Title:

qTAG: An adaptable CRISPR-based endogenous tagging protocol using optimized repair cassettes.

Authors:

Reuben Philip[1,2,4], Amit Sharma[1,4], Laura Matellan[1], Anna C. Erpf[1], Wen-Hsin Hsu[1], Johnny M. Tkach[1], Haley D. M. Wyatt[3], Laurence Pelletier[1,2]*.

Authors affiliation:

[1]Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, Ontario, M5G 1X5, Canada.
[2]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, M5S 3E1, Canada.
[3]Department of Biochemistry, University of Toronto, Toronto, Ontario, M5S 1A8, Canada.
[4]These authors contributed equally.
*Corresponding author, pelletier@lunenfeld.ca

Keywords:

Gene-editing, CRISPR, Endogenous Tagging

Abbreviations:

CRISPR – clustered regularly interspaced short palindromic repeats

sgRNA – single guide RNA

PAM – protospacer adjacent motif

NHEJ – non-homologous end joining

HDR – homology directed repair

MMEJ – microhomology-mediated end joining

DSB – double-strand break

Indels – insertion-deletion

GOI – gene of interest

bp – base pair

kb – kilobase

RNP – ribonucleoprotein

Abstract

Endogenous tagging makes it possible to study a protein's localization, dynamics, and function within its native regulatory context. This is typically accomplished by using CRISPR to introduce a sequence encoding a functional tag into the reading frame of a gene; however, the efficiency of the process is often limited. Here, we introduce the "quickTAG," or qTAG system, a versatile collection of optimized repair cassettes designed to make CRISPR-mediated tagging more accessible. By partitioning the cassette to include both a desired tag sequence linked with a selectable marker, integrations can be quickly isolated post-editing. These constructs include several key features that enhance flexibility and ease of use, such as: cassette designs for N- and C-terminus tagging; standardized cloning adaptors to simplify the incorporation of homology arms for HDR or MMEJ-based repairs; restriction sites next to each genetic element within the cassette for easy switching of tags and selectable markers; and the inclusion of lox sites flanking the selectable marker to allow for marker gene removal following integration. Our library of ~120 plasmids is available on Addgene and includes ready-to-use constructs targeting frequently overexpressed genes as well as ready-to-clone cassettes to tag your own genes of interest. These initial cassettes were engineered to harbor a variety of tags for a range of applications, including fluorescence imaging, proximity-dependent biotinylation, epitope tagging, and targeted protein degradation. This protocol can yield a selected population of edited cells within ~2-3 weeks, and a clonal-validated cell line within ~5 weeks. Alongside this protocol, the qTAG system will provide a framework to streamline endogenous tagging in addition to serving as an open resource for researchers to adapt and tailor for their own purposes.

Introduction

Gene overexpression has long been a crucial tool in molecular biology, offering insight into the complexities of gene function and their roles in cellular processes and disease. However, this method has its limitations. Mislocalization caused by excess protein abundance can lead to interactions that deviate from its native state, which can ultimately manifest in mutant phenotypes[1-5]. This underscores the significance of studying proteins at their natural expression levels, as it affords a more precise understanding of a protein's localization, dynamics, and interactions, thereby allowing for more physiologically relevant analysis. Since its discovery in 2012, the introduction of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and its programmable Cas endonucleases has brought revolutionary change in the field of molecular genetics, by providing an accessible approach to gene editing[6]. Using a single guide RNA (sgRNA), this technology enables the editing of DNA by directing endonucleases to a specific genomic locus and allowing for targeted edits to be made. This is especially relevant in the context of gene-tagging, where the insertion of large transgenes encoding functional tags into the reading frame of a gene of interest (GOI) facilitates the study of proteins, without the negative consequences associated with transgene overexpression. Cas endonucleases, particularly Cas9, exhibit a high level of precision in identifying DNA sequences to cleave by requiring the presence of a protospacer adjacent motif (PAM) within the sgRNA sequence. Once the sequence is recognized at the genomic loci, Cas9 induces a double-strand break (DSB), after which the cell's innate DNA repair mechanisms are triggered to engage repair[7,8].

In the absence of a template to guide the repair, the DSB can be resolved by joining the two ends together through nonhomologous end-joining (NHEJ)[9-11]. This mechanism frequently incorporates base pair (bp) insertions or deletions (indels), that can disrupt the gene's coding region and as such has been leveraged by the CRISPR-Cas9 system to effectively disrupt gene function[12]. While NHEJ has shown potential for targeted insertions, its precision for large insertions is restricted by the likelihood of introducing errors like indels, causing misorientation of the donor sequence during insertion, and a higher tendency for off-target integrations[13-15]. Two alternative DSB repair pathways routinely exploited for precise gene insertions are: homology-directed repair (HDR) and microhomology-mediated end joining (MMEJ) (Fig. 1). Insertions harnessing HDR requires a donor DNA template containing long stretches of homologous sequence flanking the intended modification site to mediate repair[16]. In contrast, MMEJ-based gene editing relies on the binding of microhomologies of ~5-25 bp of complementary sequence to engage repair[17,18]. Using these pathways enhances precision compared to NHEJ, but it comes at the expense of reduced efficiency, as these pathways are less prevalently active and only engaged during specific cell cycle stages. From an execution standpoint, they also necessitate the supply of loci-specific homologous sequence which can be difficult to amplify and integrate into repair constructs. Collectively, these limitations often result

in extremely low integration rates when trying insert large transgenes, making the isolation of monogenic cell lines both challenging and time consuming[14,19-21].

Instead of enhancing CRISPR or delivery components directly to address low integration rates, alternative approaches have long explored optimizing the donor construct to improve editing outcomes. This typically involves partitioning the donor cassette to include the desired tag or modification, along with the co-expression of a fluorescent or mammalian selectable marker to allow for easy enrichment of edited cells. Several methods have explored this concept[17,19,20,22-26], with donor designs being largely summarized into two groups: promoter-based and multicistronic repair cassettes. Promoter-based cassettes involve the insertion of a tag within the reading frame of a GOI followed by a distinct promoter sequence that drives the expression of the selectable marker. In contrast, multicistronic approaches make use of sequence elements like internal ribosome entry sites (IRES)[27,28] or 2A "self-cleaving" peptide sequences[29,30] to facilitate efficient co-expression of the selectable marker (Fig. 2). Implementations of these cassettes frequently exhibit one or more limitations, in that they often lack adaptability, require complex cloning to function, and are not readily accessible for labs to implement.

Here, we have generated a suite of optimized repair cassettes as an extension of these methods, which we have previously demonstrated allow for editing and enrichment of endogenously tagged cell lines[31]. Our “quickTAG” (qTAG) system provides a flexible scaffold with simple cloning that is highly adaptable to alternative tags, selection markers, and experimental schemes. To assist in tagging your own GOIs, we couple this plasmid system with a comprehensive protocol that provides detailed instructions, considerations, and potential challenges to be aware of at every stage of the endogenous tagging process.

Development of the protocol

The qTAG cassettes are built upon the core 2A-based cassette structure described previously[17,25]. The initial cassettes were designed and synthesized as gene fragments with the 2A sequence situated between a tag and resistance gene. Due to its transcriptional linkage and short sequence length, being only ~50 bp, the 2A-based approach was adopted over IRES or a short promoter for quick and accurate enrichment of knock-ins. Four core features were implemented to optimize their compatibility and usability (Fig. 3): (1) Cassette designs specific to both the N and C termini, facilitating tagging at either end of a gene. (2) To simplify cloning, we included cloning adaptors that surround the cassettes. These sequences establish a uniform cloning approach, facilitating the incorporation of homology arms regardless of the tag, resistance, or specific sequences associated with the GOI. The adaptors were also designed to enable integration of both long and short homology arms, promoting HDR or MMEJ for precise

homology-based repair. (3) To enhance cassette adaptability, we introduced a unique restriction site next to every sequence element within the cassette. This design enables easy alteration through restriction cloning and permits future customization with a variety of tags and selectable markers as needed. (4) Finally, we included lox sites that flank the selectable marker to enable the removal of the marker gene upon transient introduction of the Cre recombinase[32,33]. For this we used the mutant lox sites lox66 and lox71 to ensure irreversible deletion of the marker gene and the recovery of selection sensitivity[34]. This is particularly useful in experimental conditions where only a single selectable marker is available and can be exploited to sequentially tag genes. Along with these cassettes, we also provide a detailed protocol that demonstrates the utility of our approach including steps that describe how to approach gene design, how to clone qTAG repair constructs, guidance during cell culture and editing, as well as simple validation strategies to ascertain your edits.

Applications of this approach

Our initial available collection qTAG cassettes is outfitted with tags to tackle a wide array of applications. These include 60 variations of the core cassette structure that enable functional fluorescent and proteomic protein investigation. Along with these cassettes, supplementary Cre plasmids and a variety of ready to use gene-specific gRNA and repair plasmids are made available in Addgene (https://www.addgene.org/Laurence_Pelletier/) for immediate use by the community. For fluorescent studies, a panel of the brightest currently available monomeric fluorescent proteins spanning the green, red, and far-red emission spectra enable visualization and tracking of a target protein's endogenous localization through live or fixed cell imaging. These include mNeon[35], moxGFP[36], super-TagRFP[37], mScarlet[38], and miRFP670nano3[39]. We have also expanded our qTAG cassettes to support a wide variety of non-fluorescent tags. Highly active proximity-dependent biotinylation enzymes, such as miniTurbo and ultraID, can be used to capture the local interactome of your target protein[40,41]. For targeted-degradation studies, dTAG can be employed to induce the degradation of a specific protein, allowing for transient observation of its functionality[25]. Lastly, antibodies targeting your protein of interest can often be unreliable and yield inconsistent results. This can be especially problematic in cases where an expressed protein is low in abundance, is poorly immunogenic, or has numerous paralogs. Epitope tagging solves these issues by inserting a short sequence encoding a general epitope within the reading frame of a specific gene, allowing for detection or isolation with a corresponding epitope-specific antibody. For this, we generated cassettes that contain 3xFLAG, 3xHA, and V5 epitopes, which have high-quality reagents for detection.

Outside of the core cassettes we are providing, researchers have the freedom and flexibility to switch schemes altogether from tag-only, promoter-based methods or opt for entirely customized strategies that incorporate multiple elements of their preference. In many instances, cell lines are often restricted in which and how many mammalians selectable markers

they can be exposed to. In cases like this, the qTAG cassette can be easily modified to encompass just a tag between the flaking cloning sequences as has been traditionally done in the past[15,17,21,42]. If the tag is fluorescent, enrichment following editing can be traditionally performed via FACS. In situations where the target gene exhibits low or no expression, particularly in scenarios where gene expression fluctuates such as in differentiation or reprogramming, accumulating a sufficient level of selectable marker can pose a significant challenge in the selection for integration. To solve this, the tagging of low expression or silent genes can be carried out by transcriptional unlinking the tag and selectable marker. Replacing the "2A" element with a constitutive generalized promoter will allow for robust selection independent of the target gene expression. This approach allows for the enrichment of low or conditionally expressed genes as well as the benefit of being able to silently tag genes that have yet to be expressed due to the independent promotion of the selectable marker[19,43]. The only caveats of this strategy being that this can only be targeted for C-terminal tagging and that the final accuracy of edits within the selected cell lines may be slightly compromised, as cells may develop resistance as long as the plasmid remains within the nucleus, not necessarily at the intended target site. Finally, in cases where disruption of gene function is necessary, but clonal expansion and validation of CRISPR indels is not possible, a knockout cassette can be inserted at the N-terminus encoding region of a gene to rapidly select for targeted gene knockouts[44]. In this case, the cassette can be tailored to incorporate a selectable marker followed by a stop codon and transcriptional termination sequence, such as a poly adenylation signal, to terminate transcription.

Comparisons to other methods

A number of different strategies have been developed to address the challenges associated with endogenous tagging in human cells[17,19-22,24,45-48]. Each of these techniques offers different advantages and features with some being highly tailored for specific applications such as post-mitotic tagging or stem cell tagging. They generally fall within the three tagging strategies discussed previously, taking advantage of either HDR, MMEJ, and/or NHEJ to insert target sequences encoding a tag or a tag linked with a selectable marker into a target site. These systems contain a number of variations on the formula that bring benefits to their specific use-cases and limitations in others. To begin, although Cas9 was the primary nuclease employed across most cases, one group had opted to use Cas12 citing reasons such as its accuracy in vivo as well as its simpler CRISPR RNA structure[19]. Using alternative Cas nucleases as well as nickases can be valuable in identifying suitable sgRNAs for target insertion sites, particularly in situations where no suitable candidates were previously available. Moreover, the employment of nickases offers an advantage by significantly mitigating the signaling responses associated with DSBs[49]. The configuration of the donor DNA is a pivotal factor that varies across these systems. The choice of format is primarily determined by several key factors, such as the size of your payload

sequence, the DNA repair mechanism you wish to leverage, the desired level of initial editing success, and the overall suitability of your cells for the editing process. Different formats are available, including linear PCR cassettes, single-stranded oligodeoxynucleotides, and plasmids, each offering distinct advantages and considerations in the context of genome editing. The various methods also differ in terms of the complexity of cloning and the reagent delivery methods employed, which depend on the transfectability of the target cells. Liposomes, electroporation, and the use of viruses are all efficient methods of delivery explored in these papers, where each method often specializes in targeting specific cell types.

While all of these methods have shown their effectiveness in achieving high-efficiency cell editing or the ability to enrich for it, it's important to acknowledge that they can often be challenging to implement due to the rigidity of their systems and the availability of necessary reagents. Indeed, some donor strategies have fixed designs that do not easily accommodate changes, such as using a different tag or selectable marker. In terms of availability, some solutions are only accessible through commercial channels[45], while others may not be publicly available at all[22]. The qTAG system discussed here incorporates fundamental principles from these methods, integrating them into a user-friendly system that aims to make knock-ins more accessible. This system is made available on Addgene for the broader scientific community. The incorporation of restriction sites within the cassette not only simplifies the process of replacing tags and markers but also allows for potential modifications to the core structure to cater to diverse end-user needs.

Limitations

With the supplied qTAG cassettes however, two major limitations do exist. One drawback is the reliance on transcriptional linkage of the target gene to the selectable marker with the 2A sequence. This requires that the target GOI be expressed at moderate levels in the cell to produce enough antibiotic resistance for efficient selection. Also, while 2A peptide sequences have garnered mass adoption in molecular biology to achieve co-translation, their cleavage or 'skipping' efficiency is not always consistent in all cellular backgrounds. This can result in a minor amount of read-through[29,50]. Given the adaptable nature of the qTAG cassette scaffold, it is possible to address these limitations by terminating the tag with a stop codon and substituting the 2A sequence with a small promoter, such as hPGK. This modification would enable the selection of edits independent of the target gene's transcript levels and circumvent the constraints associated with 2A peptide sequences.

Overview of the workflow

Here, we describe the complete process of gene tagging in human cells using the qTAG cassette system and offer detailed considerations and examples to facilitate efficient gene

tagging. The procedure illustrated in Fig. 4 displays four major phases: gene design (Step 1-25), construct cloning (Steps 26–27), the delivery of editing components and cell culture (Steps 28–38), and validation assays to ascertain accurate editing (Step 39-50).

Experimental design

Considerations for choosing a gene-specific sgRNA and homology arm design

The first important consideration in creating functional endogenous fusions with minimal interference is determining whether to place the tag on the N- or C- terminus of your protein of interest. With the adoption of gene-tagging increasing more and more over the years, more gene designs detailing the location of validated sgRNAs and flanking homology is becoming more available. Prior to doing a deep dive on your GOI, check if it has previously been tagged in the literature. Two particularly useful resources for this include the Allen Institute cell catalog (https://www.allencell.org/cell-catalog.html) and the OpenCell project (https://opencell.czbiohub.org/) which are both based on systematic gene-tagging studies[21,47]. If previous gene designs are unavailable, the ideal terminus can be determined by consulting resources such as primary literature or Addgene where the GOI has been overexpressed and validated for functionality. Alternatively, to determine the most suitable terminus for tag insertion, consulting protein databases like UniProt[51] (https://www.uniprot.org/) to check the predicted domains of your protein of interest will help identify the sites for insertion with minimal disruption. In cases where no information is available, both ends can also be separately tagged and evaluated.

Crucial to the efficiency of knock-in and the minimization of off-target effects when using CRISPR is the selection of appropriate sgRNAs. Based on our experience, we found that CHOPCHOP v3 [52] (https://chopchop.cbu.uib.no/), a free web-based tool for generating sgRNAs *in-silico*, was the most intuitive when it came to choosing sgRNAs for knock-in purposes. This is because it provides a graphical display of the binding sites of sgRNAs in the terminal regions and facilitates the seamless exporting of the sequence for further design annotation. Cas9 typically cleaves 3 bp upstream of the PAM. The frequently employed Cas9 derived from Streptococcus pyogenes, known as SpCas9, targets the PAM sequence 5′-NGG-3′, where the "N" can represent any nucleotide base. The location of sgRNAs is a crucial factor to consider, and it is advisable to select a sgRNA that is within ~30 bp of the start or stop codon for 5′ or 3′ insertion, respectively. Beyond their location, predicted on-target and off-target scores should be considered when choosing and ranking potential sgRNAs. Testing the efficiency of your sgRNA *in vitro* with Sanger sequencing methods like a TIDE[53] analysis (https://tide.nki.nl/) or a T7 endonuclease assay can save time and improve tagging outcomes prior to preparing repair constructs. Lastly, it is important to ensure that the sgRNA sequence does not overlap with the donor construct. When designing homology arms around your selected sgRNA, confirm that your insertion disrupts the

sgRNA sequence or, in situations where that is not possible, introduce PAM blocking or sgRNA blocking mutations to prevent Cas9 cleavage of the donor construct upon entrance to the nucleus or re-cleavage of the genomic loci upon integration.

Another major consideration that impacts the cloning and editing phases of this method is the DNA repair mechanism exploited for tagging. HDR requires long stretches of homology, which can be synthesized as gene fragments or PCR-amplified from genomic DNA. This mechanism offers greater conservation of the original reading frame and allows for more sgRNA choices and the ability to test multiple gRNAs with a single repair construct provided that PAM blocking mutations are introduced. However, synthesizing long oligonucleotides can be expensive and challenging if the sequences are GC-rich or repetitive. In contrast, MMEJ requires shorter regions of homology, which can be added to the cassette directly via PCR, making it much more practical and cost-effective. However, MMEJ provides limited flexibility for maintaining the native reading frame, as it necessitates knock-in at or near the sgRNA cut site. This can often result in slight truncations of the protein if the selection of sgRNAs is not abundant at the terminal codons. As a result, MMEJ usually allows for only one repair construct per sgRNA. In situations where the sgRNA fails to effectively create a DSB, it becomes necessary to construct a new repair construct to evaluate a new gRNA.

Before embarking on your knock-in experiments, it is also essential to consider your target gene in the context of your specific cell line. Each cell type possesses unique properties and behaviors that can impact the efficiency and feasibility of genome editing. Factors such as growth rate, prevalence of specific DNA repair mechanisms, and susceptibility to transfection can vary among different cell lines. Additionally, the ploidy of your target cell line can influence the efficiency and stability of genomic modifications, as well as the practicality of isolating edited clones because of the number of tagged alleles to consider. Finally, evaluating the dominant isoform and expression pattern of your target gene within your target cell line is crucial for successful tagging. Searching databases like Ensembl and Uniprot can provide insight into the presence of dominant transcripts and isoforms. The Human Protein Atlas (https://www.proteinatlas.org/) can further supplement human cell-line specific RNA-sequencing information on whether your GOI is expressed in a particular cellular background. We found that it was generally best to begin with either HEK293T cells or HAP1 cells when testing sgRNA/repair pairs. HEK293T cells have high transfectability and demonstrated the greatest initial tagging rates, making them a suitable starting point to determine if your sgRNA/repair pairs were compatible. On the other hand, HAP1 cells are predominantly haploid, which allows for easy allelic verification of the knock-in. Starting with either of these cell lines would be an ideal first step in testing your gene designs prior moving on to your desired cell line.

Cloning of sgRNA and donor constructs

To create the qTAG repair constructs for MMEJ and HDR, we used standardized procedures making use of either restriction insertion or Gibson assembly, respectively. For MMEJ, flanking homology of 10-20 bp can be added to the primers used to amplify the preferred qTAG cassette. The resulting cassette fragment is then digested, and restriction cloned into a pUC19 backbone. In the case of HDR, the homology arms can be synthesized or amplified from the genome via PCR. The genomic arm sequences are flanked by the cloning adaptor sequences which act as the uniform sequence overlap required for Gibson assembly. The preferred qTAG cassette is also amplified along with the digestion of the pUC19 backbone. The four fragments of digested pUC19, left homology arm, amplified cassette, and right homology arm are then directionally cloned using Gibson assembly at a 1:3:3:3 molar ratio to form the final repair construct. In both cases, repair constructs were sequence verified prior to use for editing.

For Cas9 and sgRNA delivery, we sought to make this method as accessible as possible. Therefore, we opted to use the original Zhang protocol for SpCas9 plasmids (https://www.addgene.org/crispr/zhang/) to deliver our Cas9 and sgRNA[54]. We particularly used the MMEJ PITCh-primed variation of px330, the px330-PITCh plasmid[17,26]. The px330-PITCh plasmid encodes SpCas9 and a U6-gRNA cassette, simplifying gRNA cloning through annealing oligos that encode your gRNA sequence and performing BbsI enzyme restriction insertion cloning.

Cell culture, delivery of editing components, and enrichment of edits.

We have been typically delivering our Cas9/sgRNA plasmids along with our repair constructs with chemical transfection using Lipofectamine 3000 with good success. In this protocol we demonstrate initial tagging rates and subsequent enrichment across four human cell lines including HAP1, ARPE-19, HEK293T, U-2 OS cells. Cells were initially seeded at around 70% confluency (~500,000 cells) in 6-well plates and transfected with a 1:1 ratio of Cas9 cutting plasmid to repair plasmid totaling 3 µg. After allotting a 72 hour of editing period, we selected the cells using the appropriate mammalian selection or fluorescently sorted them to obtain a highly enriched pool of edited cells. Our quickest scenario involved tagging the *H2BC11* gene with a GFP-Puro qTAG repair construct and subsequently selecting with puromycin. We initially achieved a GFP positive rate of ~3% in HEK293T cells. Following the selection, over 90% of the cells exhibited GFP fluorescence within a week and a half of the initial transfection. For particularly sensitive cell lines where slow proliferation, low transfectability, or promoter-sensitivity are issues, ribonucleoprotein (RNP) complexes have emerged as optimal alternatives to Cas9/gRNA plasmids where the direct introduction of Cas9 protein and sgRNA are delivered as a pre-assembled complex. This method is commonly utilized when editing stem cells[20,21], and we

further showcase its applicability by successfully delivering RNPs along with qTAG repairs through electroporation to knock-in fluorescent proteins in H9 stem cells.

Validation strategies

To assess cells for precise tagging, it is necessary to perform a variety of validation steps to ascertain that your insertion occurred and inserted in-frame. These include genomic junction PCR and sequencing to ensure accurate insertion at the genomic loci has taken place. Also, if an antibody specific to your protein of interest is available, fluorescence and proteomic validations can be carried out to check if the genetic insertion manifested at the protein-level. Fixed or live-cell imaging of fluorescent or proteomic tags can corroborate the insertion via confirming the localization of your target protein, while immunoblotting can verify the genetic insertion by visualizing the target protein size differential.

Materials

Cloning Reagents

- Important sequences and primers. For a full list of sequences and primers relevant to the qTAG cassettes, see Table 1
- Plasmids, qTAG cassettes, and plasmid maps. For a full list of plasmids used in the qTAG system as well as Addgene links to access them and their corresponding plasmid maps, see Table 2
- Platinum SuperFi II Green PCR Master Mix (Invitrogen, cat. No. 12369010)
  ▲CRITICAL – A high fidelity polymerase is recommended when performing genomic/cloning amplifications. In our hands, the Platinum SuperFi II Polymerase Master Mix required the least optimization and produced the best results when supplemented with 5% DMSO to enhance amplification of plasmids DNA and genomic DNA. Other high-fidelity DNA polymerases, such as Q5 (NEB) and KAPA Hifi (Sigma-Aldrich), can alternatively be used.
- DMSO (Sigma-Aldrich, cat. No. D8418)
- QuickExtract DNA Extraction Solution (Lucigen, cat. No. LGN-QE09050)
- QIAquick PCR Purification Kit (QIAGEN, cat. No. 28106)
- QIAquick Gel Extraction Kit (QIAGEN, cat. No. 28706)
- QIAprep Spin Miniprep Kit (QIAGEN, cat. No. 27106)
- QIAGEN Plasmid Plus Midi Kit (QIAGEN, cat. No. 12945)
- FastDigest BpiI or BbsI equivalent isoschizomer (Fermentas/Thermo Fisher Scientific, cat. No. FD1014)
- FastAP (Alkaline Phosphatase) (Fermentas/Thermo Fisher Scientific, cat. No. EF0651)
- T4 Polynucleotide Kinase (NEB, cat. No. M0201S)

- T4 DNA Ligase Kit (10X Buffer + T4 Ligase) (NEB, cat. No. M0202S)
- Quick Ligation Kit (2X Buffer + Quick Ligase) (NEB, cat. No. M2200S)
- AgeI-HF (NEB, cat. No. R3552L)
- BamHI-HF (NEB, cat. No. R3136S)
- EcoRI-HF (NEB, cat. No. R3101S)
- HindIII-HF (NEB, cat. No. R3104S)
- KpnI-HF (NEB, cat. No. R3142S)
- MluI-HF (NEB, cat. No. R3198L)
- NheI-HF (NEB, cat. No. R3131L)
- XbaI (NEB, cat. No. R0145L)
- XhoI (NEB, cat. No. R0146S)
- XmaI (NEB, cat. No. R0180L)
- RedSafe (FroggaBio, cat. No. 21141)
- ddH2O
- DH5α competent cells (Invitrogen, cat. No. 18258012)
- Ampicillin (Sigma, cat. No. A5354)
- Luria–Bertani (LB) broth and LB agar (Thermo Fisher Scientific, cat. No. 12780052, 22700025)

Biological Materials and Culture Reagents

- HAP1 cells (Horizon Genomics, cat. No. C631, RRID: CVCL_Y019)
- HEK293T cells (ATCC, cat. No. CRL-3216, RRID: CVCL_0063)
- ARPE-19 cells (ATCC, cat. No. CRL-3216, RRID: CVCL_0145)
- U-2 OS cells (ATCC, cat. No. HTB-96, RRID: CVCL_0042)
- H9 cells (WiCell, cat. WA09, RRID: CVCL_9773)
- Dulbecco’s modified Eagle medium (DMEM) for HAP1, HEK293T, and HeLa cells (Gibco, cat. No. 11965084)
- McCoy's 5A medium for U-2 OS cells (Gibco, cat. No. 16600082)
- DMEM/F-12 for ARPE-19 cells (Gibco, cat. No. 11320033)
- mTeSR1 for H9 (StemCell Technologies, cat. No. 100-0276)
- CloneR for H9 (StemCell Technologies, cat. No. 05889)
- TrypLE Express Enzyme (Thermo Fisher Scientific, cat. No. 12605028)
- Matrigel (VWR, cat. No. CA89050-192)
- Fetal Bovine Serum (Gibco, cat. No. A4736301)
- 1× phosphate-buffered saline (Gibco, cat. No. 14190250)
- Trypsin-EDTA-0.25% (Gibco, cat. No. 25200056)
- GlutaMAX (Gibco, cat. No. 35050061)

- Penicillin/streptomycin (Gibco, cat. No. 15140122)
- Lipofectamine 3000 (Invitrogen, cat. No. L3000008)
- Neon Transfection Kit (Thermo Fisher Scientific, cat. No. MPK10096)
- Opti-MEM Reduced Serum Medium (Gibco, cat. No. 31985070)
- Puromycin (InvivoGen, cat. No. ant-pr-1)
- Blasticidin (InvivoGen, cat. No. ant-bl-1)
- Zeocin (InvivoGen, cat. No. ant-zn-1)

Software

- An annotatable sequence viewer
  ▲CRITICAL – Annotatable sequence viewers are highly advantageous as they simplify the visualization of the sgRNA binding site, the optimal positioning of homology arms, and the intended insertion site. SnapGene was the most intuitive and informative option in our hands. ApE and Benchling are good alternatives, while raw sequence viewers are less preferable for gene design.
- sgRNA design tool
  ▲CRITICAL – CHOPCHOP v3 is used in this protocol as it graphically displays the binding site of the sgRNAs and allows for easy export and annotation. However, any sgRNA generation tool can work.

Equipment

- Biosafety level 2 facility
- Thermal cycler (any brand).
- CO2 cell culture incubator with temperature set to 37 °C (any brand)
- Neon Transfection System (Thermo Fisher Scientific, cat. No. MPK5000)
- Fluorescent and brightfield microscopes (any brand to observe with enough magnification to observe single cells)
- Bench-top centrifuge (any brand, with a rotor for 1.5-2 mL epitubes microcentrifuge tubes at up to 13,000 rpm)
- Multipurpose centrifuge (any brand, with a rotor for 15-50 mL tube at 4,000 rpm)
- Culture dishes and plates (15cm, 10cm, 6-well, 12-well, 24-well, 48-well, 96-well; any brand as long as appropriate to culture your specific cells)
- Flow cytometer (any brand, optional)
- Fluorescence activated cell sorter (any brand, optional)

Procedure

Gene Design: Selection of sgRNAs ● TIMING ~ 1 h

1. Having considered the relevant factors involved in gene tagging, our first step is to predict and choose efficient, termini specific, sgRNAs to induce DSBs for tagging. In this protocol, we will follow the design of the histone gene *H2BC11* as an example of a 3' insertion and the actin gene *ACTB* as an example of a 5' insertion. Begin by opening CHOPCHOP on your browser of choice (https://chopchop.cbu.uib.no/).
2. Ensure that CHOPCHOP is properly configured for knock-in prior to submitting a query gene. Make these changes to the default options to maximize your sgRNA pool for knock-in (Fig. 5a). First, select 'knock-in' from the to the left drop-down menu in the 'For' tab. Next, click the 'Options' tab to view and modify the default settings. In the 'General' tab, ensure that 'All exons (inc. UTRs)' is selected for the query space to account for sgRNAs found around the sequence representing the protein's N and C termini. Also ensure the isoform consensus in the 'General' tab is determined by 'union' to account for sgRNAs across all isoforms. Finally, CHOPCHOP generates optimal primers that flank the sgRNA for subsequent amplification and Sanger sequencing indel analysis. Although these primers won't be used for knock-in purposes, the product size in the 'Primers' tab will have an impact on the exportable Genbank sequence file of the final sgRNAs we decide upon. Therefore, we recommend setting the product size to at least 1000 bp or more to allow adequate space to annotate your homology arms.
3. Select compatible sgRNAs within 30bp of the start or stop codon that demonstrates a high on-target efficiency and low off-target scores. These sgRNA sequences can be further cross referenced across other sgRNA selection tools to corroborate their predicted efficiency.

   ▲ CRITICAL – We recommend selecting more than one sgRNA to increase the chance of successful targeting and gene editing. This depends on the availability of high-quality sgRNAs for your query gene as well as the downstream repair mechanism. For example, HDR can accommodate more sgRNAs because of the long homology arms.

   ? TROUBLESHOOTING
4. Click on your ideal sgRNA to expand the view of the specific sgRNA to the base pair level within CHOPCHOP. Export the sequence information as a Genbank file and open it in SnapGene or your annotatable sequence viewer of choice.
5. Cross reference with CHOPCHOP to identify and annotate the exon information of your gene and ensure the sequence is translated such that the DNA sequences correspond to amino acids within your gene of interest. Cleary annotate your sgRNAs so that the PAM and upstream cut site are easily identifiable along with the rest of the sgRNA sequence.
6. Design two annealing oligos that encode with your sgRNA sequence (PAM excluded) according to the guidelines for pX-line SpCas9 plasmids provided by Feng Zhang's laboratory (Massachusetts Institute of Technology) in Addgene (https://www.addgene.org/crispr/zhang/).

▲CRITICAL – Any method of Cas9 and sgRNA delivery can be used. This includes lentiviral delivery or delivering the Cas9 protein precomplexed to the sgRNA as a ribonucleoprotein complex (RNP).

7. Order oligos with standard desalting, no modification and 25 µmol each.

Gene Design: Homology Arms ● TIMING ~ 1 h (Fig. 5b)

8. HDR design:
    a. Annotate 500-1000 bp of genomic sequence upstream and downstream of your insertion site to denote the left and right homology arms. For H2BC11, the insertion point that separates the two homology arms is positioned prior to the stop codon, while in the case of ACTB, it is situated immediately following the start codon.
    b. To avoid the cleavage of the donor construct or re-cleavage of successful integrations, annotate mutations that disrupt the PAM sequence of the sgRNA by introducing a silent mutation in one of the 'G' bases within the 'NGG' motif. In cases where this is not possible, such as in PAM sequences overlapping glycine or proline residues, introducing multiple silent mutations across the sgRNA sequence on the repair construct will suffice in reducing sgRNA binding[55].
    ▲CRITICAL – One set of HDR homology arms can contain mutations for testing multiple sgRNAs.
    c. Design a pair of primers to PCR-amplify the genomic junction where the insertion should take place in order to screen for successful integration.
    ▲CRITICAL – These primers should anneal outside the homology arms to ensure amplification of true allelic integrations instead of the donor plasmid. These primers can be generated via Primer BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) . We found that generating specific primers that were >20-25 bp in length with a melting temperature of ~ 60 °C worked best for these amplifications.
    d. Save this file without altering the original sequences as it will serve as your reference gene design file. Keep any mutations you intend to introduce strictly as annotations above the original sequence.
    e. To generate the HDR homology arms, copy and paste the left homology arm (LHA) and right homology arm (RHA) sequences from your design file into new sequence files.
    f. Modify the arm files to include your PAM/sgRNA-disrupting mutations.
    g. Add the universal cloning adaptors for Gibson assembly to each of the arms (Table 1).

▲CRITICAL – These are specific to 5' and 3' insertions for N and C terminus tagging cassettes, respectively.

h. Save the homology arm files. At this point, the respective arms can be synthesized as gene fragments (ThermoFisher - GeneArt Strings, IDT – gBlocks, etc.) or can be amplified from the genome by PCR. If amplifying from the genome, adhere to the previously stated genomic primer design guidelines (Step 8c), but ensure you add the cloning adapters to generate the arm products necessary for subsequent cloning.

9. MMEJ design:
   a. Annotate ~20 bp of genomic sequence flanking the cut site of your sgRNA to denote the left and right homology arms. If the sgRNA is predicted to cut inside a codon, ensure bp are added to each of the arms to maintain the reading frame after the insertion. With shorter homology arms, there is less accommodation to test multiple sgRNAs unless they are directly overlapping.
   b. Because MMEJ uses short homology arms to induce repair, homology sequences are typically very close to the sgRNA cut site and therefore successful insertion inherently disrupts the sgRNA.
   c. Annotate genomic amplification primers as described in Step 8c.
   d. Save this file as the original gene design file.
   e. To generate the MMEJ homology arms, copy and paste the LHA and RHA sequences directly into the MMEJ cassette primer scaffold shown in Table 1. These primers will be used to directly amplify your qTAG cassette of choice while integrating your MMEJ homology arms.
      ▲CRITICAL – Ensure you are copying the sense and anti-sense sequence for the forward and reverse primers, respectively.
   f. Save your homology arm files and order your oligos.

Cloning of Plasmids: Cas9 + sgRNA ● TIMING ~ 2-4 d (Fig. 6a)

10. Digest pX330 or pX330-PITCh with the FastDigest BpiI enzyme following the reaction below or manufacturer's instructions for 30 min at 37 °C.

Components	Amount
pX330/PITCh (1 µg/µL)	1 µL
FastDigest BpiI (BbsI)	1 µL
FastAP (1 U/µL)	1 µL
10X FastDigest Buffer	2 µL
ddH_2O	15 µL
Total Volume:	20 µL

11. Stain the DNA with RedSafe according to the manufacturer's instructions. Run the stained reaction on a 1% TBE gel. Excise and gel purify the digested backbone (~8000 bp) using the QIAquick Gel Extraction Kit.
12. Resuspend the sgRNA oligos obtained from Step 7 with ddH2O to generate 100 µM stocks, then set up the following reaction mix.

Components	Amount
sgRNA oligo 1 (100 µM)	1 µL
sgRNA oligo 2 (100 µM)	1 µL
10X T4 Ligation buffer	1 µL
ddH_2O	6.5 µL
T4 PNK (10 U/µL)	0.5 µL
Total Volume:	10 µL

13. Carry out the annealing reaction on a thermocycler using the following conditions.

Step	Temperature	Time	Cycle number
Incubate	37 °C	30 min	1x
Denature	95 °C	5 min	1x
Anneal	95 – 25 °C	Ramp down @ 5 °C/min	1x
Hold	4 °C	∞	1x

14. Dilute the annealed oligos 1:200 in ddH_2O
15. Perform a ligation reaction between the diluted annealed sgRNA insert and the digested pX-backbone at 22 °C for 20 min as follows or following the manufacturer's instructions:

Components	Amount
Digested pX-Backbone	50 ng
Diluted annealed sgRNA duplex	1 µL
2X Quick Ligation Buffer	5 µL
ddH_2O	up to 9 µL
Quick Ligase	1 µL
Total Volume:	10 µL

16. Add 50-100% of the reaction product to 50 µL of DH5α competent bacteria cells.
    ▲CRITICAL – Ensure you transform a vector only control (no insert) to gauge digestion efficiency.
17. For transformation, incubate the DNA/bacteria mixture on ice for 30 min, heat shock at 42 °C for 60 s and incubate back on ice for 2 min.
18. Using sterile technique, add ~500 µL of LB to the cells and allow for a recovery, incubating the cells at 37 °C, shaking at 200 rpm for 1 h.

19. Using sterile technique, spread 100 µL of bacteria on an LB plate supplemented with 50 µg/mL of ampicillin (Amp).
20. Incubate the plates overnight at 30 °C.
21. Use sterile technique to pick 4-6 colonies and inoculate each in 3-5 mL LB/Amp media. Grow the cultures at 30 °C, shaking at 200 rpm for 6 – 16 h.
    ? TROUBLESHOOTING
22. Reserve an aliquot (~100 µL) of bacteria for downstream amplification, then process the remainder for plasmid DNA extraction using the QIAprep Spin Miniprep Kit.
23. Sequence the clones with the pX sequencing primer in Table 1.
24. Once a clone is verified to contain the desired sgRNA sequence, inoculate the reserved bacteria in 50 mL of LB/Amp media and grow them at 37 °C, shaking at 200 rpm for 16 – 20 h.
25. Isolate the DNA using the QIAGEN Plasmid Plus Midi Kit for transfection grade, endotoxin-free plasmid DNA.

Cloning of Plasmids: Repair Constructs ● TIMING ~ 2-4 d (Fig. 6b, c)

26. HDR Cloning
    a. Digest any qTAG cassette plasmid with EcoRI and HindIII for 1 h at 37 °C. to release the pUC19 backbone.

Component	Amount
pUC19 (1 µg/µL)	2 µL
10X rCutsmart Buffer	2 µL
EcoRI-HF (20 U/µL)	0.5 µL
HindIII-HF (20 U/µL)	0.5 µL
ddH_2O	15 µL
Total Volume:	20 µL

    b. Stain the DNA with RedSafe according to the manufacturer's instructions. Run the stained reaction on a 1% TBE gel. Excise and gel purify the digested backbone (~2600 bp) using the QIAquick Gel Extraction Kit.
    c. Setup a PCR reaction to amplify your qTAG cassette of interest using the termini specific amplification primers in Table 1

Component	Amount
2X Platinum SuperFi II Master Mix	25µL
DMSO	2.5 µL
Fw Cassette Primer (10 µM)	1.5 µL
Rev Cassette Primer (10 µM)	1.5 µL
Template (10 ng/µL)	1 µL
ddH_2O	18.5 µL

Total Volume:	50 µL

d. Perform PCR using the following cycling conditions:
▲CRITICAL – This 2-step PCR protocol was optimized for the Platinum SuperFi II polymerase to amplify genomic and plasmid sequences with high fidelity. Ensure that you follow and optimize the cycling conditions according to the polymerase you are using.

Step	Temperature	Time	Cycle number
Initial Denature	98 °C	2 min	1x
Denature	98 °C	20 s	35x
Annealing	62 °C	15 s	
Extension	68 °C	30 s/kb	
Final Extension	68 °C	5 min	1x
Hold	4 °C	∞	1x

e. Run the PCR product on a 1% (wt/vol) agarose gel in TBE buffer at 120 V for 30 min.
f. Excise and gel purify the cassette using the QIAquick Gel Extraction Kit.
g. Perform steps 25c-f again with 100 ng of genomic DNA if you decide to amplify the HDR homology arms from genomic DNA instead of using a gene synthesis service.
h. Carry out a four-piece Gibson assembly at a ~1:3:3:3 molar vector to fragment ratio at 50 °C for 1 h as follows:

Component	Amount
2X Gibson Assembly Master Mix	10 µL
pUC19 Backbone (~2600 bp; 50 ng/µL)	50 ng
qTAG Cassette Fragment (~1500 bp; 100 ng/µL)	100 ng
Left Homology Arm (~600 bp; 40 ng/µL)	40 ng
Right Homology Arm (~600 bp; 40 ng/µL)	40 ng
ddH_2O	up to 20 µL
Total Volume:	20 µL

i. Repeat steps 16 – 22.
j. Sequence the clones with the P2A-Fw, P2A-Rev, LHA, and RHA sequencing primers in Table 1 to verify the integration of both homology arms along with the tagging cassette and ensure everything is cloned within the correct reading frame.

k. Once the qTAG donor construct is sequence verified, inoculate the reserved bacteria in 50 mL of LB/Amp media and grow them at 37 °C, shaking at 200 rpm for 16 – 20 h.
l. Isolate the DNA using the QIAGEN Plasmid Plus Midi Kit for transfection grade, endotoxin-free plasmid DNA.

27. MMEJ Cloning
    a. Digest and isolate the pUC19 backbone as instructed above (Step 25a-b)
    b. PCR the qTAG cassette of interest using the MMEJ specific amplification primers you have designed according to Table 1.

Component	Amount
2X Platinum SuperFi II Master Mix	25 µL
DMSO	2.5 µL
Fw MMEJ-Cassette Primer (10 µM)	1.5 µL
Rev MMEJ-Cassette Primer (10 µM)	1.5 µL
Template (10 ng/µL)	1 µL
ddH_2O	18.5 µL
Total Volume:	50 µL

    c. Stain the DNA with RedSafe according to the manufacturer's instructions. Run the stained reaction on a 1% TBE gel. Excise and gel purify the amplified cassette (~1500 bp) using the QIAquick Gel Extraction Kit.
    d. Digest all of the gel purified cassette with EcoRI and HindIII for 1 h at 37 °C.

Component	Amount
Purified MMEJ Cassette	up to 20 µL
10X rCutsmart Buffer	2 µL
EcoRI-HF (20 U/µL)	0.5 µL
HindIII-HF (20 U/µL)	0.5 µL
ddH_2O	up to 20 µL
Total Volume:	20 µL

    e. Purify the digested DNA with the QIAquick PCR Purification Kit.

f. Perform a ligation reaction between the MMEJ amplified qTAG cassette and the pUC19 backbone at a ~1:5 molar ratio of vector to insert at 22 °C for 20 min as follows:

Component	Amount
Digested pX-Backbone (50 ng/µL)	50 ng
MMEJ qTAG Cassette (100 ng/µL)	2 µL
2X Quickligation Buffer	5 µL
ddH_2O	up to 9 µL
Quick Ligase	1 µL
Total Volume:	10 µL

g. Follow Steps 25i-q to complete the cloning of the MMEJ repair construct.

Cell Culture: Delivery of Editing Reagents and Editing Period ● TIMING ~ 4 d

28. Chemical delivery of editing components via Lipofectamine 3000 into tissue culture cell lines
    a. Seed ~500,000 cells (or enough cells to reach 70-80% the next day) per well in a 6-well plate.
    ▲CRITICAL – Ensure a transfection control and selection control well are plated in tandem to your gene-tagging target well.
    b. The next day, co-transfect the pX and repair plasmids with Lipofectamine 3000, as follows:

Step	Tube	Components	Amount per well (6-well)
1	1	Opti-MEM	125 µL
		Lipofectamine 3000 Reagent	5 µL
2	2	Opti-MEM	125 µL
		1 µg/µL pX Cas9 Plasmid	1.5 µL
		1 µg/µL qTAG Repair Plasmid	1.5 µL
		P3000 Reagent	6 µL
3	Prepare the solutions listed in step 1 and step 2, mix and spin down		
4	Add tube 2 solution to tube 1, mix and spin down		
5	Incubate for 10 min at room temperature		
6	Add 250 µL to cells and shake to distribute		

? TROUBLESHOOTING

    c. Replace the media the next day.
    d. Allow for an editing period of 72 h post-transfection, prior to selection enrichment.

▲CRITICAL – If the cells reach confluency, Passage them to a larger vessel (6-well to 10 cm dish) to ensure room for cell growth and to prevent contact inhibition.

29. Delivery of editing components via Neon electroporation and RNPs into H9 human stem cells
    a. Prior to electroporation, dissociate H9 cells into single cells using TrypLE
    b. Collect the cells via centrifugation at 2000 rpm and resuspend them in mTeSR1.
    c. Assemble the RNP complexes on ice by incubating 5 µg purified SpCas9 (31.25 pmol) with 3.6 µg sgRNA in a total volume of 10 mL in Buffer R from the Neon electroporation kit.
    d. Add the RNP complexes to the cells and mix via gentle pipetting.
    e. Carry out the electroporation using a 100 µL Neon tip at 1200 V with a single 30 ms pulse.
       ? TROUBLESHOOTING
    f. Rapidly transfer the cells to warmed mTeSR1 media containing CloneR in a Matrigel-coated well of a 6-well plate.
    g. The following day, replace the media with fresh mTeSR1 containing CloneR.
    h. Allow for an editing period of up to 72 h prior to selection enrichment.

Cell Culture: Selection Enrichment ● TIMING ~ 3 – 15 d

30. After 72 h, replace culture media with media supplemented with the mammalian antibiotic corresponding to your tagging cassette. Below are a range of working concentrations tested for ARPE-19, HEK293T, HAP1, U-2 OS, and H9 cells.
    ▲CRITICAL – Mammalian cell sensitivity to antibiotics varies from one cell type to another. To enrich for knocked-in cells, it is important to determine the minimum concentration of antibiotic required to kill untransfected cells. Ensure you are using the optimal concentration of antibiotic reagent to fully select your cells in the timeline described in the table below.
    ▲CRITICAL – Besides mammalian selection, fluorescence markers can be analyzed using a flow cytometer to determine the efficiency of tagging. Cells expressing the fluorescence marker can be isolated through FACS if mammalian selectable markers are not an option.

	Human Tissue Culture Cell Lines (ARPE-19, HAP1, HEK293T, U-2 OS)		Human Stem Cells (H9)	
Antibiotic	Concentration	Selection Time	Concentration	Selection Time
Blasticidin	5-10 µg/mL	4-6 days	3 µg/mL	7-8 days
Puromycin	0.5-1 µg/mL	2-3 days	0.5 µg/mL	2-3 days
Zeocin	100-400 µg/mL	10-15 days	25-50 µg/mL	10-15 days

31. Once all of the untransfected cells have been killed, the cells in the transfected well(s) should represent the initial enriched tagged pool.
    ▲CRITICAL – The initial number of remaining cells can be scarce depending on the efficiency of the sgRNA and the likelihood of successful repair. In cases like these, it is recommended to allow the cells to grow for an additional 2-5 days.
    ? TROUBLESHOOTING

Cell Culture: Cre Excision ● TIMING ~ 4 d

32. Seed ~500,000 of the tagged cell line (70-80% the next day confluency) per well in a 6-well plate.
33. The next day, deliver the Cre recombinase to the cells. This can be done through a variety of means but we have provided a set of easily selectable Cre expressing plasmids (Table 2). Transfect your Cre plasmid of choice (this case Cre-Puro) with Lipofectamine 3000 as follows:

Step	Tube	Components	Amount per well (6-well)
1	1	Opti-MEM	125 µL
		Lipofectamine 3000 Reagent	5 µL
2	2	Opti-MEM	125 µL
		1 µg/µL Cre-Puro	3 µL
		P3000 Reagent	6 µL
3	Prepare the solutions listed in step 1 and step 2, mix and spin down		
4	Add tube 2 solution to tube 1, mix and spin down		
5	Incubate for 10 min at room temperature		
6	Add 250 µL to cells and shake to distribute		

    ▲CRITICAL – It is recommended to perform the Cre step, as the removal of the 2A and selectable marker sequences can minimize potential adverse effects such as 2A read through and unsuspecting fusion intermediates.
34. The following day, begin the selection of cells either through fluorescence or puromycin resistance to enrich for cells that express the Cre recombinase, leading to the formation of a tagged pool of cells with the integrated selectable marker being removed.
    ▲CRITICAL – Remove the selection immediately after the control cells have been killed as to not enrich for permanent integrations of the Cre plasmid itself. The idea is to have Cre initially expressed to delete the selectable marker then allow for the transient release of the Cre plasmid over time, opening up the opportunity to serially tag genes using one selectable marker.
    ▲CRITICAL – Retain cell pools from each step of the editing process. Proceed to genotype the WT, selected pool, and Cre pool to acertain the activity of Cre recombinase on your tagged alleles

? TROUBLESHOOTING

35. Proceed to clonal expansion and/or validation.

Cell Culture: Clonal Expansion ● TIMING ~ 14 d

36. Seed ~300-500 cells from the selected pool of edited ('tagged') cells into a 15 cm dish.
37. Allow 7 – 10 days for the single cells to form distinct colonies.
38. Use a cloning cylinder to isolate the clones then trypsinize and propagate the clones.
    ▲CRITICAL – Alternate approaches to clonal expansion can work here including limited dilution cloning as well as FACS sorting single cells.

Validation: Genotyping ● TIMING ~ 2d

39. Extract genomic DNA from wild-type and tagged cells by adding ~30 µL of the QuickExtract DNA Extraction Solution ~50, 000 cells.
40. Using primers that anneal outside of the homology arms, PCR-amplify the insertion junction from the wild-type and tagged cells.
41. Analyze the PCR products on an agarose gel, following Steps 25c – e. The size of the amplicon from the tagged cell lines should be larger than that from the wild-types.
42. Gel purify the PCR products using the QIAquick Gel Extraction Kit.
43. Sanger sequence the PCR products to ensure that the cassette was integrated correctly and that integration did not disrupt the open reading frame.

Validation: Imaging ● TIMING ~ 2d

44. If the genes were tagged with fluorescent protein cassettes, plate cells into an imaging chamber or dish for live-cell fluorescence microscopy to evaluate protein localization and behavior.
45. Alternatively, tagged cells can be plated onto coverslips, fixed, and then imaged using fluorescence microscopy.
46. Immunostaining with target-specific antibodies and tag-specific antibodies confirms accurate tagging if the signals from the target protein and tag colocalize.
    ▲CRITICAL – Antibodies specific to newer iterations of tags such as ultraID and others that continue to evolve may not always be available. In cases like these, along with the primary tag, a secondary small epitope could be directly fused with the primary tag to aid with detection. Furthermore, gene-specific antibodies can be hard to come by for protein targets with paralogs. Here, the cellular localization of the tagged protein along with the genomic validations is key to ascertaining if the gene tagging was successful.
    ? TROUBLESHOOTING

Validation: Immunoblotting ● TIMING ~ 2d

47. Plate ~500,000 – 1,000,000 (80-100% next day confluence) wild-type and tagged cells into each well of a 6-well plate.
48. Collect the cells, extract cell lysates, and prepare them for SDS-PAGE electrophoresis and immunoblotting as summarized by this protocol[56]I.
49. Transfer the protein from the gel to the membrane.
50. Immunoblot using antibodies that are specific for the target protein and the integrated tag.

    ▲CRITICAL – Antibody selections apply here the same way they apply in the imaging validation.

Timing

Steps 1-9, sgRNA selection, homology arm design: 2 hours.
Steps 10-27, cloning sgRNA plasmid, cloning repair plasmid: 2-4 days, concurrently.
Steps 28-38, delivery of editing reagents, selection enrichment, optional Cre selectable marker excision, optional clonal expansion: 2 – 5 weeks, depending on guide efficiency, selection used, and optional steps taken.
Steps 39-50, genomic, imaging, and proteomic validation: 2 days, concurrently.

Anticipated Results

Editing and enrichment of fluorescent knock-ins with qTAG cassettes

For fluorescent gene-tagging, we demonstrate the effectiveness of enriching for edits with mammalian selectable markers in human cells (Fig. 7). Targeting the C-terminus encoding region of histone protein H2BC11, we sought to insert a moxGFP-Puro cassette prior to the stop codon of the H2BC11 gene (Fig. 7a). A moxGFP-Puro qTAG repair was constructed with both long homology arms and short homology arms to demonstrate that both mechanisms could be leveraged for insertion and enrichment in HEK293T cells (Fig. 7b). When surveying cells for green fluorescence, as an indicator of cassette integration, we observed a small proportion of cells in the initial pool expressing the H2B-moxGFP fusion when tagged via HDR and MMEJ (Fig. 7c,d). After complete selection with puromycin, a strong enrichment was observed. This was indicated by the shift in the distribution of cells displaying green fluorescence that generated via HDR and MMEJ. The nuclear localization of H2B-moxGFP edited cells further indicated successful integration and correct targeting of the endogenous fluorescent fusion protein (Fig. 7e). Finally, genomic junction PCR using primers that anneal outside the homology arms confirms the presence of tagged alleles via a shift in band size proportional to the qTAG-moxGFP-Puro cassette from the untagged allele (Fig. 7f).

We additionally test this tagging workflow across human cell lines that vary in ploidy and background, using HAP1, ARPE-19, and U-2OS (Ex. Fig. 7a-d). Although varying in their initial

rates of tagging, we consistently see an enrichment of accurately tagged cells via HDR and MMEJ after selection with puromycin. Testing alternative mammalian selection markers in blasticidin and zeocin in addition to puromycin display that they all achieve the same effect in enriching for cells with tagged alleles (Ex. Fig. 8e-f). Finally, we test the ability to introduce these insertions into H9 stem cells. Repair constructs were generated to insert a qTAG-Blast-mScarlet into the N-terminus encoding region of the tubulin gene TUBA1B. For editing, electroporation with precomplexed RNPs containing purified SpCas9 and sgRNAs were used for delivery (Ex. Fig. 7g). After delivery, selection-enrichment, and clonal expansion mScarlet-TUBA1B expressing H9 cells were observed (Fig Ex. Fig. 7h).

Selectable marker gene excision with the Cre recombinase and serial gene-tagging

Next, we wanted to demonstrate the ability to recover antibiotic sensitivity and subsequently exploit a single mammalian resistance to allow for serial gene-tagging. To do this, we aimed to target two genes: the tubulin protein TUBB4B for a C-terminus encoding insertion of a qTAG-mNeon-Blast cassette and the histone protein H3C2 for a C-terminus encoding insertion of a qTAG-miRFP670nano3-Blast cassette (Fig. 8a). We began with TUBB4B and carried out the editing in ARPE-19 cells. After blasticidin selection and recovery, the selected pool was transiently transfected with a plasmid encoding Cre-2A-Puro. The cells were then selected with puromycin to form a Cre-edited pool from which clonal expansion was carried out (Fig. 8b). Genomic junction PCR from cell lines representing the various phases of cell-line editing confirms the presence of tagged alleles at the TUBB4B locus in the initial tagged pool (Fig 9c). A band shift, proportional to the loss of the blasticidin sequence, was noticed in the Cre-edited and clonal cell lines suggesting the deletion of the blasticidin gene via the Cre recombinase. To further confirm this, the TUBB4B locus was amplified from the TUBB4B-mNeon clonal cell line and Sanger sequenced. Due to the use of mutant lox sites, lox71 and lox66, if the sequence within these elements was excised correctly, a mutant, low binding lox71/66 site should form. Sanger sequencing confirmed the presence of this mutant lox site within the clonal ARPE-19 TUBB4B-mNeon cell line. We then demonstrate the cell's restored sensitivity to blasticidin after excision by using selection assays and staining with crystal violet. We confirm that a significant portion of cells in the Cre-edited and Cre clonal cell lines regained sensitivity to blasticidin (Fig. 8d). Fluorescent imaging further confirms the localization and targeting of TUBB4B-mNeon across each of the cell lines generated during the editing phases (Fig. 8e). After the marker deletion, we exploit the sensitivity of the APRE-19 TUBB4B-mNeon clonal cell line to subsequently tag the histone protein H3C2 with miRFP670nano3-Blast (Fig. 8f,g).

Non-fluorescent gene-tagging applications and validation

Finally, we provide examples of gene-tagging with qTAG cassettes harboring useful functional proteomic tags. Insertions using these tags are traditionally difficult to isolate due to

low efficiency and not having fluorescent sorting options for isolation, but by coupling it to a selectable marker, the process remains just as efficient as tagging fluorescent proteins. Here we establish and validate the tagging scheme of the nuclear membrane protein LMNB1 using a variety of non-fluorescent tags, including the proximity biotin ligases miniTurbo and ultraID, the targeted degradation tag dTAG, and epitope tags FLAG, HA, and V5 (Fig. 9a). Using HAP1 cells, we followed the experimental scheme of editing, enriching with selection, Cre-excision, and clonal expansion to isolate monogenic tagged lines for proteomic evaluation (Fig. 9b). We tested the functionality of two biotin ligases (miniTurbo and ultraID) fused to endogenous LNMB1 to capture local interactions of the target protein. Imaging of HAP1 cells expressing V5-miniTurbo-LMNB1 and V5-ultraID-LNMB1 revealed robust nuclear membrane localization upon treatment with biotin and detection via streptavidin after 1 hour (Fig. 9c). Immunoblots using a streptavidin antibody and antibodies against V5 and LMNB1 confirm further confirm the biotinylation of proximal proteins and capture of potential protein interactions within just 10 minutes, with robust biotinylation being detected after 1 and 12 hours (Fig. 9d). For targeted protein degradation, we opted to tag LMNB1 with dTAG. Imaging of HAP1 cells expressing V5-dTAG - LMNB1 and stained with anti-V5 and anti-LMNB1 antibodies revealed that V5-dTAG-LMNB1 is localized to the nuclear membrane (Fig. 10e). This signal was abolished in cells treated with dTAG-13; the small molecule responsible for inducing degradation (Fig. 10e). Immunoblots confirmed the time-dependent degradation of V5-dTAG-LMNB1 in cells treated with dTAG-13 (Fig. 10f). Finally, for epitope tagging, the general epitopes FLAG, HA, and V5 were leveraged the tagging of LMNB1. Immunofluorescence of HAP1 cells stained with epitope-specific antibodies showed that 3xFLAG-LMNB1, 3xHA-LMNB1, and V5-LMNB1 display nuclear membrane localization (Fig. 10g). Additionally, immunoblotting with epitope-specific and target-specific antibodies saw robust detection of the endogenous fusion protein confirming editing (Fig. 10d).

Acknowledgements

We thank members of the Pelletier Lab for their scientific feedback during the project. R.P. was partially funded by the Lunenfeld Tanenbaum Research Institute Studentships at Sinai Health System. A.S. and A.C.E are partially funded by a CIHR Post Doctoral Grant (#187836 and #181763, respectively). H.D.M.W is funded by a CIHR Project Grant (#156297). H.D.M.W is a Tier II Canada Research Chair in Mechanisms of Genome Stability (#950-231487). The remainder of this work was funded by CIHR Foundation (FDN # 167279) and Krembil Foundation grants to L.P. L.P. is a Tier 1 Canada Research Chair in Centrosome Biogenesis and Function. This work was supported by the Lunenfeld-Tanenbaum Research Institute Flow Cytometry Core and the by The Network Biology Collaborative Centre which are funded through Canada Foundation for Innovation, the Ontario Government, and Genome Canada and Ontario Genomics (OGI-139). The schematics and models found in this manuscript were created in part by using BioRender.

Author contributions

R.P. and A.S. developed the reagents and protocol. R.P. and A.S. designed the experiments. R.P., A.S., L.M., A.C.E., W-H.H., and J.M.T. performed the experiments. H.D.M.W prepared the purified Cas9. L.P supervised the project. R.P. and L.P. wrote the protocol. R.P., A.S., L.M., A.C.E., W-H.H., J.M.T., H.D.M.W., and L.P. revised the manuscript.

Competing Interests

The authors declare no competing interests.

Related Links

Key reference using this protocol:

Lequyer, J., Philip, R., Sharma, A., Hsu, W. H. & Pelletier, L. A fast blind zero-shot denoiser. Nat Mach Intell 4, 953-963 (2022): 10.1038/s42256-022-00547-8

References

1 Farhan, H., Weiss, M., Tani, K., Kaufman, R. J. & Hauri, H. P. Adaptation of endoplasmic reticulum exit sites to acute and chronic increases in cargo load. *EMBO J* **27**, 2043-2054 (2008). https://doi.org:10.1038/emboj.2008.136

2 Vavouri, T., Semple, J. I., Garcia-Verdugo, R. & Lehner, B. Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity. *Cell* **138**, 198-208 (2009). https://doi.org:10.1016/j.cell.2009.04.029

3 Makanae, K., Kintaka, R., Makino, T., Kitano, H. & Moriya, H. Identification of dosage-sensitive genes in Saccharomyces cerevisiae using the genetic tug-of-war method. *Genome Res* **23**, 300-311 (2013). https://doi.org:10.1101/gr.146662.112

4 Tang, Y. C. & Amon, A. Gene copy-number alterations: a cost-benefit analysis. *Cell* **152**, 394-405 (2013). https://doi.org:10.1016/j.cell.2012.11.043

5 Bolognesi, B. *et al.* A Concentration-Dependent Liquid Phase Separation Can Cause Toxicity upon Increased Protein Expression. *Cell Rep* **16**, 222-231 (2016). https://doi.org:10.1016/j.celrep.2016.05.076

6 Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012). https://doi.org:10.1126/science.1225829

7 Paques, F. & Haber, J. E. Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiol Mol Biol Rev* **63**, 349-404 (1999). https://doi.org:10.1128/MMBR.63.2.349-404.1999

8 Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* **11**, 636-646 (2010). https://doi.org:10.1038/nrg2842

9 Moore, J. K. & Haber, J. E. Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in Saccharomyces cerevisiae. *Mol Cell Biol* **16**, 2164-2173 (1996). https://doi.org:10.1128/MCB.16.5.2164

10 Bibikova, M., Golic, M., Golic, K. G. & Carroll, D. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics* **161**, 1169-1175 (2002). https://doi.org:10.1093/genetics/161.3.1169

11 Chang, H. H. Y., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol* **18**, 495-506 (2017). https://doi.org:10.1038/nrm.2017.48

12 Hart, T. *et al.* High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* **163**, 1515-1526 (2015). https://doi.org:10.1016/j.cell.2015.11.015

13 Schmid-Burgk, J. L., Honing, K., Ebert, T. S. & Hornung, V. CRISPaint allows modular base-specific gene tagging using a ligase-4-dependent mechanism. *Nat Commun* **7**, 12338 (2016). https://doi.org:10.1038/ncomms12338

14 He, X. *et al.* Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Res* **44**, e85 (2016). https://doi.org:10.1093/nar/gkw064

15 Lackner, D. H. *et al.* A generic strategy for CRISPR-Cas9-mediated gene tagging. *Nat Commun* **6**, 10237 (2015). https://doi.org:10.1038/ncomms10237

16 Jasin, M. & Rothstein, R. Repair of strand breaks by homologous recombination. *Cold Spring Harb Perspect Biol* **5**, a012740 (2013). https://doi.org:10.1101/cshperspect.a012740

17 Sakuma, T., Nakade, S., Sakane, Y., Suzuki, K. T. & Yamamoto, T. MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nat Protoc* **11**, 118-133 (2016). https://doi.org:10.1038/nprot.2015.140

18 Sfeir, A. & Symington, L. S. Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? *Trends Biochem Sci* **40**, 701-714 (2015). https://doi.org:10.1016/j.tibs.2015.08.006

19 Fueller, J. *et al.* CRISPR-Cas12a-assisted PCR tagging of mammalian genes. *J Cell Biol* **219** (2020). https://doi.org:10.1083/jcb.201910210

20 Artegiani, B. *et al.* Fast and efficient generation of knock-in human organoids using homology-independent CRISPR-Cas9 precision genome editing. *Nat Cell Biol* **22**, 321-331 (2020). https://doi.org:10.1038/s41556-020-0472-5

21 Roberts, B. *et al.* Systematic gene tagging using CRISPR/Cas9 in human stem cells to illuminate cell organization. *Mol Biol Cell* **28**, 2854-2874 (2017). https://doi.org:10.1091/mbc.E17-03-0209

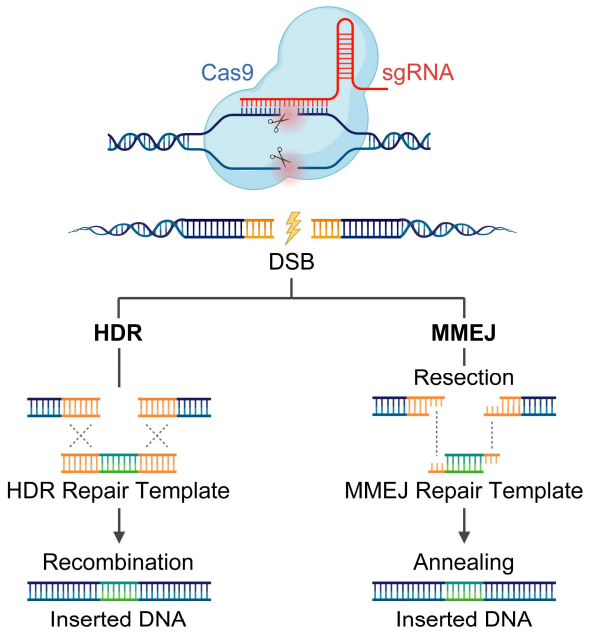
22 Perez-Leal, O. *et al.* Multiplex Gene Tagging with CRISPR-Cas9 for Live-Cell Microscopy and Application to Study the Role of SARS-CoV-2 Proteins in Autophagy, Mitochondrial Dynamics, and Cell Growth. *CRISPR J* **4**, 854-871 (2021). https://doi.org:10.1089/crispr.2021.0041

23 Supharattanasitthi, W., Carlsson, E., Sharif, U. & Paraoan, L. CRISPR/Cas9-mediated one step bi-allelic change of genomic DNA in iPSCs and human RPE cells in vitro with dual antibiotic selection. *Sci Rep* **9**, 174 (2019). https://doi.org:10.1038/s41598-018-36740-2

24 Lin, D. W. *et al.* Microhomology-based CRISPR tagging tools for protein tracking, purification, and depletion. *J Biol Chem* **294**, 10877-10885 (2019). https://doi.org:10.1074/jbc.RA119.008422

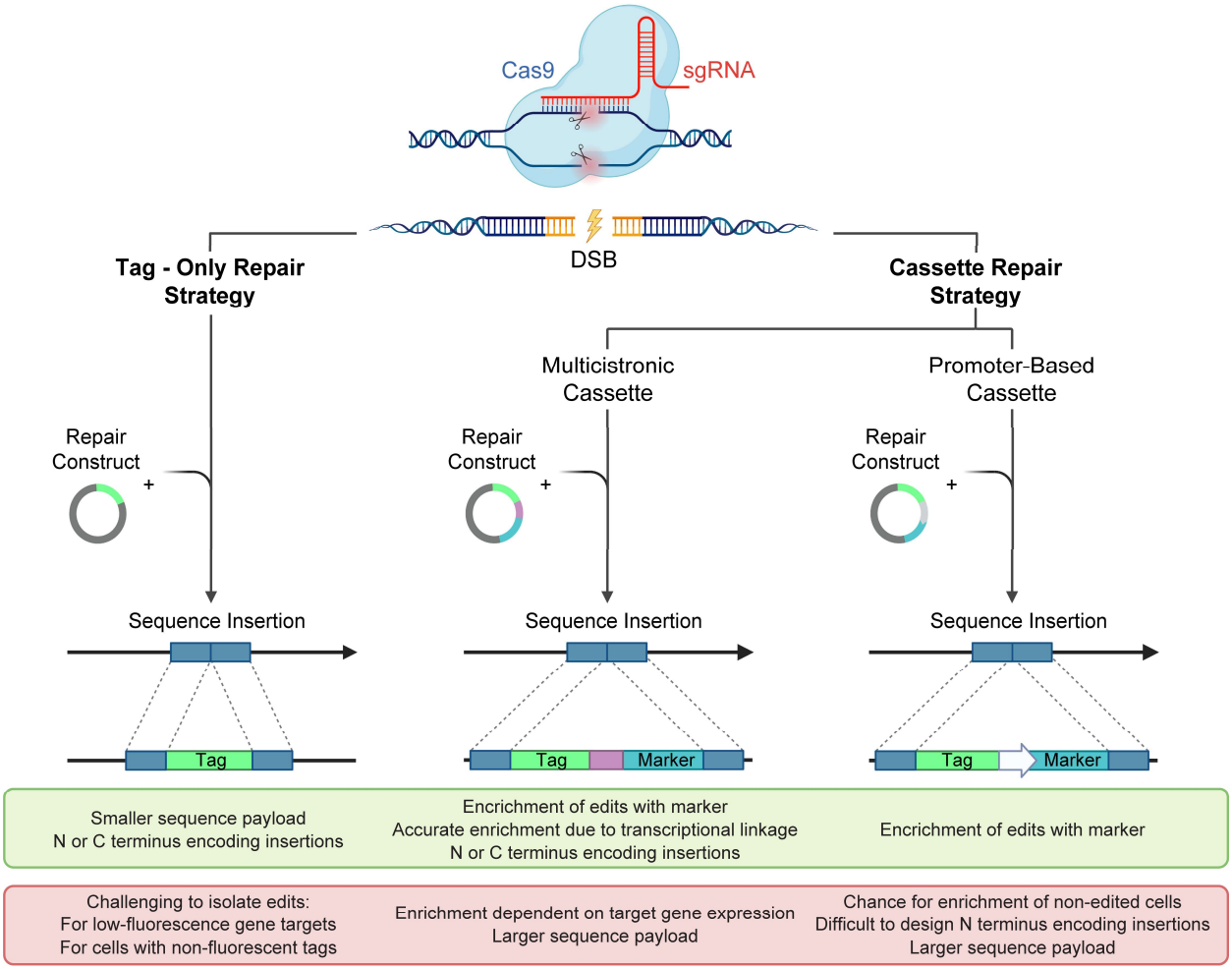
25 Nabet, B. *et al.* The dTAG system for immediate and target-specific protein degradation. *Nat Chem Biol* **14**, 431-441 (2018). https://doi.org:10.1038/s41589-018-0021-8
26 Nakamae, K. *et al.* Establishment of expanded and streamlined pipeline of PITCh knock-in - a web-based design tool for MMEJ-mediated gene knock-in, PITCh designer, and the variations of PITCh, PITCh-TG and PITCh-KIKO. *Bioengineered* **8**, 302-308 (2017). https://doi.org:10.1080/21655979.2017.1313645
27 Pelletier, J. & Sonenberg, N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**, 320-325 (1988). https://doi.org:10.1038/334320a0
28 Jang, S. K. *et al.* A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol* **62**, 2636-2643 (1988). https://doi.org:10.1128/JVI.62.8.2636-2643.1988
29 Kim, J. H. *et al.* High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* **6**, e18556 (2011). https://doi.org:10.1371/journal.pone.0018556
30 Luke, G. A. *et al.* Occurrence, function and evolutionary origins of '2A-like' sequences in virus genomes. *J Gen Virol* **89**, 1036-1042 (2008). https://doi.org:10.1099/vir.0.83428-0
31 Lequyer, J., Philip, R., Sharma, A., Hsu, W. H. & Pelletier, L. A fast blind zero-shot denoiser. *Nat Mach Intell* **4**, 953-963 (2022). https://doi.org:10.1038/s42256-022-00547-8
32 Sauer, B. & Henderson, N. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc Natl Acad Sci U S A* **85**, 5166-5170 (1988). https://doi.org:10.1073/pnas.85.14.5166
33 Albert, H., Dale, E. C., Lee, E. & Ow, D. W. Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. *Plant J* **7**, 649-659 (1995). https://doi.org:10.1046/j.1365-313x.1995.7040649.x
34 Zhang, Z. & Lutz, B. Cre recombinase-mediated inversion using lox66 and lox71: method to introduce conditional point mutations into the CREB-binding protein. *Nucleic Acids Res* **30**, e90 (2002). https://doi.org:10.1093/nar/gnf089
35 Shaner, N. C. *et al.* A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. *Nat Methods* **10**, 407-409 (2013). https://doi.org:10.1038/nmeth.2413
36 Costantini, L. M. *et al.* A palette of fluorescent proteins optimized for diverse cellular environments. *Nat Commun* **6**, 7670 (2015). https://doi.org:10.1038/ncomms8670
37 Mo, G. C. H., Posner, C., Rodriguez, E. A., Sun, T. & Zhang, J. A rationally enhanced red fluorescent protein expands the utility of FRET biosensors. *Nat Commun* **11**, 1848 (2020). https://doi.org:10.1038/s41467-020-15687-x
38 Bindels, D. S. *et al.* mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nat Methods* **14**, 53-56 (2017). https://doi.org:10.1038/nmeth.4074
39 Oliinyk, O. S. *et al.* Single-domain near-infrared protein provides a scaffold for antigen-dependent fluorescent nanobodies. *Nat Methods* **19**, 740-750 (2022). https://doi.org:10.1038/s41592-022-01467-6
40 Kubitz, L. *et al.* Engineering of ultraID, a compact and hyperactive enzyme for proximity-dependent biotinylation in living cells. *Commun Biol* **5**, 657 (2022). https://doi.org:10.1038/s42003-022-03604-5

41 Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol* **36**, 880-887 (2018). https://doi.org:10.1038/nbt.4201
42 Koch, B. *et al.* Generation and validation of homozygous fluorescent knock-in cells using CRISPR-Cas9 genome editing. *Nat Protoc* **13**, 1465-1487 (2018). https://doi.org:10.1038/nprot.2018.042
43 Roberts, B. *et al.* Fluorescent Gene Tagging of Transcriptionally Silent Genes in hiPSCs. *Stem Cell Reports* **12**, 1145-1158 (2019). https://doi.org:10.1016/j.stemcr.2019.03.001
44 Ishibashi, A. *et al.* A simple method using CRISPR-Cas9 to knock-out genes in murine cancerous cell lines. *Sci Rep* **10**, 22345 (2020). https://doi.org:10.1038/s41598-020-79303-0
45 Liang, X., Potter, J., Kumar, S., Ravinder, N. & Chesnut, J. D. Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. *J Biotechnol* **241**, 136-146 (2017). https://doi.org:10.1016/j.jbiotec.2016.11.011
46 Zhong, H. *et al.* High-fidelity, efficient, and reversible labeling of endogenous proteins using CRISPR-based designer exon insertion. *Elife* **10** (2021). https://doi.org:10.7554/eLife.64911
47 Cho, N. H. *et al.* OpenCell: Endogenous tagging for the cartography of human cellular organization. *Science* **375**, eabi6983 (2022). https://doi.org:10.1126/science.abi6983
48 Willems, J. *et al.* ORANGE: A CRISPR/Cas9-based genome editing toolbox for epitope tagging of endogenous proteins in neurons. *PLoS Biol* **18**, e3000665 (2020). https://doi.org:10.1371/journal.pbio.3000665
49 Bollen, Y. *et al.* Efficient and error-free fluorescent gene tagging in human organoids without double-strand DNA cleavage. *PLoS Biol* **20**, e3001527 (2022). https://doi.org:10.1371/journal.pbio.3001527
50 Liu, Z. *et al.* Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep* **7**, 2193 (2017). https://doi.org:10.1038/s41598-017-02460-2
51 UniProt, C. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* **51**, D523-D531 (2023). https://doi.org:10.1093/nar/gkac1052
52 Labun, K. *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res* **47**, W171-W174 (2019). https://doi.org:10.1093/nar/gkz365
53 Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* **42**, e168 (2014). https://doi.org:10.1093/nar/gku936
54 Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823 (2013). https://doi.org:10.1126/science.1231143
55 Kwart, D., Paquet, D., Teo, S. & Tessier-Lavigne, M. Precise and efficient scarless genome editing in stem cells using CORRECT. *Nat Protoc* **12**, 329-354 (2017). https://doi.org:10.1038/nprot.2016.171
56 Mahmood, T. & Yang, P. C. Western blot: technique, theory, and trouble shooting. *N Am J Med Sci* **4**, 429-434 (2012). https://doi.org:10.4103/1947-2714.100998



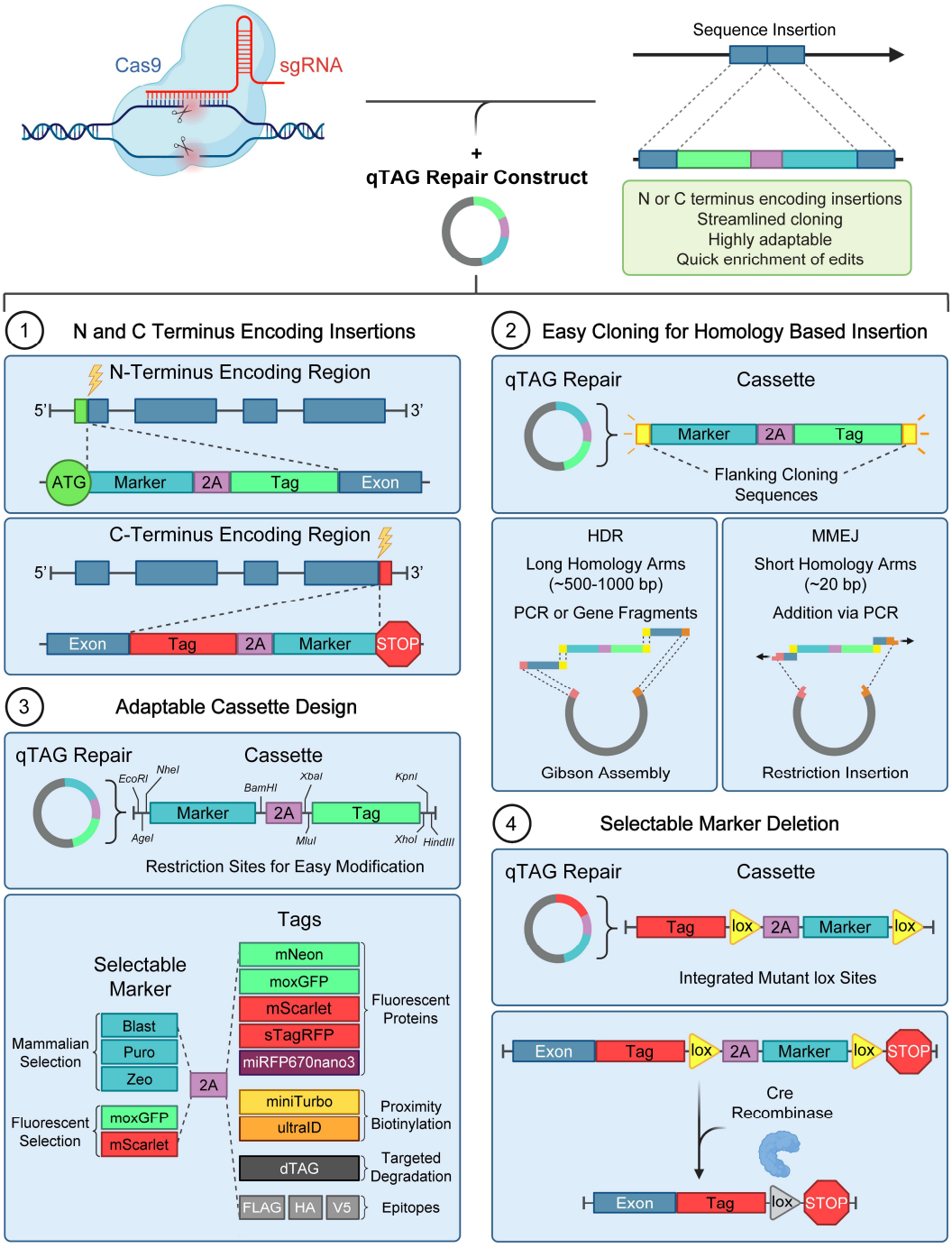
Philip et al. Figure 1

Figure 1. DNA repair mechanisms leveraged for precise endogenous tagging. Two homology-based pathways, homology-directed repair (HDR) and microhomology-mediated end joining (MMEJ), are depicted here. HDR is dependent on long stretches of homology flanking the insertion site within the donor DNA and acts via homologous recombination to insert the desired sequence. MMEJ requires short microhomology sequences that flank the insertion site within the donor DNA. A limited initial resection first occurs to reveal the homologies between the DNA and donor followed by annealing of the donor to the target site for successful insertion.



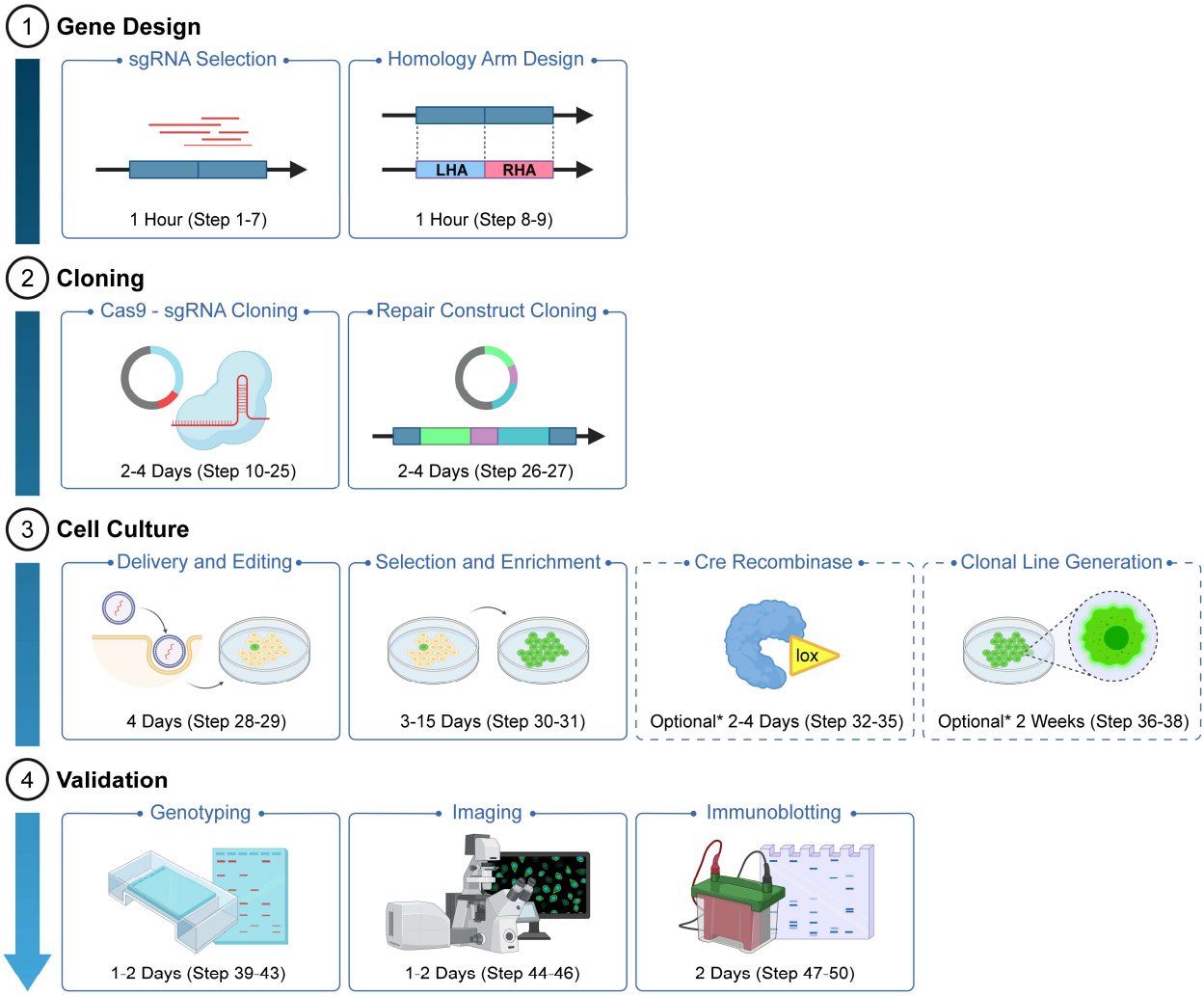
Philip et al. Figure 2

Figure 2. Comparing knock-in strategies with different cassette designs. An overview of the advantages (green) and limitations (red) of tag-only and cassette-based knock-ins including multicistronic and promoter-based designs.



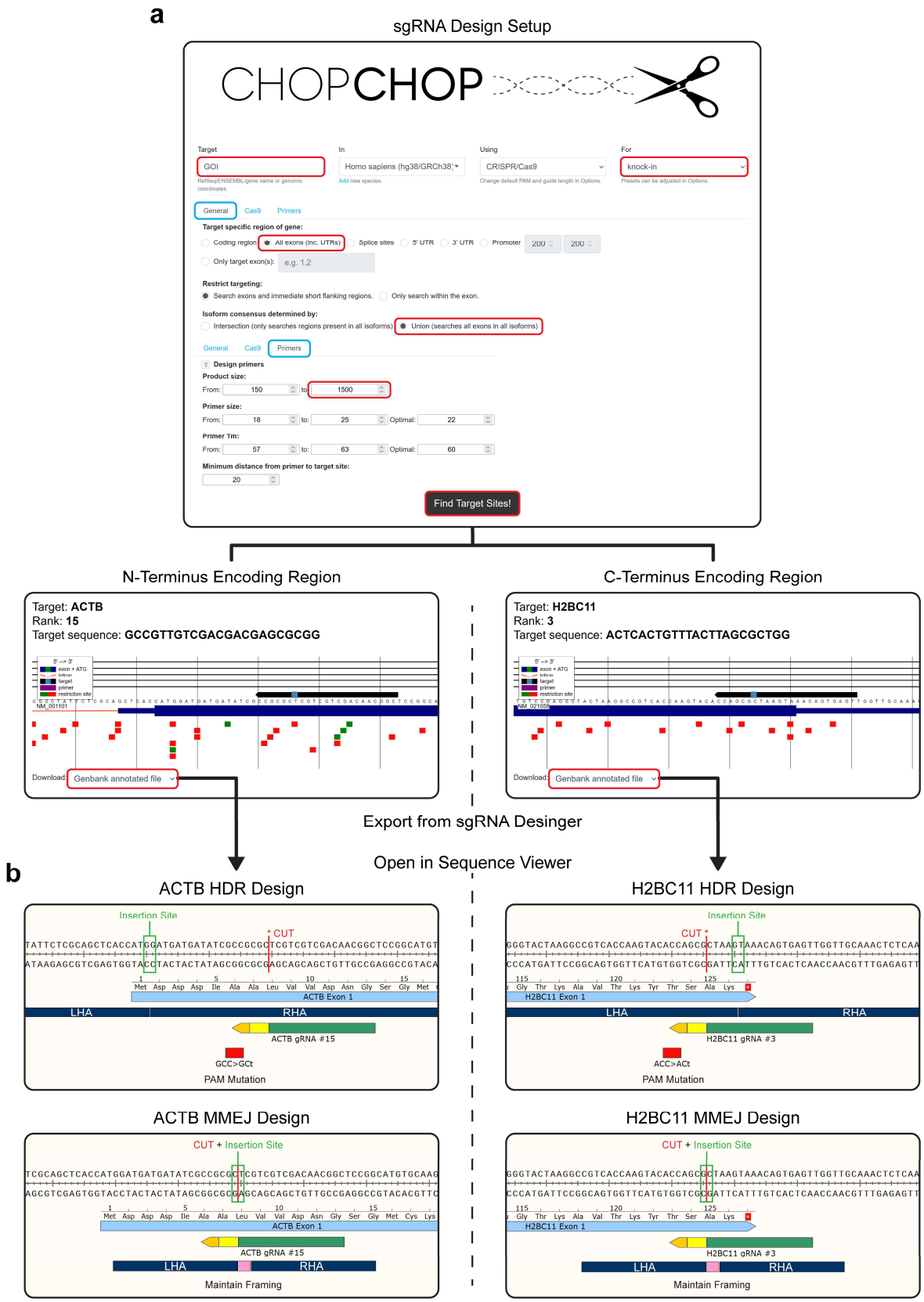
Philip et al. Figure 3

Figure 3. Features of qTAG cassettes. Four key features of the qTAG cassettes are depicted here. (1) 5' and 3' designs for N- and C-terminus encoding insertions; (2) the incorporation of universal cloning adaptors to streamline the cloning of long or short homology arms for HDR or MMEJ-based repair; (3) the inclusion of restriction sites near each genetic element in the cassette to conveniently swap tags or selectable markers; (4) the addition of lox sites on either side of the selectable marker to enable the marker gene removal post-integration.



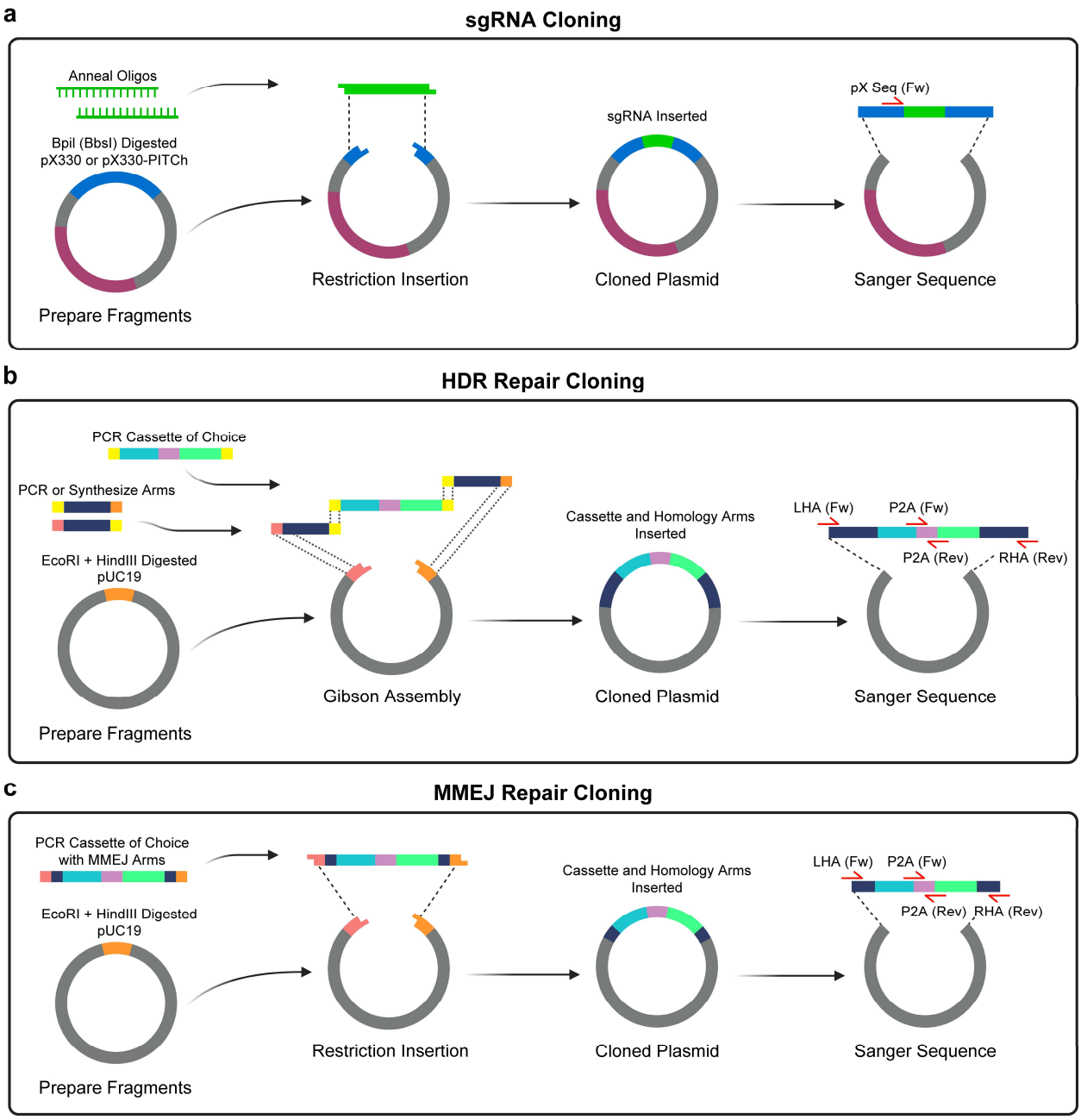
Philip et al. Figure 4

Figure 4. Overview of the endogenous tagging workflow. The process involves four stages: (1) gene design; (2) cloning; (3) cell culture; (4) validation. Each of these stages contains required and optional steps and considerations. Expected timeframes are indicated.



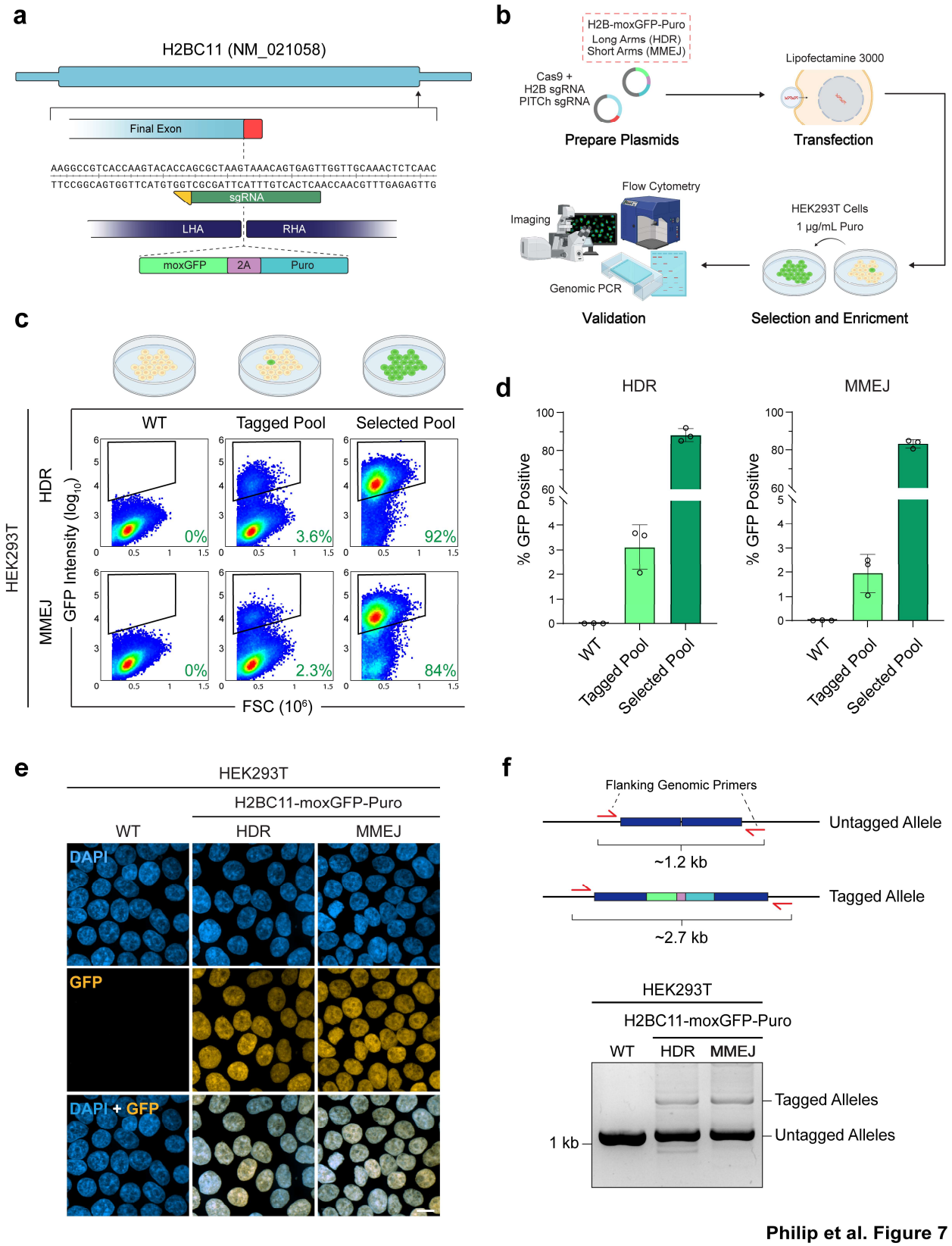
Philip et al. Figure 5

Figure 5. sgRNA and homology arm design for endogenous tagging. (a) The CHOPCHOP v3 home page is depicted where optional tabs of the interface are marked with blue rectangles. Additionally, important settings to maximize your sgRNA selection and allow for easier export and later annotation are highlighted with red rectangles. Examples of sgRNAs for N-terminus ACTB tagging and C-terminus H2BC11 tagging are depicted. **(b)** After selecting a sgRNA, the sequence can be exported from CHOPCHOP and opened in DNA sequence viewer such as SnapGene for further annotation and design. Showcased are homology arm design examples leveraging both HDR and MMEJ. These designs display the insertion site, annotations for the left homology arm (LHA) and the right homology arm (RHA) and indicate the final reading frame and sgRNA mutation necessary to enable successful endogenous tagging.



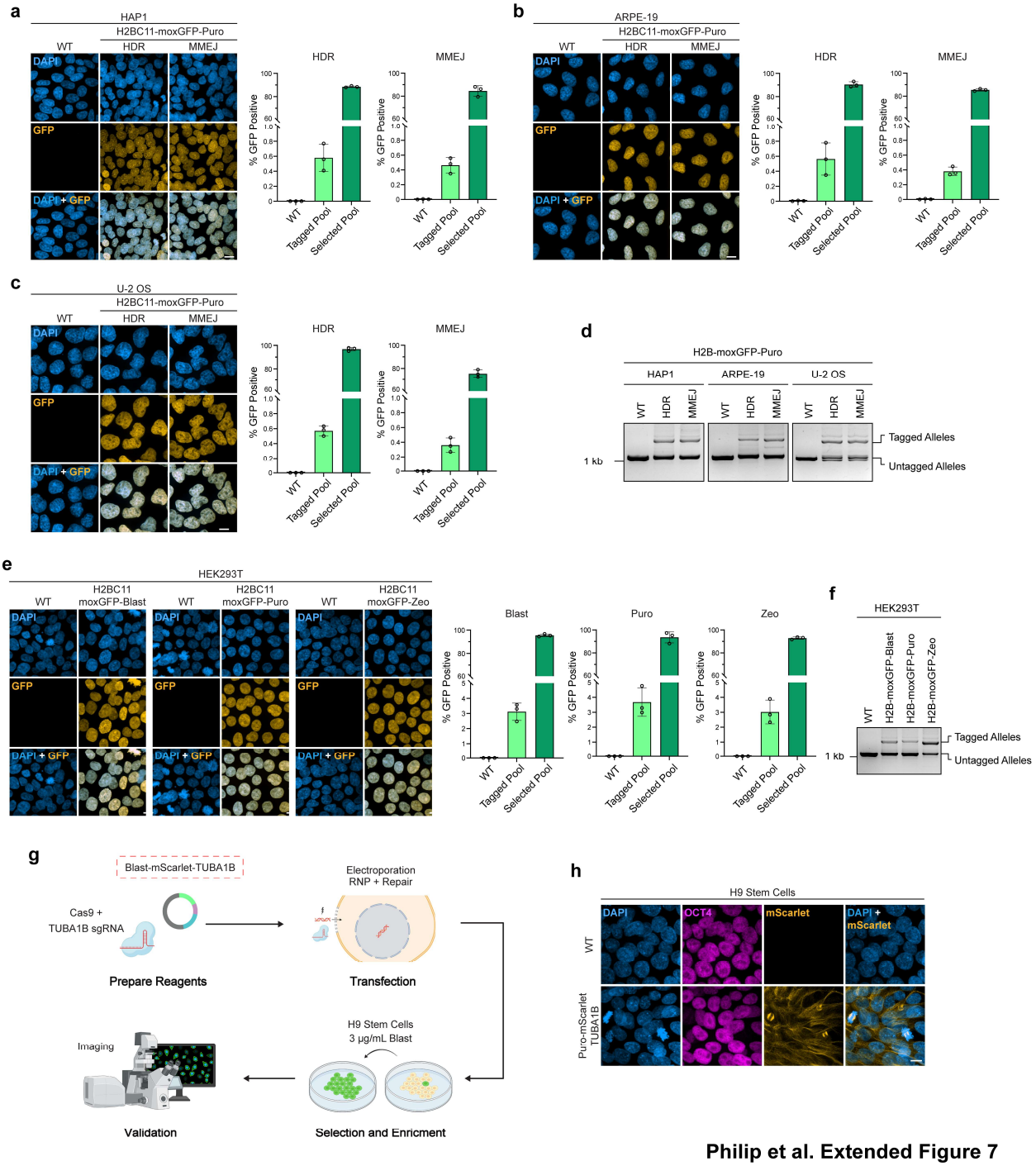
Philip et al. Figure 6

Figure 6. Visual cloning overview. A visual representation of the cloning procedure for CRISPR sgRNA cloning into Cas9 encoded backbones pX330 or pX330-PITCh **(a)**, and donor construction for HDR **(b)** or MMEJ **(c)**.



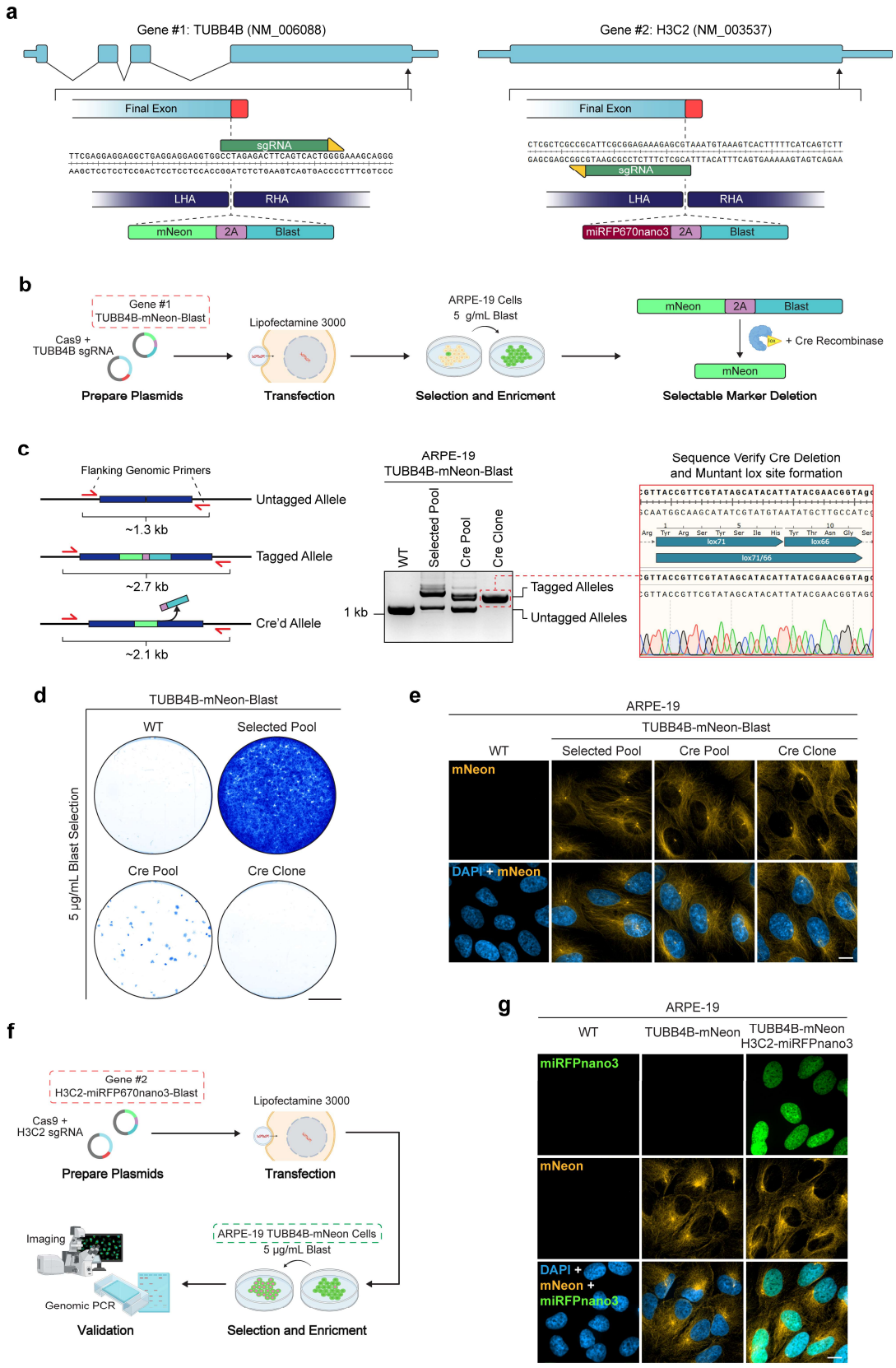
Philip et al. Figure 7

Figure 7. Editing, enrichment, and validation of fluorescent knock-ins with qTAG cassettes. (a) Schematic gene design of the H2BC11 gene displaying the binding of the sgRNA (dark green) and its PAM binding site (yellow), homology arms, and 3' targeted insertion site. **(b)** An overview of the strategy used to carry out the endogenous tagging of H2BC11 with a C-terminus encoding insertion of the qTAG-moxGFP-Puro cassette via HDR (long homology arms) and MMEJ (short homology arms). **(c)** Representative flow-cytometry plots depicting 200,000 cells from the following samples: untagged wild-type, the initial tagged H2BC11-GFP-Puro pool, and the final H2BC11-GFP-Puro selected pools of HEK293T cells obtained through both HDR and MMEJ mechanisms. Gates indicate the threshold of fluorescent green cells within the total population as an indicator of positive editing. **(d)** Flow cytometry quantifications of GFP-positive cells, based on three distinct biological replicates with each measurement encompassing 200,000 cells. **(e)** Representative images of WT, HDR-edited H2BC11-moxGFP-Puro, and MMEJ-edited H2BC11-moxGFP-Puro HEK293T cells co-stained with DAPI. (**f**) An agarose gel displaying the bands produced by genomic PCR with primers located outside the homology arms, probing for locus-specific integration of the qTAG cassette in HEK293T cells. Scale bars: 10 µm.



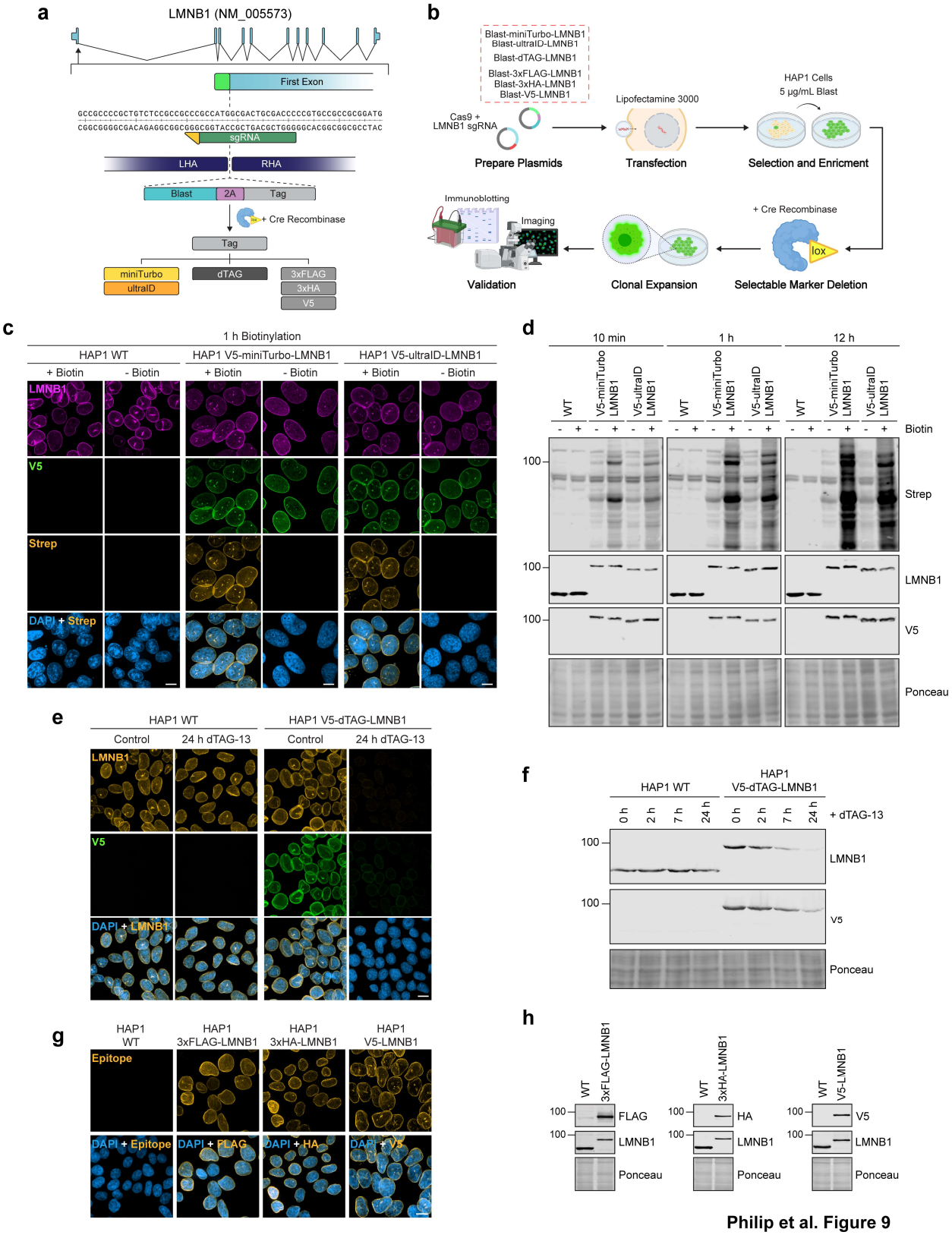
Philip et al. Extended Figure 7

Extended Figure 7. Editing, enrichment, and validation of fluorescent knock-ins with qTAG cassettes. (a, b, c) Representative images of WT, HDR-integrated, and MMEJ-integrated H2BC11-moxGFP-Puro tagged HAP1, ARPE-19, and U-2 OS cells co-stained with DAPI (Left). Flow cytometry quantifications of GFP-positive cells, based on three distinct biological replicates with each measurement encompassing 200,000 cells (Right). **(d)** Genomic PCR outside the homology arms probing for locus-specific integration of the qTAG cassette in HAP1, ARPE-19, and U-2 OS cells. **(e)** Representative images of WT, selected H2BC11-moxGFP-Blast, selected H2BC11-moxGFP-Puro, and selected H2BC11-moxGFP-Zeo HEK293T cells co-stained with DAPI (Left). Flow cytometry quantifications of GFP-positive cells, based on three distinct biological replicates with each measurement encompassing 200,000 cells (Right). **(f)** Genomic PCR outside the homology arms probing for locus-specific integration of the qTAG cassettes with alternative mammalian selectable markers in HEK293T cells. (g) Overview of an alternate strategy using electroporation and RNPs to edit the tubulin TUBA1B gene with a qTAG-Blast-mScarlet cassette in H9 stem cells. **(h)** Representative images of WT and Blast-mScarlet-TUBA1B H9 cells co-stained with DAPI and probed for pluripotency marker OCT4. Scale bars: 10 µm.



Philip et al. Figure 8

Figure 8. Selectable marker gene excision via the Cre recombinase and serial endogenous tagging. (a) Schematic gene designs of TUBB4B and H3C2. Shown are the location of the sgRNA (dark green) and its PAM binding site (yellow), homology arms, and targeted insertion site. **(b)** Overview of the strategy used to carry out the tagging of TUBB4B with a C-terminus encoding insertion of a qTAG-mNeon-Blast cassette. **(c)** An agarose gel displaying bands produced by genomic PCR with primers located outside the homology arms, probing for locus-specific integration of the qTAG cassette and Cre-edited alleles in ARPE-19 cells. Sanger sequencing of a homozygous clonal line displayed accurate excision of the marker gene as displayed by the formation of the inactive lox mutant site post Cre-editing. **(d)** Representative images of crystal violet-stained plate wells containing WT, selected pool, Cre pool, and Cre clonal cell lines of TUBB4B-mNeon APRE-19 cells selected with blasticidin. Scale bars: 1 cm. **(e)** Representative images of WT, selected pool, Cre pool, and Cre clonal cell lines of TUBB4B-mNeon APRE-19 cells co-stained with DAPI. Scale bars: 10 µm. **(f)** Overview of the strategy to carry out subsequent tagging of H3C2 with a C-terminus encoding insertion of qTAG-miRFP670nano3-Blast within the previously edited TUBB4B-mNeon ARPE-19 cells. **(g)** Representative images of WT, tagged TUBB4B-mNeon, and dual tagged TUBB4B-mNeon and H3C2-miRFP670nano3 ARPE-19 cells co-stained with DAPI. Scale bars: 10 µm.



Philip et al. Figure 9

Figure 9. Non-fluorescent gene qTAG applications and validation. (a) Schematic gene design of LMNB1 displaying the binding of the sgRNA (dark green) and its PAM binding site (yellow), homology arms, and targeted insertion site. **(b)** Overview of the strategy to carry out tagging of LMNB1 with a variety of functional proteomic qTAG cassettes. **(c)** Representative images of WT, V5-miniTurbo, and V5-ultraID tagged LMNB1 in HAP1 cells. Cells were stained with DAPI and probed with streptavidin in the absence or presence of biotin for 1 hour. Scale bars: 10 µm. **(d)** A representative immunoblot of WT, V5-miniTurbo, and V5-ultraID tagged LMNB1 in HAP1 cells probed with antibodies against streptavidin, V5, and LMNB1 in the absence or presence of biotynilation for 10 min, 1h, and 12h. **(e)** Representative images of WT and V5-dTAG tagged LMNB1 in HAP1 cells stained for DAPI and probed for V5 in the absence or presence of dTAG-13 for 24h. Scale bars: 10 µm. **(f)** A representative immunoblot of dTAG-V5 tagged LMNB1 in HAP1 cells probed for V5 and LMNB1 in the absence or presence of dTAG-13 for 2, 7 and 24h. **(g)** Representative images of WT, 3xFLAG tagged, 3xHA tagged, and V5 tagged LMNB1 in HAP1 cells stained for DAPI and probed for their respective epitopes. Scale bars: 10 µm. **(h)** Representative immunoblots of WT, 3xFLAG tagged, 3xHA tagged, and V5 tagged LMNB1 in HAP1 cells with their respective epitope-specific antibodies.

Table 1. qTAG relevant sequences and primers

Sequences (5'-3')	Purpose	Description
TTGTAAAACGACGGCCAGTGAATTC	HDR	Nterm LHA Left Gibson Adaptor Sequence
GGCAAGAGCAAGACCGGTAGC	HDR	Nterm LHA Right Gibson Adaptor Sequence
AGGAAGCGGAGGATCTGGCGGTACC	HDR	Nterm RHA Left Gibson Adaptor Sequence
AAGCTTGGCGTAATCATGGTCATAG	HDR	Nterm RHA Right Gibson Adaptor Sequence
TTGTAAAACGACGGCCAGTGAATTC	HDR	Cterm LHA Left Gibson Adaptor Sequence
GGAGGAAGCGGAGGATCTGGCGGTA	HDR	Cterm LHA Right Gibson Adaptor Sequence
CTCGAGACCGGTAAGAGCAAGTGA	HDR	Cterm RHA Left Gibson Adaptor Sequence
AAGCTTGGCGTAATCATGGTCATAG	HDR	Cterm RHA Right Gibson Adaptor Sequence
acagctGAATTCGCATCGTACGCGTACGTG TTTGGNNNNNNNNNNNNNNNNNNNNGGC AAGAGCAAGACCGGTAGC	MMEJ	Nterm MMEJ Fw Cassette Primer (Replace 'N')
agctgtAAGCTTGCATCGTACGCGTACGTG TTTGGNNNNNNNNNNNNNNNNNNNNGGTA CCGCCAGATCCTCCG	MMEJ	Nterm MMEJ Rev Cassette Primer (Replace 'N' - double check reverse strand)
acagctGAATTCGCATCGTACGCGTACGTG TTTGGNNNNNNNNNNNNNNNNNNNNGGA GGAAGCGGAGGATCTGGC	MMEJ	Cterm MMEJ Fw Cassette Primer (Replace 'N')
agctgtAAGCTTGCATCGTACGCGTACGTG TTTGGNNNNNNNNNNNNNNNNNNNNTCAC TTGCTCTTACCGGTCTCG	MMEJ	Cterm MMEJ Rev Cassette Primer (Replace 'N' - double check reverse strand)
GGCAAGAGCAAGACCGGTAGC	Amplification + Sequencing	Nterm Cassette Fw Amplification & Sequencing Primer
GGTACCGCCAGATCCTCCG	Amplification + Sequencing	Nterm Cassette Rev Amplification & Sequencing Primer
GGAGGAAGCGGAGGATCTGGC	Amplification + Sequencing	Cterm Cassette Fw Amplification & Sequencing Primer
TCACTTGCTCTTACCGGTCTCG	Amplification + Sequencing	Cterm Cassette Rev Amplification & Sequencing Primer
GCTACTAACTTCAGCCTGCTGAAGC	Sequencing	P2A Fw Sequencing Primer
AGGTCCAGGGTTCTCCTCCAC	Sequencing	P2A Rev Sequencing Primer
CTGCAAGGCGATTAAGTTGGGTAAC	Sequencing	LHA Fw Sequencing Primer
GGCTCGTATGTTGTGTGGAATTGT	Sequencing	RHA Rev Sequencing Primer
GCCTTTTTACGGTTCCTGGC	Sequencing	pX-line gRNA Sequencing Primer

Table 2. Plasmids for the qTAG system

Plasmid Name	Terminus	Tag	Marker	Plasmid and Map
qTAG-N-Blast-mNeon	N	mNeonGreen	Blasticidin	https://www.addgene.org/207676/
qTAG-N-Blast-moxGFP	N	moxGFP	Blasticidin	https://www.addgene.org/207677/
qTAG-N-Blast-mScarlet	N	mScarlet	Blasticidin	https://www.addgene.org/207678/
qTAG-N-Blast-sTagRFP	N	super-TagRFP	Blasticidin	https://www.addgene.org/207679/
qTAG-N-Blast-miRFP670nano3	N	miRFP670nano3	Blasticidin	https://www.addgene.org/207680/
qTAG-N-Blast-miniTurbo	N	miniTurbo	Blasticidin	https://www.addgene.org/207681/
qTAG-N-Blast-ultraID	N	ultraID	Blasticidin	https://www.addgene.org/207682/
qTAG-N-Blast-dTAG	N	dTAG	Blasticidin	https://www.addgene.org/207683/
qTAG-N-Blast-3xFLAG	N	3xFLAG	Blasticidin	https://www.addgene.org/207684/
qTAG-N-Blast-3xHA	N	3xHA	Blasticidin	https://www.addgene.org/207685/
qTAG-N-Blast-V5	N	V5	Blasticidin	https://www.addgene.org/207686/
qTAG-N-Puro-mNeon	N	mNeonGreen	Puromycin	https://www.addgene.org/207687/
qTAG-N-Puro-moxGFP	N	moxGFP	Puromycin	https://www.addgene.org/207688/
qTAG-N-Puro-mScarlet	N	mScarlet	Puromycin	https://www.addgene.org/207689/
qTAG-N-Puro-sTagRFP	N	super-TagRFP	Puromycin	https://www.addgene.org/207690/
qTAG-N-Puro-miRFP670nano3	N	miRFP670nano3	Puromycin	https://www.addgene.org/207691/
qTAG-N-Puro-miniTurbo	N	miniTurbo	Puromycin	https://www.addgene.org/207692/
qTAG-N-Puro-ultraID	N	ultraID	Puromycin	https://www.addgene.org/207693/
qTAG-N-Puro-dTAG	N	dTAG	Puromycin	https://www.addgene.org/207694/
qTAG-N-Puro-3xFLAG	N	3xFLAG	Puromycin	https://www.addgene.org/207695/
qTAG-N-Puro-3xHA	N	3xHA	Puromycin	https://www.addgene.org/207696/
qTAG-N-Puro-V5	N	V5	Puromycin	https://www.addgene.org/207697/
qTAG-N-Zeo-mNeon	N	mNeonGreen	Zeocin	https://www.addgene.org/207698/
qTAG-N-Zeo-moxGFP	N	moxGFP	Zeocin	https://www.addgene.org/207699/
qTAG-N-Zeo-mScarlet	N	mScarlet	Zeocin	https://www.addgene.org/207700/
qTAG-N-Zeo-sTagRFP	N	super-TagRFP	Zeocin	https://www.addgene.org/207701/
qTAG-N-Zeo-miRFP670nano3	N	miRFP670nano3	Zeocin	https://www.addgene.org/207702/
qTAG-N-Zeo-miniTurbo	N	miniTurbo	Zeocin	https://www.addgene.org/207703/
qTAG-N-Zeo-ultraID	N	ultraID	Zeocin	https://www.addgene.org/207704/
qTAG-N-Zeo-dTAG	N	dTAG	Zeocin	https://www.addgene.org/207705/
qTAG-N-Zeo-3xFLAG	N	3xFLAG	Zeocin	https://www.addgene.org/207706/
qTAG-N-Zeo-3xHA	N	3xHA	Zeocin	https://www.addgene.org/207707/
qTAG-N-Zeo-V5	N	V5	Zeocin	https://www.addgene.org/207708/
qTAG-C-mNeon-Blast	C	mNeonGreen	Blasticidin	https://www.addgene.org/207709/
qTAG-C-moxGFP-Blast	C	moxGFP	Blasticidin	https://www.addgene.org/207710/
qTAG-C-mScarlet-Blast	C	mScarlet	Blasticidin	https://www.addgene.org/207711/
qTAG-C-sTagRFP-Blast	C	super-TagRFP	Blasticidin	https://www.addgene.org/207712/
qTAG-C-miRFP670nano3-Blast	C	miRFP670nano3	Blasticidin	https://www.addgene.org/207713/
qTAG-C-miniTurbo-Blast	C	miniTurbo	Blasticidin	https://www.addgene.org/207714/
qTAG-C-ultraID-Blast	C	ultraID	Blasticidin	https://www.addgene.org/207715/
qTAG-C-dTAG-Blast	C	dTAG	Blasticidin	https://www.addgene.org/207716/
qTAG-C-3xFLAG-Blast	C	3xFLAG	Blasticidin	https://www.addgene.org/207717/
qTAG-C-3xHA-Blast	C	3xHA	Blasticidin	https://www.addgene.org/207718/
qTAG-C-V5-Blast	C	V5	Blasticidin	https://www.addgene.org/207719/
qTAG-C-mNeon-Puro	C	mNeonGreen	Puromycin	https://www.addgene.org/207720/
qTAG-C-moxGFP-Puro	C	moxGFP	Puromycin	https://www.addgene.org/207721/
qTAG-C-mScarlet-Puro	C	mScarlet	Puromycin	https://www.addgene.org/207722/
qTAG-C-sTagRFP-Puro	C	super-TagRFP	Puromycin	https://www.addgene.org/207723/
qTAG-C-miRFP670nano3-Puro	C	miRFP670nano3	Puromycin	https://www.addgene.org/207724/
qTAG-C-miniTurbo-Puro	C	miniTurbo	Puromycin	https://www.addgene.org/207725/
qTAG-C-ultraID-Puro	C	ultraID	Puromycin	https://www.addgene.org/207726/
qTAG-C-dTAG-Puro	C	dTAG	Puromycin	https://www.addgene.org/207727/
qTAG-C-3xFLAG-Puro	C	3xFLAG	Puromycin	https://www.addgene.org/207728/
qTAG-C-3xHA-Puro	C	3xHA	Puromycin	https://www.addgene.org/207729/
qTAG-C-V5-Puro	C	V5	Puromycin	https://www.addgene.org/207730/
qTAG-C-mNeon-Zeo	C	mNeonGreen	Zeocin	https://www.addgene.org/207731/
qTAG-C-moxGFP-Zeo	C	moxGFP	Zeocin	https://www.addgene.org/207732/
qTAG-C-mScarlet-Zeo	C	mScarlet	Zeocin	https://www.addgene.org/207733/
qTAG-C-sTagRFP-Zeo	C	super-TagRFP	Zeocin	https://www.addgene.org/207734/
qTAG-C-miRFP670nano3-Zeo	C	miRFP670nano3	Zeocin	https://www.addgene.org/207735/
qTAG-C-miniTurbo-Zeo	C	miniTurbo	Zeocin	https://www.addgene.org/207736/
qTAG-C-ultraID-Zeo	C	ultraID	Zeocin	https://www.addgene.org/207737/
qTAG-C-dTAG-Zeo	C	dTAG	Zeocin	https://www.addgene.org/207738/

qTAG-C-3xFLAG-Zeo	C	3xFLAG	Zeocin	https://www.addgene.org/207739/
qTAG-C-3xHA-Zeo	C	3xHA	Zeocin	https://www.addgene.org/207740/
qTAG-C-V5-Zeo	C	V5	Zeocin	https://www.addgene.org/207741/
pEF1α-Cre-PGK-Puro	-	Cre Recombinase	Puromycin	https://www.addgene.org/207742/
pEF1α-Cre-2A-Puro	-	Cre Recombinase	Puromycin	https://www.addgene.org/207743/
pEF1α-Cre-2A-moxGFP	-	Cre Recombinase	moxGFP	https://www.addgene.org/207744/
pEF1α-Cre-2A-mScarlet	-	Cre Recombinase	mScarlet	https://www.addgene.org/207745/
pEF1α-Cre-2A-mTagBFP2	-	Cre Recombinase	mTagBFP2	https://www.addgene.org/207746/
pEF1α-Cre-2A-miRFP670nano3	-	Cre Recombinase	miRFP670nano3	https://www.addgene.org/207747/
pX330	-	Cas9 + sgRNA	-	https://www.addgene.org/42230/
pX330-PITCh	-	Cas9 + sgRNA + PITCh sgRNA	-	https://www.addgene.org/127875/
pX330-PITCh-ACTB	-	Cas9 + sgRNA + PITCh sgRNA	-	https://www.addgene.org/207748/
qTAG-N-Solo-mNeon-ACTB	N	mNeonGreen	-	https://www.addgene.org/207749/
qTAG-N-Solo-mScarlet-ACTB	N	mScarlet	-	https://www.addgene.org/207750/
qTAG-N-Blast-mNeon-ACTB	N	mNeonGreen	Blasticidin	https://www.addgene.org/207751/
qTAG-N-Puro-mNeon-ACTB	N	mNeonGreen	Puromycin	https://www.addgene.org/207752/
qTAG-N-Blast-mScarlet-ACTB	N	mScarlet	Blasticidin	https://www.addgene.org/207753/
qTAG-N-Puro-mScarlet-ACTB	N	mScarlet	Puromycin	https://www.addgene.org/207754/
pX330-PITCh-H2BC11	-	Cas9 + sgRNA + PITCh sgRNA	-	https://www.addgene.org/207755/
qTAG-C-Solo-mScarlet-H2BC11	C	mScarlet	-	https://www.addgene.org/207756/
qTAG-C-moxGFP-Blast-H2BC11	C	moxGFP	Blasticidin	https://www.addgene.org/207757/
qTAG-C-moxGFP-Puro-H2BC11	C	moxGFP	Puromycin	https://www.addgene.org/207758/
qTAG-C-moxGFP-Zeo-H2BC11	C	moxGFP	Zeocin	https://www.addgene.org/207759/
qTAG-C-moxGFP-Puro-H2BC11-MMEJ	C	moxGFP	Puromycin	https://www.addgene.org/207760/
qTAG-C-mScarlet-Puro-H2BC11	C	mScarlet	Puromycin	https://www.addgene.org/207761/
qTAG-C-miRFP670nano3-Blast-H2BC11	C	miRFP670nano3	Blasticidin	https://www.addgene.org/207762/
pX330-PITCh-TUBA1B	-	Cas9 + sgRNA + PITCh sgRNA	-	https://www.addgene.org/207763/
qTAG-N-Solo-mScarlet-TUBA1B	N	mScarlet	-	https://www.addgene.org/207764/
qTAG-N-Solo-sTagRFP-TUBA1B	N	super-TagRFP	-	https://www.addgene.org/207765/
qTAG-N-Blast-moxGFP-TUBA1B	N	moxGFP	Blasticidin	https://www.addgene.org/207766/
qTAG-N-Puro-moxGFP-TUBA1B	N	moxGFP	Puromycin	https://www.addgene.org/207767/
qTAG-N-Blast-sTagRFP-TUBA1B	N	super-TagRFP	Blasticidin	https://www.addgene.org/207768/
qTAG-N-Puro-sTagRFP-TUBA1B	N	super-TagRFP	Puromycin	https://www.addgene.org/207769/
pX330-PITCh-LMNB1	-	Cas9 + sgRNA + PITCh sgRNA	-	https://www.addgene.org/207770/
qTAG-N-Solo-mScarlet-LMNB1	N	mScarlet	-	https://www.addgene.org/207771/
qTAG-N-Blast-moxGFP-LMNB1	N	moxGFP	Blasticidin	https://www.addgene.org/207772/
qTAG-N-Puro-moxGFP-LMNB1	N	moxGFP	Puromycin	https://www.addgene.org/207773/
qTAG-N-Blast-miniTurbo-LMNB1	N	miniTurbo	Blasticidin	https://www.addgene.org/207774/
qTAG-N-Blast-ultraID-LMNB1	N	ultraID	Blasticidin	https://www.addgene.org/207775/
qTAG-N-Blast-dTAG-LMNB1	N	dTAG	Blasticidin	https://www.addgene.org/207776/
qTAG-N-Blast-3xFLAG-LMNB1	N	3xFLAG	Blasticidin	https://www.addgene.org/207777/
qTAG-N-Blast-3xHA-LMNB1	N	3xHA	Blasticidin	https://www.addgene.org/207778/
qTAG-N-Blast-V5-LMNB1	N	V5	Blasticidin	https://www.addgene.org/207779/
pX330-PITCh-H3C2	-	Cas9 + sgRNA + PITCh sgRNA	-	https://www.addgene.org/207780/
qTAG-C-Solo-mScarlet-H3C2	C	mScarlet	-	https://www.addgene.org/207781/
qTAG-C-mNeon-Blast-H3C2	C	mNeonGreen	Blasticidin	https://www.addgene.org/207782/
qTAG-C-moxGFP-Puro-H3C2	C	moxGFP	Puromycin	https://www.addgene.org/207783/
qTAG-C-miRFP670nano3-Blast-H3C2	C	miRFP670nano3	Blasticidin	https://www.addgene.org/207784/
pX330-PITCh-TUBB4B	-	Cas9 + sgRNA + PITCh sgRNA	-	https://www.addgene.org/207785/
qTAG-C-mNeon-Blast-TUBB4B	C	mNeonGreen	Blasticidin	https://www.addgene.org/207786/
pX330-PITCh-LAMP1	-	Cas9 + sgRNA + PITCh sgRNA	-	https://www.addgene.org/207787/
qTAG-C-moxGFP-Puro-LAMP1	C	moxGFP	Puromycin	https://www.addgene.org/207788/
pX330-PITCh-TOMM20	-	Cas9 + sgRNA + PITCh sgRNA	-	https://www.addgene.org/207789/
qTAG-C-mNeon-Puro-TOMM20	C	mNeonGreen	Puromycin	https://www.addgene.org/207790/
pX330-PITCh-GOLGA2	-	Cas9 + sgRNA + PITCh sgRNA	-	https://www.addgene.org/207791/
qTAG-N-Puro-moxGFP-GOLGA2	N	moxGFP	Puromycin	https://www.addgene.org/207792/
pX330-PITCh-MAPRE1	-	Cas9 + sgRNA + PITCh sgRNA	-	https://www.addgene.org/207793/
qTAG-C-moxGFP-Puro-MAPRE1	C	moxGFP	Puromycin	https://www.addgene.org/207794/

Table 3. Troubleshooting table

Step	Problem	Possible Reason	Solution
3	Scarcity in viable sgRNA	- No sgRNAs near insertion site - Only low efficiency sgRNAs available - Only unspecific sgRNAs available due to high sequence homology paralogs.	- If no suitable sgRNAs are found within approximately 30 base pairs of the desired insertion site, it is feasible to expand the search range to around 50-100 base pairs, although this may result in a decrease in efficiency. - It's important to remain cautious about potential interference with critical elements in the 3' and 5' regions, such as transcription start sites and polyadenylation signals. - Another approach could involve making slight truncations of the tagged protein by identifying viable sgRNAs located within exons. However, it's crucial to exercise caution to avoid affecting functional domains that could disrupt protein function.
21	No bacterial colonies after cloning of sgRNA or qTAG repair constructs	- Old competent bacterial cells or cloning reagents - Ligations or assembly not prepared with ideal conditions	- Ensure to always perform a ligation/assembly positive and negative control along with a transformation control. - Make sure to adhere to the cloning conditions outlined in the protocol. When performing both ligations and Gibson assembly, it is recommended to begin with a minimum of 50 ng of the cut backbone and then adjust the amounts of the other fragments to maintain a 1:3 pmol ratio of backbone to fragment. If you encounter issues, consider raising the base amount of vector used and transitioning to a 1:5 pmol ratio, which may enhance the overall efficiency. Ensure fragments are pure. - Longer incubation times can also produce colonies. Conducting an overnight ligation at room temperature or extending the Gibson assembly time to 2 hours can improve clone yield.
28b, 29e	Cell death after transfection	- Cell-line specific transfection toxicity	- Try to find and adapt a transfection protocol from existing literature that is tailored to your specific cell line. - If conditions are unavailable for your cell line, conduct a transfection optimization. When using chemical transfection reagents like Lipofectamine 3000, adjust parameters such as cell quantity, DNA quantity, Lipofectamine reagent amount, and P3000 reagent amount to achieve the best transfection results while minimizing cell death. In the case of electroporation, it is essential to fine-tune the cell number, voltage, number of pulses, and duration of the pulse for each cell line to mitigate cell death. You can implement both of these optimization processes using one of the Cre-2A-fluorescent protein plasmids to visually evaluate transfection efficiency.
31	No cells after selection enrichment	- sgRNA efficiency - incompatibility of sgRNA and repair construct - Gene is lowly or conditionally expressed or is silent - transfection efficiency - tagging impairs viability	- Allow for up to one week of recovery after all the control cells have died. Low efficiency sgRNAs may still form resistant colonies after a prolonged recovery period. - Assess the sgRNA's cutting efficiency individually using a TIDE analysis or a T7 endonuclease assay to confirm its ability to bind and cleave at the intended target site. - Verify that the arms designed to target the insertion site are in close proximity to the selected sgRNA. Ensure that the essential modifications have been made to prevent unwanted cleavage of the donor or re-cleavage of the integrated material.

			- Employ qPCR or immunoblotting to determine the expression status of the target gene in your cell line. If the gene is not expressed, a promoter-based strategy may be more suitable than a multicistronic approach. - Confirm the successful delivery of plasmids into the cells by implementing a fluorescent transfection control with the Cre-2A-Fluorescent-protein plasmids. - If tagging causes a growth defect or cell death, reevaluate the choice of the gene for tagging. Alternatively, you may explore other approaches, such as attempting tagging at the opposite terminus or considering the tagging of paralogous genes.
34	Introduction of Cre does not excise the selectable marker sequence or inserts sequence	- The Cre recombinase is not entering the nucleus - Leftover lox sites from serial or parallel tagging	- Check with imaging to verify the nuclear entry of Cre by counter-staining with DAPI. - Conduct junction PCR on the alleles modified by Cre across all the genes you are targeting and subsequently sequence them to confirm their tagging and in-frame Cre-deletion.
46	No fluorescence or unspecific signal is observed after tagging with Fluorescent protein	- Low expression - Off-target integration	- Perform immunoblotting and genomic PCR to confirm that the genetic integration is not compromised. - Swap the fluorescent protein tag sequence with a higher performance fluorescent protein. - Execute junction PCR at both the integration site and the potential off-target binding sites of your sgRNA to ensure there are no unintended integrations. Sequencing can also be employed to verify this aspect.