 19.10.1–21 (2010).
36. Andrews, S. FastQC: a quality control tool for high throughput sequence data. (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>).
37. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina paired-end read merger. *Bioinformatics* **30**, 614–620 (2014).
38. Pearson, W. R., Wood, T., Zhang, Z. & Miller, W. Comparison of DNA sequences with protein sequences. *Genomics* **46**, 24–36 (1997).
39. Blankenberg, D. *et al.* Manipulation of FASTQ data with Galaxy. *Bioinformatics* **26**, 1783–1785 (2010).
40. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
41. Tsai, S. Q. *et al.* Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nature Biotechnol.* **32**, 569–576 (2014).
42. Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M. D. C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nature Biotechnol.* **32**, 267–273 (2014).
43. R Core Team. R: A language and environment for statistical computing. (R Foundation for Statistical Computing, 2015).
44. Zhu, Z., González, F. & Huangfu, D. The iCRISPR platform for rapid genome editing in human pluripotent stem cells. *Methods Enzymol.* **546**, 215–250 (2014).
45. Hill, J. T. *et al.* Poly peak parser: method and software for identification of unknown indels using Sanger sequencing of polymerase chain reaction products. *Dev. Dyn.* **243**, 1632–1636 (2014).
46. Cradick, T. J., Qiu, P., Lee, C. M., Fine, E. J. & Bao, G. COSMID: a web-based tool for identifying and validating CRISPR/Cas off-target sites. *Mol. Ther. Nucleic Acids* **3**, e214 (2014).
47. Shi, Y., Kirwan, P. & Livesey, F. J. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nature Protocols* **7**, 1836–1846 (2012).
48. Ran, F. A. *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380–1389 (2013).
49. Jonsson, T. *et al.* A mutation in *APP* protects against Alzheimer's disease and age-related cognitive decline. *Nature* **488**, 96–99 (2012).



Extended Data Figure 1 | *In vitro* and *in vivo* characterization of the wild-type 7889SA human iPS cell line. a, Immunofluorescence staining of pluripotent stem cell markers. **b, iPS cells possess a normal human male karyotype. **c,** Nanostring expression analysis of pluripotent stem cell genes**

in reprogrammed iPS cells compared to HUES9. **d,** *In vivo* differentiation and analysis of iPS-cell-derived teratoma containing tissues of all germ cell layers. Scale bars, 100 μ m.

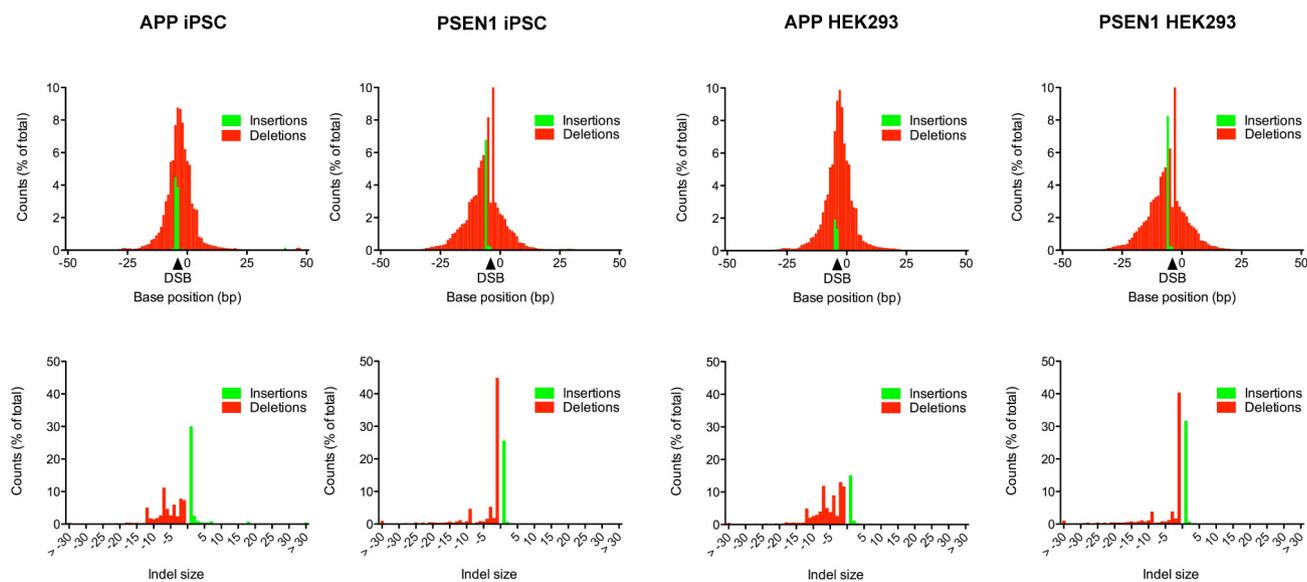


Extended Data Figure 2 | CRISPR/Cas-blocking mutations increase HDR accuracy by preventing re-editing, are incorporated in multiple rounds of re-editing and can also be applied to scarless editing using CORRECT. a, b, HDR reads from five unpooled templates containing intended pathogenic and CRISPR/Cas-blocking or non-blocking control mutations. Percentages of accurate HDR for reads containing blocking (B) or control (C) mutations at the *APP* (a) and *PSEN1* (b) locus in HEK293 cells. Values represent mean \pm s.e.m. ($n = 3$). ND, not detected. *** $P < 0.001$, ** $P < 0.01$, one-way ANOVA. **c, d**, Proportion of next-generation sequencing reads containing putative single, double, or triple HDR events (left) for *APP* (c) and *PSEN1* (d). Putative ‘double HDR’ examples of the most frequent reads that either contain a non-blocking control mutation C with an additional CRISPR/Cas-blocking mutation B, or do not contain C and have two different CRISPR/Cas-blocking mutations (middle). Reads that contain the non-blocking mutation (C+) are more frequently re-edited to incorporate a CRISPR/Cas-blocking mutation (‘double HDR’) than reads containing a blocking mutation B instead of the non-blocking mutation C (C-). See Fig. 1c for legend. To facilitate data analysis, all replicates were pooled to increase read numbers for rare events. **e, f**, Schematics depicting details of the two tested CORRECT approaches: in step 1 of re-guide (e), the APP^{Swe} mutation

was introduced together with a CRISPR/Cas-blocking guide RNA target mutation, which was then removed again in step 2 using a re-sgRNA specific for the mutated sequence and wild-type Cas9. In step 1 of re-Cas (f), the APP^{A673T} mutation was introduced together with a CRISPR/Cas-blocking PAM-altering NGCG mutation, which was then removed in step 2 using the VRER Cas9 variant, which specifically detects the NGCG PAM. We chose to use the very active APP-sgRNA12 to test CORRECT by re-Cas, which was also used in Fig. 3c and 3d to generate APP^{Swe} mutant lines. However, as the APP^{Swe} mutation is located in the target sequence of this sgRNA, it may block re-editing by CRISPR/Cas and could therefore complicate the interpretation of results. We therefore decided to knock-in the protective APP^{A673T} mutation⁴⁹ instead, which lies outside of the target sequence. In both cases, the blocking mutations, which restored the original sequence at the site of the blocking mutation (which blocks further re-cutting in this step), but retained the intended APP mutation. Note that due to repeated editing, CORRECT may increase the probability of off-target effects, but presumably not the number of potential off-target sites, as the same (for re-Cas) or a very similar (for re-guide) guide RNAs are used in both editing steps.

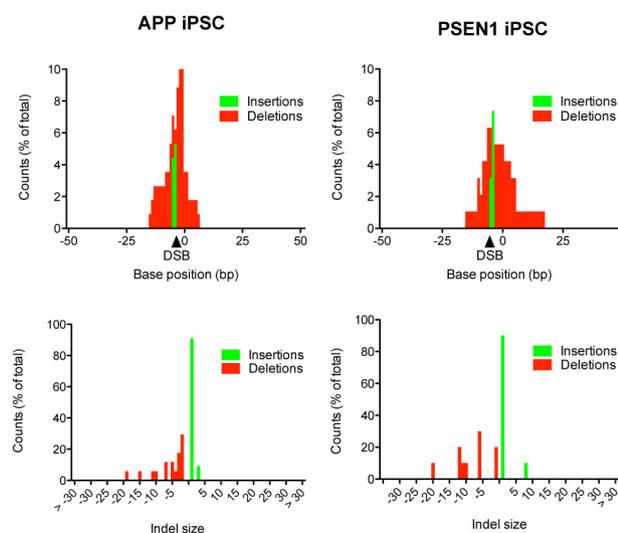
a

NGS



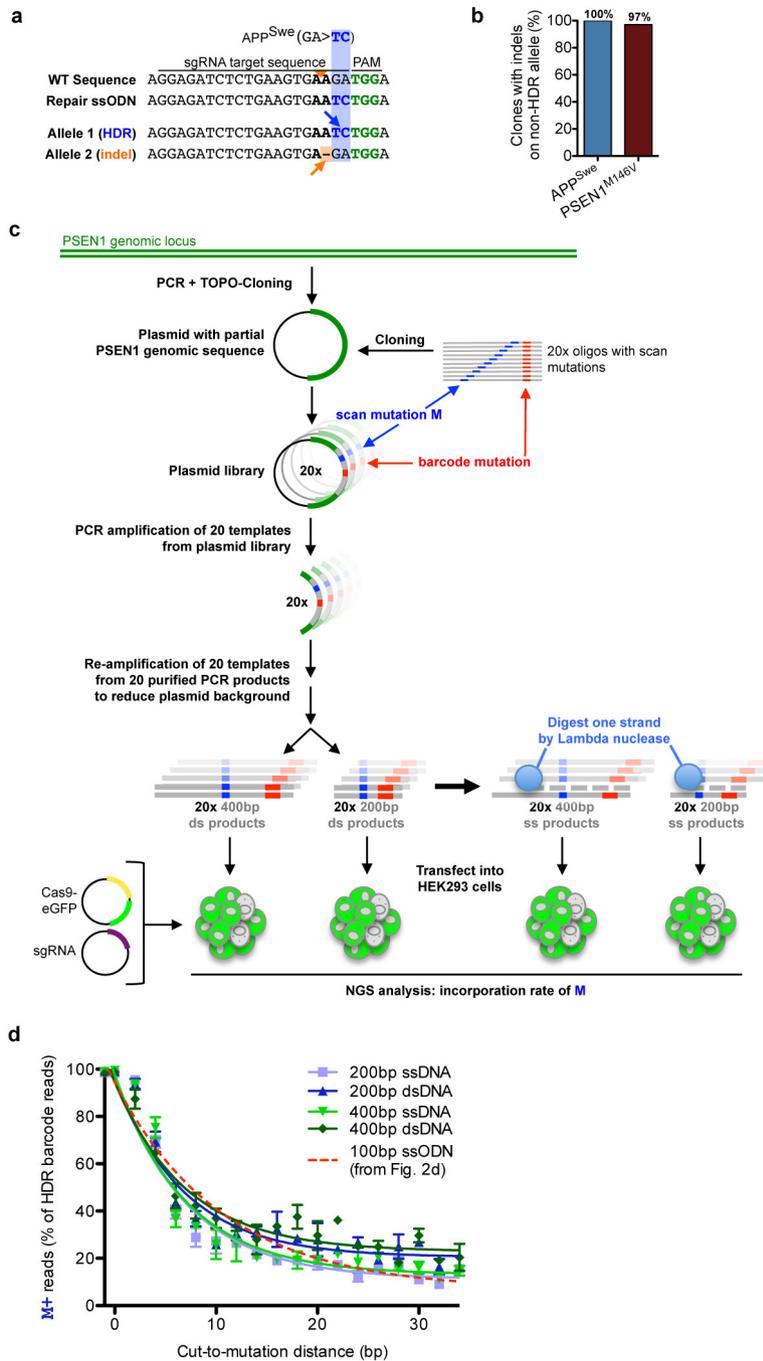
b

Single Cell Clones



Extended Data Figure 3 | Analysis of CRISPR/Cas9-induced indels in gene edited iPSC cells and HEK293 cells. a, Plot depicting frequency of indels at each position around the targeted locus in all next-generation sequencing reads with editing events from the analysis shown in Fig. 1. Insertions are plotted at the location where they begin, and deletions

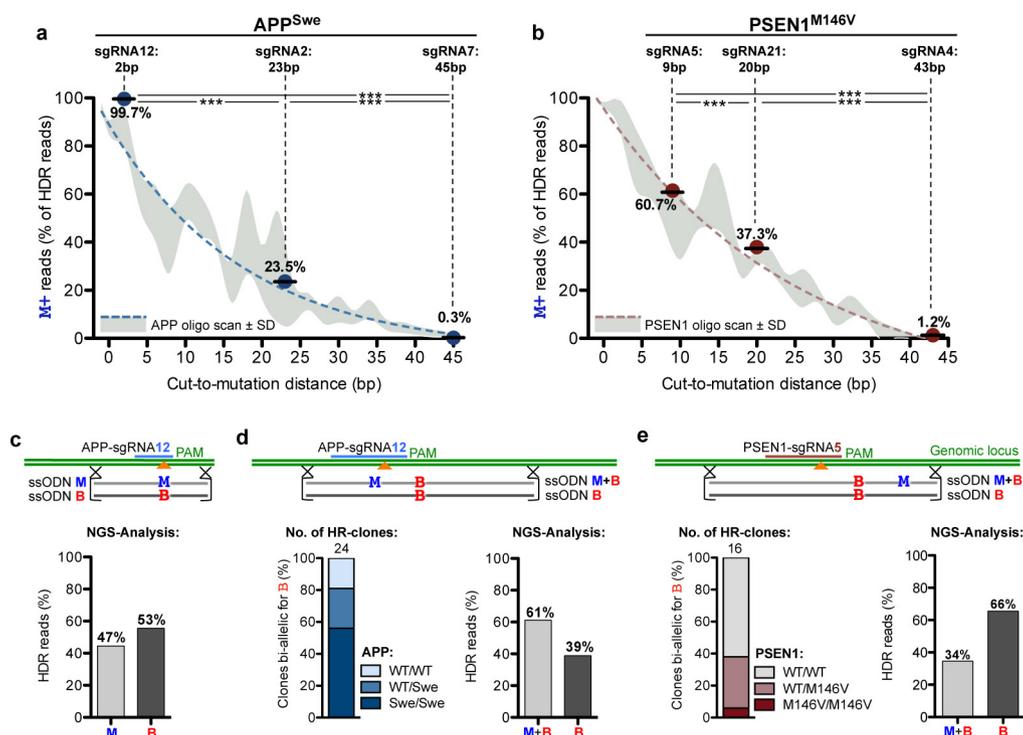
are plotted across all deleted base positions (top). Histogram illustrating distribution of indel sizes (bottom). **b,** Indel position (top) and size (bottom) of indel-containing alleles from single-cell clones analysed in Extended Data Fig. 4a, b.



Extended Data Figure 4 | Heterozygous clones with HDR on one allele almost always contain indels on the non-HDR allele, and longer ssDNA or dsDNA HDR repair templates do not influence mutation incorporation probabilities related to cut-to-mutation distance.

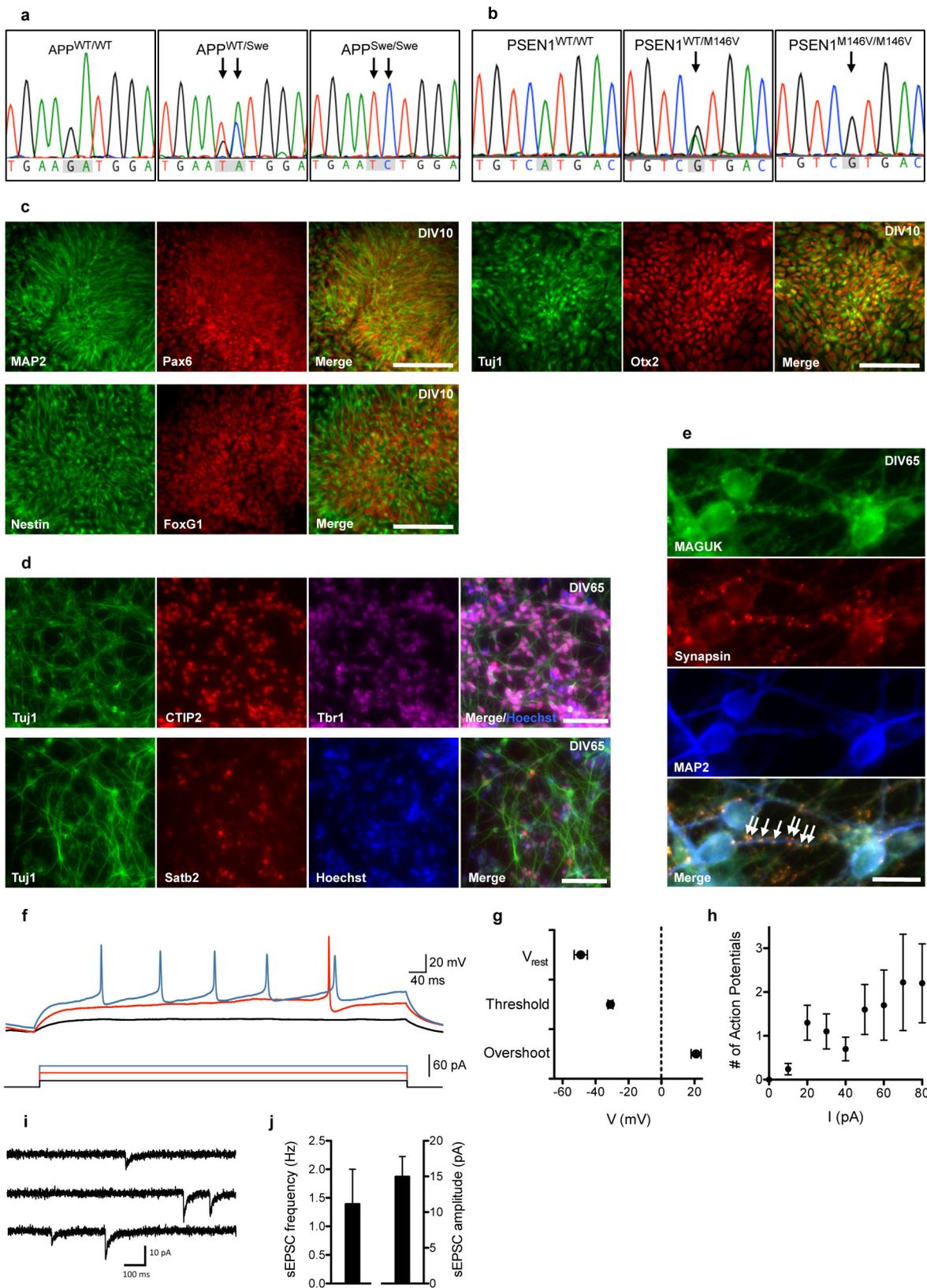
a, Sanger sequencing reads of both *APP* alleles of a single-cell clone with mono-allelic HDR (blue arrow). The non-HDR allele is altered by NHEJ in the guide RNA target sequence (orange arrow). **b**, Single-cell clones with HDR on one allele are mostly altered by NHEJ on the non-HDR

allele (*APP*, $n = 26$; *PSEN1*, $n = 34$). **c**, Schematic describing the generation of large ssDNA and dsDNA HDR repair templates for the *PSEN1* locus (see Methods for details). **d**, The monotonic relationship between incorporation of intended mutations (*M*) by HDR and cut-to-mutation distance is not altered by providing longer ssDNA and dsDNA templates ($n = 2$). Red dashed trend line shows previously determined 100-nt oligonucleotide scan result (from Fig. 2d) for comparison.



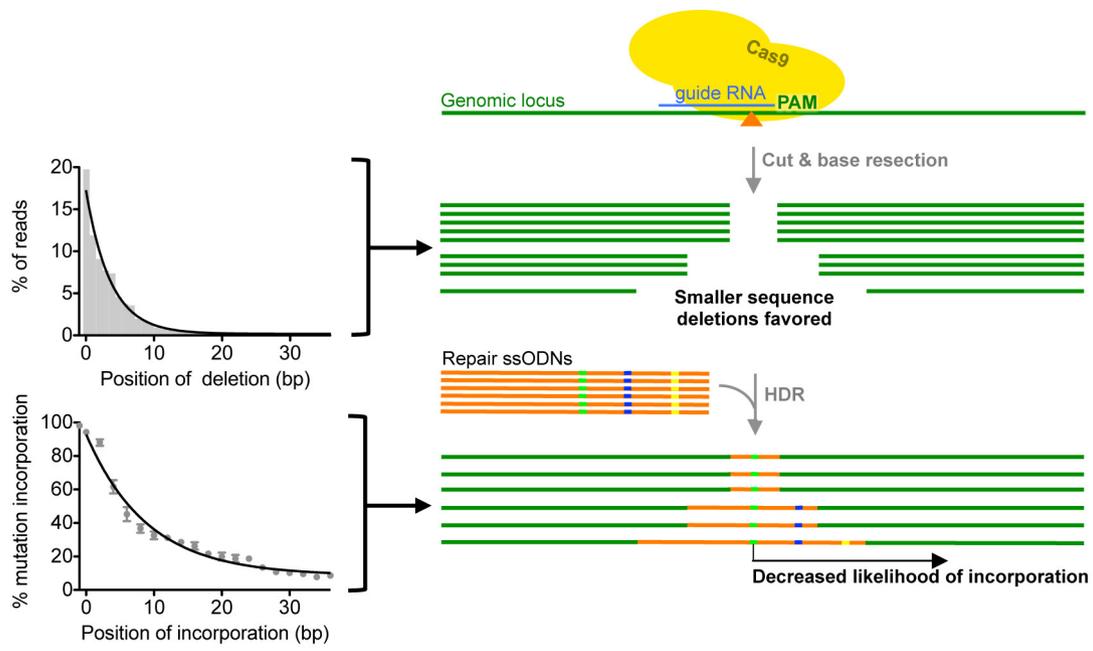
Extended Data Figure 5 | Mutation incorporation rates at various cut-to-mutation distances follow the distance effect, and mixed repair templates as a strategy to generate heterozygous iPS cell single-cell clones. **a, b**, Incorporation rate of *APP* and *PSEN1* pathogenic mutations at increasing distance from the cut site targeted by three distinct sgRNA/ssODN pairs is governed by distance. Incorporation rates (solid dots represent mean \pm s.e.m., note s.e.m. is too small to be visible, ($n = 3$)) match almost exactly the curves for each locus previously determined by oligonucleotide scan (dashed trend line \pm s.d. of raw data from Fig. 2c, d). *** $P < 0.001$, one-way ANOVA. **c, d**, Mixed ssODN editing approach at the *APP* locus with blocking mutations in one (**c**) or both (**d**) ssODNs (top); zygosity quantification of single-cell clones (**d**, bottom left) and incorporation rates of CRISPR/Cas-blocking mutation B and pathogenic mutation M determined by next-generation sequencing analysis (**d**, bottom right). Note that for the M/B approach in **c**, both oligonucleotides are incorporated at equal levels, as they have similar blocking activities, whereas for the M+B/B approach in **d**, the M+B ssODN

is preferentially incorporated, presumably due to a synergistic blocking effect of both M and B. For the clone quantification in Fig. 3d, the rate of wild-type clones was not assessed, because the silent mutation did not introduce a restriction site. However, given the $\sim 50\%$ ssODN incorporation rates determined by deep sequencing, about 25% of HDR clones are predicted to be wild type. **e**, Mixed ssODN editing approach at the *PSEN1^{M146V}* locus (top). Using an sgRNA with the smallest possible cut-to-mutation distance (*PSEN1*-sgRNA5), two ssODNs were provided, each containing the same silent PAM-altering CRISPR/Cas-blocking mutation B, but only one containing the pathogenic mutation M. Frequencies of pathogenic mutation genotypes in single-cell clones with bi-allelic HDR of B (bottom left) and incorporation rates of CRISPR/Cas-blocking and pathogenic mutations by next-generation sequencing (bottom right). Note that due to the 9 bp distance to the cleavage site, the incorporation of M is lower than 50% (as expected from the distance effect).



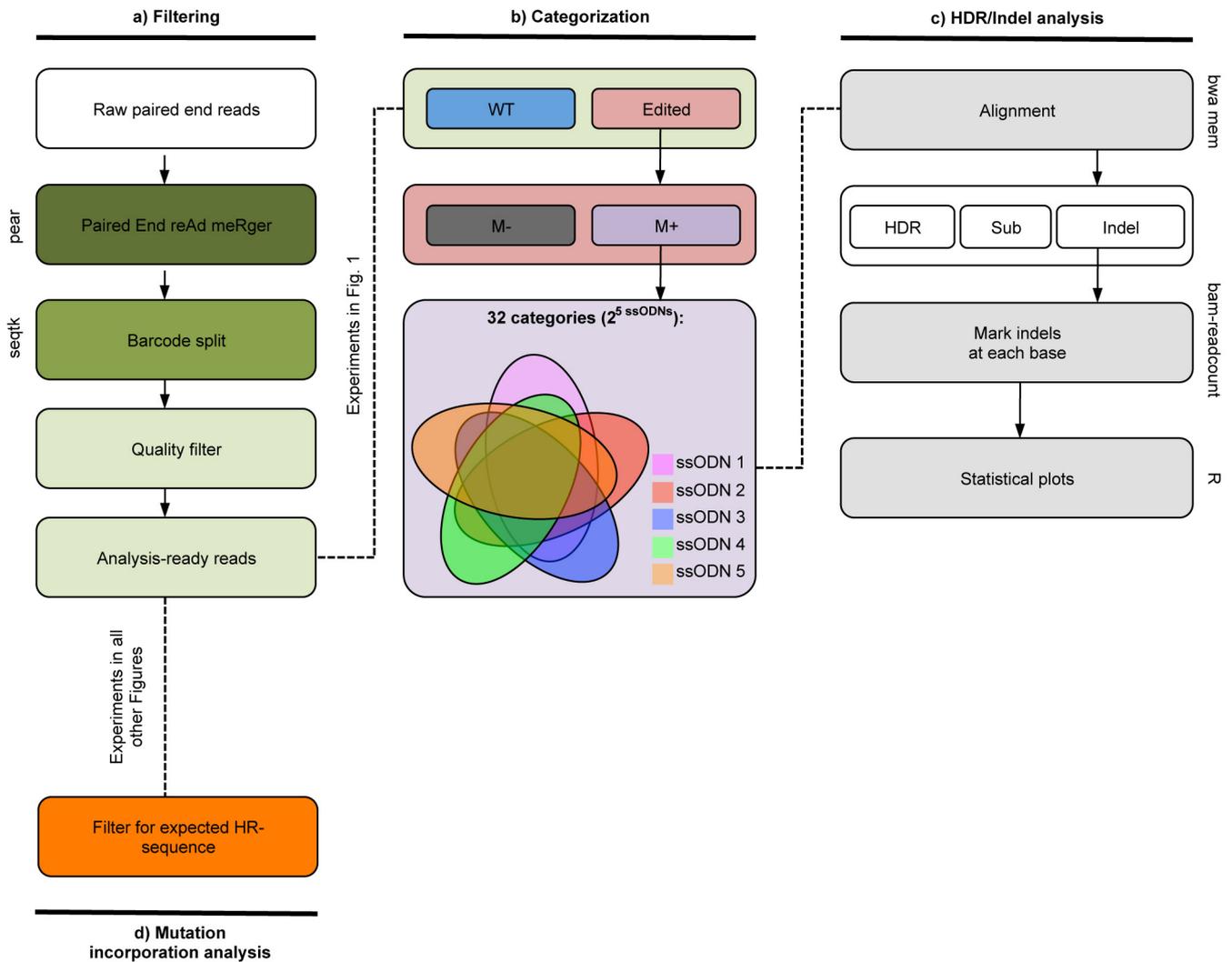
Extended Data Figure 6 | Characterization of iPS-cell-derived cortical neurons. **a, b**, Sanger sequencing reads of *APP*^{Swe} and *PSEN1*^{M146V} gene edited iPS cell lines. **c–e**, Immunofluorescence staining of markers for neural precursors at DIV10 (**c**), cortical neurons at DIV65 (**d**) and functional synapses at DIV65 (**e**). Scale bars; 100 μ m (**c, d**), 10 μ m (**e**). **f**, Evoked action potentials recorded in a neuron current-clamped to -65 mV. **g**, Mean (\pm s.e.m.) resting membrane potential (V_{rest}), action

potential threshold and action potential overshoot (DIV 71–85; $n = 18$). Properties of the largest action potential elicited in each cell were measured. **h**, Mean number of evoked action potentials increases with increasing stimulus strength. **i**, Spontaneous synaptic activity recorded in a neuron voltage-clamped to -70 mV. **j**, Mean (\pm s.e.m.) frequency and amplitude of spontaneous excitatory postsynaptic currents (sEPSCs) (DIV 71–85; $n = 8$).



Extended Data Figure 7 | Possible mechanism underlying the distance effect for HDR-mediated mutation incorporation with CRISPR/Cas9. CRISPR/Cas9 causes a DSB at a genomic locus, which leads to variable size deletions or strand resections in different cells. Genomes with small deletions or resections are more common than large ones, which is

reflected in the distribution of deleted bases after NHEJ (top left). During HDR, only the part of the repair template overlapping this deletion may be used, which results in fewer mutations incorporations more distal to the cleavage site (bottom left, data pooled for *APP* and *PSEN1* from Fig. 2d).



Extended Data Figure 8 | Next-generation sequencing data analysis pipeline for HDR and indel detection. **a**, For all next-generation sequencing experiments, raw forward and reverse paired next-generation sequencing reads were first merged to obtain single high-quality reads (tool: PEAR), de-multiplexed to separate experiment-specific barcoded reads (seqtk) then filtered to remove low-quality reads. **b**, For experiments using pooled oligonucleotides containing CRISPR/Cas-blocking mutations (displayed in Fig. 1), reads were separated into wild-type (WT) and edited reads, which were then filtered to include only reads that had incorporated the pathogenic mutation (M+) (that is, containing a pathogenic and

CRISPR/Cas-blocking mutation). To account for multiple HDR events after re-editing, reads were then separated into 32 unique categories covering every possible combination of CRISPR/Cas-blocking mutations. **c**, Reads were aligned (bwa mem) and accurate HDR (perfect alignment) or indel distribution was reported (bam-readcount, R). For analysis in Extended Data Fig. 2c, d, reads that had incorporated multiple CRISPR/Cas-blocking mutation were separately analysed. **d**, For the mutation incorporation analyses performed in all other figures reads were filtered for the expected sequence and counted.

Extended Data Table 1 | List of HDR rates determined by next-generation sequencing and single-cell clone analysis

a

| Figure | Locus | sgRNA | Template type | Cell type | % HDR of total reads (SD) | % HDR of edited reads (SD) |
|--------|------------------------|-------|---------------|-----------|---------------------------|----------------------------|
| 1d | APP ^{Swe} | 2 | ssODN | iPSC | 4.3 (0.7) | 10.4 (1.2) |
| 1d | APP ^{Swe} | 2 | ssODN | HEK293 | 4.5 (0.2) | 10.6 (0.8) |
| 1e | PSEN1 ^{M146V} | 22 | ssODN | iPSC | 2.2 (0.2) | 2.8 (0.2) |
| 1e | PSEN1 ^{M146V} | 22 | ssODN | HEK293 | 1.2 (0.1) | 2.6 (0.4) |
| 1i | APP ^{Swe} | 2 | ssODN | iPSC | 3.8 (0.1) | 15.7 (1.7) |
| 1k | APP ^{A673T} | 12 | ssODN | iPSC | 0.3 (0.1) | 3.5 (0.4) |
| ED2a | APP ^{Swe} | 2 | ssODN | HEK293 | 4.9 (0.1) | 8.2 (0.2) |
| ED2b | PSEN1 ^{M146V} | 22 | ssODN | HEK293 | 1.2 (0.1) | 1.9 (0.1) |
| 2c | APP ^{Swe} | 2 | ssODN | iPSC | 2.1 (0.7) | 5.7 (3.1) |
| 2c | PSEN1 ^{M146V} | 22 | ssODN | iPSC | 3.2 (1.5) | 4.1 (2.0) |
| 2d | APP ^{Swe} | 2 | ssODN | HEK293 | 4.1 (0.2) | 9.6 (0.4) |
| 2d | PSEN1 ^{M146V} | 22 | ssODN | HEK293 | 4.3 (0.2) | 9.1 (0.1) |
| ED4d | PSEN1 ^{M146V} | 22 | 200 ssDNA | HEK293 | 6.0 (0.04) | 13.8 (0.1) |
| ED4d | PSEN1 ^{M146V} | 22 | 200 dsDNA | HEK293 | 3.6 (0.1) | 9.5 (0.3) |
| ED4d | PSEN1 ^{M146V} | 22 | 400 ssDNA | HEK293 | 5.1 (0.05) | 11.6 (0.3) |
| ED4d | PSEN1 ^{M146V} | 22 | 400 dsDNA | HEK293 | 3.1 (0.1) | 8.3 (0.1) |
| ED5a | APP ^{Swe} | 12 | ssODN | iPSC | 5.9 (0.3) | 10.6 (0.4) |
| ED5a | APP ^{Swe} | 2 | ssODN | iPSC | 6.7 (1.5) | 10.9 (0.3) |
| ED5a | APP ^{Swe} | 7 | ssODN | iPSC | 0.4 (0.1) | 1.1 (0.2) |
| ED5b | PSEN1 ^{M146V} | 5 | ssODN | iPSC | 1.8 (0.8) | 4.8 (0.1) |
| ED5b | PSEN1 ^{M146V} | 22 | ssODN | iPSC | 2.0 (0.3) | 3.4 (0.6) |
| ED5b | PSEN1 ^{M146V} | 4 | ssODN | iPSC | 1.6 (0.3) | 6.8 (1.4) |
| ED5c | APP ^{Swe} | 12 | ssODN | iPSC | 3.5 (ND) | 5.7 (ND) |
| ED5d | APP ^{Swe} | 12 | ssODN | iPSC | 3.6 (ND) | 8.6 (ND) |
| ED5e | PSEN1 ^{M146V} | 5 | ssODN | iPSC | 2.1 (ND) | 4.4 (ND) |

b

| Figure | Locus | sgRNA | Template type | Picked clones | Bi-allelic HDR clones | % bi-allelic HDR clones |
|--------|------------------------|-------|---------------|---------------|-----------------------|-------------------------|
| 3c | APP ^{Swe} | 12 | ssODN | 720 | 24 | 3.3 |
| 3c | APP ^{Swe} | 2 | ssODN | 912 | 20 | 2.2 |
| 3c | APP ^{Swe} | 7 | ssODN | 1623 | 6 | 0.4 |
| 3c | PSEN1 ^{M146V} | 5 | ssODN | 912 | 22 | 2.4 |
| 3c | PSEN1 ^{M146V} | 21 | ssODN | 960 | 21 | 2.2 |
| 3c | PSEN1 ^{M146V} | 4 | ssODN | 1056 | 20 | 1.9 |
| 3d | APP ^{Swe} | 12 | ssODN | 768 | 16 | 2.1 |
| ED5d | APP ^{Swe} | 12 | ssODN | 1056 | 24 | 2.3 |
| ED5e | PSEN1 ^{M146V} | 5 | ssODN | 192 | 16 | 8.3 |

Extended Data Table 2 | Off-target analysis of knock-in APP^{Swe} and PSEN1^{M146V} iPS cell lines

a

| Gene | gRNA | Tool | ID | Sequence | Type | Mis-match | Location | Strand | COSMID score | ZHANG score | Indels in het |
|------|------|--------|------|--------------------------|----------|-----------|-----------------------|--------|--------------|-------------|---------------|
| APP | 2 | | | GCAGAAATTCGACATGACTCAGG | Target | 0 | 21:25897598-25897620 | - | | | |
| APP | 2 | COSMID | OT1 | ACA-AATTCCAACATGACTCTGG | Del 17 | 2 | 16:48796036-48796057 | - | 1.5 | | None |
| APP | 2 | COSMID | OT2 | CCA-AATTCCAACATGACTCTGG | Del 17 | 2 | 1:147436251-147436272 | + | 1.5 | | None |
| APP | 2 | COSMID | OT3 | CCA-AATTCCAACATGACTCTGG | Del 17 | 2 | 3:103798317-103798338 | - | 1.5 | | None |
| APP | 2 | COSMID | OT4 | TCA-AATTCACAACATGACTCTGG | Del 17 | 2 | 5:62074737-62074758 | - | 1.5 | | None |
| APP | 2 | COSMID | OT5 | GCA-AATTCACAACATGACTCTGG | Del 17 | 2 | X:4446208-4446229 | + | 2.68 | | None |
| APP | 2 | ZHANG | OT6 | TCACAATTCTGACATGACTCAGG | No indel | 3 | 9:80909384-80909406 | - | | 2.525 | None |
| APP | 2 | ZHANG | OT7 | GCAGGACTCCAACATGACTCAGG | No indel | 3 | 12:63304976-63304998 | + | | 0.964 | None |
| APP | 2 | ZHANG | OT8 | GCAGAAATCCAACATGACTCAAG | No indel | 3 | 3:108345664-108345686 | + | | 0.838 | None |
| APP | 2 | ZHANG | OT9 | TGTGAATTCACAACATGACTCAAG | No indel | 4 | 2:156152304-156152326 | + | | 0.795 | None |
| APP | 2 | ZHANG | OT10 | GGTGAATTCACAACATGACTCAAG | No indel | 4 | 1:207331941-207331963 | + | | 0.782 | None |

b

| Gene | gRNA | Tool | ID | Sequence | Type | Mis-match | Location | Strand | COSMID score | ZHANG score | Indels in homo |
|------|------|--------|------|---------------------------|----------|-----------|------------------------|--------|--------------|-------------|----------------|
| APP | 12 | | | GGAGATCTCTGAAGTGAAGATGG | Target | 0 | 21:25897623-25897642 | - | | | |
| APP | 12 | COSMID | OT1 | TG-GATCTCTGAAGTGAAGATGG | Del 18 | 1 | 1:25967504-25967525 | + | 0.78 | | None |
| APP | 12 | COSMID | OT2 | AGTGA-CTCTGAAGTGAAGATGG | Del 15 | 2 | 9:21442538-21442559 | + | 0.99 | | None |
| APP | 12 | COSMID | OT3 | GG-GTTGCTCTGAAGTGAAGAGGG | Del 18 | 2 | 20:32252244-32252265 | + | 1.08 | | None |
| APP | 12 | COSMID | OT4 | GGACTTC-CTGAAGTGAAGAGGG | Del 13 | 2 | 10:1347959-13479615 | - | 1.14 | | None |
| APP | 12 | COSMID | OT5 | GAAATTC-TGAAGTGAAGAGGG | Del 12 | 2 | 6:102732573-102732594 | + | 1.16 | | None |
| APP | 12 | ZHANG | OT6 | GGAAATCTCTAAAGTGAAGAAAAG | No indel | 2 | 6:21801477-21801499 | + | | 3.934 | None |
| APP | 12 | ZHANG | OT7 | GGACCTACTGAAGTGAAGAAAAG | No indel | 3 | 1:85276298-85276320 | - | | 2.426 | None |
| APP | 12 | ZHANG | OT8 | TAAGATTCTCTGAAGTGAAGAAAAG | No indel | 3 | 11:128019181-128019203 | - | | 1.737 | None |
| APP | 12 | ZHANG | OT9 | GAAAGATCCTGGAGTGAAGACAG | No indel | 3 | 4:84241317-84241339 | - | | 1.384 | None |
| APP | 12 | ZHANG | OT10 | GCAGATCTCGGAAGTGAAGCTAG | No indel | 3 | 2:464569-464591 | + | | 1.374 | None |

c

| Gene | gRNA | Tool | ID | Sequence | Type | Mis-match | Location | Strand | COSMID score | ZHANG score | Indels in het | Indels in homo |
|-------|------|--------|------|--------------------------|----------|-----------|------------------------|--------|--------------|-------------|---------------|----------------|
| PSEN1 | 5 | | | TGTGTGTCATGACTATCCTCCTGG | Target | 0 | 14:73173656-73173678 | + | | | | |
| PSEN1 | 5 | COSMID | OT1 | TG-TATCATCGCTATCCTCCCGG | Del 18 | 2 | 17:71454517-71454538 | + | 1.55 | | None | None |
| PSEN1 | 5 | COSMID | OT2 | TGCTGACATGA-TATCCTCCAGG | Del 9 | 2 | 17:22640654-22640675 | + | 1.67 | | None | None |
| PSEN1 | 5 | COSMID | OT3 | TATGTTCAT-AAATCCTCCTGG | Del 11 | 2 | 12:26052306-26052327 | - | 1.94 | | None | None |
| PSEN1 | 5 | COSMID | OT4 | TGTTTAAATGAC-ATCCTCCAGG | Del 8 | 2 | 14:19342489-19342510 | + | 2.03 | | None | None |
| PSEN1 | 5 | COSMID | OT5 | TGTTTAAATGAC-ATCCTCCAGG | Del 8 | 2 | 15:20767818-20767839 | + | 2.03 | | None | None |
| PSEN1 | 5 | ZHANG | OT6 | TGATGTCACTACTATCCTCCAAG | No Indel | 2 | 13:62696807-62696829 | + | | 6.438 | None | None |
| PSEN1 | 5 | ZHANG | OT7 | TFTTTCTCTGACTATCCTCCAG | No Indel | 3 | 9:36047217-36047239 | + | | 2.543 | None | None |
| PSEN1 | 5 | ZHANG | OT8 | TGATGTCTGACTATCCTCAAAG | No Indel | 3 | 4:153813094-153813116 | + | | 1.422 | None | None |
| PSEN1 | 5 | ZHANG | OT9 | AGATATCATCACTATCCTCCTAG | No Indel | 4 | 12:105761175-105761197 | + | | 1.299 | None | None |
| PSEN1 | 5 | ZHANG | OT10 | TATCTAATGACTATCCTCCAG | No Indel | 4 | 12:29848003-29848025 | - | | 0.905 | None | None |

a-c. List of properties of the five most similar off-target sites predicted each for APP-sgRNA2 used for heterozygous APP^{Swe} lines (a), APP-sgRNA12 used for homozygous APP^{Swe} lines (b) and PSEN1-sgRNA5 used for both heterozygous and homozygous PSEN1^{M146V} lines (c) using COSMID or the Zhang laboratory CRISPR design tool. Red bases indicate sequence differences from target sequence. No off-target indels were identified.