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:

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	CTTGTCGTTCTGCTCCTTGA CTTGTCGTTCTGCTCCTTGA	AGG AGt	-/+	-53	chr12:120212316-120212317	chr12:120212167 C/T (het); chr12:120212618 C/T (het)
	Cr2	1.15	GCACCTA/GCAGAAGAGCTTG GCACCTA/GCAGAAGAGCTTG	AGG AGt	-/+	-10		
	Cr3	1.05	TCTAGGTCACAGTCGCAGTT TCTAGGTCACAGTCGCAGTT	GGG GGt	-/+	51		
SEC61B AICS-10	Cr1	24	CCCTCATCTCCAAT/ATGGTA CaCTCATCTCCAAT/ATGGTA	TGG TGG	+/+	3	chr9:99222363-99222364	none
	Cr2	0.88	GCCATACCAT/ATTGGAGATG GCCATACCAT/ATTGGAGATG	AGG AGt	+/-	-7		
TOMM20 AICS-11	**Cr1**	0.34	AATTGTAAGTGCTCAGAGCT AATTGTAAGTGCaCAGtcCT	TGG TGG	-/-	-23	chr1:235112066-235112067	none
	Cr2	0.051	TGGTAGTTGAGCAGCTCTGG TGGTAGTTGAGCAGCTCTGG	GGG GGt	-/-	72		
TUBA1B AICS-12	**Cr1**	0.052	GATGCACTCACG/CTGCGGGA GATGCACTC---/CTGCGGtA	AGG AGt	-/+	-8	chr12:49129719-49129720	none
	Cr2	0.023	AGAGATAAGGTCTGTCGCCC AGAGATAAGGTCTGTCGCCC	AGG AGt	-/+	-110		
LMNB1 AICS-13	**Cr1**	0.95	GGGGTCGCAGTCGCCAT/GGC GGGGTCGCAGTCGC---/GGC	GGG GGG	+/-	-3	chr5:126777511-126777512	none
	Cr2	0.52	GTCGCAGTCGCCAT/GGCGGG GTCGCAGTCGC---/GGCGGG	CGG CGG	+/-	-6		
FBL AICS-14	**Cr1**	4.18	AAC/TGAAGTTCAGCGCTGTC AAC/TGAAGTTCAGCcCTGag	AGG cGG	-/-	14	chr19:39834540-39834541	none
	Cr2	2.01	CA/GTTCTTCACCTTGGGGGG CA/GTTCTTCACtTTaGGaGG	TGG TGG	-/+	-15		
ACTB AICS-16	Cr1	3.17	GCTATTCTCGCAGCTCACCA GCTATTCTCGCAaCTgACaa	TG/G TG/G	-/-	-5	chr7:55529654-5529655	chr7:5528862 A/G (hom)
	Cr2	0.52	GCCGTTGTCGACGACGAGCG GCCGTTGTCGACGACGAGCG	CGG CtG	-/+	19		
DSP AICS-17	Cr1	0.039	TCATTTAGCAGTAGTTCTAT TCATTTAGCAGTAGTagcAT	TGG TGG	+/+	-10	chr6:7585875-7585876	chr6:7585734 G/C (hom)
	Cr2	0.17	AGAACTACTGCTAAATGAGT gctACTACTGCTAAATGAGT	AGG AtG	+/-	-26		
TJP1 AICS-23	Cr1	0.3	CTTGGCGGCCGCAGCTCTGG CTTtGCGGCCGCAGCTCTGG	CGG Cac	-/+	7	chr15:29822025-29822026	none
	Cr2	0.26	TCTCTCTCCAGCGCCGCGCG TCTCTCTCCAGCGCCGCGCG	AGG caa	-/+	-24		
	Cr3	2.09	GGCCGCGGAGGCGCTCACCT GGCCGCGGAGGttCTCACCT	TGG TtG	-/+	25		
MYH10 AICS-24	Cr1	0.14	TTTACAATG/GCGCAGAGAAC TTTACAATG/GCaCAaAGgAC	TGG aGG	-/-	8	chr17:8623243-8623244	none
	Cr2	0.045	G/GCGCAGAGAACTGGACTCG G/GCaCAaAGgACaGGgCTGG	AGG AGG	-/-	16		
	Cr4	0.11	GTTCTCTGCGC/CATTGTAAA GTcCTtTGtGC/CATTGTAAA	TGG TGG	-/+	-6		
GALT AICS-19	Cr1	0.00069	CGCC/TGACCACGCCGACCAC CGCC/TGACCACGgaGACCcC	AGG AGG	+/+	+13	chr9:34650446-34650447	none
	Cr2	0.00055	TCAAGGCCCTGTGGTCGGCG TCAAGGCCCTGgGGTCtcCG	TGG TGG	+/-	+9		
TUBG1 AICS-18	Cr1	0.031	AGTCTGGCCGTGTGGCCGCA AGTCTGGtCtaGTaGCCGCA	TGG TGa	+/-	-43	chr17:42615038-42615039	chr17:42614528 C/G (het); *
	Cr2	0.011	GGAGATGTAGTCTGGCCGTG GGAGATGTAGTCTGGtCtaG	TGG TaG	+/-	-35		
	Cr3	0.02	AGGGCTTGGGCCAACCAGTA AGGGCTTGGGCCAACCAGTA	AGG AGt	+/-	+36		

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	5' Junctional Rev	3' Junctional Fwd	3' Junctional Rev
PXN-EGFP	TGTGCAGTGGCACGATCTTGG	ACTTCAGGGTCAGCTTGCCG	AAGACCCCAACGAGAAG	TCAGTGAAGAGCTTGCTGGC
SEC61B-mEGFP	TATCTACCTCGGAATCACCC	AAGTCGATGCCCTTCAGCTCG	GTGAGCAAGGGCGAGGAGCTG	TGGGCGACAGAGTGAGATTCC
TOMM20-mEGFP	AGCGTGTCTGTTACAAGTGTTG	AAGTCGATGCCCTTCAGCTCG	GTGAGCAAGGGCGAGGAGCTG	CCCACCTGCTCCACTCTTTT
TUBA1B-mEGFP	GACTAGGGCTACAGGGC	GCAGATGAACTTCAGGGTCA	AAGACCCCAACGAGAAG	TTAGTGTAGGTTGGGCGCTC
LMNB1-mEGFP	TTCAAGACGCACAGATCTCAC	AAGTCGATGCCCTTCAGCTCG	GTGAGCAAGGGCGAGGAGCTG	ACACATTTCCCCAGAGAAAGC
FBL-mEGFP	ATTACAGGCACGAGCCACTGC	AAGTCGATGCCCTTCAGCTCG	GTGAGCAAGGGCGAGGAGCTG	ACGCGGGGGAAGAGTAGAGC
ACTB-mEGFP	AGAAGTCCACCGAGTCCTGC	AAGTCGATGCCCTTCAGCTCG	GTGAGCAAGGGCGAGGAGCTG	GTGAAGCTGTAGCGCGCTC
DSP-mEGFP	ACCCTCAGGAAGCGTAGAGT	TTGCCGTCCTCCTTGAAGTC	GAGCAAAGACCCCAACGAGA	TGCCAATGCTTTGTTGTCGG
TJP1-mEGFP	GGTCTAATGTGGGGTGTGGG	AAGTCGATGCCCTTCAGCTCG	GTGAGCAAGGGCGAGGAGCTG	TTCTCCCAGCCAGCAAACAA
MYH10-mEGFP	GGGCCATTGTGCCCAGAAGT	GACACGCTGAACTTGTGGC	AAGACCCCAACGAGAAG	CACCGTTCCAACCCTGTGGC

Cell Line	Untagged Allele Fwd	Untagged Allele Rev
PXN-EGFP	GTGACCTCAGTAGCTGCATG	CAGGGGTGAAGACAAGCAG
SEC61B-mEGFP	TCAGTTAGGCCACATCAGCG	GTGCCCTAAACTGAGCAACG
TOMM20-mEGFP	TCTGCCTCCTTTGTTAACTTGAC	TGCTCAGTTTTCACAAACACAGT
TUBA1B-mEGFP	GTCTTGGTCTGGAAGGAGGC	CAAGAGAAGCCCCTGGACAG
LMNB1-mEGFP	CTCGTCTTGCATTTTCCCGC	GACCGAGACCCTGTTCCTTC
FBL-mEGFP	GCCAACTGCATTGACTCCAC	AGCAAAATGGCGACCACAAC
ACTB-mEGFP	CTGGGACTCAAGGCGCTAAC	CGATGGGGTACTTCAGGGTG
DSP-mEGFP	AGGTCTTGTTGACCCGGAAGT	ACGCACTGCATCCAAGTGTACT
TJP1-mEGFP	CCGAGTTGAATTCCCTCCCC	CTATGCACCTGCCCAGTACG
MYH10-mEGFP	TGTGGTGAGGGTGAAAGAGGA	AGACATGGGTAAGCAAGCAACA

Cell Line	Full Allele Fwd	Full Allele Rev
PXN-EGFP	TGTGCAGTGGCACGATCTTGG	TCAGTGAAGAGCTTGCTGGC
SEC61B-mEGFP	TATCTACCTCGGAATCACCC	TGGGCGACAGAGTGAGATTCC
TOMM20-mEGFP	AGCGTGTCTGTTACAAGTGTTG	CCCACCTGCTCCACTCTTTT
TUBA1B-mEGFP	GACTAGGGCTACAGGGC	TTAGTGTAGGTTGGGCGCTC
LMNB1-mEGFP	TTCAAGACGCACAGATCTCAC	ACACATTTCCCCAGAGAAAGC
FBL-mEGFP	ATTACAGGCACGAGCCACTGC	ACGCGGGGGAAGAGTAGAGC
ACTB-mEGFP	AGAAGTCCACCGAGTCCTGC	GTGAAGCTGTAGCGCGCTC
DSP-mEGFP	ACCCTCAGGAAGCGTAGAGT	TGCCAATGCTTTGTTGTCGG
TJP1-mEGFP	GGTCTAATGTGGGGTGTGGG	TTCTCCCAGCCAGCAAACAA
MYH10-mEGFP	GGGCCATTGTGCCCAGAAGT	CACCGTTCCAACCCTGTGGC

All primers are listed in 5′ to 3′ orientation.

Table S3. Antibodies used in western blot, immunofluorescence, and flow cytometry experiments.

Primary Antibodies

Antibody	Type	Source	Application	Secondary Antibody
Alpha tubulin	mouse monoclonal, clone DM1A	ThermoFisher #62204	WB: 1:10,000 IF: 1:250	goat anti-mouse goat anti-mouse
Beta actin	mouse monoclonal, clone GT5512	GeneTex #GTX629630	WB: 1:10,000	goat anti-mouse
Desmoplakin	rabbit polyclonal NW6 rabbit polyclonal 1G4	Kathleen Green, Northwestern University Kathleen Green, Northwestern University	WB, 1:1,000 IF: 1:200	goat anti-rabbit goat anti-rabbit
Fibrillarin	rabbit polyclonal	Abcam #ab5821	WB: 1:800 IF: 1:100	goat anti-rabbit goat anti-rabbit
Lamin B1 (Nuclear lamin B1)	rabbit polyclonal	Abcam #ab16048	WB: 1:2,000 IF: 1:500	goat anti-rabbit goat anti-rabbit
Myosin IIB (Non-muscle myosin heavy chain IIB)	rabbit polyclonal	Cell Signaling Technology #3404	WB, 1:1,000 IF: 1:200	goat anti-rabbit goat anti-rabbit
Paxillin	mouse monocolonal, clone 349	BD Biosciences #610051	WB: 1:10,000 IF: 1:750	goat anti-mouse goat anti-mouse
Sec61 beta rabbit polyclonal	rabbit polyclonal rabbit polyclonal	Abcam #ab15576 Sigma Aldrich #HPA049407	WB: 1:10,000 IF 1:250 IF: 1:25	goat anti-rabbit goat anti-rabbit goat anti-rabbit
Tight junction protein ZO1 (ZO1)	rabbit polyclonal	ThermoFisher #617300	WB: 1:500 IF: 1:50	goat anti-rabbit goat anti-rabbit
Tom20	mouse monoclonal, clone F10	Santa Cruz Biotechnologies #sc17764	WB: 1:250 IF: 1:100	goat anti-mouse goat anti-mouse
Beta actin (loading control)	mouse monoclonal, clone BA3R	ThermoFisher #MA515739	WB: 1:10,000	goat anti-mouse
Alpha actinin (loading control)	mouse monoclonal, clone 0.T.02	ThermoFisher #MA191860	WB: 1:2,000	goat anti-mouse
GFP	mouse monoclonal mix, clones 7.1 and 13.1	Sigma Aldrich #11814460001	WB: 1:250	goat anti-mouse

Directly Conjugated Primary Antibodies

Antibody	Type	Source	Application
TRA-1-60, Brilliant Violet™ 510	mouse monoclonal, clone TRA-1-60	BD Biosciences #563188	Flow: 1:40
SSEA-3, Alexa Fluor® 647	rat monoclonal, clone MC-631	BD Biosciences #561145	Flow: 1:160
SSEA-1, Brilliant Violet™ 421	mouse monoclonal, clone MC480	BD Biosciences #562705	Flow: 1:2,500
Nanog, Alexa Fluor® 647	mouse monoclonal, clone N31-355	BD Biosciences #561300	Flow: 1:67
Oct3/4, Brilliant Violet™ 510	mouse monoclonal, clone 40/Oct-3	BD Biosciences #563524	Flow: 1:67
Sox2, V450	mouse monoclonal, clone O30-678	BD Biosciences #561610	Flow: 1:80
TRA-1-60, PerCP-Cy™5.5	mouse monoclonal, clone TRA-1-60	BD Biosciences #561573	Flow: 1:25
SSEA-3, PE	rat monoclonal, clone MC-631	BD Biosciences #560237	Flow: 1:10
SSEA-1, Brilliant Violet™ 650	mouse monoclonal, clone HI98	BD Biosciences #564232	Flow: 1:10
Oct3/4, PerCP-Cy™5.5	mouse monoclonal, clone 40/Oct-3	BD Biosciences #560794	Flow: 1:5
Sox2, Alexa Fluor® 647	mouse monoclonal, clone 245610	BD Biosciences #560294	Flow: 1:25
Nanog, PE	mouse monoclonal, clone N31-355	BD Biosciences #560483	Flow: 1:2.5
Cardiac Troponin T, Alexa Fluor® 647	mouse monoclonal, clone 13-11	BD Biosciences #565744	Flow: 1:200 IF: 1:200
Mouse IgG1 k Isotype Control, Alexa Fluor® 647	mouse monoclonal, clone MOPC-21	BD Biosciences #565571	Flow: 1:200
Brachyury, APC	goat polyclonal	R&D Systems #IC2085A	Flow: 1:40
Goat IgG Isotype Control, APC	goat polyclonal	R&D Systems #IC108A	Flow: 1:40
Sox17, APC	goat polyclonal	R&D Systems #IC1924A	Flow: 1:80
Goat IgG Isotype Control, APC	goat polyclonal	R&D Systems #IC108A	Flow: 1:80
Pax6, Alexa Fluor® 647	mouse monoclonal, clone O18-1330	BD Biosciences #562249	Flow: 1:20
Mouse IgG2a k Isotype Control, Alexa Fluor® 647	mouse monoclonal, clone MOPC-173	BD Biosciences #558053	Flow: 1:10

Secondary Antibodies

Antibody	Type	Source	Application
goat anti-mouse IgG (H+L), Alexa Fluor® 647 conjugate	goat polyclonal	ThermoFisher #A21236	WB: 1:10,000 IF: 1:500
goat anti-rabbit IgG (H+L), Alexa Fluor® 647 conjugate	goat polyclonal	ThermoFisher #A21245	WB: 1:10,000 IF: 1:500

Table of antibodies used in western blots (WB), immunofluorescence (IF), and flow cytometry (Flow) experiments showing dilutions used per application.

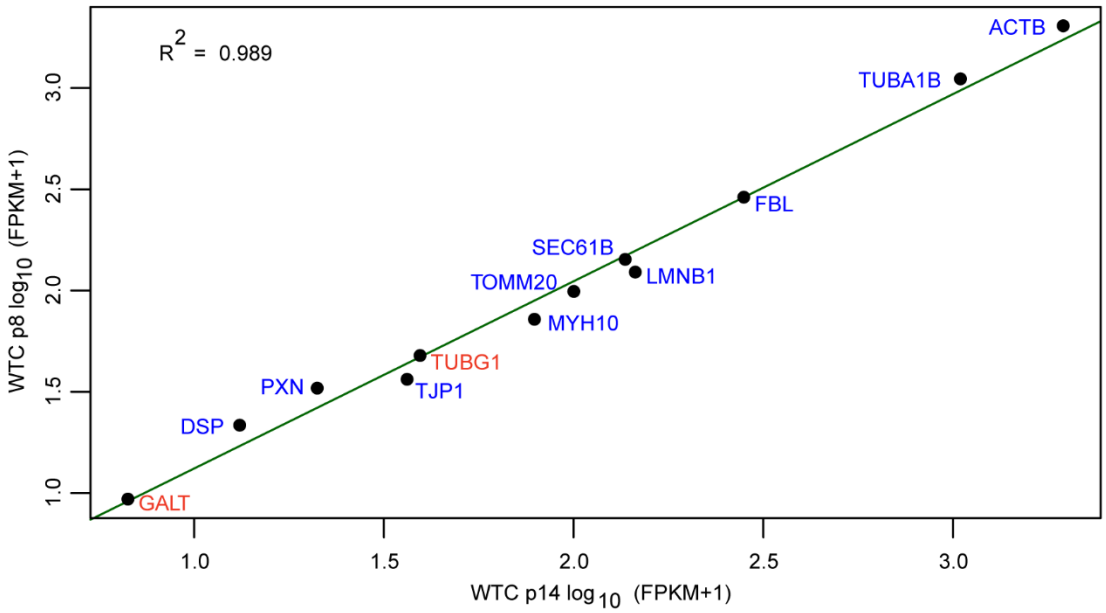


Figure S1. Expression levels of the 12 genes attempted for genome editing in the WTC parental cell line. Transcript abundance for each gene was estimated from RNA-Seq data. Samples were derived from the WTC parental line after 8 passages (p8) in culture and 14 passages (p14) in culture, as indicated. 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.

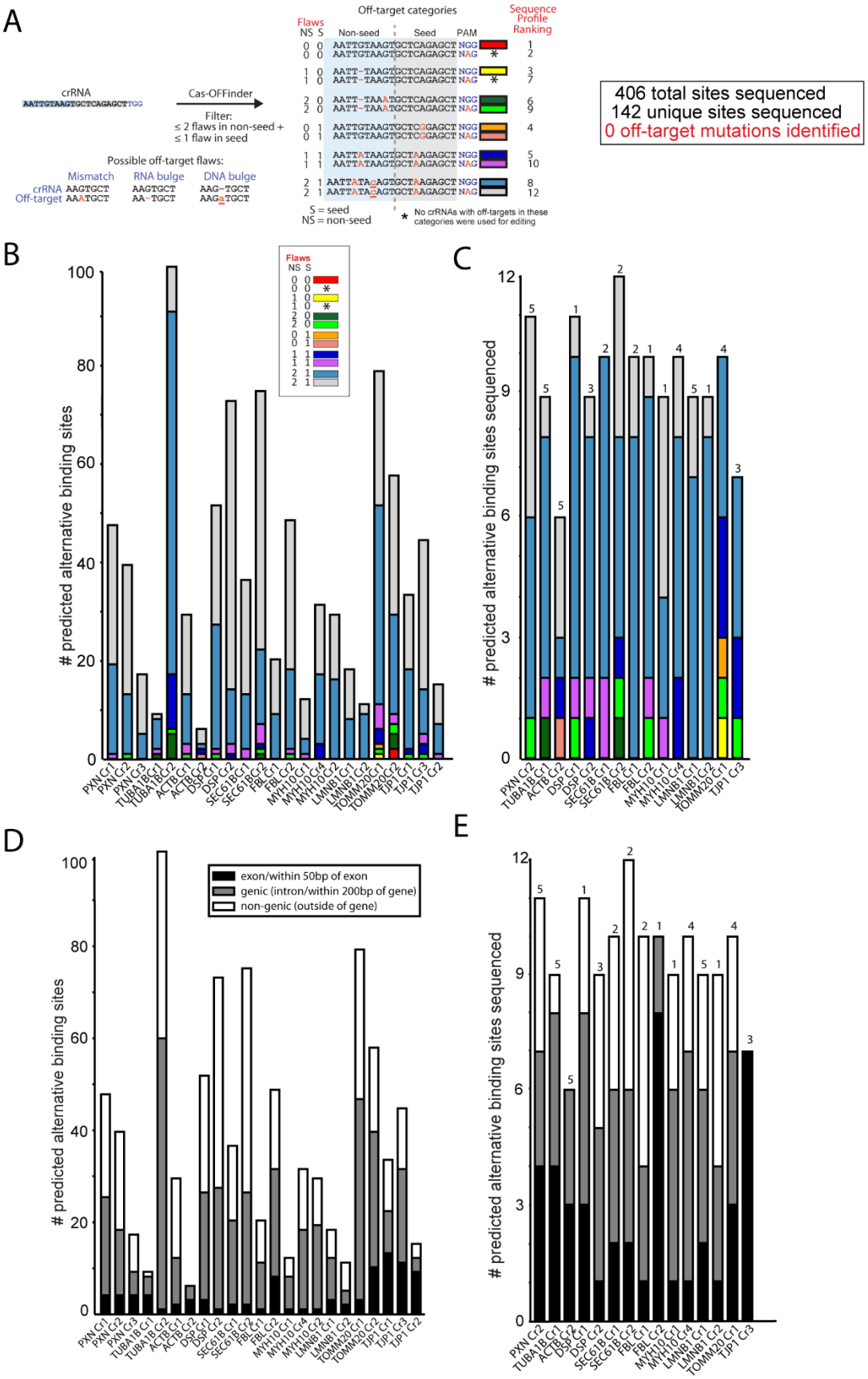

A
crRNA
AATTGTAAGTGCTCAGAGCTTGG
Cas-OFFinder
Filter:
≤ 2 flaws in non-seed +
≤ 1 flaw in seed
Possible off-target flaws:
Mismatch
RNA bulge
DNA bulge
crRNA AAGTGCT
Off-target AAATGCT
AAGTGCT
AA-TGCT
AAG-TGCT
AAGATGCT
Off-target categories
Flaws
NS S
Non-seed
Seed
PAM
Sequence Profile Ranking
S = seed
NS = non-seed
* No crRNAs with off-targets in these categories were used for editing
406 total sites sequenced
142 unique sites sequenced
0 off-target mutations identified
B
predicted alternative binding sites
C
predicted alternative binding sites sequenced
D
exon/within 50bp of exon
genic (intron/within 200bp of gene)
non-genic (outside of gene)
predicted alternative binding sites
E
predicted alternative binding sites sequenced

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 ≤2 mismatches/bulges in the non-seed and/or ≤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; exon: inside exon or within 50 bp of exon; genic: in intron (but >50 bp from an exon) or within 200 bp of an annotated gene; 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.

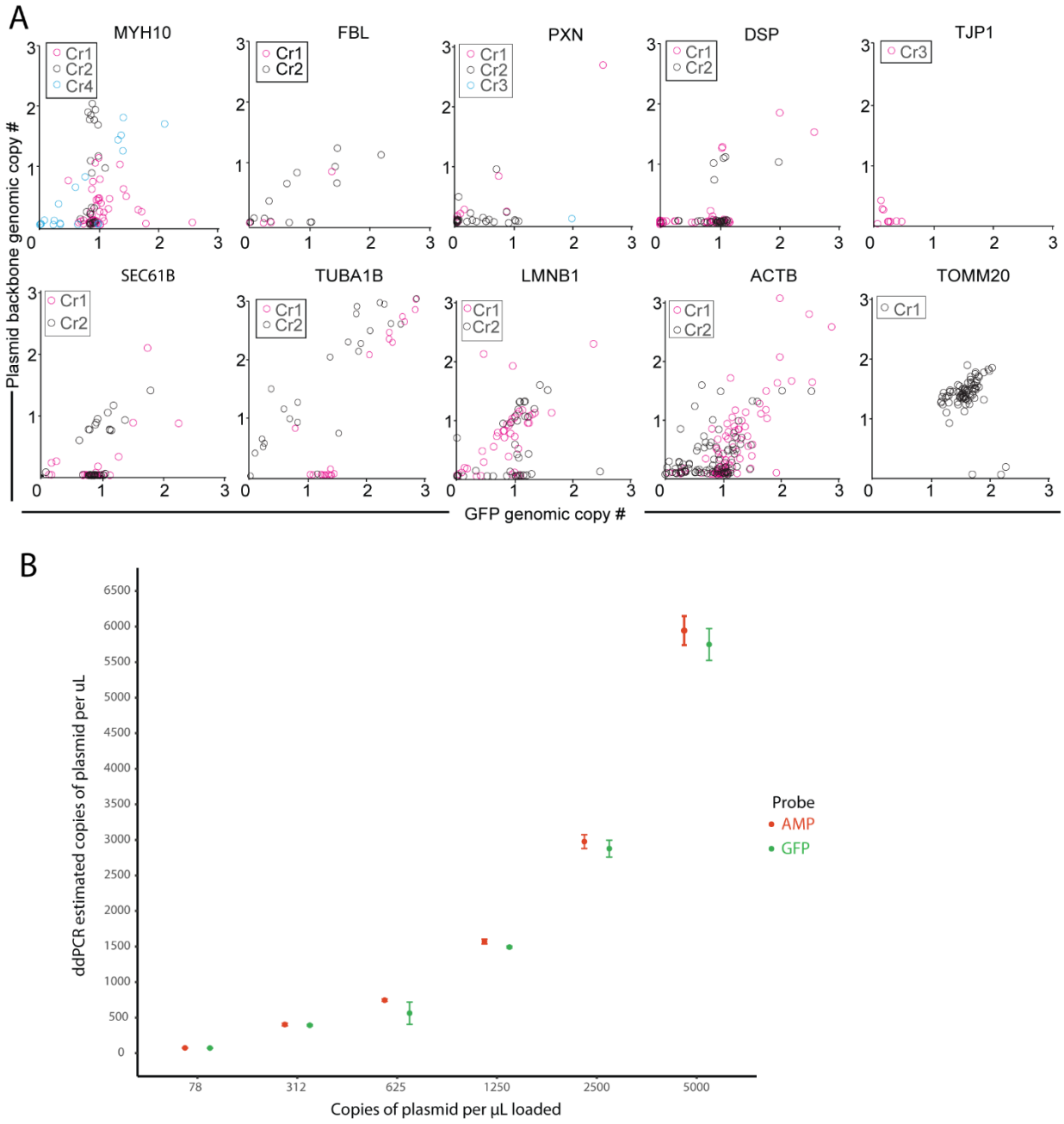


Figure S3. ddPCR screening data. (A) ddPCR screening data for all experiments (Fig. 2A step 1). Each data point represents one clone. Clones with GFP genomic copy number of ~1 to ~2 and plasmid backbone genomic copy number <0.2 were typically considered for further analysis. TJP1 clones consistently 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.

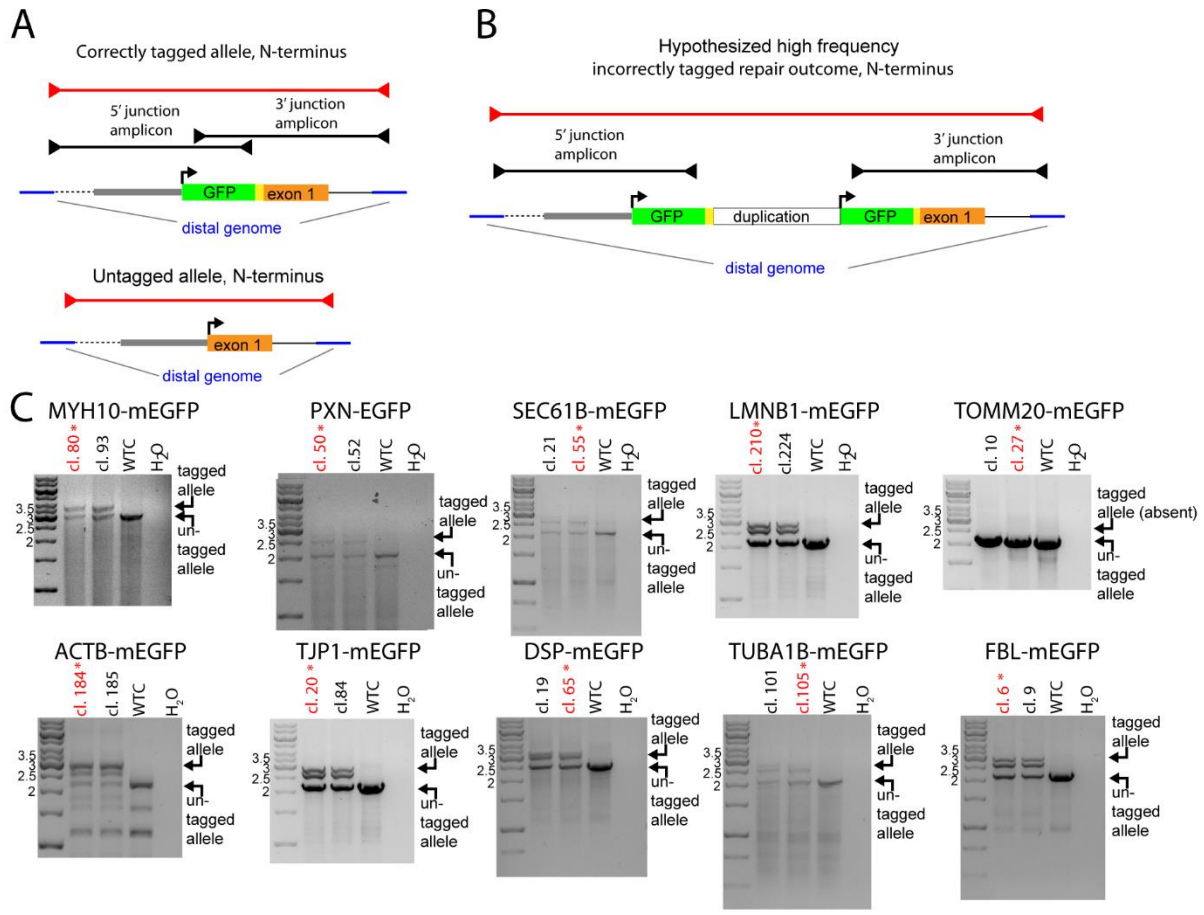


Figure S4. Amplification of complete junctional (non-tiled) PCR products to demonstrate presence of the allele anticipated from tiled junctional PCR product data. (A) Junctional PCR primers complementary to sequences flanking the homology arms in the distal genome (also used in tiled junctional PCR assays, 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.

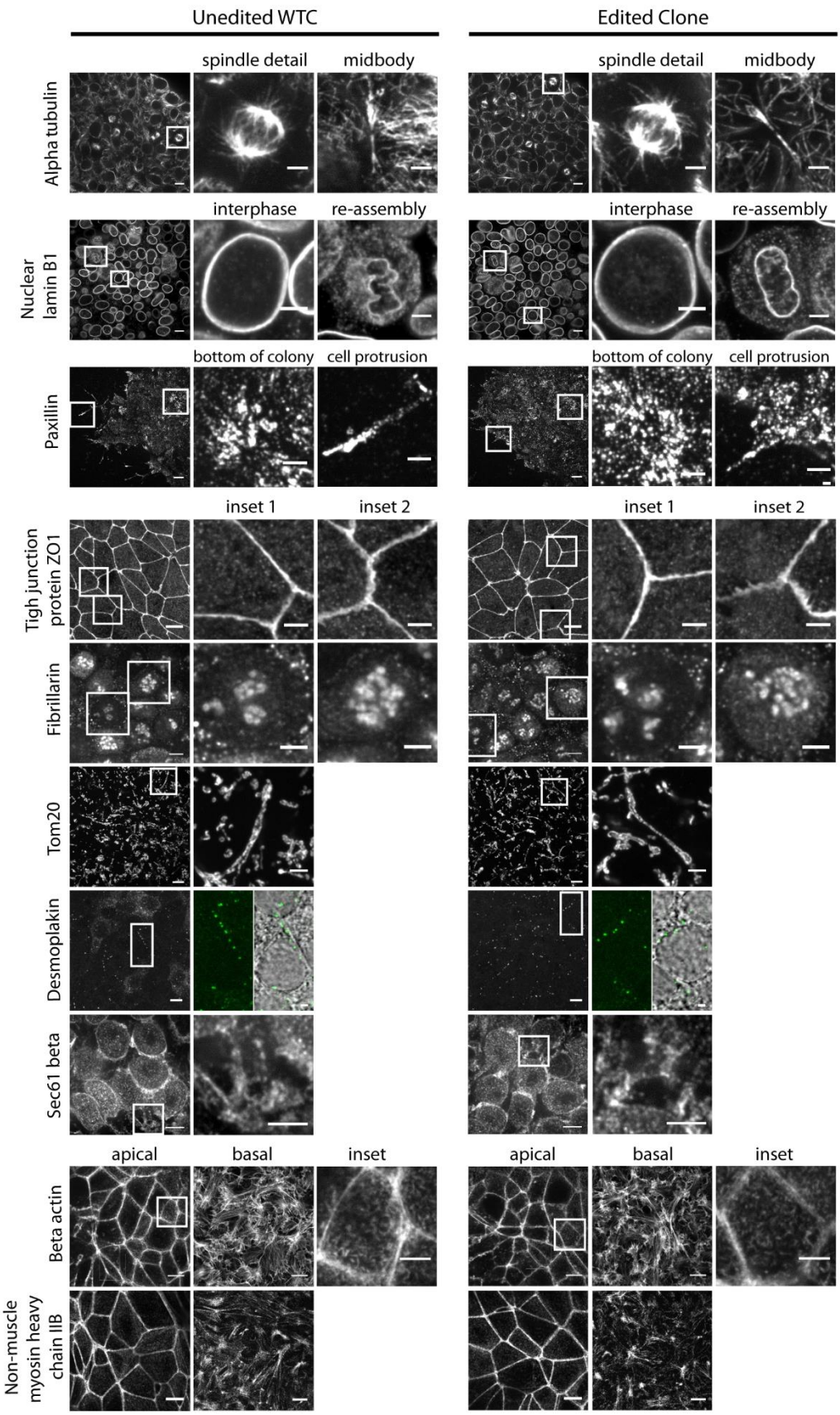

Unedited WTC
Edited Clone
spindle detail
midbody
Alpha tubulin
interphase
re-assembly
Nuclear lamin B1
bottom of colony
cell protrusion
Paxillin
inset 1
inset 2
Tigh junction protein ZO1
Fibrillarin
Tom20
Desmoplakin
Sec61 beta
apical
basal
inset
Beta actin
Non-muscle myosin heavy chain IIB

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).

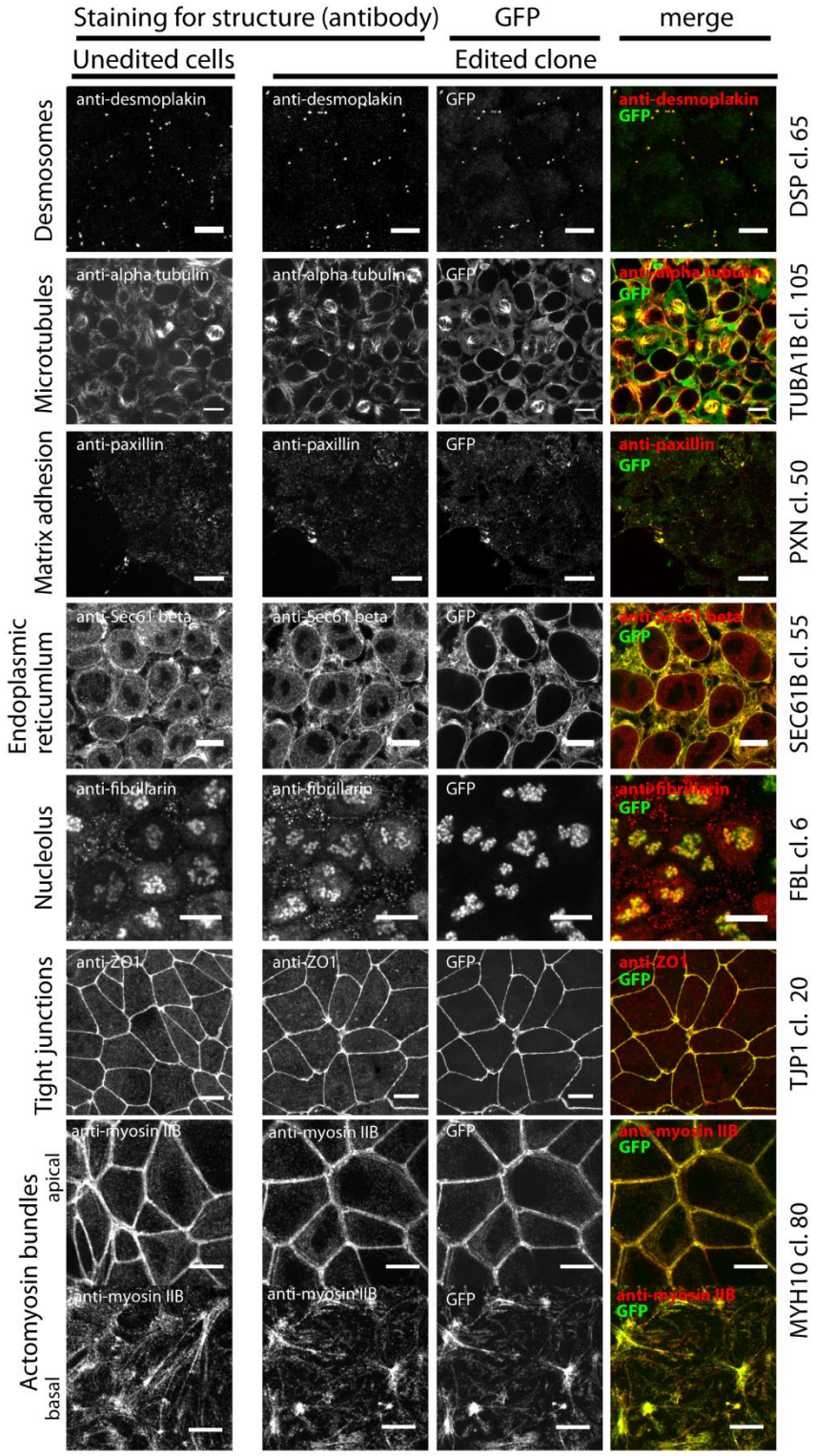


Figure S6. Comparison of GFP tag localization and endogenous protein stain in edited cell lines. Antibodies raised against the tagged protein were used to stain unedited and edited cells, as indicated. Labels on left of images indicate the tagged structure, and labels on the right indicate tagged gene 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).

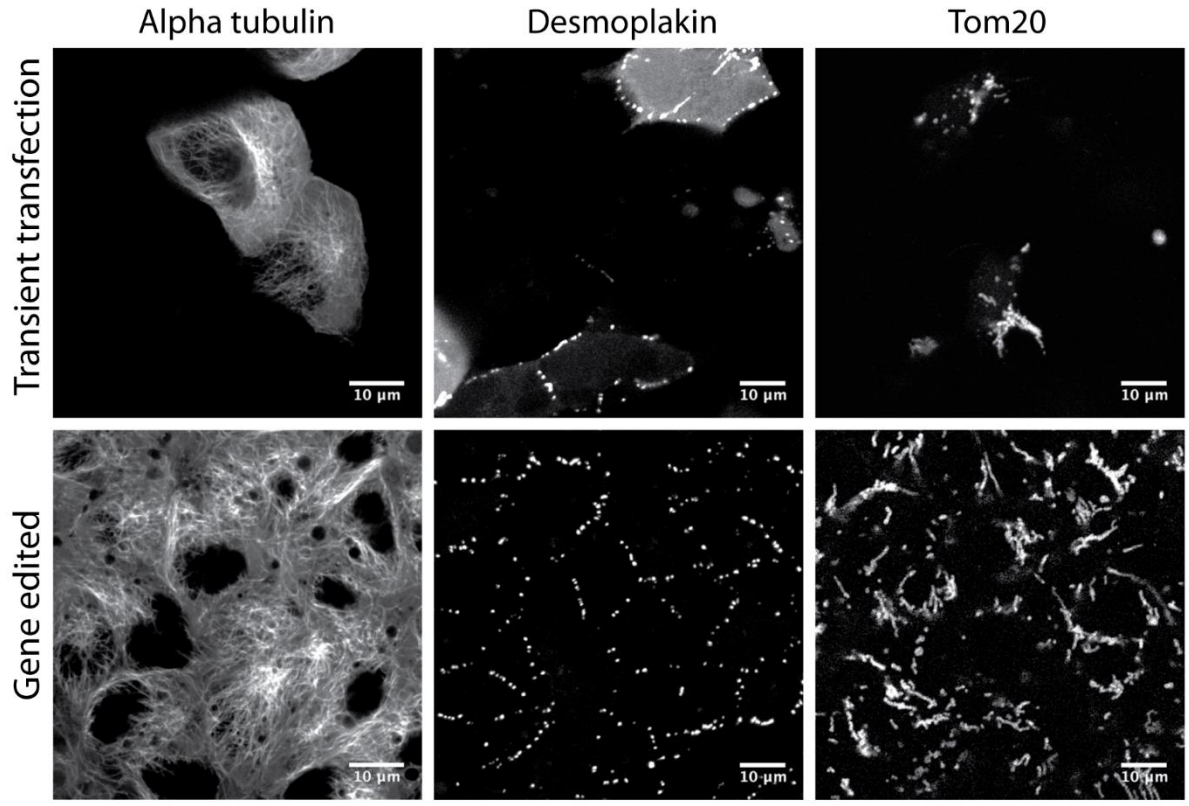
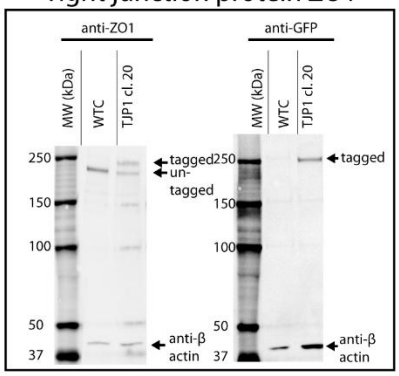


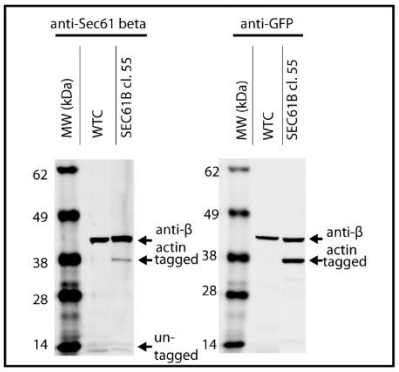
Figure S7. Live cell imaging comparison of transiently transfected cells and genome edited cells.

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).

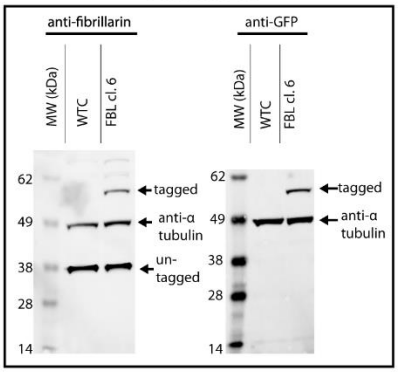
Tight junction protein ZO1



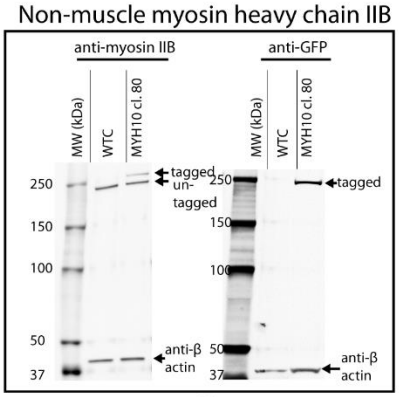
Sec61 beta



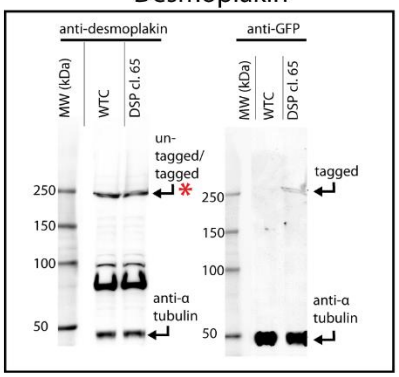
Fibrillarin



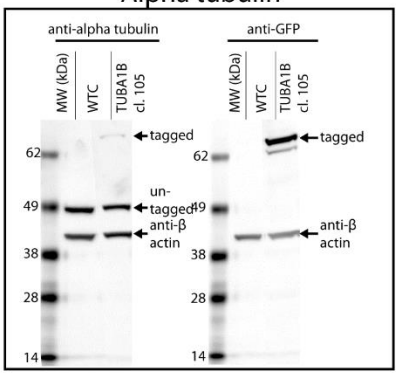
Non-muscle myosin heavy chain IIB



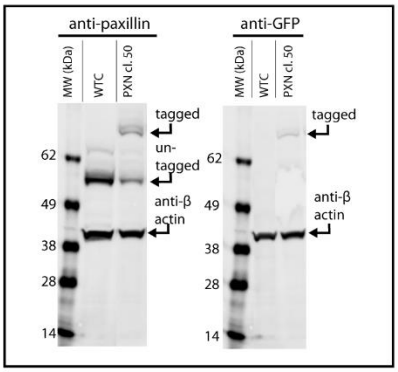
Desmoplakin



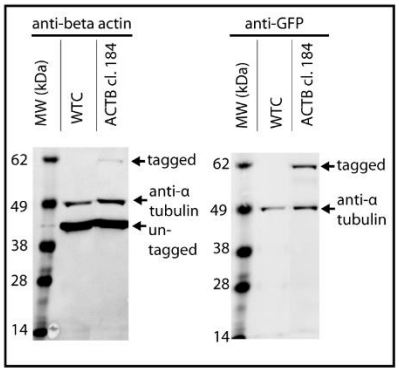
Alpha tubulin



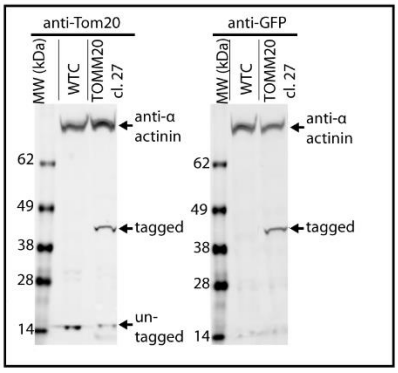
Paxillin



Beta actin



Tom20



Nuclear lamin B1

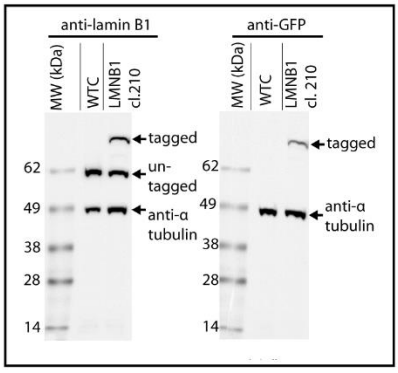
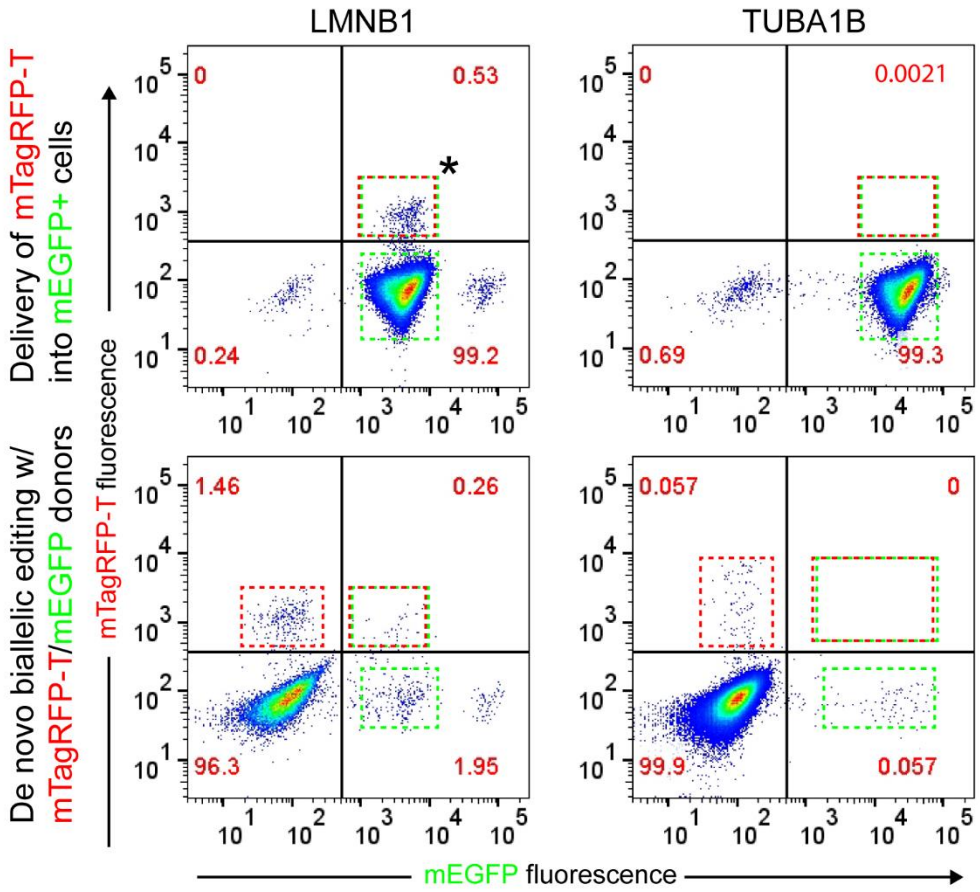


Figure S8. Western blot analysis of all 10 edited clonal lines. Western blot analyses for all experiments are presented as in Fig. 4B. Proteins and antibodies used on different blots are as indicated. In all cases, blots with antibodies against the respective target proteins are shown in the left blot 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.

A


LMNB1
TUBA1B
Delivery of mTagRFP-T into mEGFP+ cells
De novo biallelic editing w/ mTagRFP-T/mEGFP donors
mTagRFP-T fluorescence
mEGFP fluorescence
0
0.53
0.24
99.2
0
0.0021
0.69
99.3
1.46
0.26
96.3
1.95
0.057
0
99.9
0.057
*

B

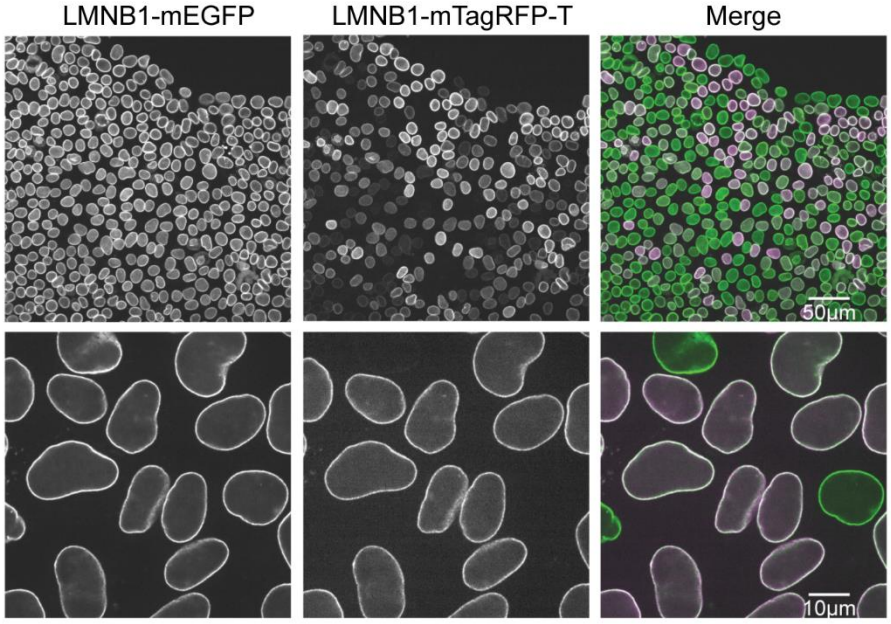

LMNB1-mEGFP
LMNB1-mTagRFP-T
Merge
50μm
10μm

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.

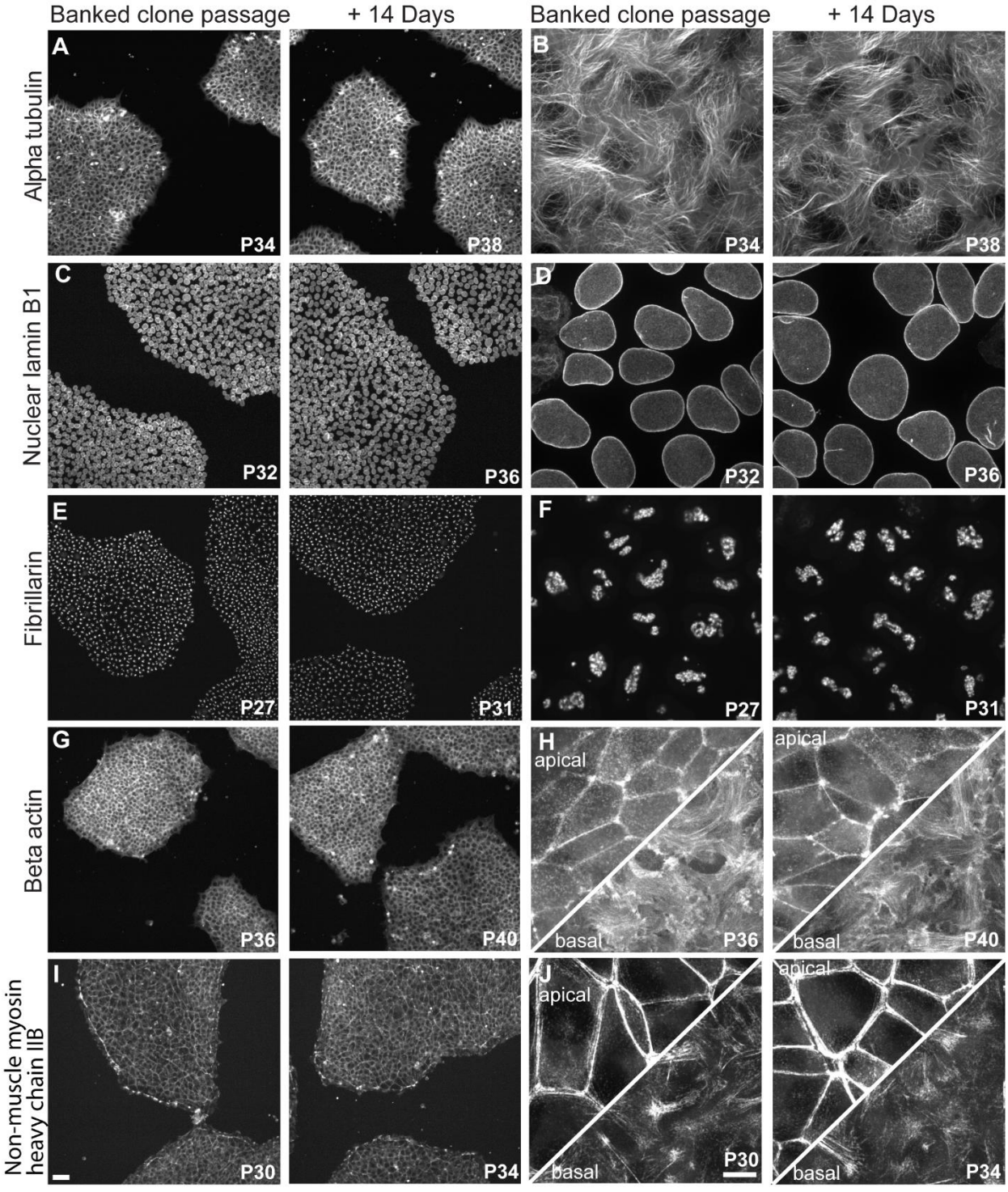
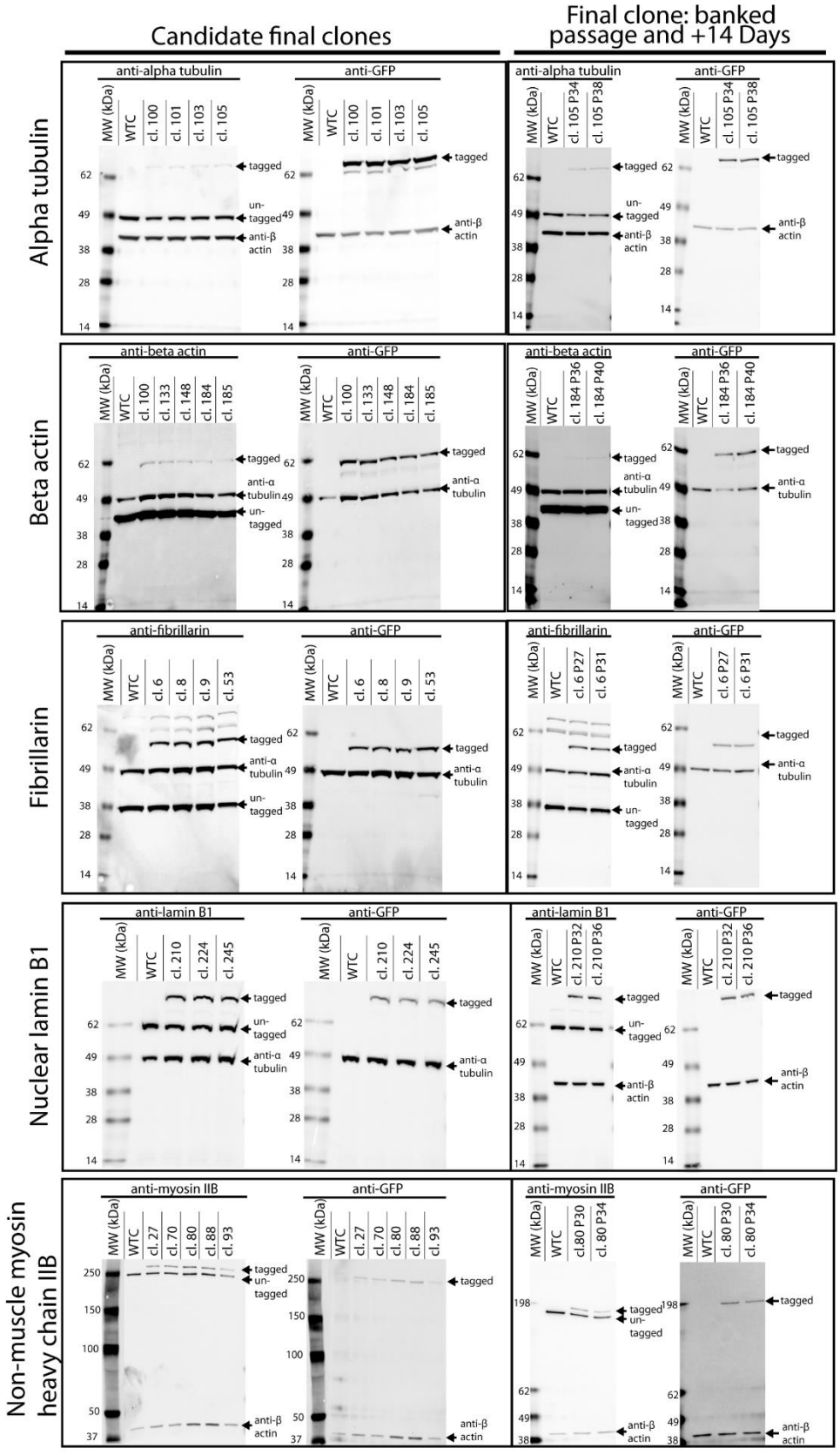

Banked clone passage
+ 14 Days
Banked clone passage
+ 14 Days
Alpha tubulin
Nuclear lamin B1
Fibrillarin
Beta actin
Non-muscle myosin heavy chain IIB
A
B
C
D
E
F
G
H
I
J
P34
P38
P34
P38
P32
P36
P32
P36
P27
P31
P27
P31
P36
P40
apical
basal
P36
apical
basal
P40
P30
P34
apical
basal
P30
apical
basal
P34

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.



clone	% GFP tagged
cl. 100	65
cl. 101	62
cl. 103	58
cl. 105	55
cl. 105 p34	53
cl. 105 p38	52

cl. 100	5
cl . 133	5
cl. 148	6
cl. 185	6
cl. 184	5
cl. 184 p36	4
cl. 184 p40	4

cl. 8	20
cl. 9	21
cl. 53	32
cl. 6	26
cl. 6 p27	29
cl. 6 p31	26

cl. 224	36
cl. 245	31
cl. 210	33
cl. 210 p32	29
cl. 210 p36	32

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.

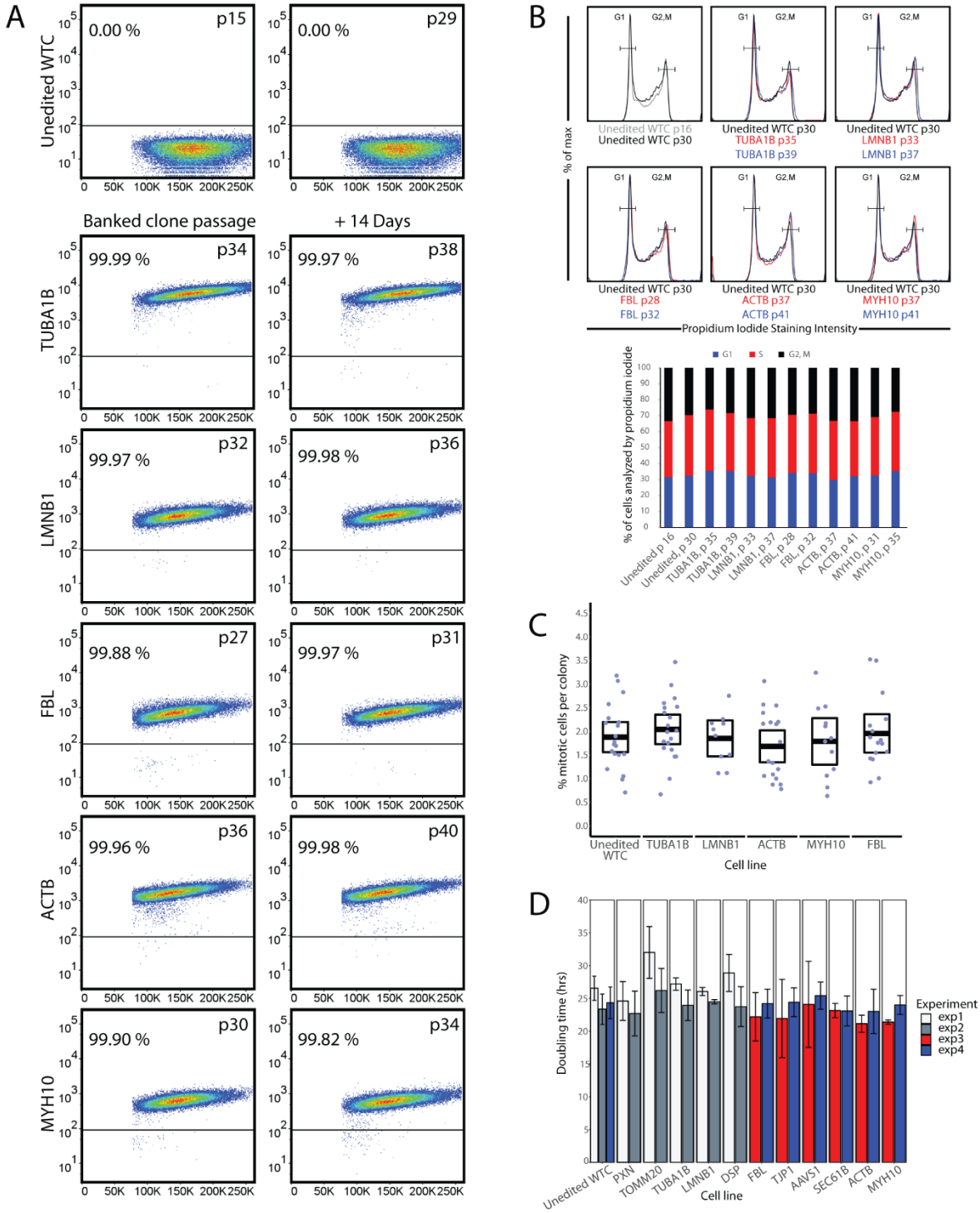


Figure S12. Flow cytometry analysis of GFP tag expression stability, flow cytometry analysis of cell cycle dynamics, microscopy analysis of mitotic index, and culture growth assays. (A) Endogenous GFP signal in final edited clones was compared in otherwise identical cultures separated by four passages (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/