PEAR: a flexible fluorescent reporter for the identification and enrichment of successfully prime edited cells

Dorottya A. Simon[1,2,†], András Tálas[1,3,†,*], Péter I. Kulcsár[1,4,5], Ervin Welker[1,5,*]

[1]Institute of Enzymology, Research Centre for Natural Sciences, Budapest, Hungary; [2]ProteoScientia Ltd., Hungary; [3]School of Ph.D. Studies, Semmelweis University, Budapest, Hungary; [4]Biospiral-2006 Ltd., Szeged, Hungary; [5]Institute of Biochemistry, Biological Research Centre, Szeged, Hungary

*Corresponding author. E-mail:talas.andras@ttk.hu, welker.ervin@ttk.hu

[†]These authors contributed equally.

ABSTRACT

Prime editing is a recently developed gene engineering tool that allows the introduction of short insertions, deletions or substitutions into the genome. However, the efficiency of prime editing, generally reaching around 10-30% editing, has not resembled its versatility. Here, Prime Editor Activity Reporter (PEAR), a sensitive fluorescent tool is introduced for the identification of single cells with prime editing activity. Possessing no background fluorescence, PEAR specifically reports on prime editing events in individual cells. By design, it ensures unrestricted flexibility for sequence variations in the full length of the target sequence. The application of PEAR as an enrichment marker of prime editing can increase the edited population by up to 70% and alleviate the burden of the otherwise time and labour consuming process of cloning of the correctly edited cells, therefore considerably improving the applicability of prime editing in fundamental research and biotechnological uses.

Introduction

Discovery of CRISPR systems in bacteria and archaea has not only drastically increased the spectrum of organisms that we can genetically modify but it has also equipped us with a simple tool for the introduction of a variety of different types of modifications into the genome [1-3]. Many of the CRISPR-based approaches rely on generating double strand breaks that is usually accompanied by genome-wide off-target editing [4-7]. Recently, more precise CRISPR tools; base [8,9] and prime [10] editors have been developed, that can introduce modifications into the DNA in a base pair resolution without the requirement of donor DNA templates or the introduction of double-strand DNA breaks. While current base editor (BE) variants provide efficient editing, they are restricted to certain substitution mutations [8,9,11,12]. In contrast, prime editors (PEs) can introduce all types of substitutions and/or precise indels, however, their efficiency lags behind of BEs'. PEs consist of a nickase version of *Streptococcus pyogenes* Cas9 (SpCas9) fused to a reverse transcriptase enzyme [10]. The fused reverse transcriptase can extend the nicked DNA strand using an RNA template that is located on the 3' terminus of an extended, single guide RNA (sgRNA), called the prime editing sgRNA (pegRNA). By careful design of the pegRNA, modifications can be introduced downstream of the nick generated on the non-targeted DNA strand. The 3' end of the pegRNA, that is complementary to the non-targeted strand, consists of the reverse transcriptase template (RT) and the primer binding site (PBS). The PBS is complementary to the 3' end of the nicked non-targeted DNA strand, with which it forms a hybrid DNA-RNA helix. The RT, which contains the mutation(s) to be introduced, also comprises the protospacer adjacent motif (PAM). The optimal length of the PBS and the RT varies from target to target and requires extensive optimisation for each different target. The efficiency of prime editing can be increased by nicking the non-edited strand (prime editor 3 – PE3) at the expense of generating unwanted indels at the targeted locus [10]. The effect of the nicking varies depending on its position and it requires further optimisation. In some cases, the efficiency of prime editing can be further increased by also altering the PAM sequence with the desired mutation (PE3b); this may prevent the SpCas9 from cleaving the sequence again. This method can reduce the chance of co-introducing unwanted indels with the intended mutations, thus increasing the occurrence of precise modifications [10].

The development of improved PE variants can be forecasted based on the history of base editors that are now available with highly improved features and efficiency [13-19]. The aim of our study was to develop a reporter system for an easy, fluorescence-based detection of prime editing outcomes, that allows maximum flexibility for the target sequences and pegRNA designs to be tested on it. Several systems have been developed for reporting base editing that use BEs for the generation or alteration of a start or stop codon[20,21], or to rescue a disruptive amino acid and subsequently recover functions of antibiotic resistance genes or fluorescence proteins [22]. Alternatively, a non-synonymous mutation in the chromophore of a fluorescent protein that induces fluorescence spectral change has also been explored as a method to monitor base editing activity [23,24]. Reporter systems have also been demonstrated for prime editing, but they are either restricted to few target sequences [25,26] and/or showed a rather low signal [20].

We also wanted to find out whether prime edited mammalian cells could be identified and enriched using a plasmid-based surrogate marker for chromosomal DNA modifications, as demonstrated before with Cas nucleases [27,28] and base editors [20,23,29,30].

Results

We aimed to develop a reporter system, that possesses several key features. It should be a transient plasmid-based system that is not restricted to one or a few cell lines nor does it require extensive work, such as the generation of cell lines. It needs to be based on a gain-of-function fluorescent signal with minimal background to be capable of detecting the signal in the timeframe of a transient system. The sequence requirements of efficient prime editing are not yet fully understood [10,31]; thus, it is crucial for the widespread application of a reporter system, that the sequence of the target and the flanking nucleotides of the position to be edited can be freely interchanged. We have recently developed a base editor activity reporter (BEAR), which meets these criteria [30], and it has the potential to be converted into a suitable tool for prime editing. BEAR is based on a split GFP protein separated by the last intron of the mouse *Vim* gene. The sequence of the functional 5' splice site is altered in such way, that splicing and therefore the GFP fluorescence is disrupted, however, they can be restored by applying base editors (**Figure S1a, b**). In this system the G-GT-RAGT sequence, which ensures optimal splicing at the 5' splice site, can be altered by one nucleotide without a significant decrease in the splicing efficiency. A systematic investigation of the sequence requirements for efficient splicing has revealed, that the activity of several inactive splice variants could be recovered by restoring just one nucleotide in its sequence, using either adenine base editors (ABEs) or in other cases, cytosine base editors (CBEs). Using the information acquired during the development of the BEAR system [30] and exploiting the ability of prime editing to alter more than one nucleotide simultaneously, we have designed a plasmid that contains an inactive splice site, which can be activated by the action of a prime editor harbouring an appropriately designed pegRNA (**Figure 1a,b**). Cells containing plasmids with an active splice site sequence will then be able to efficiently express GFP, which can be quantified by flow cytometry. Hence, we name assays that exploit this design as Prime Editor Activity Reporters (PEARs).
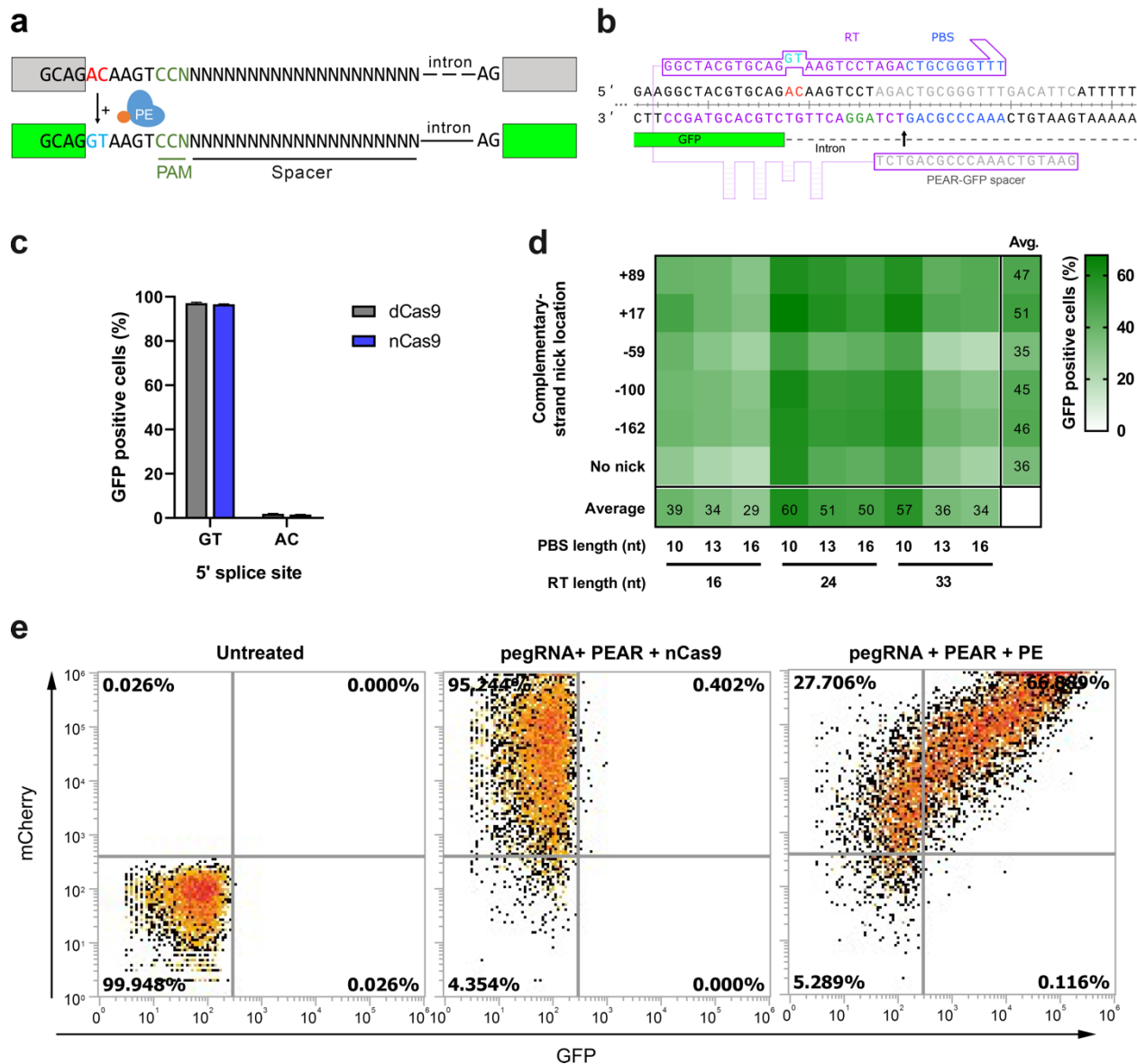
Figure 1 – Principle of the Prime Editor Activity Reporter (PEAR) assay

(**a**) Schematic of the Prime Editor Activity Reporter (PEAR). The mechanism of PEAR is based on the same concept as BEAR, and it contains the same inactive splice site, shown in panel (a). PE can revert the 'G-AC-AAGT' sequence to the canonical 'G-GT-AAGT' splice site. Unlike BEAR, here prime editing occurs on the antisense strand of the DNA, hence, this method enables us to position the spacer sequence within the intron. Here, the entire length of the spacer is freely adjustable (shown as 'N'-s). The altered bases of the splice site are shown in red, and the edited bases are shown in blue. The PAM sequence is dark green, nCas9 is blue and the fused reverse transcriptase is orange.

(**b**) Detailed view of the 5' splice site and the surrounding sequences of the PEAR-GFP plasmid. The 3' end of the first exon of GFP is shown in green, the intron is shown as a dashed line. The spacer sequence in the pegRNA and the target sequence in the DNA are shown in grey, the PAM is green, and the inactive splice site is red. The RT (purple) and the PBS (blue) sequences in the pegRNA and the targeted sequences are also coloured. The Cas9 nick site is indicated by a black arrow.

(**c**) PEAR-GFP plasmids with an active or inactive 5' splice site. Flow cytometry measurements of GFP positive HEK293T cells co-transfected with a PEAR-GFP reporter plasmid harboring a pre-edited active

(GT) or an inactive (AC) 5' splice site and nuclease (nCas9, blue) or dead Cas9 (dCas9, grey) constructs as indicated in the figure.

(**d**) Optimisation of PBS, RT and complementary DNA strand nicking on the PEAR-GFP plasmid. The heatmap shows the average percentage of GFP positive cells of three replicates of transfections with the PEAR-GFP plasmid and prime editor, using different pegRNAs and secondary nick locations. The position of the second nick is given in relation to the first nick. Positive values indicate 3', negative values indicate 5' direction on the targeted DNA. When no second nick was introduced it is indicated as "no nick".

(**e**) Flow cytometry measurements of control and prime edited HEK293T cells using the PEAR-GFP plasmid. Density dot plots of the non-transfected (left), nickase Cas9 transfected (middle) and PE transfected (right) cells. The pegRNA plasmid also expresses mCherry for monitoring transfection. In the plots in the middle and on the right, a pegRNA with an RT length of 24 and PBS length of 10 was used, and the second nick was introduced in position +17.

The editing windows of base editors exploited for the BEAR system are located PAM distal, 5' of the nick (**Figure S1a,b**). PEs on the other hand can introduce mutations 3' of the nick, a region that contains the PAM sequence (**Figure 1a,b**). To ensure maximum flexibility for the SpCas9 target, we kept the target sequence in the intron by moving the PAM to the complementary strand. As a result, the target sequence can be freely adjusted in its entire length (**Figure 1a,b**). Thus, while BEAR could be used with a few millions of different target sequences, PEAR offers unrestricted sequence variations for the SpCas9 target. Additionally, PEAR preserves the flexibility of BEAR to interchange the targeted and the flanking nucleotides [30]. No other fluorescence-based reporter systems have demonstrated such degree of flexibility before [20,25,26].

HEK293T cells were transfected with either an inactive or a nickase SpCas9 (dSpCas9 or nSpCas9, respectively) along with plasmids containing either the inactive splice site sequence selected to be edited (PEAR-GFP), or a functioning one, which would result from the action of a PE. The latter plasmid harbouring the functional 5' splice site is hereafter referred to as pre-edited. Cells showed high fluorescence when dSpCas9 (as a negative control) was co-transfected with the pre-edited plasmid while the PEAR-GFP with disrupted 5' splice site resulted in no expression of the fluorescent protein (**Figure 1c**). When the two plasmids were co-transfected with the H840A nickase Cas9 (a component of the prime editor) it did not alter the fluorescence of cells compared to the negative control (**Figure 1c**). These results are in line with former observations that indels cannot, only substitution mutations in the inactive 5' splice site can restore fluorescence [30] and suggest that this system will likely be suitable to report specifically on the action of prime editors.

To see whether PE can indeed recover fluorescence using the PEAR-GFP plasmid, several pegRNAs combining various PBS- and RT-lengths and complementary nicking positions were tested. The heatmap in **Figure 1d** shows that the PEAR-GFP plasmid can in fact be efficiently edited under several conditions. The editing shown in **Figure 1e** applying ten nucleotide-long PBS (PBS-10) and the 24 nucleotide-long RT region (RT-24) gave the highest efficiency among the tested combinations. The sixth row of the heatmap in **Figure 1d** shows the editing efficiency when the complementary strand is not nicked. In line with the expectations, these values are generally smaller, than when the complementary strand is nicked [10]. The best complementary strand nicking site is position +17 (in relation to the first nick) in the case of most PBS and RT length combinations. Interestingly, in all conditions PBS-10 gives greater values than PBS-13 and PBS-16 when combined with any of the RTs. Comparing the RTs, the efficiency of the editing is in the order of RT-24>RT-33>RT-16 in most conditions (16 out of 18). These results are consistent with the concept that the effect of the length of PBS and RT on editing efficiency is primarily independent from one another [32]. It is also in line with the current optimisation practice, where first, the length of the PBS is optimised using a given RT length, and then the length of the RT is optimised using the PBS selected in the first step. Thus, these experiments support the hypothesis, that PE efficiency is governed by the same factors when using a PEAR plasmid, as demonstrated earlier on chromosomal targets [10,31].

In order to validate this idea, we compared prime editing efficiency on PEAR plasmids with two HEK293T cell lines (that were generated in ref. [30]); now we named them as HEK-PEAR-GFP and HEK-PEAR-mScarlet, that contain a chromosomal copy of a PEAR plasmid (PEAR-GFP-2) harbouring either an interrupted GFP or an mScarlet sequence, respectively. To facilitate the comparison, we exploited the corresponding plasmids (PEAR-GFP-2 and PEAR-mScarlet plasmids, respectively) harbouring the same sequence surrounding the inactive splice site in the cell lines. Since the sequences to be edited in the two cell lines were designed to be used with base editing that works 5', upstream of the PAM sequence, we needed to identify another PAM site that would allow the application of prime editing. **Figure 2 and Figure S2** shows the layout of PEAR-GFP-2 (**Figure 2a**) and PEAR-mScarlet (**Figure 2d, Figure S2a**) with the target and secondary nick positions clearly indicated. The two nucleotides of the inactive splice site to be corrected are in positions +16-17 in relation to the new PAM

in HEK-PEAR-GFP. In the HEK-PEAR-mScarlet cell line they are in positions +26-27 or +33-34 in relation to the two PAMs.
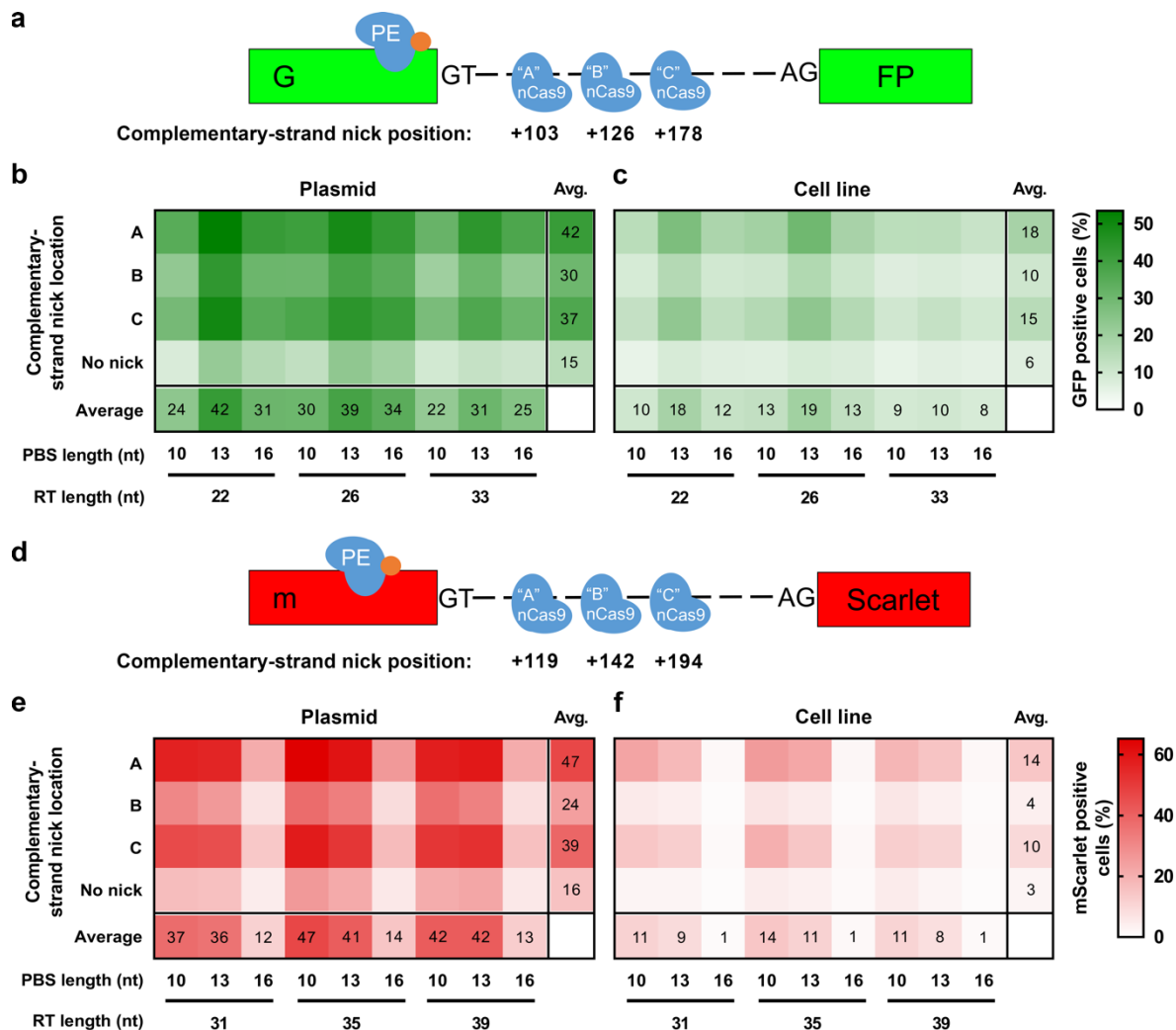

Figure 2 – Prime editing with PEAR in a genomic context

Schematic diagram of the PEAR-GFP-2 (**a**) and the PEAR-mScarlet (**d**) coding sequence is shown alongside the cell lines in panels (**c** and **f**), which contain the corresponding plasmid. The prime editor corrects the inactive splice site sequence and restores GFP or mScarlet fluorescence, respectively. The second nick positions that enhance prime editing are the same in (**a**) and (**d**), and are indicated as A, B or C.

The heatmaps show the average percentage of GFP (**b** and **c**) and mScarlet (**e** and **f**) positive cells derived from three replicates, where the prime editor was transfected into cells alongside different pegRNAs and sgRNAs with different complementary strand nick locations. In panel (**b**) PEAR-GFP-2 and in panel (**e**) PEAR-mScarlet plasmid is co-transfected. In panel (**c**) HEK-PEAR-GFP and in panel (**f**) HEK-PEAR-mScarlet is the transfected cell line. On the left side of the heatmaps the positions of the second nicks are indicated with A, B and C letters, corresponding to those in panel (**a**) and (**d**). When no second nick was introduced, it is indicated as "no nick".

In case of the GFP expressing cells, we tested 36 different conditions for both the cell line and the PEAR-GFP plasmid: all possible PBS – RT – second nick site combinations. The heatmap in **Figure 2b** shows that PBS-13 and RT-22 or RT-26 in combination with the second nick site at position 103 results in maximum efficiency for the PEAR-GFP-2 plasmid. In lack of the second nick, the efficiency of the editing is considerably lower. Comparing these results with the efficiency of prime editing in the cell line (**Figure 2c**), the most ideal conditions are the same. The overall pattern of the two heatmaps is indisputably similar, exhibiting strong correlation (r=0.89). In **Figure 2c** a somewhat lower efficiency is apparent in general compared to that apparent with the plasmid in **Figure 2b**. This difference is likely due to the much higher copies of the plasmids that are present in the cell during the experiments in **Figure 2b**.

In case of the mScarlet expressing cells, first, we compared the efficiencies of prime editing with target 1 (**Figure S2b)** and target 2 (**Figure S2c**), by exploring 36 different combinations of PBS, RT and second nick site for each target employing only the corresponding PEAR-mScarlet plasmid. Target 1 gave higher prime editing efficiencies and with both targets PBS 10 and 13 was optimal, but PBS 16 had a negative effect on prime editing efficiency.

With target 1 the top working PBS is 10 nucleotides with both the fluorescent cell line (**Figure 2f**) and the PEAR-mScarlet plasmid (**Figure 2e**), the differences between the various combinations are more apparent with the cell line. The two heatmaps exhibit a remarkably similar pattern and a strong correlation can be seen between the editing efficiency using the cell line and the PEAR plasmid (r=0.93). Similar results (r=0.92) can be achieved with target 2 when comparing plasmid (**Figure S2d**) and genomic (**Figure S2e**) results. Collectively, these results strongly imply that the main features of prime editing and factors affecting its efficiency are accurately reflected within our system.

Post-transfection selections of edited cells by FACS or antibiotics have proven to be effective methods for the enrichment of genetically modified cells within a population, provided that the Cas9 plasmid co-expresses a protein that is fluorescent, or it bears antibiotic resistance [20,24,27-30]. Experiments also showed that selecting for cells with a surrogate marker subjected to the same type of genetic modification (knockout, HDR or base editing), results in higher enrichment of successfully edited cells, than when solely selecting for transfection markers [20,23,27-30] suggesting that some cells allow for

higher editing efficiency. This phenomenon could be due to the nuclease being able to exhibit higher activity, or the activity of the DNA repair system could be considerably more engaged in these cells. We proposed that the level of enrichment, which can be achieved using the transfection marker alone, can be exceeded by enriching cells in which a marker protein in a plasmid is simultaneously prime edited. The PEAR-GFP plasmid was used in these experiments as a surrogate marker with previously optimised RT and PBS conditions (**Figure 1d**). To test this principle, first, we used the PEAR-mScarlet cell line and measured the ratio of the mScarlet positive cells in (1) all measured cells, (2) within cells that either express the blue fluorescent protein (BFP) for transfection enrichment or (3) express both BFP and GFP for PEAR enrichment. For this purpose, we chose two efficient previously tested PEAR-GFP targeting pegRNAs (**Figure 1d**). **Figure 3a** shows that transfection marker enrichments resulted in a significant increase in the ratio of edited cells to non-edited cells. Gating for the BFP-expressing cells increased the percentage of the edited population from 18% to 24% with both pegRNAs tested. PEAR-enrichment resulted in a considerable increase compared to the transfection enrichments in case of both pegRNAs (57 and 76% mScarlet expressing cells, respectively). These results demonstrate that prime edited cells can be substantially enriched when using PEAR plasmid as a surrogate marker.
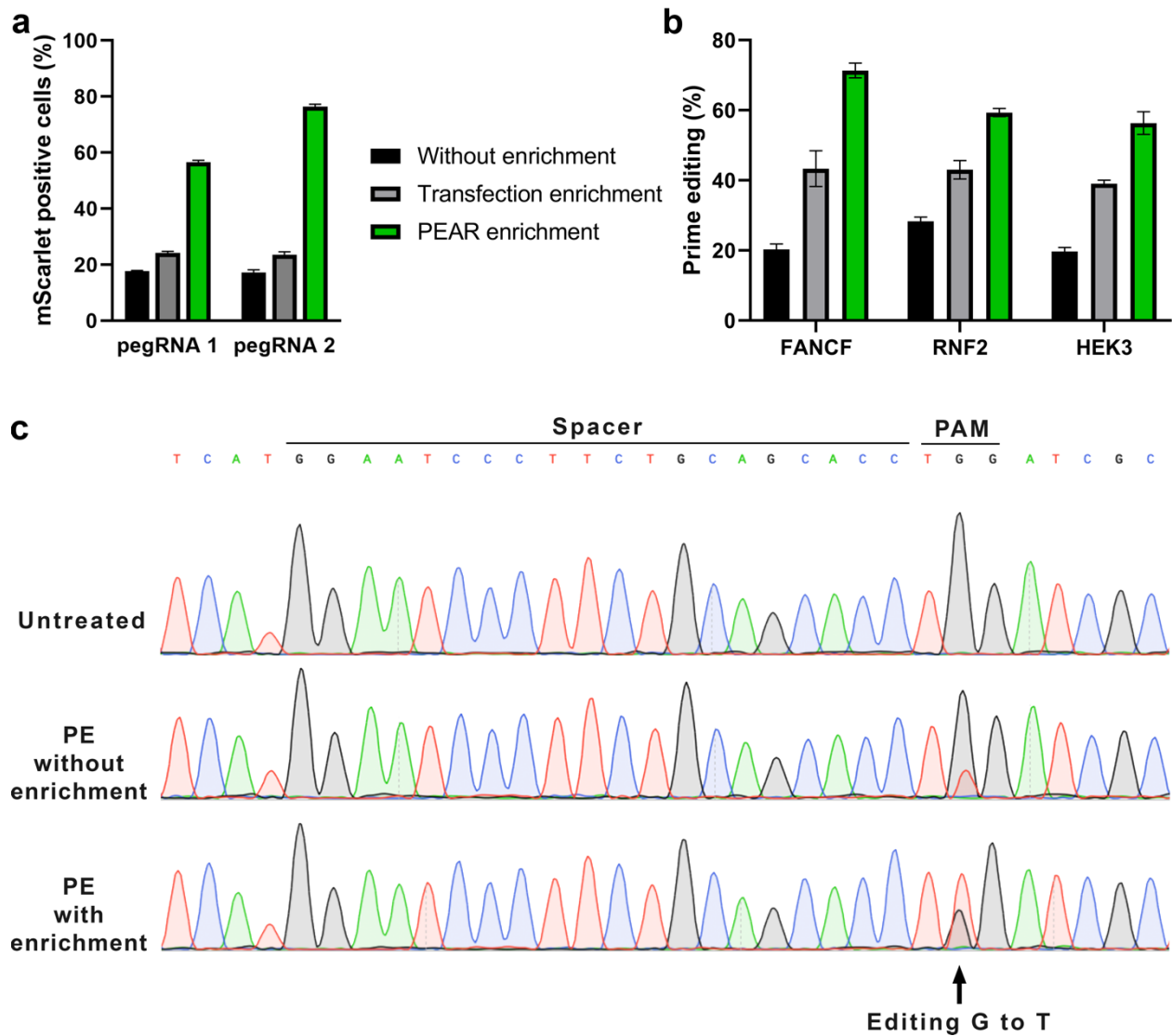
Figure 3 – PEAR enriches for prime edited cells

(**a**) Flow cytometry measurements of prime edited (i.e., mScarlet positive) cells in the HEK-PEAR-mScarlet cell line. The PEAR-GFP plasmid, prime editor and two pegRNAs targeting either the plasmid or the genomic PEAR sequences were co-transfected with sgRNAs (for complementary strand nicking). Two efficient pegRNAs were selected from a previous experiment (**Figure 1d**) to edit the PEAR-GFP plasmid (pegRNA 1 had an RT of 24 and a PBS of 10, pegRNA 2 had an RT of 33 and a PBS of 10). The mScarlet positive cell count was determined from cells that were gated for all live single cells (black bars), for BFP positive cells (transfection enrichment, gray bars), and for both BFP and GFP positive cells (PEAR enrichment, green bars).

(**b**) FACS enrichment of cells bearing prime edited targets. EditR and Sanger sequencing were used to assess genome editing efficiency. Percentages of the desired base editing outcome is shown. The PEAR-GFP plasmid and endogenous genomic targets (FANCF, RNF2 and HEK3) were simultaneously edited by PE. Edited cells were sorted to 3 fractions: all cells (without enrichment, black), mCherry positive cells (transfection enrichment, grey), and both mCherry and GFP positive cells (PEAR enrichment – green). Columns represent means +/- SD of three replicates.

(**c**) Chromatograms of Sanger sequencing of the FANCF genomic site are shown with or without prime editing and/or PEAR enrichment. The expected G to T modification's position 5 nucleotides from the nick is indicated by an arrow.

To test the applicability of PEAR on enriching the prime edited population in endogenous targets, we chose three genomic targets (FANCF, RNF2, HEK3) for which the optimal pegRNAs are already known [10], and using these targets, editing is expected to be readily detectable with Sanger sequencing and EditR, even without enrichment (**Figure 3b, c**). Here, mCherry was used as the indicator of transfection efficiency, and the PEAR-GFP plasmid was employed here too as a surrogate marker. On the third day from transfection, cells were sorted into three fractions: (1) all single, living cells; (2) mCherry positive cells, and (3) mCherry and GFP positive cells. After genomic DNA purification from the populations and PCR amplification of the targeted amplicon, the editing percentage was determined using EditR: G to T, C to A and T to A editing were detected at the FANCF, RNF2 and HEK3 sites, respectively. The editing rates have increased using transfection enrichment (from 20% to 43%, 28% to 43% and 19% to 39%, respectively), however, they have all showed a greater increase when the PEAR enrichment was used (to 71%, 59% and 56%, respectively). These results demonstrate that PEAR can substantially enrich for prime edited cells in transfected populations, reaching about 70% prime-edited cells.

Discussion

Prime editing provides a great flexibility for genomic modifications without having to supply a donor DNA or risking a double strand DNA break [10]. This technology could revolutionise the rapidly expanding universe of genome editing. The prime editing technology has readily been translated from mammalian models [10,26,32-34] to plants [25,35-37] and *Drosophila* [38]. Despite its great potential, we have found surprisingly few studies that had demonstrated the practical use of prime editing in mammalian cells [26,34] which is a vast underestimation of the true flexibility, that this method promises. This might be due to the extensive optimisation it requires and the relatively low editing efficiency that can be achieved at a subset of targets. In this aspect, any approach that can substantially increase the editing efficiency is critical to aid its general use. Here, we have developed PEAR, a sensitive fluorescent assay for the monitoring of prime editing activity in mammalian cells, and we have also shown that prime editing occurs at much higher frequencies in those individual cells, where the PEAR plasmid is co-

edited. In fact, the enrichment with a transfection marker doubled, while PEAR enrichments tripled the number of edited cells in our experiments (**Figure 3a-c**). Thus, the application of PEAR as a surrogate marker could substantially contribute to the more widespread use of prime editing, by considerably increasing the number of potential targets, which may now be edited with appropriate efficiency.

PEAR is also a versatile tool that can be utilised for the characterisation of prime editors. We examined more than 250 conditions combining different sgRNAs, PBS- and RT-lengths and nick sites, and our results confirmed that the efficiency of prime editing to modify PEAR plasmids is governed by the same factors as prime editing on chromosomal targets (**Figure 2**). The tolerance of the 5' splice site for substitutions [30] makes the sequences of target region of the PEAR plasmid easily adjustable. This versatility allows the user to examine the effect of sequence features and other factors affecting the efficiency of prime editing in a systematic manner that cannot be explored otherwise using genomic targets. Our approach beyond its general use may also assist the development of more efficient prime editors or variants with enhanced characteristics in the future.

Owning to the relatively low efficiency of prime editing with many potential target sequences, PEAR is expected to become a widespread tool to be used in a wide variety of biomedical and biotechnological applications. PEAR may be an especially useful tool to be used with hard-to-transfect cells or where the efficiency of PEs is particularly compromised.

Materials and Methods

Materials

Restriction enzymes, T4 ligase, Dulbecco's modified Eagle medium (DMEM), fetal bovine serum, Turbofect and penicillin/streptomycin were purchased from Thermo Fischer Scientific Inc. DNA oligonucleotides and the GenElute HP Plasmid Miniprep and Midiprep kit used for plasmid purifications were acquired from Sigma-Aldrich. Q5 High-Fidelity DNA Polymerase, NEB5-alpha competent cells, HiFi Assembly Master Mix were purchased from New England Biolabs Inc.

Plasmid construction

The PEAR-GFP plasmid (pDAS12125_PEAR-GFP) was constructed by Esp3I digestion of the pAT9624-BEAR-cloning plasmid [30], followed by one-pot cloning of the linker oligonucleotides *(12125-L1 and –L2)*. The reaction included 2 units of Esp3I enzyme, 1.5 units of T4 DNA ligase, 1 mM DTT, 500 µM ATP, 50 ng vector and 5-5 µM of target-coding oligonucleotides. Components were mixed in with 1x Tango buffer, and the mixture was incubated at 37°C for 30 minutes before being transformed into NEB5-alpha competent cells. The pre-edited PEAR-GFP plasmid (pDAS12124_PEAR-GFP-preedited) was constructed by the one-pot cloning protocol above, using linker oligonucleotides *12124-L1 and –L2*. Previously used plasmids pAT9651-BEAR-GFP and pAT9752-BEAR-mScarlet [30] were renamed here, for consistent nomenclature, to PEAR-GFP-2 and PEAR-mScarlet, respectively, since these plasmids were used here for the evaluation of prime editing.

To monitor transfection efficiency, fluorescent protein (mCherry or TagBFP) expressing pegRNA cloning plasmids were constructed (pDAS12069-U6-pegRNA-mCherry, pDAS12070-U6-pegRNA-BFP). To construct both plasmids, an exchangeable cassette from a previously made plasmid was cloned into pAT9658 and pAT9679 to NdeI and EcoRI sites. Both pegRNA cloning plasmids bear a spacer cloning site (between BpiI sites) and a PBS-RT cloning site (between Esp3I sites).

For the construction of pegRNAs, a one-step pegRNA cloning protocol was described by Anzalone et al. [10], however, it resulted in frequent deletions and/or mutations in the constructs, therefore we also used alternative strategies. The most effective was a two-step cloning procedure: first, the spacer coding linkers were cloned into pDAS12069

or pDAS12070 plasmids between BpiI sites using 2 units of the BpiI enzyme, 1.5 units of T4 DNA ligase, 500 μM ATP, 1x Green buffer, 50 ng vector and 5-5 μM of each annealed oligonucleotides. In the second cloning step the PBS-RT bearing oligonucleotides were cloned into between Esp3I sites of the plasmids created in the first cloning step with the above method using 1 mM DTT and 1X Tango buffer. The RNF2 targeting pegRNA in Figure 3 has a C base in position 83.

For the second nicking of a plasmid or the genome, all sgRNA targets were cloned into pAT9658-sgRNA-mCherry or pAT9679-sgRNA-BFP plasmids between BpiI sites via one pot cloning, as described above.

For detailed plasmid (**Table S1**) and primer (**Table S2**) information see Supporting Information. The sequences of all plasmid constructs were confirmed by Sanger sequencing (Microsynth AG).

Plasmids acquired from the non-profit plasmid distribution service Addgene were the following: pCMV-PE2 (#132775, [10]), pAT9624-BEAR-cloning (#162986), pAT9651-BEAR-GFP (#162989), pAT9752-BEAR-mScarlet (#162991), pAT9658-sgRNA-mCherry (#162987), pAT9679-sgRNA-BFP (#162988) - [30], pX330-Flag-wtSpCas9-H840A (#80453), pX330-Flag-dSpCas9 (#92113).

The following plasmids developed in this study are available from Addgene: pDAS12125_PEAR-GFP (#00000), pDAS12124_PEAR-GFP-preedited (#00000), pDAS12069-U6-pegRNA-mCherry (#00000), pDAS12070-U6-pegRNA-BFP (#00000), pDAS12230_pegRNA-PEAR-GFP(10PBS-24RT)-mCherry (#00000), pDAS12137_sgRNA-PEAR-GFP-nick(+17)-mCherry (#00000).

Cell culturing and transfection

HEK293T (ATCC, CRL-1573) cells were grown at 37°C in a humidified atmosphere of 5% CO_2 in DMEM medium supplemented with 10% heat-inactivated fetal bovine serum with 100 units/mL penicillin and 100 μg/mL streptomycin.

HEK293T cells were seeded on 48-well plates a day before transfection, at a density of 5×10^4 cells/well. A total of 565 ng DNA was used: 55 ng of PEAR target plasmid, 153 ng of pegRNA-mCherry (or pegRNA-BFP in the case of PEAR-mScarlet), 49 ng of sgRNA-mCherry (or sgRNA-BFP in the case of PEAR-mScarlet), and 308 ng of PE2 coding plasmid. These were mixed with 1 μl Turbofect reagent diluted in 50 μl serum-

free DMEM, and the mixture was incubated for 30 minutes at RT before adding it to the cells. Each transfection was performed in three replicates. Cells were analysed by flow cytometry on day three, after the transfection.

The BEAR-GFP and BEAR-mScarlet cell lines were constructed as described earlier in ref. [30]. The previously used BEAR-GFP and BEAR-mScarlet cell lines were renamed here, for consistent nomenclature, to PEAR-GFP and PEAR-mScarlet cell lines, since these were used here for the evaluation of prime editing. These stable cell lines were transfected with a total of 565 ng DNA: 170 ng of pegRNA-mCherry (or pegRNA-BFP in the case of PEAR-mScarlet), 54 ng of sgRNA-mCherry (or sgRNA-BFP in the case of PEAR-mScarlet) and 340 ng of PE2 coding plasmid, that was all mixed with 1 µl Turbofect reagent diluted in 50 µl serum-free DMEM. The mixture then was incubated for 30 minutes at RT before adding it to the cells.

In the enrichment experiments, where fluorescence activated cell sorting was used, HEK293T cells cultured on T-25 flasks were seeded a day before transfection at a density of 1.3×10^6 cells/flask. 7924 ng total DNA: 1754 ng of PEAR-GFP plasmid, 1330-1330 ng of pegRNA-mCherry plasmids (one targeting the genome and one targeting the PEAR plasmid), 425-425 ng of sgRNA-mCherry plasmids (one targeting the genome and one targeting the PEAR plasmid) and 2660 ng of PE2 coding plasmid were mixed with 16 µl Turbofect reagent diluted in 800 µl serum-free DMEM. The mixture was then incubated for 30 minutes before it was added to the cells. Three days after transfection cells were trypsinised and sorted directly into genomic lysis buffer, which was followed by genomic DNA extraction.

In all experiments the pCMV-PE2 [10] plasmid was used as the prime editor expressing plasmid and pX330-Flag-wtSpCas9-H840A as the nSpCas9 and pX330-Flag-dSpCas9 as the dSpCas expressing plasmid. When no second nick was introduced a mock sgRNA expressing plasmid (pAT9922-mCherry or pAT9762-BFP) was used.

Flow cytometry and cell sorting

Flow cytometry analysis was carried out using an Attune NxT Acoustic Focusing Cytometer (Applied Biosystems by Life Technologies). In all experiments, a minimum of 10,000 viable single cells were acquired by gating based on the side and forward light-scatter parameters. BFP, GFP, mCherry and mScarlet signals were detected

using the 405 (for BFP), 488 (for GFP) and 561 nm (for mCherry and mScarlet) diode laser for excitation, and the 440/50 (BFP), 530/30 (GFP), 620/15 (mCherry) and 585/16 nm (mScarlet) filter for emission. Attune Cytometric Software v.2.1.0 was used for data analysis. To compare prime editing in the cell line or on a plasmid Pearson's correlation coefficient was used.

Cell sorting was carried out on a FACSAria III cell sorter (BD Biosciences). The live single cell fraction was acquired by gating based on the side and forward light-scatter parameters. GFP and mCherry signals were detected using the 488 and 561 nm diode laser for excitation and the 530/30 and 610/20 nm filter for emission, respectively. A minimum of 50,000 cells were sorted in all experiments.

Genomic DNA purification, genomic PCR and EditR analysis

Genomic DNA from FACS or other experiments was extracted according to the Puregene DNA Purification protocol (Gentra Systems Inc.). The purified genomic DNA was subjected to PCR analysis, conducted with Q5 polymerase and locus specific primers (see **Supplementary Table S2** in Supporting Information). PCR products were gel purified via NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) and were Sanger sequenced. Editing efficiencies were quantified by EditR 1.0.9 (https://moriaritylab.shinyapps.io/editr_v10/ [39].

Acknowledgements

Author contributions

A.T. and E.W. conceived and designed the experiments, interpreted the results and wrote the manuscript with input from all authors. D.A.S performed experiments and interpreted the results. A.T. contributed to molecular cloning and performed experiments on mammalian cells. P. I. K. contributed to molecular cloning and created the PEAR cell lines. E.W. supervised the research. All authors read and approved the final manuscript.

Competing interests

None declared.

References

1. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W. and Marraffini, L.A. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819-823.

2. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823-826.

3. Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F. and Jaenisch, R. (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *cell*, **153**, 910-918.

4. Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J. and Le, L.P. (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nature biotechnology*, **33**, 187-197.

5. Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H.R., Hwang, J., Kim, J.-I. and Kim, J.-S. (2015) Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nature methods*, **12**, 237-243.

6. Kleinstiver, B.P., Tsai, S.Q., Prew, M.S., Nguyen, N.T., Welch, M.M., Lopez, J.M., McCaw, Z.R., Aryee, M.J. and Joung, J.K. (2016) Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nature biotechnology*, **34**, 869-874.

7. Tsai, S.Q., Nguyen, N.T., Malagon-Lopez, J., Topkar, V.V., Aryee, M.J. and Joung, J.K. (2017) CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR–Cas9 nuclease off-targets. *Nature methods*, **14**, 607.

8. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A. and Liu, D.R. (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*, **533**, 420-424.

9. Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I. and Liu, D.R. (2017) Programmable base editing of A• T to G• C in genomic DNA without DNA cleavage. *Nature*, **551**, 464-471.

10. Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J., Wilson, C., Newby, G.A. and Raguram, A. (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, **576**, 149-157.

11. Kurt, I.C., Zhou, R., Iyer, S., Garcia, S.P., Miller, B.R., Langner, L.M., Grünewald, J. and Joung, J.K. (2020) CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nature Biotechnology*, 1-6.

12. Zhao, D., Li, J., Li, S., Xin, X., Hu, M., Price, M.A., Rosser, S.J., Bi, C. and Zhang, X. (2020) Glycosylase base editors enable C-to-A and C-to-G base changes. *Nature biotechnology*, 1-6.

13. Kim, Y.B., Komor, A.C., Levy, J.M., Packer, M.S., Zhao, K.T. and Liu, D.R. (2017) Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nature biotechnology*, **35**, 371-376.

14. Komor, A.C., Zhao, K.T., Packer, M.S., Gaudelli, N.M., Waterbury, A.L., Koblan, L.W., Kim, Y.B., Badran, A.H. and Liu, D.R. (2017) Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C: G-to-T: A base

editors with higher efficiency and product purity. *Science advances*, **3**, eaao4774.

15. Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maianti, J.P., Raguram, A. and Liu, D.R. (2018) Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nature biotechnology*, **36**, 843-846.

16. Huang, T.P., Zhao, K.T., Miller, S.M., Gaudelli, N.M., Oakes, B.L., Fellmann, C., Savage, D.F. and Liu, D.R. (2019) Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nature biotechnology*, **37**, 626-631.

17. Gehrke, J.M., Cervantes, O., Clement, M.K., Wu, Y., Zeng, J., Bauer, D.E., Pinello, L. and Joung, J.K. (2018) An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nature biotechnology*, **36**, 977.

18. Zafra, M.P., Schatoff, E.M., Katti, A., Foronda, M., Breinig, M., Schweitzer, A.Y., Simon, A., Han, T., Goswami, S. and Montgomery, E. (2018) Optimized base editors enable efficient editing in cells, organoids and mice. *Nature biotechnology*, **36**, 888-893.

19. Gaudelli, N.M., Lam, D.K., Rees, H.A., Solá-Esteves, N.M., Barrera, L.A., Born, D.A., Edwards, A., Gehrke, J.M., Lee, S.-J. and Liquori, A.J. (2020) Directed evolution of adenine base editors with increased activity and therapeutic application. *Nature biotechnology*, **38**, 892-900.

20. Katti, A., Foronda, M., Zimmerman, J., Diaz, B., Zafra, M.P., Goswami, S. and Dow, L.E. (2020) GO: a functional reporter system to identify and enrich base editing activity. *Nucleic acids research*, **48**, 2841-2852.

21. Wang, P., Xu, L., Gao, Y. and Han, R. (2020) BEON: A functional fluorescence reporter for quantification and enrichment of adenine base editing activity. *Molecular Therapy*.

22. Martin, A.S., Salamango, D., Serebrenik, A.A., Shaban, N.M., Brown, W.L. and Harris, R. (2019) A panel of eGFP reporters for single base editing by APOBEC-Cas9 editosome complexes. *Scientific reports*, **9**, 1-8.

23. Coelho, M.A., Li, S., Pane, L.S., Firth, M., Ciotta, G., Wrigley, J.D., Cuomo, M.E., Maresca, M. and Taylor, B.J. (2018) BE-FLARE: a fluorescent reporter of base editing activity reveals editing characteristics of APOBEC3A and APOBEC3B. *BMC biology*, **16**, 150.

24. Standage-Beier, K., Tekel, S.J., Brookhouser, N., Schwarz, G., Nguyen, T., Wang, X. and Brafman, D.A. (2019) A transient reporter for editing enrichment (TREE) in human cells. *Nucleic acids research*, **47**, e120-e120.

25. Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., Wang, Y., Anzalone, A.V., Raguram, A. and Doman, J.L. (2020) Prime genome editing in rice and wheat. *Nature biotechnology*, **38**, 582-585.

26. Sürün, D., Schneider, A., Mircetic, J., Neumann, K., Lansing, F., Paszkowski-Rogacz, M., Hänchen, V., Lee-Kirsch, M. and Buchholz, F. (2020) Efficient generation and correction of mutations in human iPS cells utilizing mRNAs of CRISPR base editors and prime editors. *Genes*, **11**, 511.

27. Ramakrishna, S., Cho, S.W., Kim, S., Song, M., Gopalappa, R., Kim, J.-S. and Kim, H. (2014) Surrogate reporter-based enrichment of cells containing RNA-guided Cas9 nuclease-induced mutations. *Nature communications*, **5**, 1-10.

28. Grav, L.M., Lee, J.S., Gerling, S., Kallehauge, T.B., Hansen, A.H., Kol, S., Lee, G.M., Pedersen, L.E. and Kildegaard, H.F. (2015) One-step generation of triple

knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. *Biotechnology journal*, **10**, 1446-1456.

29. St. Martin, A., Salamango, D., Serebrenik, A., Shaban, N., Brown, W.L., Donati, F., Munagala, U., Conticello, S.G. and Harris, R.S. (2018) A fluorescent reporter for quantification and enrichment of DNA editing by APOBEC–Cas9 or cleavage by Cas9 in living cells. *Nucleic acids research*, **46**, e84-e84.

30. Tálas, A., Simon, D., Kulcsár, P., Varga, É. and Welker, E. (2020) BEAR reveals that increased fidelity variants can successfully reduce the mismatch-tolerance of adenine but not cytosine base editors. *Research Square*, **10.21203/rs.3.rs-105846/v1**.

31. Kim, H.K., Yu, G., Park, J., Min, S., Lee, S., Yoon, S. and Kim, H.H. (2020) Predicting the efficiency of prime editing guide RNAs in human cells. *Nature Biotechnology*, 1-9.

32. Kim, D.Y., Moon, S.B., Ko, J.-H., Kim, Y.-S. and Kim, D. (2020) Unbiased investigation of specificities of prime editing systems in human cells. *Nucleic acids research*, **48**, 10576-10589.

33. Liu, Y., Li, X., He, S., Huang, S., Li, C., Chen, Y., Liu, Z., Huang, X. and Wang, X. (2020) Efficient generation of mouse models with the prime editing system. *Cell discovery*, **6**, 1-4.

34. Schene, I.F., Joore, I.P., Oka, R., Mokry, M., van Vugt, A.H., van Boxtel, R., van der Doef, H.P., van der Laan, L.J., Verstegen, M.M. and van Hasselt, P.M. (2020) Prime editing for functional repair in patient-derived disease models. *Nature communications*, **11**, 1-8.

35. Jiang, Y.-Y., Chai, Y.-P., Lu, M.-H., Han, X.-L., Lin, Q., Zhang, Y., Zhang, Q., Zhou, Y., Wang, X.-C. and Gao, C. (2020) Prime editing efficiently generates W542L and S621I double mutations in two ALS genes in maize. *Genome biology*, **21**, 1-10.

36. Lu, Y., Tian, Y., Shen, R., Yao, Q., Zhong, D., Zhang, X. and Zhu, J.K. (2020) Precise genome modification in tomato using an improved prime editing system. *Plant Biotechnology Journal*.

37. Xu, W., Zhang, C., Yang, Y., Zhao, S., Kang, G., He, X., Song, J. and Yang, J. (2020) Versatile nucleotides substitution in plant using an improved prime editing system. *Molecular plant*, **13**, 675-678.

38. Bosch, J.A., Birchak, G. and Perrimon, N. (2021) Precise genome engineering in Drosophila using prime editing. *Proceedings of the National Academy of Sciences*, **118**.

39. Kluesner, M.G., Nedveck, D.A., Lahr, W.S., Garbe, J.R., Abrahante, J.E., Webber, B.R. and Moriarity, B.S. (2018) EditR: a method to quantify base editing from Sanger sequencing. *The CRISPR journal*, **1**, 239-250.

Supporting information

for

PEAR: a flexible fluorescent reporter for the identification and enrichment of successfully prime edited cells

Dorottya A. Simon[†], András Tálas[†,*], Péter I. Kulcsár, Ervin Welker[*]

*Corresponding author. E-mail:talas.andras@ttk.hu, welker.ervin@ttk.hu

[†]These authors contributed equally.

Table of contents

Figure S1 – Principle of the Base Editor Activity Reporter (BEAR) assay

(**a**) Schematic representation of the Base Editor Activity Reporter (BEAR). BEAR consists of a split GFP coding sequence separated by an intron, of which the 5' splice donor site (G-**GT**-AAGT) is altered, resulting in an inactive splice site (dashed line) and a dysfunctional protein (grey). ABE converts the inactive splice site into a functional one. Here the 'G-**AC**-AAGT' inactive splice site is illustrated, that can be modified by ABE to 'G-**GC**-AAGT', which is a functional non-canonic splice site, and hence, restores GFP expression (green). In this assay, base editors act on the sense strand of the DNA. The altered bases of the splice site are shown in red, the edited base is shown in blue and the variable bases in the sequence of the spacer are shown as 'N'-s. The PAM sequence is dark green, nCas9 is blue and the fused tadA deaminase is purple.

(**b**) Detailed view of the 5' splice site and the surrounding sequences of the BEAR-GFP plasmid. The 3' end of the first exon of GFP is shown in green, the intron is shown as a dashed line. The spacer sequence and the target sequence are shown in grey, the PAM is green and the inactive splice site is red. The Cas9 nick site is indicated by a black arrow.
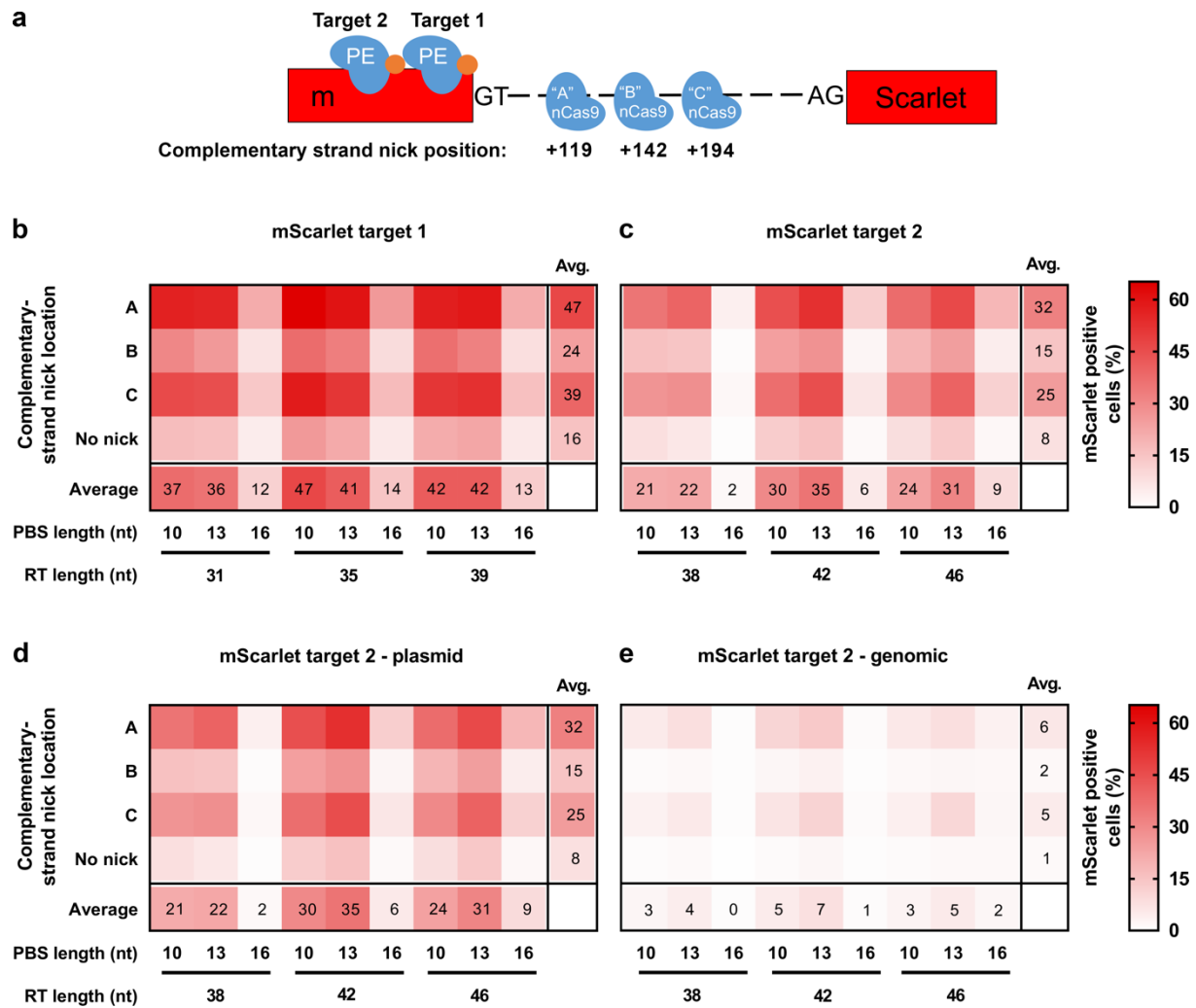
Figure S2 – Prime editing of the PEAR-mScarlet plasmid and cell line with two targets

(**a**) Schematic diagram of the PEAR-mScarlet plasmid. The prime editor corrects the inactive splice site and restores mScarlet fluorescence. The second nick positions that enhance prime editing are the same in (**b-e**) and indicated in the figures as A, B and C. The nick positions are given relative to the first nick of target 1, positive values indicate 3' direction on the targeted DNA. The spacer sequence is directed to either target1 (**b**) or target2 (**c-e**).

(**b-e**) The heatmaps show the average percentage of mScarlet positive cells derived from three replicates of the co-transfections of PE with different pegRNAs and nicking sgRNAs when targeting either the PEAR-mScarlet plasmid (**b-d**) or the PEAR-mScarlet cell line (**e**). The spacer sequence is directed to either target 1 (**b**) or target 2 (**c-e**). The position of the second nick is indicated on the left side of the heatmaps with A, B and C letters. When no second nick was introduced, it is indicated as "no nick".

Table S1 – List of plasmids used in this study

Plasmid name	Reference	Comment
General		
pCMV-PE2	Anzalone et al.	Addgene #132775
pX330-Flag-wtSpCas9-H840A	Unpublished	Welker Lab, Addgene #80453
pX330-Flag-dSpCas9	Kulcsar et al.	Addgene #92113
PEAR target plasmids		
pDAS12124_PEAR-GFP-1-preedited	This study	
pDAS12125_PEAR-GFP-1	This study	
PEAR-GFP-2	Tálas et al.	pAT9651-BEAR-GFP (Addgene #162989)
PEAR-mScaret	Tálas et al.	pAT9752-BEAR-mScarlet (Addgene #162991)
pAT9624-BEAR-cloning	Tálas et al.	Addgene #162986
pegRNA plasmids		
12067-U6-pegRNA-mCherry	This study	target cloning plasmid
12068-U6-pegRNA-BFP	This study	target cloning plasmid
12227_U6-pegRNA-PEAR-GFP_10PBS-16RT_mCherry	This study	targeting the PEAR-GFP plasmid
12228_U6-pegRNA-PEAR-GFP_13PBS-16RT_mCherry	This study	targeting the PEAR-GFP plasmid
12229_U6-pegRNA-PEAR-GFP_16PBS-16RT_mCherry	This study	targeting the PEAR-GFP plasmid
12230_U6-pegRNA-PEAR-GFP_10PBS-24RT_mCherry	This study	targeting the PEAR-GFP plasmid
12231_U6-pegRNA-PEAR-GFP_13PBS-24RT_mCherry	This study	targeting the PEAR-GFP plasmid
12232_U6-pegRNA-PEAR-GFP_16PBS-24RT_mCherry	This study	targeting the PEAR-GFP plasmid
12233_U6-pegRNA-PEAR-GFP_10PBS-33RT_mCherry	This study	targeting the PEAR-GFP plasmid
12234_U6-pegRNA-PEAR-GFP_13PBS-33RT_mCherry	This study	targeting the PEAR-GFP plasmid
12235_U6-pegRNA-PEAR-GFP_16PBS-33RT_mCherry	This study	targeting the PEAR-GFP plasmid
12248_U6-pegRNA-PEAR-GFP-2_10PBS-22RT_mCherry	This study	targeting the PEAR-GFP-2 plasmid and the PEAR-GFP cell line
12249_U6-pegRNA-PEAR-GFP-2_13PBS-22RT_mCherry	This study	targeting the PEAR-GFP-2 plasmid and the PEAR-GFP cell line
12250_U6-pegRNA-PEAR-GFP-2_16PBS-22RT_mCherry	This study	targeting the PEAR-GFP-2 plasmid and the PEAR-GFP cell line
12251_U6-pegRNA-PEAR-GFP-2_10PBS-26RT_mCherry	This study	targeting the PEAR-GFP-2 plasmid and the PEAR-GFP cell line
12252_U6-pegRNA-PEAR-GFP-2_13PBS-26RT_mCherry	This study	targeting the PEAR-GFP-2 plasmid and the PEAR-GFP cell line
12253_U6-pegRNA-PEAR-GFP-2_16PBS-26RT_mCherry	This study	targeting the PEAR-GFP-2 plasmid and the PEAR-GFP cell line
12254_U6-pegRNA-PEAR-GFP-2_10PBS-33RT_mCherry	This study	targeting the PEAR-GFP-2 plasmid and the PEAR-GFP cell line
12255_U6-pegRNA-PEAR-GFP-2_13PBS-33RT_mCherry	This study	targeting the PEAR-GFP-2 plasmid and the PEAR-GFP cell line
12256_U6-pegRNA-PEAR-GFP-2_16PBS-33RT_mCherry	This study	targeting the PEAR-GFP-2 plasmid and the PEAR-GFP cell line
12257_U6-pegRNA-PEAR-mScarlet-target1_10PBS-31RT_BFP	This study	targeting the PEAR-mScarlet - target1 plasmid and the PEAR-mScarlet cell line
12258_U6-pegRNA-PEAR-mScarlet-target1_13PBS-31RT_BFP	This study	targeting the PEAR-mScarlet - target1 plasmid and the PEAR-mScarlet cell line

12259_U6-pegRNA-PEAR-mScarlet-target1_16PBS-31RT_BFP	This study	targeting the PEAR-mScarlet - target1 plasmid and the PEAR-mScarlet cell line
12260_U6-pegRNA-PEAR-mScarlet-target1_10PBS-35RT_BFP	This study	targeting the PEAR-mScarlet - target1 plasmid and the PEAR-mScarlet cell line
12261_U6-pegRNA-PEAR-mScarlet-target1_13PBS-35RT_BFP	This study	targeting the PEAR-mScarlet - target1 plasmid and the PEAR-mScarlet cell line
12262_U6-pegRNA-PEAR-mScarlet-target1_16PBS-35RT_BFP	This study	targeting the PEAR-mScarlet - target1 plasmid and the PEAR-mScarlet cell line
12263_U6-pegRNA-PEAR-mScarlet-target1_10PBS-39RT_BFP	This study	targeting the PEAR-mScarlet - target1 plasmid and the PEAR-mScarlet cell line
12264_U6-pegRNA-PEAR-mScarlet-target1_13PBS-39RT_BFP	This study	targeting the PEAR-mScarlet - target1 plasmid and the PEAR-mScarlet cell line
12265_U6-pegRNA-PEAR-mScarlet-target1_16PBS-39RT_BFP	This study	targeting the PEAR-mScarlet - target1 plasmid and the PEAR-mScarlet cell line
12266_U6-pegRNA-PEAR-mScarlet-target2_10PBS-38RT_BFP	This study	targeting the PEAR-mScarlet - target2
12267_U6-pegRNA-PEAR-mScarlet-target2_13PBS-38RT_BFP	This study	targeting the PEAR-mScarlet - target2
12268_U6-pegRNA-PEAR-mScarlet-target2_16PBS-38RT_BFP	This study	targeting the PEAR-mScarlet - target2
12269_U6-pegRNA-PEAR-mScarlet-target2_10PBS-42RT_BFP	This study	targeting the PEAR-mScarlet - target2
12270_U6-pegRNA-PEAR-mScarlet-target2_13PBS-42RT_BFP	This study	targeting the PEAR-mScarlet - target2
12271_U6-pegRNA-PEAR-mScarlet-target2_16PBS-42RT_BFP	This study	targeting the PEAR-mScarlet - target2
12272_U6-pegRNA-PEAR-mScarlet-target2_10PBS-46RT_BFP	This study	targeting the PEAR-mScarlet - target2
12273_U6-pegRNA-PEAR-mScarlet-target2_13PBS-46RT_BFP	This study	targeting the PEAR-mScarlet - target2
12274_U6-pegRNA-PEAR-mScarlet-target2_16PBS-46RT_BFP	This study	targeting the PEAR-mScarlet - target2
12239_U6-pegRNA-RNF2_mCherry	This study	Targeting RNF2 genomic target
12241_U6-pegRNA-FANCF_mCherry	This study	Targeting FANCF genomic target
12244_U6-pegRNA-HEK3_mCherry	This study	Targeting HEK3 genomic target
sgRNA plasmids		
12136_U6-sgRNA-PEAR-GFP_nick(+89)_mCherry	This study	Secondary nick on the PEAR-GFP plasmid
12137_U6-sgRNA-PEAR-GFP_nick(+17)_mCherry	This study	Secondary nick on the PEAR-GFP plasmid
12138_U6-sgRNA-PEAR-GFP_nick(-59)_mCherry	This study	Secondary nick on the PEAR-GFP plasmid
12139_U6-sgRNA-PEAR-GFP_nick(-100)_mCherry	This study	Secondary nick on the PEAR-GFP plasmid
12140_U6-sgRNA-PEAR-GFP_nick(-162)_mCherry	This study	Secondary nick on the PEAR-GFP plasmid
12198_U6-sgRNA-PEAR-GFP-2_nick(+103)_mCherry	This study	Secondary nick on the PEAR-GFP-2 plasmid, and in the PEAR-GFP cell line
12199_U6-sgRNA-PEAR-GFP-2_nick(+126)_mCherry	This study	Secondary nick on the PEAR-GFP-2 plasmid, and in the PEAR-GFP cell line
12200_U6-sgRNA-PEAR-GFP-2_nick(+ 17)_mCherry	This study	Secondary nick on the PEAR-GFP-2 plasmid, and in the PEAR-GFP cell line
12210_U6-sgRNA-PEAR-mScarlet_nick(+103)_BFP	This study	Secondary nick on the PEAR-mScarlet plasmid and in the PEAR-mScarlet cell line
12211_U6-sgRNA-PEAR-mScarlet_nick(+126)_BFP	This study	Secondary nick on the PEAR-mScarlet plasmid and in the PEAR-mScarlet cell line
12212_U6-sgRNA-PEAR-mScarlet_nick(+ 17)_BFP	This study	Secondary nick on the PEAR-mScarlet plasmid and in the PEAR-mScarlet cell line
12156_U6-sgRNA-RNF2_nick(+41)_mCherry	This study	Secondary nick in the genome
12177_U6-sgRNA-FANCF_nick(+48)_mCherry	This study	Secondary nick in the genome
12178_U6-sgRNA-HEK3_nick(+90)_mCherry	This study	Secondary nick in the genome
pAT9922_U6-sgRNA-mock-mCherry	This study	mock sgRNA used in when not nicking the plasmid/genome
9762_U6-sgRNA-mock-TagBFP	This study	mock sgRNA used in when not nicking the plasmid/genome

Table S2 – List of oligonucleotides used in this study

Oligonucleotide name	Oligonucleotide sequence
12124_L1	CGTGCAGGTAAGTCCTAGACTGCGGGTTTGACATTC
12124_L2	AAATGAATGTCAAACCCGCAGTCTAGGACTTACCTG
12125_L1	CGTGCAGACAAGTCCTAGACTGCGGGTTTGACATTC
12125_L2	AAATGAATGTCAAACCCGCAGTCTAGGACTTGTCTG
12067-for	ATCTTGTGGAAAGGACGAAACACCGGGTCTTCGTAGTCAAAAGCCTCCGG
12067-rev	CGGGCACCGGAGCCACTCGAATTCAAAAAAAAGGTCTTCTCCCTATCAGTG
pegRNA_gRNA-scaffold-L1	[Phos]AGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG
pegRNA_gRNA-scaffold-L2	[Phos]GCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAG
12227-12235-spacer-L1	CACCGAATGTCAAACCCGCAGTCT
12227-12235-spacer-L2	AAACAGACTGCGGGTTTGACATTC
12227_16RT_10PBS_for	GTGCGCAGGTAAGTCCTAGACTGCGGGTTT
12227_16RT_10PBS_rev	AAAATCTAGGACTTACCTGCAAACCCGCAG
12228_16RT_13PBS_for	GTGCGCAGGTAAGTCCTAGACTGCGGGTTTGAC
12228_16RT_13PBS_rev	AAAATCTAGGACTTACCTGCGTCAAACCCGCAG
12229_16RT_16PBS_for	GTGCGCAGGTAAGTCCTAGACTGCGGGTTTGACATT
12229_16RT_16PBS_rev	AAAATCTAGGACTTACCTGCAATGTCAAACCCGCAG
12230_24RT_10PBS_for	GTGCGGCTACGTGCAGGTAAGTCCTAGACTGCGGGTTT
12230_24RT_10PBS_rev	AAAATCTAGGACTTACCTGCACGTAGCCAAACCCGCAG
12231_24RT_13PBS_for	GTGCGGCTACGTGCAGGTAAGTCCTAGACTGCGGGTTTGAC
12231_24RT_13PBS_rev	AAAATCTAGGACTTACCTGCACGTAGCCGTCAAACCCGCAG
12232_24RT_16PBS_for	GTGCGGCTACGTGCAGGTAAGTCCTAGACTGCGGGTTTGACATT
12232_24RT_16PBS_rev	AAAATCTAGGACTTACCTGCACGTAGCCAATGTCAAACCCGCAG
12233_33RT_10PBS_for	GTGCATGCCCGAAGGCTACGTGCAGGTAAGTCCTAGACTGCGGGTTT
12233_33RT_10PBS_rev	AAAATCTAGGACTTACCTGCACGTAGCCTTCGGGCATAAACCCGCAG
12234_33RT_13PBS_for	GTGCATGCCCGAAGGCTACGTGCAGGTAAGTCCTAGACTGCGGGTTTGAC
12234_33RT_13PBS_rev	AAAATCTAGGACTTACCTGCACGTAGCCTTCGGGCATGTCAAACCCGCAG
12235_33RT_16PBS_for	GTGCATGCCCGAAGGCTACGTGCAGGTAAGTCCTAGACTGCGGGTTTGACATT
12235_33RT_16PBS_rev	AAAATCTAGGACTTACCTGCACGTAGCCTTCGGGCATAATGTCAAACCCGCAG
12248-12256-spacer-L1	CACCGTTCAAGTCCGCCATGCCCGA
12248-12256-spacer-L2	AAACTCGGGCATGGCGGACTTGAAC
12248_peg1	GTGCACTTACCTGCACGTAGCCTTCGTCAAGTCCGC
12248_peg2	AAAAGCGGACTTGACGAAGGCTACGTGCAGGTAAGT
12249_peg1	GTGCACTTACCTGCACGTAGCCTTCGTCAAGTCCGCCAT
12249_peg2	AAAAATGGCGGACTTGACGAAGGCTACGTGCAGGTAAGT
12250_peg1	GTGCACTTACCTGCACGTAGCCTTCGTCAAGTCCGCCATGCC
12250_peg2	AAAAGGCATGGCGGACTTGACGAAGGCTACGTGCAGGTAAGT
12251_peg1	GTGCATGCACTTACCTGCACGTAGCCTTCGTCAAGTCCGC

12251_peg2	AAAAGCGGACTTGACGAAGGCTACGTGCAGGTAAGTGCAT
12252_peg1	GTGCATGCACTTACCTGCACGTAGCCTTCGTCAAGTCCGCCAT
12252_peg2	AAAAATGGCGGACTTGACGAAGGCTACGTGCAGGTAAGTGCAT
12253_peg1	GTGCATGCACTTACCTGCACGTAGCCTTCGTCAAGTCCGCCATGCC
12253_peg2	AAAAGGCATGGCGGACTTGACGAAGGCTACGTGCAGGTAAGTGCAT
12254_peg1	GTGCGCAGTCTATGCACTTACCTGCACGTAGCCTTCGTCAAGTCCGC
12254_peg2	AAAAGCGGACTTGACGAAGGCTACGTGCAGGTAAGTGCATAGACTGC
12255_peg1	GTGCGCAGTCTATGCACTTACCTGCACGTAGCCTTCGTCAAGTCCGCCAT
12255_peg2	AAAAATGGCGGACTTGACGAAGGCTACGTGCAGGTAAGTGCATAGACTGC
12256_peg1	GTGCGCAGTCTATGCACTTACCTGCACGTAGCCTTCGTCAAGTCCGCCATGCC
12256_peg2	AAAAGGCATGGCGGACTTGACGAAGGCTACGTGCAGGTAAGTGCATAGACTGC
12257-12265-spacer-L1	CACCGTGATGAACTTCGAGGACGG
12257-12265-spacer-L2	AAACCCGTCCTCGAAGTTCATCAC
12257_peg1	GTGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGT
12257_peg2	AAAAACTTCGAGGACGGCGGCGCCCGTGACCGTGACGCAGGTAAGT
12258_peg1	GTGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCA
12258_peg2	AAAATGAACTTCGAGGACGGCGGCGCCGTGACCGTGACGCAGGTAAGT
12259_peg1	GTGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCATCA
12259_peg2	AAAATGATGAACTTCGAGGACGGCGGCGCCCGTGACCGTGACGCAGGTAAGT
12260_peg1	GTGCATGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGT
12260_peg2	AAAAACTTCGAGGACGGCGGCGCCCGTGACCGTGACGCAGGTAAGTGCAT
12261_peg1	GTGCATGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCA
12261_peg2	AAAATGAACTTCGAGGACGGCGGCGCCGTGACCGTGACGCAGGTAAGTGCAT
12262_peg1	GTGCATGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCATCA
12262_peg2	AAAATGATGAACTTCGAGGACGGCGGCGCCCGTGACCGTGACGCAGGTAAGTGCAT
12263_peg1	GTGCGTCTATGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGT
12263_peg2	AAAAACTTCGAGGACGGCGGCGCCCGTGACCGTGACGCAGGTAAGTGCATAGAC
12264_peg1	GTGCGTCTATGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCA
12264_peg2	AAAATGAACTTCGAGGACGGCGGCGCCGTGACCGTGACGCAGGTAAGTGCATAGAC
12265_peg1	GTGCGTCTATGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCATCA
12265_peg2	AAAATGATGAACTTCGAGGACGGCGGCGCCCGTGACCGTGACGCAGGTAAGTGCATAGAC
12266-12274-spacer-L1	CACCGGAGCGCGTGATGAACTTCG
12266-12274-spacer-L2	AAACCGAAGTTCATCACGCGCTCC
12266_peg1	GTGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCATCAC
12266_peg2	AAAAGTGATGAACTTCGAGGACGGCGGCGCCCGTGACCGTGACGCAGGTAAGT
12267_peg1	GTGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCATCACGCG
12267_peg2	AAAACGCGTGATGAACTTCGAGGACGGCGGCGCCCGTGACCGTGACGCAGGTAAGT
12268_peg1	GTGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCATCACGCGCTC
12268_peg2	AAAAGAGCGCGTGATGAACTTCGAGGACGGCGGCGCCCGTGACCGTGACGCAGGTAAGT
12269_peg1	GTGCATGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCATCAC
12269_peg2	AAAAGTGATGAACTTCGAGGACGGCGGCGCCCGTGACCGTGACGCAGGTAAGTGCAT
12270_peg1	GTGCATGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCATCACGCG
12270_peg2	AAAACGCGTGATGAACTTCGAGGACGGCGGCGCCCGTGACCGTGACGCAGGTAAGTGCAT

12271_peg1	GTGCATGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCATCACGCGCTC
12271_peg2	AAAAGAGCGCGTGATGAACTTCGAGGACGGCGGCGCCGTGACCGTGACGCAGGTAAGTGCAT
12272_peg1	GTGCGTCTATGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCATCAC
12272_peg2	AAAAGTGATGAACTTCGAGGACGGCGGCGCCGTGACCGTGACGCAGGTAAGTGCATAGAC
12273_peg1	GTGCGTCTATGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCATCACGCG
12273_peg2	AAAACGCGTGATGAACTTCGAGGACGGCGGCGCCGTGACCGTGACGCAGGTAAGTGCATAGAC
12274_peg1	GTGCGTCTATGCACTTACCTGCGTCACGGTCACGGCGCCGCCGTCCTCGAAGTTCATCACGCGCTC
12274_peg2	AAAAGAGCGCGTGATGAACTTCGAGGACGGCGGCGCCGTGACCGTGACGCAGGTAAGTGCATAGAC
12239_RNF2-spacer-L1	CACCGTCATCTTAGTCATTACCTG
12239_RNF2-spacer-L2	AAACCAGGTAATGACTAAGATGAC
12239_peg1	GTGCAACGAACACCTCATGTAATGACTAAGATG
12239_peg2	AAAACATCTTAGTCATTACATGAGGTGTTCGTT
12241_FANCF-spacer-L1	CACCGGAATCCCTTCTGCAGCACC
12241_FANCF-spacer-L2	AAACGGTGCTGCAGAAGGGATTCC
12241_peg1	GTGCGGAAAAGCGATCAAGGTGCTGCAGA
12241_peg2	AAAATCTGCAGCACCTTGATCGCTTTTCC
12244_HEK3-spacer-L1	CACCGGCCCAGACTGAGCACGTGA
12244_HEK3-spacer-L2	AAACTCACGTGCTCAGTCTGGGCC
12244_peg1	GTGCTCTGCCATCTCGTGCTCAGTCTG
12244_peg2	AAAACAGACTGAGCACGAGATGGCAGA
12136_L1	CACCGCTCGTGACCACCCTGACCTA
12136_L2	AAACTAGGTCAGGGTGGTCACGAGC
12137_L1	CACCGTTCAAGTCCGCCATGCCCGA
12137_L2	AAACTCGGGCATGGCGGACTTGAAC
12138_L1	CACCGATATGGGAAAGTTTAAAAGA
12138_L2	AAACTCTTTTAAACTTTCCCATATC
12139_L1	CACCGCAACCACTTTAAATATGTGT
12139_L2	AAACACACATATTTAAAGTGGTTGC
12140_L1	CACCGCTAACTTAAGGAGTCCCCG
12140_L2	AAACCGGGGACTCCTTAAGTTAGC
12198_L1	CACCGGACCAACACATATTTAAAG
12198_L2	AAACCTTTAAATATGTGTTGGTCC
12199_L1	CACCGATATCTGCTTGCTGACTTAA
12199_L2	AAACTTAAGTCAGCAAGCAGATATC
12200_L1	CACCGCTATTCATCCAGTATCCAC
12200_L2	AAACGTGGATACTGGATGAATAGC
12210_L1	same as: 12198_L1
12210_L2	same as: 12198_L2
12211_L1	same as: 12199_L1
12211_L2	same as: 12199_L2
12212_L1	same as: 12200_L1
12212_L2	same as: 12200_L2
12156_L1	CACCGTCAACCATTAAGCAAAACAT

12156_L2	AAACATGTTTTGCTTAATGGTTGAC
12177_L1	CACCGGGGTCCCAGGTGCTGACGT
12177_L2	AAACACGTCAGCACCTGGGACCCC
12178_L1	CACCGTCAACCAGTATCCCGGTGC
12178_L2	AAACGCACCGGGATACTGGTTGAC
9922_L1	CACCGCAGACAAGTAGGGTGGGCC
9922_L2	AAACGGCCCACCCTACTTGTCTGC
9762_L1	CACCGCAGTCAAGTGCTGGAGGTG
9762_L2	AAACCACCTCCAGCACTTGACTGC
Primers for genomic PCR	
HEK3_for	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATGTGGGCTGCCTAGAAAGG
HEK3_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCCAGCCAAACTTGTCAACC
RNF2_for	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACGTCTCATATGCCCCTTGG
RNF2_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACGTAGGAATTTTGGTGGGACA
FANCF2-for	GGGCCGGGAAAGAGTTGCTG
FANCF2-rev	GCCCTACATCTGCTCTCCCTCC