Supplemental Materials

Molecular Biology of the Cell

Roberts et al.
Supplementary information

Systematic gene tagging using CRISPR/Cas9 in human stem cells to illuminate cell organization

Brock Roberts*,†, Amanda Haupt*, Andrew Tucker, Tanya Grancharova, Joy Arakaki, Margaret A. Fuqua, Angelique Nelson, Caroline Hookway, Susan A. Ludmann, Irina A. Mueller, Ruian Yang, Alan R. Horwitz, Susanne M. Rafelski, and Ruwanthi N. Gunawardane†

Allen Institute for Cell Science, 615 Westlake Ave North, Seattle, WA 98109

† Address correspondence to: Brock Roberts (brockr@alleninstitute.org), Ruwanthi Gunawardane (rug@alleninstitute.org)
* These authors contributed equally

Contents:
Supplementary Table S1
Supplementary Table S2
Supplementary Table S3
Supplementary Figure S1
Supplementary Figure S2
Supplementary Figure S3
Supplementary Figure S4
Supplementary Figure S5
Supplementary Figure S6
Supplementary Figure S7
Supplementary Figure S8
Supplementary Figure S9
Supplementary Figure S10
Supplementary Figure S11
Supplementary Figure S12
| Gene Cell Line | crRNA | % HDR | crRNA sequence (5’ - 3’) | PAM | Gene strand/crRNA strand | bp between insertion site and PAM -3 | Coordinates of insertion site | WTC specific variants within 1kb of insertion position |
|---------------|-------|------|--------------------------|-----|-------------------------|-----------------------------------|--------------------------------|---------------------------------------------|
| PXN AICS-5    | Cr1   | 0.9  | CTTCCTCGCTCCTGCTCCTTGA   | AGG | +/-                     | -53                               | chr12:12-1202121               | chr12:120212167 C/T (het); chr12:120212818 C/T (het) |
|               | Cr2   | 1.15 | GCCACCTA/GGAGAAGCTCCTG   | AGG | +/-                     | -10                               | chr9:992                      | none                                          |
|               | Cr3   | 1.05 | TCTAGTCCTAGTGCTAGCTT    | GGG | +//-                    | 51                                | chr12:12-1202121              | chr12:120212167 C/T (het); chr12:120212818 C/T (het) |
| SEC61B AICS-10| Cr1   | 24   | CACTACATCAATACTGCTGTA   | TGG | +/-                     | 3                                 | chr1:125                      | none                                          |
|               | Cr2   | 0.88 | GCCATCACATTAGGAGAATG    | AGG | +/-                     | -7                                | chr1:125                      | none                                          |
| TOMM20 AICS-11| Cr1   | 0.34 | ATTTGAAGTCACTCCATGG    | TGG | +/-                     | -23                               | chr9:992                      | none                                          |
|               | Cr2   | 0.51 | TGGGTGTAGTCCTGCCTTGTAGTGCTG | GGG | +/-                     | 72                                | chr1:125                      | none                                          |
| TUBA1B AICS-12| Cr1   | 0.52 | GCTGGCTGCGGCTTCCAATG    | GGG | +/-                     | -3                                | chr12:49                      | 129719-49129720                             |
|               | Cr2   | 0.23 | AGAGAGATGAGCCAATGCTGC   | AGG | +/-                     | -110                              | chr12:49                      | 129719-49129720                             |
| LMNB1 AICS-13 | Cr1   | 0.95 | GGGTGTGCGGCTGTCGCTG    | GGG | +/-                     | 3                                 | chr5:126                      | 775715-12677751                            |
|               | Cr2   | 0.52 | GCTGGCTGCGGCTGTCGCTG   | GGG | +/-                     | -3                                | chr5:126                      | 775715-12677751                            |
| FBL AICS-14   | Cr1   | 4.18 | AAC/TAAGAGTCGCAGATCGTG | AGG | +/-                     | -14                               | chr19:39                      | 834540-39834541                            |
|               | Cr2   | 2.01 | CA/GGTTTCTAATAGGGGG    | CCG | +/-                     | -15                               | chr19:39                      | 834540-39834541                            |
| ACTB AICS-16  | Cr1   | 3.17 | GCTATATCGGCGGCTCAGCA   | TG/G | +/-                     | -5                                | chr7:552                      | 9564-55288652/A/G (hom)                     |
|               | Cr2   | 0.52 | GCCGCTGCTAGGGCGGAGC    | AGG | +/-                     | -19                               | chr7:552                      | 9564-55288652/A/G (hom)                     |
| DSP AICS-17   | Cr1   | 0.039| TCTTTGACTGATGCTACTAT   | TGG | +/-                     | -10                               | chr6:758                      | 5875-7585786                              |
|               | Cr2   | 0.17 | AGAAGCCTAGCCAAAAGTAC   | AGG | +/-                     | -26                               | chr6:758                      | 5875-7585786                              |
| TJP1 AICS-23  | Cr1   | 0.3  | CTGGGGGCCTGGAGGGCCTGG  | CCG | +/-                     | 7                                 | chr15:29                      | 822025-29822026                            |
|               | Cr2   | 0.26 | TCTCTCTAAGGCGGCGGAGG   | AGG | +/-                     | -24                               | chr15:29                      | 822025-29822026                            |
|               | Cr3   | 2.09 | GCCGGCGGGAGGGCCCTCACT   | TCC | +/-                     | 25                                | chr15:29                      | 822025-29822026                            |
| MYH10 AICS-24 | Cr1   | 0.14 | TTTACAAATAG/CCGAGAAGAAC | TGG | +/-                     | 8                                 | chr17:86                      | 23243-8623244                             |
|               | Cr2   | 0.045| G/GGAGAAGAATGAGGACCTC   | AGG | +/-                     | 16                                | chr17:86                      | 23243-8623244                             |
|               | Cr4   | 0.11 | GTCTTCTTGGC/CATTGTTAA   | TGG | +/-                     | -6                                | chr17:86                      | 23243-8623244                             |
| GALT AICS-19  | Cr1   | 0.00069| GCCCTGACCGACCGACCGAC   | AGG | +/-                     | +13                               | chr9:346                      | 50446-34650447                            |
|               | Cr2   | 0.00055| TCAAAGCCGCTGGGCTGGCCG  | AGG | +/-                     | +9                                | chr9:346                      | 50446-34650447                            |
| TUBG1 AICS-18 | Cr1   | 0.031| AGTGTGGCTGTGGCAGTGG    | TGG | +/-                     | -43                               | chr17:42                      | 610538-42615039                            |
|               | Cr2   | 0.011| GAGAGTGTGCTGATTGAGCG   | TGG | +/-                     | -35                               | chr17:42                      | 610538-42615039                            |
|               | Cr3   | 0.02 | AGGCCCTGCGCCTACACCAAGTA | AGG | +/-                     | +36                               | chr17:42                      | 610538-42615039                            |
Table S1. Design features for crRNAs used in editing experiments shown in Fig. 1D.

Name of the targeted gene (and AICS cell line identifier used in the cell line catalog at Allen Cell Explorer and the Allen Cell Collection at Coriell), crRNA number, HDR efficiency, and binding sequence are shown. Percent HDR was determined by FACS and is shown as a percentage of GFP+ cells within the gated cell population in each experiment. The crRNA used to create the final clone chosen for expansion and distribution for each gene is bolded and underlined. The non-complementary DNA strand corresponding to the crRNA binding site and PAM in the WTC genome is shown in black. The non-complementary DNA strand corresponding to the crRNA binding site in the donor plasmid and PAM is shown in red. Mutations introduced into the donor plasmid to eliminate Cas9 cleavage are indicated by lower case (point mutations), dashes (deletions), or forward slash (where the tag and linker sequence interrupts the crRNA binding site). The distance between the intended insertion site and the PAM -3 site (where double strand breaks are anticipated) is indicated for each crRNA. Distances are negative when the double strand break is anticipated 5′ of the insertion site and positive when the double strand break is anticipated 3′ of the insertion site relative to gene orientation. Gene orientation and crRNA orientation are defined according to strand in the GRCh38 reference genome. Genomic coordinates are indicated for the site of integration, and for single nucleotide polymorphisms (SNPs) and insertions or deletions (INDELs) specific to the WTC genome within the homology arm region of the plasmid. In cases where the WTC-specific SNP was heterozygous, the reference genome variant was used in the homology arm. Coordinates are from the GRCh38 (GCA 000001405.15) assembly, NCBI annotation 107. *TUBG1 heterozygous SNP was changed to WTC variant in donor plasmid.
### Table S2. PCR primers used in experiments.

| Cell Line | 5’ Junctional Fwd | 3’ Junctional Rev |
|-----------|-------------------|-------------------|
| PXN-EGFP  | TGTGCAAGGCAAGCCCTG | TCAGTAAGAGGCCATCC |
| SE0618-mEGFP | ACTTCTACTGAGAGAATCT | CTGAGAACGAGGAGGATC |
| TOMM20-mEGFP | ACCTAGGAGAGGAGGAG | GACAGATAGAGAGGAGG |
| LN981-mEGFP | GAACCTGAGGAGGAGGAG | GACAGATAGAGAGGAGG |
| BFL-mEGFP  | ATTCGAGGAGGAGGAGGAG | GACAGATAGAGAGGAGG |
| ACTB-EGFP  | AGCGAGGAGGAGGAGGAG | GACAGATAGAGAGGAGG |
| DSB-mEGFP  | ATTCGAGGAGGAGGAGGAG | GACAGATAGAGAGGAGG |
| Tj1-mEGFP  | ATTCGAGGAGGAGGAGGAG | GACAGATAGAGAGGAGG |
| MYH10-mEGFP | ATTCGAGGAGGAGGAGGAG | GACAGATAGAGAGGAGG |

All primers are listed in 5’ to 3’ orientation.
Table S3. Antibodies used in western blot, immunofluorescence, and flow cytometry experiments.

| Primary Antibodies | Secondary Antibody |
|--------------------|-------------------|
| Antibody           | Type              | Source | Application | Secondary Antibody |
| Alpha tubulin      | mouse monoclonal, clone DM1A | ThermoFisher #2204 | WB: 1:10,000 | goat anti-mouse |
| Beta actin         | mouse monoclonal, clone GT512 | GeneTex A17622850 | WB: 1:10,000 | goat anti-mouse |
| Desmin (mouse)    | rabbit polyclonal | Invitrogen | WB: 1:1,000 | goat anti-rabbit |
| Ki67               | rabbit polyclonal | ThermoFisher #1024 | WB: 1:10,000 | goat anti-rabbit |
| Lamin B1 (nuclear lamin B1) | rabbit polyclonal | Abcam #218646 | WB: 1:1,000 | goat anti-rabbit |
| Myosin IB (non-muscle myosin heavy chain IB) | rabbit polyclonal | Cell Signaling Technology #3405 | IF: 1:500 | goat anti-mouse |
| Paxillin           | rabbit monoclonal, clone 349 | BD Biosciences #10051 | WB: 1:10,000 | goat anti-rabbit |
| Sedar1 beta        | rabbit polyclonal | Abcam #21576 | WB: 1:10,000 | goat anti-rabbit |
| Tight junction protein ZO1 (ZO1) | rabbit polyclonal | ThermoFisher #617300 | IF: 1:50 | goat anti-rabbit |
| Tom20              | mouse monoclonal, clone F10 | Sigma Aldrich #4A03407 | WB: 1:250 | goat anti-mouse |
| Beta actin (loading control) | mouse monoclonal, clone BA9R | ThermoFisher #4551730 | WB: 1:10,000 | goat anti-mouse |
| Alpha actin (loading control) | mouse monoclonal, clone 0.T.02 | ThermoFisher #FM191860 | WB: 1:2,000 | goat anti-mouse |
| GFAP               | mouse monoclonal mix, clones 7.1 and 13.1 | Sigma Aldrich #1131445001 | WB: 1:250 | goat anti-mouse |

| Directly Conjugated Primary Antibodies | Type | Source | Application |
|---------------------------------------|------|--------|-------------|
| TRA-1-81, Brilliant Violet™ 510      | mouse monoclonal, clone TRA-1-81 | BD Biosciences #43188 | Flow: 1:60 |
| SSEA-1, Alexa Fluor® 647             | rat monoclonal, clone MC431 | BD Biosciences #461145 | Flow: 1:60 |
| SSEA-1, Brilliant Violet™ 421        | mouse monoclonal, clone MC480 | BD Biosciences #462705 | Flow: 1:60 |
| Nagnag, Alexa Fluor® 647             | mouse monoclonal, clone N511055 | BD Biosciences #461300 | Flow: 1:60 |
| Oct4/4, Brilliant Violet™ 510        | mouse monoclonal, clone OX20T-3 | BD Biosciences #463524 | Flow: 1:60 |
| Sox2, Alexa Fluor® 488               | mouse monoclonal, clone OX2450 | BD Biosciences #461310 | Flow: 1:60 |
| TRA-1-81, PeroCP-Cy5.5               | mouse monoclonal, clone TRA-1-81 | BD Biosciences #461078 | Flow: 1:60 |
| SSEA-3, PE                           | rat monoclonal, clone MC431 | BD Biosciences #460237 | Flow: 1:60 |
| SSEA-1, Brilliant Violet™ 650        | mouse monoclonal, clone 2H8 | BD Biosciences #460232 | Flow: 1:60 |
| Oct4/4, PeroCP-Cy5.5                 | mouse monoclonal, clone OX20T-3 | BD Biosciences #460704 | Flow: 1:60 |
| Sox2, Alexa Fluor® 647               | mouse monoclonal, clone OX2450 | BD Biosciences #460204 | Flow: 1:60 |
| Nagnag, PE                           | mouse monoclonal, clone N511055 | BD Biosciences #460483 | Flow: 1:60 |
| Tejon, Alexa Fluor® 647              | mouse monoclonal, clone OX2450 | BD Biosciences #460483 | Flow: 1:60 |
| Mouse IgG1 k Lac Z Control, Alexa Fluor® 647 | mouse monoclonal, clone MOPC-21 | BD Biosciences #465571 | Flow: 1:200 |
| Brachyury, APC                        | goat polyclonal | R&D Systems #C3065A | Flow: 1:40 |
| Goat IgG1 k Lac Z Control, APC        | goat polyclonal | R&D Systems #C108A | Flow: 1:40 |
| Sox 17, APC                           | goat polyclonal | R&D Systems #C108A | Flow: 1:80 |
| Goat IgG1 k Lac Z Control, APC        | goat polyclonal | R&D Systems #C108A | Flow: 1:60 |
| Pansy, Alexa Fluor® 647              | mouse monoclonal, clone OX18-1330 | BD Biosciences #462240 | Flow: 1:20 |
| Mouse IgG2a k Lac Z Control, Alexa Fluor® 647 | mouse monoclonal, clone MOPC-172 | BD Biosciences #458003 | Flow: 1:10 |

Table of antibodies used in western blots (WB), immunofluorescence (IF), and flow cytometry (Flow) experiments showing dilutions used per application.
Figure S1. Expression levels of the 12 genes attempted for genome editing in the WTC parental cell line. Transcript abundance for each gene was estimated from RNA-Seq data. Samples were derived from the WTC parental line after 8 passages (p8) in culture and 14 passages (p14) in culture, as indicated. Transcript abundances were calculated in units of fragments per kilobase of transcript per million fragments mapped (FPKM). Log_{10} (FPKM+1) transcript abundances from parental WTC p8 and p14 samples were plotted against each other and were highly correlated (R^2 = 0.989). The two genes (TUBG1 and GALT) that were not successfully edited are highlighted in red.
A

Possible off-target flaws:
- Mismatch
- cRNA
- RNA bulge
- NTP bulge
- Off-target

Flaws
NS: 5
- 2 flaws in non-seed +
- 1 flaw in seed

Off-target categories
- Non-seed
- Seed
- PAM

Sequencing Profile
- Ranking

406 total sites sequenced
142 unique sites sequenced
0 off-target mutations identified

B

C

D

E

- # predicted alternative binding sites
- # predicted alternative binding sites sequenced

- exon/within 500 bp of exon
- exonic (inter/intronic within 2000 bp of gene)
- non-genic (outside of gene)
Figure S2. Predicted genome wide CRISPR/Cas9 alternative binding sites, categorized according to sequence profile and location with respect to genes. (A) Predicted alternative CRISPR/Cas9 binding sites are categorized for each crRNA used. Each predicted off-target sequence was categorized according to its sequence profile (the number of mismatches and RNA or DNA bulges it contains relative to the crRNA used in the experiment and their position relative to the PAM). Cas-OFFinder was used to identify all alternative sites genome-wide with \( \leq 2 \) mismatches/bulges in the non-seed and/or \( \leq 1 \) mismatch/bulge in the seed region, with an NGG or NAG PAM. As indicated, the seed and non-seed region of a crRNA binding sequence was defined with respect to its proximity to the PAM sequence. Overlapping Cas-OFFinder results with the same double strand break site were collapsed into one category using sequence profile ranking (see Methods). (B) Predicted off-target sequence breakdown based on sequence profile (colors refer to categories defined in (A). A subset of CRISPR/Cas9 alternative binding sites identified by Cas-OFFinder were selected for sequencing. (C) Breakdown of sequenced off-target sites by sequence profile. (D) All predicted off-target sites were additionally categorized according to their location with respect to annotated genes. Genomic location was defined as follows: \textit{exon}: inside exon or within 50 bp of exon; \textit{genic}: in intron (but \( >50 \) bp from an exon) or within 200 bp of an annotated gene; \textit{non-genic}: \( >200 \) bp from an annotated gene. (E) Breakdown of sequenced off-target sites by genomic location with respect to annotated genes. Numbers above bars represent the number of clones sequenced for each experiment. All 406 sequenced sites were found to be wild type.
Figure S3. ddPCR screening data. (A) ddPCR screening data for all experiments (Fig. 2A step 1). Each data point represents one clone. Clones with GFP genomic copy number of ~1 to ~2 and plasmid backbone genomic copy number <0.2 were typically considered for further analysis. TJP1 clones consistently produced GFP copy number values <1 despite validation by junctional PCR, imaging and western blot as putative mono-allelic clones. This result is unresolved and under investigation. (B) A dilution series of the donor plasmid used for the PXN-EGFP tagging experiment was used to confirm equivalent amplification of the AMP and GFP sequences in two-channel ddPCR assays.
Figure S4. Amplification of complete junctional (non-tiled) PCR products to demonstrate presence of the allele anticipated from tiled junctional PCR product data. (A) Junctional PCR primers complementary to sequences flanking the homology arms in the distal genome (also used in tiled junctional PCR assays, shown in black), were used together to co-amplify tagged and untagged alleles (red). N-terminal tag shown as an example. (B) This assay served to rule out anticipated DNA repair outcomes where tiled junctional PCR data leads to a misleading result because the GFP tag sequence has been duplicated during HDR, as indicated by the schematic. An N-terminal tag duplication is shown as an example. (C) Molecular weight markers are as indicated (kb). Two final clones (indicated by “cl. #”) are represented for each experiment. Asterisk indicates the final clone chosen for distribution and imaging. A band intermediate in size between the anticipated tagged and untagged allele products is consistently observed, which we hypothesized corresponds to a heteroduplex of the tagged and untagged allele products.
Figure S5. Comparison of unedited versus edited cells by immunofluorescence. Labeled structures in unedited WTC parental cells and edited cell lines are compared. Whole field of views (FOVs) shown on the left, with insets highlighted with white boxes. Alpha tubulin panel: anti-alpha tubulin antibody staining FOV with insets highlighting a spindle and a midbody (midbody inset was obtained from an apical image slice, not shown in the FOV). Images represent single z-section slices. FOV scale bar is 10 µm, insets are 3 µm. Nuclear lamin B1 panel: anti-lamin B1 antibody staining FOV with insets illustrating interphase and nuclear envelope re-assembly. Images represent maximum intensity projections of 3 apical z-sections. FOV scale bar is 10 µm, insets are 3 µm. Paxillin panel: anti-paxillin antibody staining FOV with insets highlighting the basal cell surface and cell protrusions in detail. Images represent maximum intensity projections of 3 basal z-sections. FOV scale bar is 10 µm, insets are 3 µm. Tight junction protein ZO1 panel: anti-ZO1 antibody staining FOV with two insets. Images represent maximum intensity projections of 10 apical z-sections. FOV scale bar is 10 µm, insets are 3 µm. Fibrillarin panel: anti-fibrillarin antibody staining FOV with two insets illustrating variation in nucleolar staining. Images represent a single apical z-section. FOV scale bar is 5 µm, insets are 3 µm. Tom20 panel: anti-Tom20 antibody staining FOV with one inset highlighting a single mitochondrial tubule. Images represent maximum intensity projections of 4 basal z-sections. FOV scale bar is 10 µm, inset is 3 µm. Desmoplakin panel: anti-desmoplakin staining FOV with one inset showing the GFP channel and transmitted light image overlay to show desmoplakin puncta localization at the cell-cell boundaries. Images represent maximum intensity projections of z-sections spanning the entire colony and single z plane for the transmitted light image. FOV scale bar is 10 µm, inset is 1 µm. Sec61 beta panel: anti-Sec61 beta antibody staining FOV with one inset. Images represent maximum intensity projections of 3 z-sections near the middle of the cell colony. FOV scale bar is 8 µm, inset is 4 µm. Beta actin panel: Phalloidin-Rhodamine staining showing apical and basal FOVs, and an apical region inset. Images represent maximum intensity projections of either apical or basal z-sections. Apical and basal image scale bars are 10 µm, inset is 4 µm. Non-muscle myosin heavy chain IIB panel: anti-myosin IIB antibody staining FOV showing apical and basal regions. Images represent maximum intensity projections of 4 apical or basal z-sections of the cell colony. Scale bars are 10 µm. All images acquired on a spinning disk confocal microscope except panels shown for desmoplakin, which was acquired on a laser scanning confocal microscope. Antibody and method details are available in Table S3 and the Allen Cell Explorer (Allen Institute for Cell Science, 2017).
Figure S6. Comparison of GFP tag localization and endogenous protein stain in edited cell lines. Antibodies raised against the tagged protein were used to stain unedited and edited cells, as indicated. Labels on left of images indicate the tagged structure, and labels on the right indicate tagged gene and clone. In edited cells, imaging of the GFP tag in fixed cells was performed.
simultaneously, and co-localization of the GFP tag and antibody stain is indicated in the merged panels, as indicated. Scale bars are 10 µm. Additional immunofluorescence data is available at the Allen Cell Explorer (Allen Institute for Cell Science, 2017).
Figure S7. Live cell imaging comparison of transiently transfected cells and genome edited cells.

Top panels depict transiently transfected WTC cells and bottom panels depict gene edited clonal lines. **Left:** WTC transfected with EGFP-tagged alpha tubulin construct compared to the TUBA1B-mEGFP edited cell line. Images are a single apical frame. **Middle:** WTC transfected with EGFP-tagged desmoplakin construct compared to the DSP-mEGFP edited cell line. Images are maximum intensity projections of apical 4 z-frames. **Right:** WTC transfected with mCherry-tagged Tom20 construct compared to the TOMM20-mEGFP edited cell line. Images are single basal frames of the cell. All imaging was performed in 3D on live cells using laser-scanning confocal microscope. Movie versions of these z-stacks can be found at the Allen Cell Explorer (Allen Institute for Cell Science, 2017).
Figure S8. Western blot analysis of all 10 edited clonal lines. Western blot analyses for all experiments are presented as in Fig. 4B. Proteins and antibodies used on different blots are as indicated. In all cases, blots with antibodies against the respective target proteins are shown in the left blot and show the tagged and untagged protein products. Separation of untagged and tagged protein versions from the mEGFP-tagged desmoplakin clone was not possible due to the large size of the target protein (asterisk). Blots with anti-GFP antibodies showing only the tagged protein are shown in the right blot, as indicated. Alpha actinin, beta actin, and alpha tubulin were used as loading controls, as indicated. Lysates from unedited cells and the edited clonal lines are as indicated, as are bands corresponding to the labeled, predicted proteins. Antibody information is available in Table S3.
Figure S9. Editing experiments testing the feasibility of biallelic editing of the LMNB1 and TUBA1B loci. (A) Final clones LMNB1-mEGFP and TUBA1B-mEGFP were transfected using the standard editing protocol with a donor cassette targeting the untagged allele of the tagged locus, encoding mTagRFP-T (sequential delivery, top row). Additionally, unedited cells were transfected with editing reagents according to the standard editing protocol, using a 1:1 mix of the mEGFP and mTagRFP-T donor plasmids (simultaneous delivery, bottom row). Flow cytometry was used to identify cells with mono-allelic edits (either tag), as well as cells with biallelic editing (both tags). Frequency of editing with mTagRFP-T was quantified by flow cytometry. mTagRFp-T+ LMNB1-mEGFP cells were isolated by FACS (asterisk denotes sorted population). (B) The sorted population from (A) (indicated by asterisk) revealed similar subcellular localization of GFP and mTagRFP-T signal to the nuclear envelope in the majority of cells, suggesting successful biallelic tagging. Scale bars are as indicated in the merged panels. Cells were fixed in 4% paraformaldehyde before imaging. Low magnification images (top row) reveal that sorting significantly enriches the population for mTagRFP-T+ cells, which vary in mTagRFP-T intensity. This pattern was also seen with LMNB1-mEGFP+ sorted cells (Fig. 1E) before clones were selected.
Figure S10. Live imaging analysis at two culture time points of TUBA1B-mEGFP edited cells and the four final edited clones that displayed a low abundance of tagged protein. Endogenous GFP signal in final edited clones was compared in live imaging experiments (or fixed samples for LMNB1-mEGFP) in otherwise identical cultures separated by four passages (14 days) of culture time. Similar intensity levels of mEGFP-tagged structures before and after four passages suggests the transgene is not silenced over time. (A, C, E, G, I): Low magnification images show similar intensity levels within and between colonies at final banked passage and after 4 passages (14 days). (A) and (E) are single z-slices at the bottom and middle of the cell height, respectively. (C), (G), and (I) are maximum intensity projections through z (scale bar, 50 μm). (B, D, F, H, J): High magnification images show similar intensity levels in structures at greater detail. Panels (H) and (J) are split to show the apical and basal localization of mEGFP-tagged beta actin (H) and non-muscle myosin heavy chain IIB (J): The apical images are maximum intensity projections of the top 10 z-slices through the cells, and the basal image is the bottom z-slice. (B) and (F) are single z-slices taken at the bottom and middle of the cell height, respectively. (D) is a maximum intensity projection through the entire cell. Scale bar, 10 μm. Contrast and brightness adjustments are identical for each early/late pair at each magnification so that intensities can be compared directly.
### Candidate final clones

| Protein         | Clone | % GFP tagged |
|-----------------|-------|--------------|
| Alpha tubulin   | cl. 100 | 65           |
|                 | cl. 101 | 62           |
|                 | cl. 103 | 58           |
|                 | cl. 105 | 55           |
|                 | cl. 105 p34 | 53         |
|                 | cl. 105 p38 | 52         |

### Final clone: banked passage and +14 Days

| Protein         | Clone | % GFP tagged |
|-----------------|-------|--------------|
| Alpha tubulin   | cl. 100 | 5            |
|                 | cl. 133 | 5            |
|                 | cl. 148 | 6            |
|                 | cl. 185 | 6            |
|                 | cl. 184 | 5            |
|                 | cl. 184 p36 | 4          |
|                 | cl. 184 p40 | 4          |

### Fibrillarin

| Clone | % GFP tagged |
|-------|--------------|
| cl. 8 | 20           |
| cl. 9 | 21           |
| cl. 53| 32           |
| cl. 6 | 26           |
| cl. 6 p27 | 29       |
| cl. 6 p31 | 26       |

### Nuclear lamina B1

| Clone | % GFP tagged |
|-------|--------------|
| cl. 224 | 36          |
| cl. 245 | 31          |
| cl. 210 | 33          |
| cl. 210 p32 | 29       |
| cl. 210 p36 | 32       |

### Non-muscle myosin heavy chain IIb

| Clone | % GFP tagged |
|-------|--------------|
| cl. 27 | 35           |
| cl. 70 | 36           |
| cl. 88 | 36           |
| cl. 93 | 38           |
| cl. 80 | 37           |
| cl. 80 p30 | 34      |
| cl. 80 p34 | 33      |
Figure S11. Western blot analysis of candidate clones at one culture time point and final clones at two culture time points from editing experiments that displayed a low abundance of tagged protein. Final tagged clones from four experiments in which the tagged protein copy displayed diminished abundance relative to the untagged copy, in addition to TUBA1B-mEGFP clones, were compared to independently derived clones from the same experiment that were also validated as correctly edited. All clones were blotted both with anti-GFP and with antibodies recognizing the targeted protein, as indicated. Additionally, the final clone from each experiment was analyzed by immunoblot in the same manner in otherwise identical cultures separated by 4 passages (14 days) of culture time. The fraction of GFP-tagged protein, relative to total, is indicated.
Figure S12. Flow cytometry analysis of GFP tag expression stability, flow cytometry analysis of cell cycle dynamics, microscopy analysis of mitotic index, and culture growth assays. (A) Endogenous GFP signal in final edited clones was compared in otherwise identical cultures separated by four passages (14 days) of culturing time (indicated). Forward scatter is
shown on the x-axis and GFP intensity is shown on the y-axis. Unedited cells are included as negative controls, as indicated. **(B)** Propidium iodide staining and flow cytometry were used to quantify numbers of cells in G1 (indicated), S phase (indicated) and G2/M phase (indicated) in final edited clones. Cultures of unedited cells at low passage (p16) and high passage (p30), chosen to approximate the final passage number of edited and expanded clones were compared in the upper left plot. Banked final clones (passage indicated), and same clones after 4 passages (14 days) in culture (indicated), were co-analyzed. Plots for each clone at both passages are shown in overlays, along with unedited cells at p30 (top). Gating was used to define the fraction of cells in G1 and G2/M, as indicated, with cells intermediate between peaks defined as S phase. Fractions of cells in each phase of the cell cycle are displayed as percentages (bottom), as indicated. **(C)** DAPI staining of colonies from each of the same five clonal lines was additionally used to quantify the numbers of mitotic cells per colony, as indicated. DAPI staining was only performed on colonies from each experiment at the lower passage number. Plot shows individual colony data points and mean percent mitotic cells per colony for each cell line with 95% confidence intervals. One-way ANOVA found no significant difference in percent mitotic cells per colony between cell lines (F(5,91)=0.606, p=0.696). **(D)** ATP quantitation was used as an indirect measure of cell growth. Two independent experiments were performed for each cell line; within each experiment cell lines were plated in triplicate. Doubling time was calculated using counts at time of seeding and at 96 hours after seeding (see Methods). Bars represent the average doubling time for each experiment with 95% confidence intervals (three wells for each experiment). One-way ANOVA found no significant difference in doubling time between cell lines (F(11,13)=1.794, p=0.157).

Allen Institute for Cell Science. (2017). Allen Cell Explorer. Available at: http://www.allencell.org/