HHS Public Access

Author manuscript

Nat Biotechnol. Author manuscript; available in PMC 2022 February 16.

Published in final edited form as:

Nat Biotechnol. 2022 February ; 40(2): 189–193. doi:10.1038/s41587-021-00901-y.

CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells

Karl Petri[1,2,*], **Weiting Zhang**[3,4,*], **Junyan Ma**[3,4,5,*], **Andrea Schmidts**[4,6,*], **Hyunho Lee**[1,2], **Joy E. Horng**[1,2], **Daniel Y. Kim**[1,2], **Ibrahim C. Kurt**[1,2], **Kendell Clement**[1,2], **Jonathan Y. Hsu**[1,2], **Luca Pinello**[1,2], **Marcela V. Maus**[4,6], **J. Keith Joung**[1,2,8], **Jing-Ruey Joanna Yeh**[3,4,8]

[1]Molecular Pathology Unit and Center for Cancer Research, Massachusetts General Hospital, Charlestown, MA 02129, USA

[2]Department of Pathology, Harvard Medical School, Boston, MA 02115, USA

[3]Cardiovascular Research Center, Massachusetts General Hospital, Charlestown, MA 02129, USA

[4]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA

[5]Medical College, Dalian University, Dalian 116622, China

[6]Cellular Immunotherapy Program, Massachusetts General Hospital Cancer Center, Charlestown, MA 02114, USA

Abstract

Prime Editors have been delivered using DNA or mRNA vectors. Here we demonstrate prime editing (PE) with purified ribonucleoprotein (RNP) complexes. We introduced somatic mutations in zebrafish embryos with frequencies as high as 30% and demonstrate germline transmission. We also observed unintended insertions, deletions and pegRNA scaffold incorporations. In HEK293T

[8]These authors jointly supervised this work: J. Keith Joung and Jing-Ruey Joanna Yeh. jjoung@mgh.harvard.edu; jyeh1@mgh.harvard.edu. **Correspondence and requests for materials** should be addressed to J.K.J. or J.-R. J.Y..

[*]These authors contributed equally to this work

Author Contributions

K.P., W.Z., J.M., H.L., A.S., J.E.H., I.C.K., J.K.J. and J.-R.J.Y. designed the project; K.P., W.Z., J.M., H.L., A.S., J.E.H. and D.Y.K. performed the experiments; K.P., W.Z., J.M., H.L., A.S., J.E.H., I.C.K. and J.Y.H. developed the methods; K.P., H.L., K.C. and L.P. performed informatic analysis; M.V.M., J.K.J. and J.-R.J.Y. provided resources and oversight; K.P., W.Z., J.M., A.S., H.L., J.K.J., and J.-R.J.Y. wrote the manuscript with input from all the authors.

Competing Financial Interests Statement

J.K.J. has financial interests in Beam Therapeutics, Chroma Medicine (f/k/a YKY, Inc.), Editas Medicine, Excelsior Genomics, Pairwise Plants, Poseida Therapeutics, SeQure Dx, Transposagen Biopharmaceuticals, and Verve Therapeutics (f/k/a Endcadia). K.P. has a financial interest in SeQure Dx, Inc.. L.P. has financial interests in Edilytics, SeQure Dx, Inc. and Excelsior Genomics. K.P. and D.Y.K. are paid consultants to Verve Therapeutics. K.C. is an employee, shareholder, and officer of Edilytics, Inc. J.K.J.'s, L.P.'s, K.C.'s, K.P.'s, and D.Y.K.'s interests were reviewed and are managed by Massachusetts General Hospital and Partners HealthCare in accordance with their conflict of interest policies. M.V.M. is an inventor on patents related to adoptive cell therapies, held by Massachusetts General Hospital and University of Pennsylvania (some licensed to Novartis). M.V.M. holds equity in TCR2 and Century Therapeutics, and has served as a consultant for multiple companies involved in cell therapies. The remaining authors declare no competing interests.

Additional information

Supplementary information is available for this paper.

and primary human T cells, PE RNP complexes introduced desired edits with frequencies of up to 21% and 7.5%, respectively.

Editorial Summary

Prime editors are delivered as ribonucleoproteins to zebrafish embryos and human primary cells.

The recently described prime editing (**PE**) method can install all 12 nucleotide substitutions, short insertions, and short deletions[1,2]. PE employs a prime editing guide RNA (**pegRNA**) and a prime editor protein (**PE2**) consisting of a Cas9 nickase (**nCas9**) fused to an engineered M-MLV reverse transcriptase (**RT**) (Fig. 1a and Supplementary Note 1). PE has been used to edit the genomes of mammalian cells and organoids in culture, plants, *Drosophila*, and somatic cells in mice with the PE components delivered as DNAs or RNAs[1,3–7]. Here we demonstrate that purified PE RNP complexes can induce somatic and germline-transmissible mutations in zebrafish and efficiently edit the genomes of human cell lines and human primary cells.

We first constructed bacterial and mammalian expression plasmids encoding PE2 without or with a polyhistidine tag on its N- or C-terminal ends (named PE2, His-PE2 and PE2-His, respectively) and found that the untagged and tagged proteins elicited similar editing efficiencies or spectra of editing outcomes when expressed in human HEK293T cells (Supplementary Fig. 1a–b). Subsequently, we overexpressed and purified full-length PE2-His protein from *E. coli* (Fig. 1b) and verified that purified PE2-His together with *in vitro* transcribed pegRNA could direct site-specific DNA nicks and carry out pegRNA-templated reverse transcription *in vitro* (Fig. 1c and Supplementary Figs. 2a–c). Amplification and sequencing of these reaction products showed that PE2 can reverse transcribe the entire pegRNA, including pegRNA scaffold and spacer sequences (Supplementary Figs. 2c–e).

Next, we tested whether PE RNP complexes with PE2-His could mediate targeted point mutations in zebrafish embryos at the *tyr_1* and *tyr_2* sites in the zebrafish *tyr* gene using the **PE3** strategy (Supplementary Note 1). We used pegRNAs harboring primer binding sites (**PBS**s) and RT templates (**RTT**s) of various lengths and, for the *tyr_1* pegRNAs, two different pegRNA architectures (C9[8] and C9E[9]), together with their matched nicking gRNAs (**ngRNA**) (Supplementary Note 2). RNP complexes were injected into single cell zebrafish embryos incubated at 28.5 °C or 32 °C and assessed for on-target prime editing one day following injection using targeted amplicon next-generation sequencing (**NGS**) (Fig. 1d). We observed three types of on-target edits: “pure prime edits” (**PPE**s) — alleles with only the intended edit, “impure prime edits” (**IPE**s) — alleles with both the intended edit and additional mutations, and “byproduct edits” — alleles without the intended edit but with other mutations (Supplementary Note 3, Supplementary Table 1 and 2). The mean PPE and IPE frequencies were 0.53% and 0.45% for *tyr_1*, and 0.25% and 0.48% for *tyr_2*, respectively, though the rates varied across the different conditions tested (Supplementary Fig. 3a). In comparison, byproduct frequencies were more uniform across the different conditions (generally < 10% for *tyr*_1 and < 5% for *tyr*_2) (Supplementary Fig. 3a). IPE alleles had a wide range of unintended edits including pegRNA scaffold incorporations, deletions, potential RTT duplications/insertions, and other

insertions/substitutions (Supplementary Note 3). Notably, deletions surrounding the 3' boundary of the RT-synthesized DNA flap were common in IPE alleles (Fig. 1e and Supplementary Fig. 3b). By contrast, byproduct edits mostly harbored deletions centered around the nicking site of PE2 (Fig. 1e and Supplementary Fig. 3b). The C9E pegRNA scaffold generally yielded higher PPE and IPE frequencies without substantially increased byproduct edits. Moreover, edit frequencies of RTTs of 13–15 nts were comparable, or in some cases preferable, to longer lengths, whereas the effects of PBS length and temperature were more variable (Supplementary Note 2).

Next, we compared the PE2 and PE3 strategies (Supplementary Note 1) in zebrafish, with PBS lengths of 10 and 13 nts, and incubation temperatures of 28.5 and 32 °C (Fig. 1f). The pegRNAs were designed to introduce single G → C or G → T point mutations into eight different zebrafish genes. All the pegRNAs used the C9E scaffold, had 14 or 15-nt RTT lengths, and targeted sites that we had determined could be efficiently mutated with Cas9 protein (Supplementary Fig. 4). We observed PPEs and IPEs at variable frequencies (Supplementary Fig. 5) and byproduct edits at generally higher frequencies than PPEs and IPEs across all conditions tested at the target sites (Supplementary Fig. 5). Comparative analyses of these data revealed that certain conditions modestly improved PPE frequencies without proportionally increasing IPE frequencies or appreciably impacting byproduct edit frequencies: incubation at 32 °C was better than 28.5 °C and pegRNA PBS length of 10 nts was better than 13 nts (Fig. 1g). Outcomes with the PE3 strategy did not appear to be consistently better than with PE2 (Figs. 1g and 1h), contrary to what has been reported in human cell lines[1]. Varying the pegRNA:ngRNA ratio or deploying the PE3b strategy (Supplementary Note 1) also did not consistently improve PPE frequencies relative to PE2 (Supplementary Fig. 6a–c and Supplementary Note 4). We speculate that the effects of PE3 might depend on cell division rate and cellular DNA repair machinery differences between zebrafish embryos and human cells (Supplementary Note 4). Overall, pegRNAs with PBS length of 10 nts and embryos incubated at 32 °C generally yielded higher PPE frequencies (mean ranges of 0.46% - 4.01% and 0.28% - 2.06% with PE2 and PE3, respectively; Fig. 1h) without proportional increases in undesired IPE or byproduct frequencies.

To investigate the efficiency of PE2 for creating short defined-length deletions and insertions in zebrafish, we used pegRNAs designed to introduce 5-bp or 10-bp deletions at three sites and insertions of 3 – 18 bps at two sites. We obtained mean PPE frequencies ranging from 4.13% to 33.61% for precise deletions (Fig. 1i) and from 0.10% to 18.00% for precise insertions (Fig. 1j), which were higher than PPE frequencies for point mutations at the same target sites (Fig. 1h). Because the intended edits in these experiments were deletions or insertions, we could not always distinguish between IPEs and the insertions and deletions (indels) included in byproduct edits but the frequencies of all of these "non-PPE" alleles were comparable to or higher than those of the PPEs (Supplementary Figs. 7a and 8a–b). The majority of the non-PPE alleles are larger deletions or deletions close to the pegRNA nicking site or the 3' end of the DNA sequence complementary to the RTT (Supplementary Figs. 7b and 8c).

Next, we demonstrated that PE-induced mutations could be passed through the zebrafish germline by installing two human pathogenic variants that cannot be readily created using

existing CRISPR base editors (BEs): *TYR* P301L[10,11], associated with oculocutaneous albinism, and oncogene *KRAS* G12V[12]. Previous attempt to create the corresponding *TYR* mutation in the zebrafish *tyr* gene (P302L), which requires a C→T substitution (CCC→CTC), using a CRISPR cytosine BE was not successful[11] probably because this position is embedded in a string of cytosines. The G→T transversion required for installing the *kras* G12V mutation in zebrafish cannot be generated by existing BEs. Using PE, we successfully introduced the *tyr* P302L and the *kras* G12V substitutions into zebrafish embryos with mean somatic cell PPE frequencies up to 3.33% and 6.53%, respectively (Fig. 2a). As expected, we also observed associated IPE and byproduct edits (Fig. 2a). Incubation at a higher temperature (34.5 °C) did not yield consistently higher PPE frequencies (Fig. 2a) nor did variation of pegRNA:PE2 protein ratios (Supplementary Fig. 9). For the *tyr* P302L mutation, we screened 14 adult F0 fish and identified one founder that transmitted the desired PPE to 8.3% of its offspring and a second founder that transmitted an IPE to 7.1% of its offspring (Fig. 2b–2c and Supplementary Table 3). Similarly, for the *kras* G12V mutation, we screened 14 adult F0 fish and identified one founder that transmitted the desired PPE to 12.3% of its offspring and a second founder that transmitted an IPE to 7.1% of its offspring (Fig. 2b–2c and Supplementary Table 4). Interestingly, both germline-transmitted IPE alleles harbored a single C-to-G or A-to-G substitution at the first base 3’ to the RTT-encoded sequence, perhaps reflecting provenance from pegRNA scaffold incorporations (Supplementary Note 3). Efficient germline transmission of both of these pathogenic variants that we attempted to install demonstrate that PE can be used to create zebrafish models of human disease.

Because gene-editing nuclease-induced homology-directed repair (**HDR**) could also be used to create the same modifications that we installed using PE, we compared the efficiencies and precision of PE and HDR for introducing targeted base substitutions, insertions, and deletions at two sites. The results show that PE induced all three types of intended mutations at frequencies comparable to or higher than Cas9-mediated HDR (Supplementary Fig. 10). PE consistently induced fewer unwanted edits than Cas9-mediated HDR (Supplementary Fig. 10), demonstrating the relatively higher precision of the PE method.

To determine whether PE could induce unwanted off-target mutations in zebrafish, we used targeted amplicon NGS to assess all sites in the zebrafish reference genome (GRCz11) with three or fewer mismatches relative to the on-target sites of five pegRNAs (Methods, Supplementary Table 5). Among the 29 mismatched sites examined, two sites for the *kras* pegRNA showed significant evidence of indel mutations with mean indel frequencies of 1.04% and 0.69% (mean on-target indel frequency was 12.79% in the same samples). The majority of the editing outcomes at the two off-target sites (*kras* OT1 and OT2) consisted of RTT insertions and pegRNA nick-centered deletions, although deletions at the 3’ end of the sequence homologous to the RTT and pegRNA scaffold incorporations were also observed (Supplementary Fig. 11). Of note, *kras* OT1 and OT2 had homology not only to the spacer sequence of the *kras* pegRNA but their adjacent sequences also showed homology to the RTT sequence (Supplementary Fig. 11).

Finally, we tested whether PE RNP could be used to generate targeted edits in human cells. We nucleofected HEK293T cells with highly concentrated preparations of PE2 protein

(Methods) and synthetic pegRNAs/ngRNAs with 5'- and 3'-end stabilizing modifications (Supplementary Table 6). PE RNP introduced a G-to-C substitution, a 3-bp insertion, and a 5-bp deletion using three pegRNA/ngRNA combinations with mean PPE efficiencies of 5.95%, 14.59%, and 20.61%, respectively (Fig. 2d). The same pegRNA/ngRNA combinations induced PPE efficiencies ranging from 1.49% to 7.54% in primary human T cells (Fig. 2e). The 3-bp insertion pegRNA with stabilization modifications at both the 5' and 3' ends gave higher PPE frequencies than its counterpart with modifications only at 3' ends in both cell types (Supplementary Fig. 12). Although the structures and types of IPEs and/or byproduct edits observed in human cells were similar to those in zebrafish (Supplementary Fig. 13), the relative frequencies of these unwanted edits were lower. The lower frequencies of PPEs we observed in human primary cells vs human cell line are consistent with previous studies using DNA or RNA-encoded PE components[1,5].

This report demonstrates the ability of PE to introduce a broad range of targeted genetic modifications with high efficiencies in zebrafish. A subset of these genetic modifications could also be induced with BEs and further studies are needed to compare the efficiencies between PE and various BE platforms for such BE-inducible modifications[2]. In addition to the conditions tested in this study, other strategies, such as use of synthetic pegRNAs harboring stabilizing modifications could potentially further increase PPE frequencies in zebrafish (Supplementary Fig. 14). If the efficiency of PE can be improved, it may become possible to introduce phenotypes in F0 animals as has been previously shown with CRISPR Cas9 and BE[11,13]. Based on this study, we have prepared some general guidelines for using PE RNP in zebrafish (Supplementary Note 5). The successful germline transmission of PE-induced mutations in zebrafish also suggests the utility of PE for creating disease models in other vertebrates.

Our demonstration of PE using RNPs could broaden therapeutic applications of this platform given that the large size of PE2 presents challenges for delivery by viral vectors (even using "split inteins"[14]) or its efficient synthesis as an mRNA transcript. PE RNP will also likely induce fewer off-target effects due to its shorter duration of activity (as has been observed with CRISPR Cas9 RNP delivery[15]) and avoids the possibility of vector DNA integration into the genome. The findings presented here could facilitate the use of PE in a wide variety of experimental systems and therapeutic applications.

METHODS

Generation of pegRNAs and gRNAs

All pegRNAs and gRNAs used in this study were produced by *in vitro* transcription, except for the synthetic pegRNAs and nicking gRNAs obtained from Synthego and Agilent Technologies (Supplementary Table 6). The DNA templates for *in vitro* transcription were constructed in two steps as described below. For *tyr*_1 and *tyr*_2 pegRNAs and gRNAs in the standard CRISPR-Cas9 (C9) scaffold[8], first SP6 promoter and spacer sequences were added to the C9 scaffold derived from the plasmid OC1758 (Supplementary Table 7) *via* PCR using primers carrying these sequences (Supplementary Tables 8–9). In the second step, the final DNA template with a 3' extension containing the RT template and the primer binding sequence (PBS) for pegRNAs or without the 3' extension was generated by PCR.

For pegRNAs and gRNAs in the enhanced CRISPR-Cas9 (C9E) scaffold[9], first a target site-specific oligonucleotide was annealed to the C9E constant oligonucleotide (Supplementary Tables 8–10) and a fill-in reaction was performed using T4 DNA polymerase (New England Biolabs, M0203L) as previously described[16]. This was followed by the generation of the final DNA template by PCR as detailed above for the C9 scaffold RNAs. The sequences of pegRNA and gRNA target sites and 3' extensions of pegRNAs used in the study are in Supplementary Table 6 and 11. The PCR reactions were performed with Phusion High-Fidelity DNA polymerase (New England Biolabs, M0530L) using the following cycling conditions: 98 °C for 30 s, followed by 35 cycles of: 98 °C for 10 s, 65 °C for 30 s, and 72 °C for 10 s, followed by a final 72 °C extension for 5 min. The products from the first PCR or construction step described above were either purified using the Monarch® PCR & DNA Cleanup Kit (New England Biolabs, T1030L) or diluted 20-fold with deionized water before being used as templates for the PCR in the second step. *In vitro* transcription of pegRNAs or gRNAs was performed using the HiScribe™ T7 High Yield RNA synthesis Kit (New England Biolabs, E2040S) or HiScribe™ SP6 RNA Synthesis Kit (New England Biolabs, E2070S) and purified using the Monarch® RNA Cleanup Kit (New England Biolabs, T2030L).

Comparison of the editing efficiencies of PE2 protein with and without the His-tag in human cells

Using standard Gibson cloning methods, we added a 6xHis-tag and a short Glycine-Serine (GS) linker sequence to the 5' or 3' of the PE2 ORF within pCMV-PE2 parental plasmid (addgene #132775) to generate mammalian expression construct pCMV-His-PE2 and pCMV-PE2-His respectively (Supplementary Table 7). The editing efficiencies of PE2 with and without the His-tag were compared by transfecting HEK293T cells with pCMV-His-PE2, pCMV-PE2-His, and pCMV-PE2 (addgene #132775) (Supplementary Figure 1). HEK293T (ATCC® CRL-3216™) cells were obtained from ATCC. HEK293T cells (1.25×10^5 /well) were seeded in 24-well plates and transfected 24 hrs later with 375 ng of the PE2 plasmid, 125 ng of the pegRNA plasmid (IK1849, IK1856, or IK1872), 41.5 ng of the nicking gRNA plasmid (IK1900, IK1904, or IK1909) (Supplementary Table 7), and 1.5 μL of the TransIT-X2 reagent (Mirus Bio, MIR6005) following the manufacturer's protocol. The genomic DNA was extracted 72 hrs post-transfection using the QIAmp DNA Mini Kit (Qiagen, 51306) and the target loci were amplified through touchdown PCR using Phusion High Fidelity Polymerase (NEB, M0530S) and the primers listed in Supplementary Table 12, sequenced on the MiSeq system (Illumina) (v2 sequencing kit, 2×150 bp) and analyzed with CRISPResso2[17].

Purification of PE2-His protein

The C-terminal His-tagged PE2 protein was overexpressed in *Escherichia coli* and purified as described below. We generated a bacterial codon-optimized SpCas9 (H840A) sequence within pET-28b backbone (Addgene #73018) and fused it with bipartite NLS, a 33-amino acid linker and the engineered M-MLV RT from the parental pCMV-PE2 plasmid (Addgene #132775). To this construct, we inserted a 6xHis-tag and a short GS linker to 5' or 3' of the PE2 ORF to generate the pET-His-PE2 and pET-PE2-His constructs, respectively (Supplementary Table 7). The amino acid sequences of His-tagged PE2 in these constructs

are shown in Supplementary Table 13. Chemically competent *E. coli* Rosetta (DE3) cells (Novagen, 70954) were transformed with pET-PE2-His or pET-His-PE2 by heat shock following the manufacturer's instructions. 10 ml of overnight culture grown from a single colony in Luria Bertani (LB) medium with 50 μg/ml kanamycin was used to inoculate 1 liter of the same medium and cultured at 37 °C. Once OD_{600} of the culture reached 0.6, PE2-His expression was induced with 1 mM isopropyl ß-D-1-thiogalactopyranoside (IPTG) at 18 °C for 5 hrs, followed by 26 °C for 12 hrs. Cells were collected by centrifugation and lysed in 20 mM Tris-HCl (pH 8.0), 500 mM NaCl (concentration used in Fig. 1b "final purification"), 10 mM imidazole, 1 mM phenylmethylsulfonyl fluoride, 10 μg/ml lysozyme, 10% CHAPS and 5 units/ml Benzonase® Nuclease (Millipore, E1014) at room temperature for 30 mins or till the lysates were no longer viscous. After centrifugation at 6000 × g for 30 mins at 4 °C, 50 ml of supernatant was mixed with 3 ml of Ni Sepharose® Excel resin (GE Healthcare, GE17371201) pre-equilibrated with the wash buffer containing 20 mM Tris-HCl, pH 8.0, 500 mM NaCl and 10 mM imidazole and was kept on a rotator at 4 °C for 2 hrs. Subsequently, the resin with the supernatant was loaded onto the Econo-Column® chromatography column (Bio-Rad, 7374011), washed with the wash buffer, followed by a wash in the same buffer containing 20 mM imidazole. Protein was eluted with 20 mM Tris-HCl (pH 8.0), 300 mM NaCl, and 250 mM imidazole, analyzed for purity by SDS-PAGE and the elution buffer replaced with 20 mM Tris, pH 8.0, 200 mM KCl, 10 mM $MgCl_2$, 10–15% glycerol by dialysis using Slide-A-Lyzer™ G2 Dialysis Cassettes with 20-kDa cutoff (Thermo Fisher Scientific, 87737) at 4 °C overnight. Protein concentration was determined with A_{280} on NanoDrop spectrometer or Quick Start™ Bradford Protein Assay (Bio-Rad, 5000201) and the purified proteins were stored at −80 °C.

In vitro testing of PE2-His in nicking and reverse transcription assays

All *in vitro* nicking assays were performed using a 51-basepair (bp) dsDNA substrate containing a single 5'FAM label such that a nicking event would yield a 5'FAM-labeled product of 34 nt. The dsDNA substrate was generated by annealing the corresponding oligonucleotides (Integrated DNA Technologies, Supplementary Table 14) by heating at 95 °C for 3 mins and slow cooling to room temperature. The templates for gRNAs and pegRNAs (Supplementary Table 14) were generated by cloning hybridized oligonucleotides (Integrated DNA Technologies) into MSP3485 (a bacterial gRNA expression vector with T7 promoter) and pHL10 (a PE adapted MSP3485 vector) linearized by *Bsa*I (Supplementary Table 7). After linearization of the plasmids with BbsI-HF, gRNAs and pegRNAs were produced through *in vitro* transcription using T7 RiboMAX Express Large Scale RNA Production kit (Promega, P1320) followed by column-based purification using the MEGAclear Transcription Clean-up Kit (Thermo Fisher, AM1908). PE2-His protein was prepared as described above, whereas Alt-R S.p. Cas9 H840A Nickase V3 (nCas9) was purchased from Integrated DNA Technologies (#1081058). Each reaction used 5 pmol of dsDNA substrate and was performed in 20μl total volume, unless indicated otherwise. *In vitro* reactions assessing nicking contained nCas9 (nicking control) at a molar ratio of 1:10:16 (dsDNA:nCas9:(pegRNA or gRNA)), or PE2-His at a ratio varied between 1:6:8, 1:10:16, and 1:16:20 (dsDNA:PE2-His:(pegRNA or gRNA)). nCas9 or PE2-His was complexed with gRNA/pegRNA (RNP) at room temperature for 15 mins, then mixed with the reaction buffer (5% glycerol, 100 mM KCl, 10 mM Hepes pH 7.5, 0.2 mM EDTA, 3

mM MgCl2, 5 mM DTT final concentration) and the dsDNA substrate. *In vitro* reactions assessing both nicking and reverse transcription were supplemented with 1 μl of 10 mM dNTPs (New England Biolabs, N0446S). The reactions were incubated at 37 °C for 1 hr (Supplementary Fig. 2b) or 15 min (Supplementary Fig. 2c), stopped by adding 1 μl of Proteinase K (New England Biolabs, P8107S) at 37 °C for 20 – 45 min followed by 2x urea loading buffer (NAT Diagnostics, EC-857) and boiling the reactions at 95 °C for 3 mins. The reaction products were separated on a heated 15% denaturing TBE-PAGE (NAT Diagnostics, EC-833) and visualized with a transilluminator (VWR, 490006–532).

Preparation of PE2-His concentrate for nucleofection

Rosetta (DE3) competent cells (Novagen, 70954) or T7 Express competent cells (New England Biolabs, C2566H) were transformed with pET-PE2-His following the manufacturers' instructions. 30 ml of the overnight culture grown from a single colony in LB with 50 μg/ml kanamycin was transferred into 3 liters of the same medium with 10 mM $MgCl_2$ and cultured at 37 °C. Once OD_{600} of the culture reached 0.7–0.8, PE2-His expression was induced with 0.5 mM IPTG for 18 hrs at 18 °C. Cells were collected and sonicated in lysis buffer consisting of 100 mM Tris-HCl, pH 8.0,1 M NaCl, 10 mM imidazole, 20% glycerol, 0.1% Triton X-100, 5 mM 2-mercaptoethanol, and the cOmplete™ Protease Inhibitor Cocktail (Roche, 11697498001) on ice using Branson 450 Analog Sonifier. After centrifugation of the lysates at 24,676 × g for 30 mins at 4 °C, the supernatant was mixed with the Ni Sepharose™ 6 Fast Flow resin (GE Healthcare, GE17–5318-01) pre-equilibrated with wash buffer I (100 mM Tris-HCl, pH 8.0, 1 M NaCl, 10 mM imidazole, 20% glycerol, 0.1% Triton X-100, and 5 mM 2-mercaptoethanol) and was kept on a rotator at 4 °C for 1 hr. Then the resin was loaded onto the Econo-Column® chromatography column (Bio-Rad, 7371022) and washed with wash buffers I and II (100 mM Tris-HCl, pH 8.0, 1 M NaCl, 20 mM imidazole, 20% glycerol, 5 mM 2-mercaptoethanol). PE2-His protein was eluted from the column with 100 mM Tris-HCl, pH 8.0, 500 mM NaCl, 300 mM imidazole, 20% glycerol, and 2 mM tris(2-carboxyethyl) phosphine) and analyzed for purity by SDS-PAGE. The protein was concentrated and elution buffer replaced with 50 mM Tris-HCl, pH 7.5, 300 mM KCl and 10% glycerol using Amicon® Centrifugal Filter Units with 30-kDa cutoffs (EMD Millipore, UFC910008 or UFC510008) at 4 °C. Protein concentration was determined with A_{280} on NanoDrop spectrometer and purified protein stored in aliquots at −80 °C.

Prime editing RNP experiments in HEK293T cells

HEK293T (ATCC® CRL-3216™) cells were obtained from ATCC. HEK293T cells were cultured in DMEM with 10% heat-inactivated FBS, 2 mM GlutaMax (ThermoFisher, 35050–061), 50 U/ml penicillin-streptomycin (ThermoFisher, 15070063) at 37 °C and 5% CO_2 and seeded at 3×10^5 cells/mL 24 hrs before nucleofection. To prepare the PE2 protein-pegRNA/gRNA complex (RNP), approximately 600 pmol of the PE2 protein was incubated with 800 pmol of the pegRNA and 267 pmol of the nicking guide RNA (Agilent Technologies, Supplementary Table 6) in 320 mM KCl in 8.6 μl total volume for 15 min at room temperature. HEK293T cells (2×10^5) were nucleofected with 2 μl of the RNP using SF Cell Line 4D-Nucleofector™ X Kit S (Lonza, V4XC-2032) with the ED-130 program

and genomic DNA extracted 72 hrs later using QIAmp DNA Mini Kit (Qiagen, 51306) according to the manufacturer's protocol.

Prime editing RNP experiments in primary human T cells

Leukopaks from healthy donors were obtained from the Massachusetts General Hospital Blood Transfusion service under an institutional review board–approved protocol. T cells were isolated from the leukopaks using RosetteSep Human T Cell Enrichment Cocktail (Stemcell Technologies, 15061) followed by Ficoll density gradient centrifugation (Ficoll-Paque Plus, GE Healthcare, 17–1440-03) according to the manufacturers' instruction. The purified T cells were cryopreserved in FBS (Corning, 35–010-CV) with 10% DMSO (Thermo Scientific, 85190).

For PE experiments, the T cells (1×10^6 cells/mL) were cultured in RPMI 1640 medium containing 2 mM GlutaMAX (Thermo Scientific, 72400–120), 10% FBS, 1% penicillin-streptomycin (100 IU/ml; Thermo Scientific, 15140122) and 20 IU/mL recombinant human interleukin-2 (Peprotech, 200–02) for 24 hrs and activated with 1.5% phytohemagglutinin M (Thermo Scientific, 10576015). Approximately 48 hrs later, 5×10^5 cells were electroporated with RNP (prepared as above) via the EH-115 program using the P3 primary cell nucleofector kit (Lonza, V4XP-3024) and a 4D-Nucleofector. Electroporation conditions for primary T cells were previously optimized using the pmaxGFP vector (Lonza) and flow cytometry analysis (Supplementary Data 1). The cells were cultured at 1×10^6 cells/mL for approximately 72 hrs post electroporation, after which dead cells were removed by Ficoll density gradient centrifugation[18], and remaining live cells washed thrice in PBS + 2% FBS and genomic DNA extracted using QIAmp DNA Mini Kit (Qiagen, 51306).

Zebrafish husbandry

All zebrafish husbandry and experiments were approved by the Massachusetts General Hospital Subcommittee on Research Animal Care, and performed in accordance with the guidelines of the Institutional Animal Care and Use Committee at the Massachusetts General Hospital.

Zebrafish prime editing

Microinjections were performed using 1-cell stage of TuAB zebrafish embryos. For PE experiments, each embryo was injected with 2 nl of the RNP mixture (see Supplementary Table 15 for formulation) and immediately transferred to incubators at designated temperatures. To determine CRISPR editing efficiency at the selected target sites (Supplementary Fig. 4), the embryos were injected with approximately 2 nl of solution containing Cas9 protein (500 ng/μl) and 20 ng/μl each of the gRNA targeting the ten sites and incubated at 28.5 °C. One day post fertilization, 5–10 embryos that developed normally from each condition were pooled and lysed in 10 mM Tris-HCl (pH 8.0), 2 mM EDTA (pH 8.0), 0.2% Triton X-100, and 100 μg/ml Proteinase K (5–6 μl lysis buffer/embryo). Lysates were incubated at 50 °C overnight with occasional mixing, heated at 95 °C for 10 min to inactive Proteinase K, and stored at 4 °C for deep sequencing.

Zebrafish homology-directed repair (HDR)

To determine the efficiencies of HDR by Cas9 RNP complexes and single-stranded oligodeoxynucleotide (ssODN) templates, we designed 120-bp ssODNs encompassing the intended edits (Supplementary Table 16). The sequences of the oligonucleotides corresponded to the sequences of the non-target DNA strand at the Cas9 target loci. The embryos were injected with approximately 2 nl of solution containing high dose Cas9 RNP complexes (500 ng/μl Cas9 protein and 200 ng/μl gRNA) or low dose Cas9 RNP complexes (250 ng/μl Cas9 protein and 100 ng/μl gRNA) with 1 μM ssODNs and incubated at 28.5 °C. Lysates from embryos were prepared as for PE.

Screening of founders for PE-mediated edits

Adults raised from embryos that had been injected with *tyrP302L* or *krasG12V* pegRNAs were screened for founders that could transmit the PE-mediated edits through germline. Potential founders were outcrossed with wild-type fish and their embryos screened in 6–12 pools of five embryos per pool by allele-specific PCR using the following primers: Tyr302_AS_1F, GGGCCTTTACTGCGCAACCT; Tyr302_1R TTGTACCTTCCAGCGCGTTC; zfKrG12V_AS_2F, TTGTGGTCGTGGGAGCTAT; zfKrG12V_1R, CTCACCTCTATAGTTGGGTC. The DNA from the pools that showed positive PCR results were subjected to targeted next-generation amplicon sequencing. Next, individual embryos from the founders that showed transmission of pure PE or impure PE alleles were screened by allele-specific PCR and the mutant sequences were confirmed by deep sequencing and/or Sanger sequencing. Frequency and sequence of potential byproduct alleles were analyzed only in the offspring of the identified PE founders.

Targeted deep sequencing

PE target site regions were amplified from 1 μl of the zebrafish embryo lysate using touchdown PCR with Phusion High Fidelity Polymerase (NEB, M0530S) and the primers containing partial Illumina sequencing adapters (Supplementary Table 12). The PCR products were purified using paramagnetic beads (1:1 beads to sample), using the same purification protocol as with AMPure XP beads (Beckman Coulter, B37419AB), and product purity was assessed via capillary electrophoresis on a QIAxcel instrument (Qiagen). Illumina barcodes and P5/P7 sequences were attached to the purified product (100 ng) *via* amplification with NEBNext Multiplex Oligos for Illumina (NEB, E7600S). The resulting PCR products were purified and purity assessed as described above. The resulting sequencing libraries were quantified via droplet digital PCR (Bio-Rad) and sequenced using MiSeq system (Illumina) (v2 kit, 2×150bp).

Deep sequencing analysis

The sequencing data were analyzed with CRISPResso2[17] using HDR mode. CRISPResso (HDR mode) assigns sequencing reads to three categories: ‘HDR’, ‘reference’, and ‘ambiguous’. The HDR category consists of reads that have higher sequence similarity to the edited than to the unedited amplicons. The reference category contains reads that have higher sequence similarity to the unedited amplicons than with the edited amplicons. Reads in the ambiguous category align equally well to the edited and unedited amplicon (e.g. if

the site of intended edit is deleted in the read). The HDR category represents all reads with PE activity, i.e. pure PE + impure PE = total PE. To discern pure PE from impure PE, we defined an editing window ranging from at least one bp before the putative PE2 nicking position to at least four bp after the end of the RT template. For Cas9 and ssODN-mediated HDR the editing window was set to at least four bp up-/ and downstream of the sequence homologous to the ssODN. If in addition to the intended edit, any other mutation was found in this window, the read was classified as impure PE, otherwise as pure PE. Reads with additional mutations but without the intended edit were reported by CRISPResso2 either as ambiguous (if the location of the intended edit was deleted) or as NHEJ (if the location of the intended edit was intact but an edit was observed in the editing window). The reads of both groups ("ambiguous" and "NHEJ") were interpreted as undesired PE byproducts. CRISPResso HDR was run with quality filtering (only reads with an average Phred score >= 30 were considered).

Generation of allele plots

The replicate with the highest PPE rate was chosen to be presented as an allele plot. The alleles and their read percentage were obtained from the 'Alleles_frequency_table' in the CRISPResso output. The allele frequencies are reported for the full-length amplicons. For pegRNA-encoded substitutions, the top five alleles of both categories are presented. In the case of pegRNA-encoded deletions and insertions that were more difficult to discern from indels, the top 10 alleles of the union of IPE and byproduct alleles are presented. A window of 15 bp upstream to 50 bp downstream from the pegRNA nicking site was chosen, except for the human cell line plots which had an extended window that covered the nicking gRNA target.

PE off-target analysis in zebrafish.

The zebrafish reference genome (GRCz11) was computationally searched for genomic sites that closely-matched the target sites of five pegRNAs used in this study using Cas-OFFinder[19]. All closely-matched sites with up to three mismatches on canonical chromosomes were considered and subsequently amplified by PCR with primers specified in Supplementary Table 12. Indel frequencies were calculated (Supplementary Table 5) using CRISPResso2 with the same editing window as described above. Two off-target sites failed to amplify consistently, and one site was excluded due to high background in the control samples. Significance of editing was determined using Fisher's exact test and Benjamini-Hochberg correction for multiple comparison (Statistical Analysis).

Data Availability

Deep sequencing data will be deposited in the NCBI Sequence Read Archive (project number: PRJNA713914). Uncropped gel images of Fig. 1b (left panel) and Supplementary Fig. 2e can be found in the supplementary materials.

Statistical analysis

For all the bar graphs, mean and s.e.m. (only for samples with n>2) were calculated and plotted using GraphPad Prism8. For Fig. 1g, median and confidence interval

(95%) were calculated and plotted using GraphPad Prism8. For Supplementary Table 5, statistical analysis (significance level = 0.01) with two-sided Fisher's exact test was performed in python using the scipy.stats.fisher_exact function. Benjamini-Hochberg correction for multiple comparison was performed in python with the function statsmodels.stats.multitest.multipletests (with the argument "method='fdr_bh'").

Life sciences reporting summary

Further information about research design and statistics can be found in the Nature Research Reporting Summary that is appended to this article.

Code availability

The authors will make all custom computer code that was used in this work available upon request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by the Hassenfeld Scholar Award (to J.-R.J.Y.), NIH R01 GM134069 (to J.-R.J.Y.), NIH RM1 HG009490 (to J.K.J. and L.P.), NIH R35 GM118158 (to J.K.J.), the Defense Advanced Research Projects Agency (DARPA) Safe Genes Program grant no. HR0011-17-2-0042 (to J.K.J.), and National Human Genome Research Institute (NHGRI) Genomic Innovator Award R35 HG010717 (to L.P.). K.P. was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 417577129. J.M. received support from the China Scholarship Council (CSC, 201808210354). A.S. received support from a John Hansen Research Grant by the DKMS (DKMS-SLS-JHRG-2020-04). We thank K.K. Lam for technical assistance and K.K. Lam and J. Grünewald for discussions and technical advice. We thank L. Paul-Pottenplackel for help with revising the manuscript.

References

1. Anzalone AV et al. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature 576, 149–157, doi:10.1038/s41586-019-1711-4 (2019). [PubMed: 31634902]
2. Anzalone AV, Koblan LW & Liu DR Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. Nat Biotechnol 38, 824–844, doi:10.1038/s41587-020-0561-9 (2020). [PubMed: 32572269]
3. Lin Q et al. Prime genome editing in rice and wheat. Nat Biotechnol 38, 582–585, doi:10.1038/s41587-020-0455-x (2020). [PubMed: 32393904]
4. Liu Y et al. Efficient generation of mouse models with the prime editing system. Cell Discov 6, 27, doi:10.1038/s41421-020-0165-z (2020).
5. Surun D et al. Efficient Generation and Correction of Mutations in Human iPS Cells Utilizing mRNAs of CRISPR Base Editors and Prime Editors. Genes (Basel) 11, doi:10.3390/genes11050511 (2020).
6. Bosch JA, Birchak G & Perrimon N Precise genome engineering in Drosophila using prime editing. Proc Natl Acad Sci U S A 118, e2021996118, doi:https://doi-org.ezp-prod1.hul.harvard.edu/10.1073/pnas.2021996118 (2021). [PubMed: 33443210]
7. Schene IF et al. Prime editing for functional repair in patient-derived disease models. Nat Commun 11, 5352, doi:10.1038/s41467-020-19136-7 (2020). [PubMed: 33097693]
8. Hwang WY et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 31, 227–229, doi:10.1038/nbt.2501 (2013). [PubMed: 23360964]

9. Chen B et al. Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. Cell 155, 1479–1491, doi:10.1016/j.cell.2013.12.001S0092-8674(13)01531-6 [pii] (2013). [PubMed: 24360272]

10. Gronskov K, Ek J & Brondum-Nielsen K Oculocutaneous albinism. Orphanet J Rare Dis 2, 43, doi:10.1186/1750-1172-2-43 (2007). [PubMed: 17980020]

11. Zhang Y et al. Programmable base editing of zebrafish genome using a modified CRISPR-Cas9 system. Nat Commun 8, 118, doi:10.1038/s41467-017-00175-6 (2017). [PubMed: 28740134]

12. Munoz-Maldonado C, Zimmer Y & Medova M A Comparative Analysis of Individual RAS Mutations in Cancer Biology. Front Oncol 9, 1088, doi:10.3389/fonc.2019.01088 (2019). [PubMed: 31681616]

13. Jao LE, Wente SR & Chen W Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc Natl Acad Sci U S A 110, 13904–13909, doi:10.1073/pnas.1308335110 (2013). [PubMed: 23918387]

14. Levy JM et al. Cytosine and adenine base editing of the brain, liver, retina, heart and skeletal muscle of mice via adeno-associated viruses. Nat Biomed Eng 4, 97–110, doi:10.1038/s41551-019-0501-5 (2020). [PubMed: 31937940]

15. Kim S, Kim D, Cho SW, Kim J & Kim JS Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res 24, 1012–1019, doi:10.1101/gr.171322.113gr.171322.113 [pii] (2014). [PubMed: 24696461]

16. Gagnon JA et al. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. PLoS One 9, e98186, doi:10.1371/journal.pone.0098186 (2014). [PubMed: 24873830]

References

17. Clement K et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat Biotechnol 37, 224–226, doi:10.1038/s41587-019-0032-3 (2019). [PubMed: 30809026]

18. Davidson WF & Parish CR A procedure for removing red cells and dead cells from lymphoid cell suspensions. J Immunol Methods 7, 291–300 (1975). [PubMed: 167077]

19. Bae S, Park J & Kim JS Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 30, 1473–1475, doi:10.1093/bioinformatics/btu048 (2014). [PubMed: 24463181]

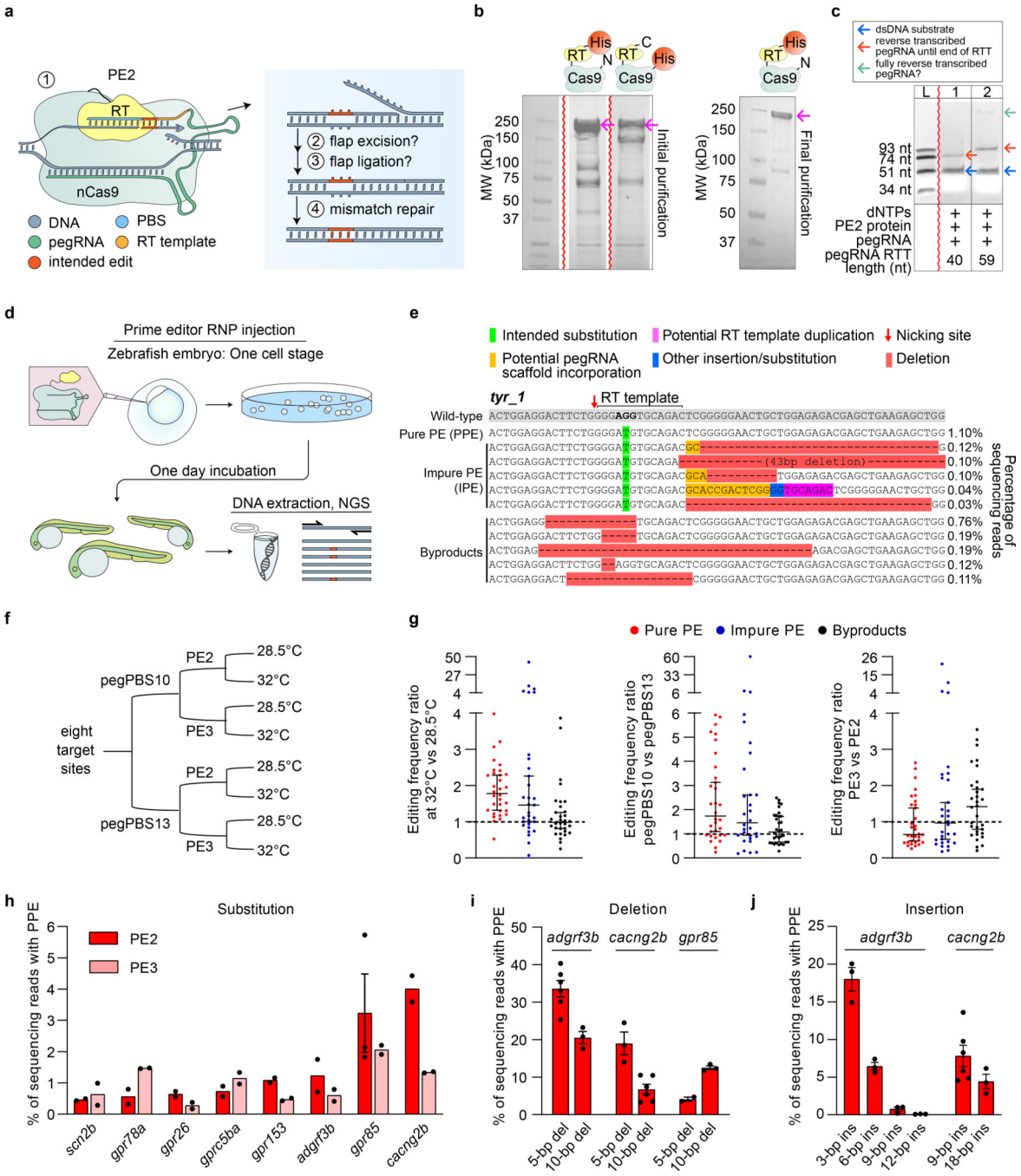

Fig. 1 |. Ribonucleoprotein-mediated prime editing in zebrafish.

a, Schematic illustrating the proposed PE mechanism[1]. **b**, TGX Stain-free™ gel image showing purification products of N- and C-terminally His-tagged PE2 protein. Magenta arrows indicate the desired protein product. Initial (left panel) and final (right panel) purification (representation of two experiments) are shown. Full-length PE2-His was isolated with higher yield and purity. **c**, Urea-polyacrylamide gel showing *in vitro* editing of a dsDNA substrate by the PE2-His protein (lanes 1 and 2; representation of two experiments). Unnicked and unextended oligonucleotide migrates at 51 nts (blue arrows).

RT-extended oligonucleotide migrates at 74 or 93 nts (red arrows). Larger extension product (green arrow) potentially represents RT-mediated polymerization of the full pegRNA scaffold (see Supplementary Fig. 2d–e). Red zigzag line, gel cropping boundary (see full gel in Supplementary Fig. 2c). L, ssDNA size ladder. **d**, Workflow for RNP injections into zebrafish embryos for prime editing. **e**, Prime editing outcomes at the *tyr*_1 site using pegRNA with 12-nt PBS, 13-nt RTT and C9E scaffold, PE3 strategy, and incubation at 32 °C. Protospacer adjacent motif (PAM) sequence is shown in bold. Red arrow, prime editor nicking location. The top five alleles with the highest frequencies in each category are shown. 10 embryos were pooled for the experiment. **f**, Experimental conditions tested at eight zebrafish PE target sites. Specified temperatures denote embryo incubation temperatures. pegPBS10, pegRNA with 10-nt PBS; pegPBS13, pegRNA with 13-nt PBS. **g**, Ratios of pure PE, impure PE and byproduct frequencies between 32 ºC and 28.5 ºC (left panel), between pegPBS10 and pegPBS13 (middle panel) and between PE3 and PE2 (right panel). Dots represent ratios calculated using the results from eight target sites (Supplementary Fig. 5). The bars represent the median and confidence interval (95%). n=32 different comparisons. **h**, Pure PE frequencies for eight target sites in zebrafish using pegPBS10 to induce +5 G→C or +5 G→T edits at 32 °C. **i,** PE2 for specified 5-bp and 10-bp deletions (del). **j,** PE2 for varying lengths of insertions (ins). (**h-j**) Dots, individual data points (n=2–6 biologically independent replicates, 5–10 embryos per replicate); Bars, mean; Error bars, ± s.e.m. (only for n >2).

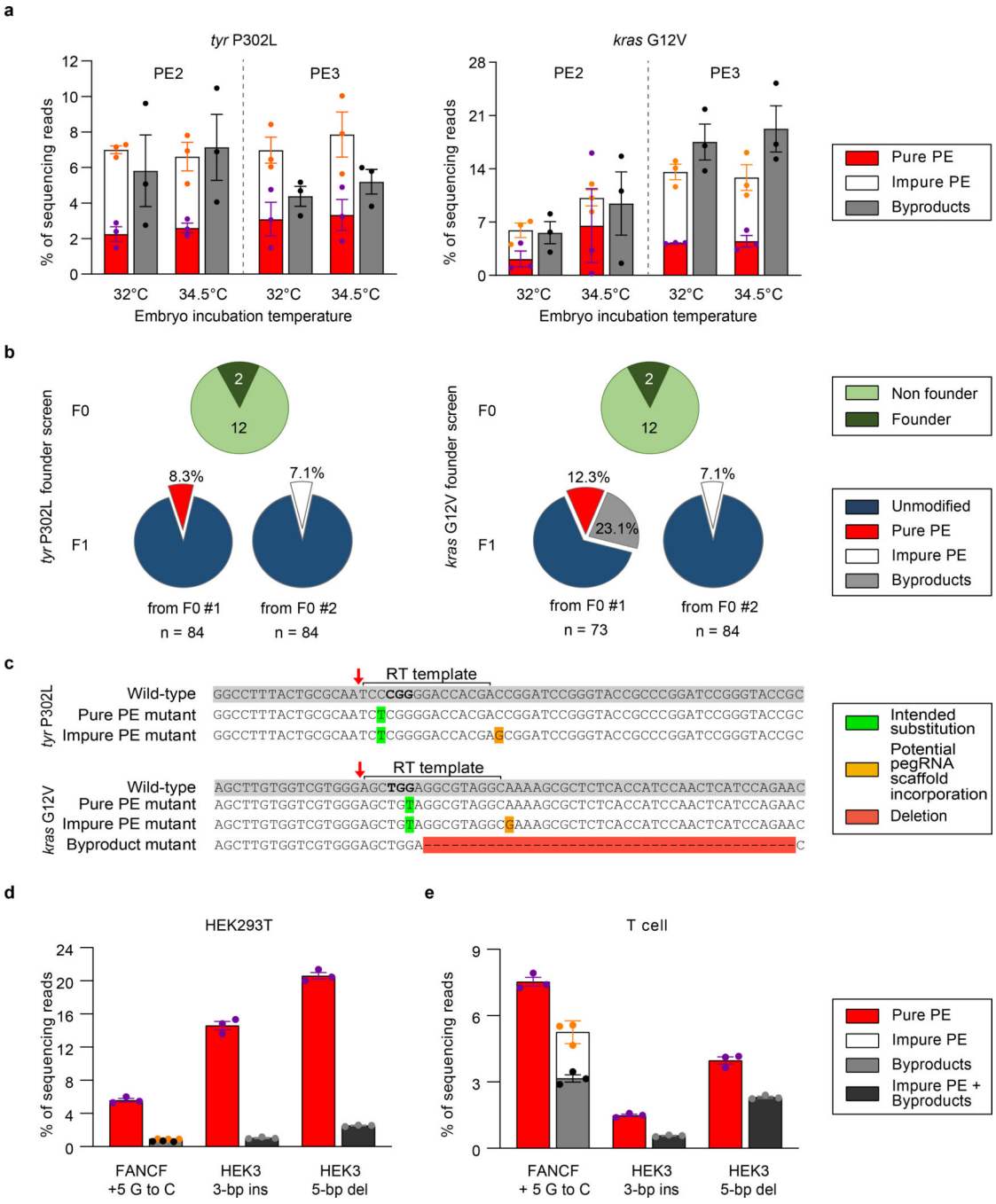

Figure 2 |. Germline transmission of prime editor-mediated edits corresponding to human pathogenic variants in zebrafish and RNP-mediated prime editing in human cells.

a, Installation of the *tyr* P302L mutation (corresponding to human *TYR* P301L) and the *kras* G12V mutation (corresponding to human *KRAS* G12V) in zebrafish using PE2 and PE3 and varying embryo incubation temperatures as noted. The molar ratio of pegRNA:ngRNA in PE3 was 10:1. Dots represent individual data points (PPE, purple; IPE, orange; byproducts, black). Intended edits were +3 C→T and +6 G→T for *tyr* P302L and *kras* G12V, respectively (Intended edits are annotated as: edit position (counting from

the pegRNA-mediated nick), reference base to edited base.). Bars and error bars represent the mean ± s.e.m. of 3 biologically independent replicates (n=3), where each replicate was a pool of 5–10 embryos. Mean frequencies of impure PE are shown above that of pure PE. **b**, Top panels show founder (F0) screens for *tyr* P302L (left) and *kras* G12V (right). Bottom panels show the fraction of F1 offspring carrying PE induced mutations. **c**, Wild-type and mutant sequences transmitted through germline by the *tyr* P302L (top) and *kras* G12V (bottom) founders. Protospacer adjacent motif (PAM) sequence is shown in bold. RT, reverse transcriptase. **d-e**, RNP-mediated PE3 editing efficiencies in HEK293T cells (**d**) and human primary T cells (**e**). Synthetic pegRNAs/ngRNAs with stabilizing 5' and 3' modifications were used. Dots represent individual data points (PPE, purple; IPE, orange; byproducts, black; IPE + byproducts, gray) and bars and error bars represent mean ± s.e.m. (n=3 biologically independent replicates). For prime editor-specified substitution at the FANCF site, impure PE data points are shown above byproduct frequencies. For prime editor-specified insertion and deletion at the HEK3 site, impure PE and byproducts are combined because they cannot always be distinguished (Supplementary Fig. 13).