Creating an open source collection of GFP-tagged human iPSC lines to model stem cell organization and dynamics

W-2137

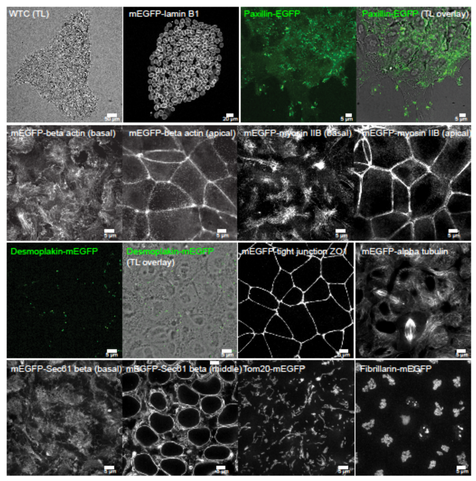
Brock Roberts, Angelique Nelson, Joy Arakaki, Kaytlyn Gerbin, Amanda Haupt, Andrew Tucker, Tanya Grancharova, Margaret A. Fuqua, Caroline Hookway, Susan A. Ludmann, Irina A. Mueller, Ruian Yang, Alan R. Horwitz, Susanne M. Rafelski, and Ruwanthi N. Gunawardane and the Allen Institute for Cell Science

ALLEN INSTITUTE for CELL SCIENCE

Abstract

The Allen Institute for Cell Science (AICS) is creating a dynamic visual model of hiPSC organization to aid in understanding and predicting normal and pathological cell states. Our approach utilizes CRISPR/Cas9 gene editing to introduce fluorescent tags via homology driven repair (HDR) into genomic loci whose products localize to specific organelles. Editing yields isogenic hiPSC lines expressing fusion proteins unique to each cell line under endogenous regulation. Live cell imaging, image analysis and modeling, and open distribution to the scientific community of each unique cell line defines our endeavor. Because we will perform systematic editing at numerous genomic loci, our data has begun and will continue to elucidate variables and trends important for gene editing in stem cells. Here we present our CRISPCR/Cas9- based gene editing protocol and workflow to introduce fluorescent tags into the genome of stem cells and our initial progress and conclusions from the generation of ~1000 clones spanning 10 different targets. We will describe our screening strategy to identify clones harboring precisely incorporated GFP tags at the genomic loci and demonstrate the various consequences of imprecise editing. We will also present our quality control assays including the characterization of stem cell properties, off-target analysis, karyotyping, directed differentiation into cardiomyocytes, and next generation sequencing. Furthermore, we will present data supporting the correct subcellular localization of the tagged proteins from imaging studies. In experiments initiated to date we have generated hiPSC lines for ~15 major cellular structures including cell-matrix adhesions, the actin and microtubule cytoskeleton, mitochondria, desmosomes, endoplasmic reticulum, and nuclear envelope.

The Allen Cell Collection consists of openly available hiPSC lines with gene tags marking organelles

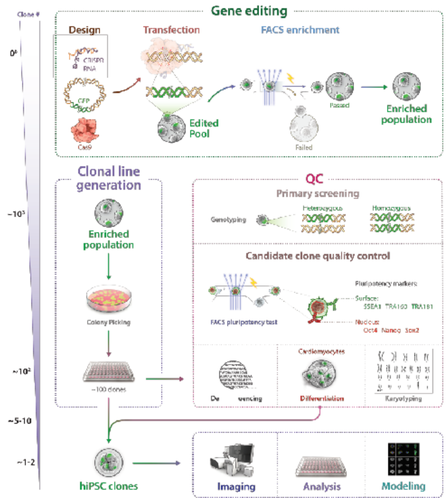


Gene	Protein	Cellular structure	Terminus tagged	FP
Paxillin	Paxillin (PXN)	Matrix adhesions	C-terminus	EGFP
Sec61 beta	Sec61 translocon beta subunit (SEC61B)	Endoplasmic reticulum	N-terminus	mEGFP
TOM20	Translocase of outer mitochondrial membrane 20 (TOMM20)	Mitochondria	C-terminus	mEGFP
Alpha tubulin	Tubulin-alpha 1b (TUBA1B)	Microtubules	N-terminus	mEGFP
Nuclear lamin B1	Lamin B1 (LMNB1)	Nuclear envelope	N-terminus	mEGFP
Fibrillarin	Fibrillarin (FBL)	Nucleolus	C-terminus	mEGFP
Beta actin	Actin beta (ACTB)	Actin filaments	N-terminus	mEGFP
Desmoplakin	Desmoplakin (DSP)	Desmosomes	C-terminus	mEGFP
Tight junction protein ZO1	Tight junction protein 1 (TJP1)	Tight junctions	N-terminus	mEGFP
Non-muscle myosin heavy chain IIB	Myosin heavy chain 10 (MYH10)	Actomyosin bundles	N-terminus	mEGFP
NA	Safe harbor locus (AAVS1)	Cytoplasm	NA	mEGFP

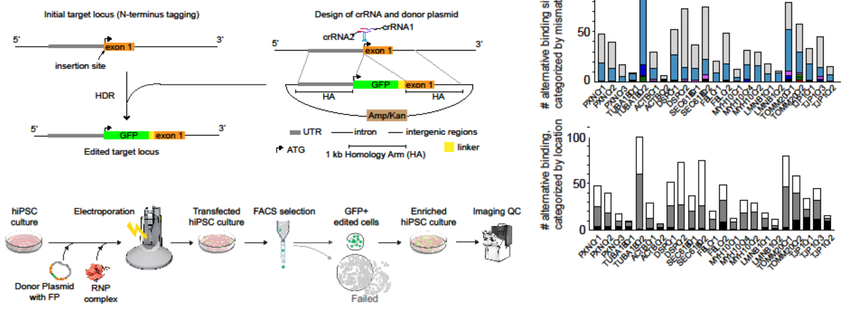
Under development and likely ready for release in 2017:

LAMP1	Lysosomal associated membrane protein 1 (LAMP1)	Lysosome	C-terminus	mEGFP
Beta-galactoside alpha-2,6-sialyltransferase 1	ST6 beta-galactoside alpha-2,6-sialyltransferase 1 (ST6GAL1)	Golgi	C-terminus	mEGFP
LaminB1	Lamin B1 (LMNB1)	Nucleus	N-terminus	tdTomato
LC3	Microtubule associated protein 1 light chain 3 beta (MAP1LC3B)	Autophagosomes	N-terminus	mEGFP
Alpha tubulin	Tubulin-alpha 1b (TUBA1B)	Microtubules	N-terminus	mtagRFP-T
Centrin-2	Centrin 2 (CETN2)	Centrosome	N-terminus	mtagRFP-T
Peroxisomal membrane protein PMP34	Solute carrier family 25 member 17 (SLC25A17)	Peroxisomes	C-terminus	mEGFP
Ras-related protein Rab-5A	RAB5A, member RAS oncogene family (RAB5A)	Endosomes	N-terminus	mEGFP
Connexin-43	Gap junction protein alpha 1 (GJA1)	Gap junction	C-terminus	mEGFP
NA	Safe harbor locus (AAVS1)	Plasma membrane	NA	CaaX-mtagRFP-T

Gene tagging requires design, selection, screening and QC



Designed donor plasmids are co-electroporated with Cas9 protein complexed with low-promiscuity crRNAs



-RNP delivery of CRISRP/Cas9 and donor plasmids with Neon transfection leads to successful editing in most experiments
-crRNAs are chosen based both on proximity to tag site and off-target profile

crRNA: AATTGTAAGTGCTCAGAGCTTGG

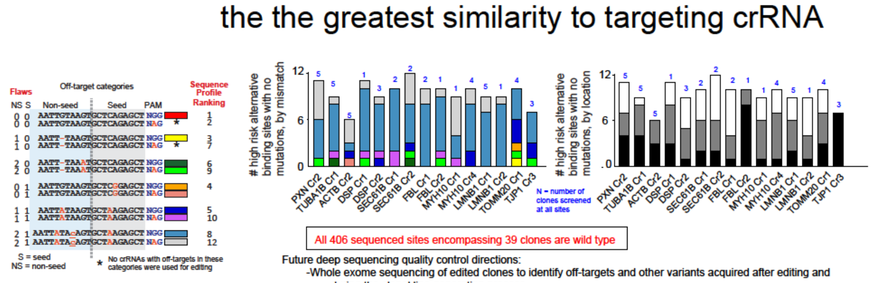
Possible off-target flaws:

	Mismatch	RNA bulge	DNA bulge
crRNA	AAGTGCT	AAGTGCT	AAG−TGCT
Off-target	AAATGCT	AA−TGCT	AAGaTGCT

Cas-OFFinder filter:

≤ 2 flaws in non-seed + ≤ 1 flaw in seed
NGG and NAG PAMs considered

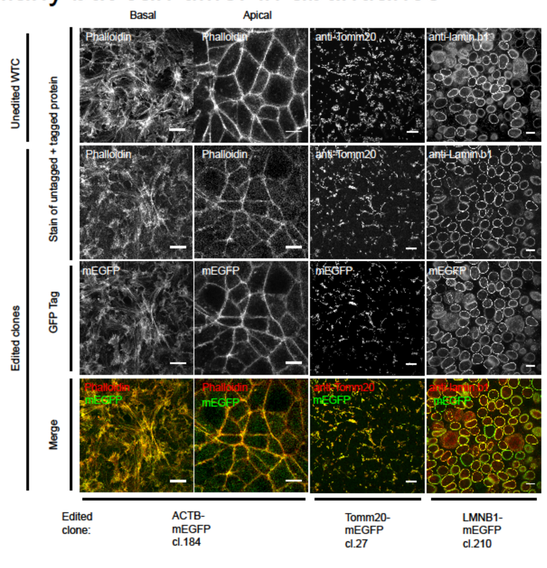
No mutations are observed in sequences with the the greatest similarity to targeting crRNA



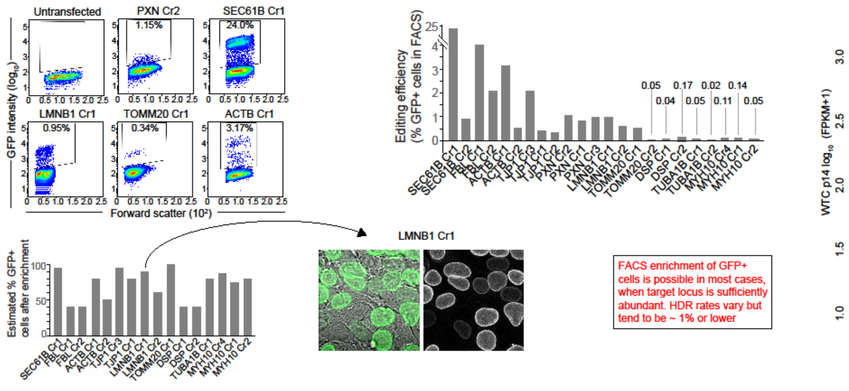
All 406 sequenced sites encompassing 39 clones are wild type

Future deep sequencing quality control directions:
-Whole exome sequencing of edited clones to identify off-targets and other variants acquired after editing and during the clonal line generation process
-Transcriptome sequencing to assess effect of editing/culturing on transcriptional profiles of edited stem cells

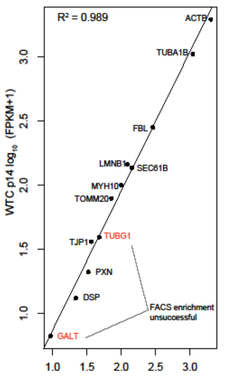
Tagged and untagged protein copies localize similarly but can differ in abundance



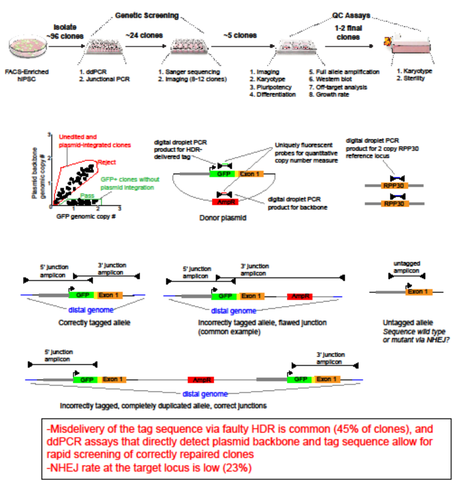
FACS is used to enrich for edited cells and determine HDR

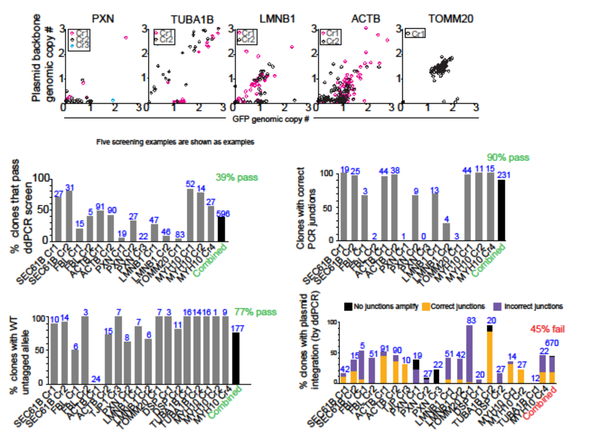


FACS enrichment of GFP+ cells is possible in most cases, when target locus is sufficiently abundant. HDR rates vary but tend to be ~ 1% or lower



Individual clones from FACS-enriched cultures are screened with PCR

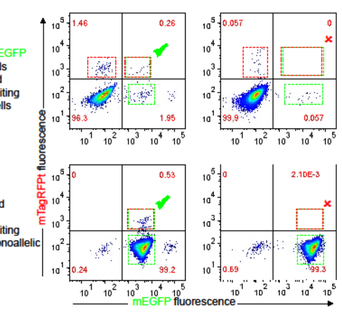


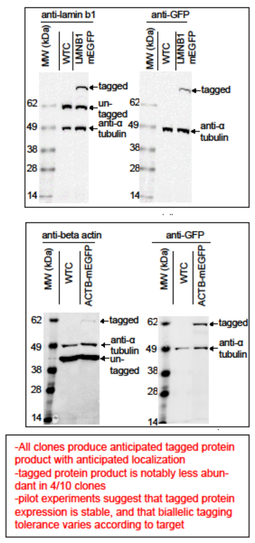


-Misdelivery of the tag sequence via faulty HDR is common (45% of clones), and ddPCR assays that directly detect plasmid backbone and tag sequence allow for rapid screening of correctly repaired clones
-NHEJ rate at the target locus is low (23%)

mTagRFPt/mEGFP donor plasmids co-transfected for biallelic editing in unedited cells

mTagRFPt donor plasmid transfected for biallelic editing in mEGFP monoallelic cell line



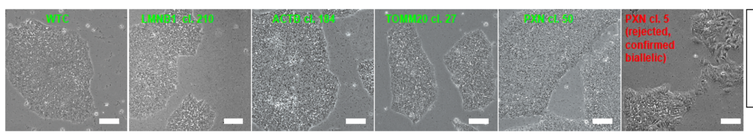


Protein	Clone	Relative level of total protein	Relative level of tagged protein
	WTC	1.00	N/A
Alpha tubulin	cl. 105	0.83	0.55
Beta actin	cl. 184	1.35	0.05
Desmoplakin	cl. 65	n.a	n.a
Fibrillarin	cl. 6	0.95	0.23
Nuclear lamin B1	cl. 210	0.80	0.33
Non-muscle myosin heavy chain IIB	cl. 80	1.15	0.33
Paxillin	cl. 50	1.38	0.49
Sec61 beta	cl. 55	1.47	0.53
Tight junction protein ZO1	cl. 20	1.07	0.50
Tom20	cl. 27	1.44	0.54

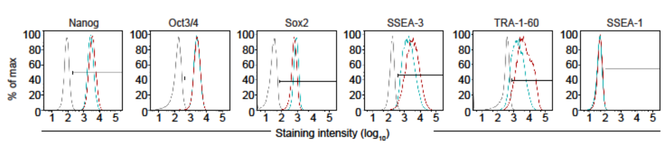
N/A = not applicable n.a. = not quantified

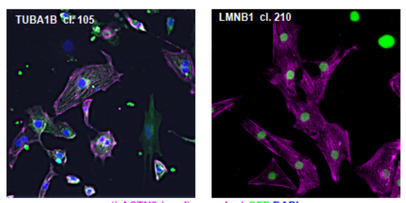
-All clones produce anticipated tagged protein product with anticipated localization
-tagged protein product is notably less abundant in 4/10 clones
-pilot experiments suggest that tagged protein expression is stable, and that biallelic tagging tolerance varies according to target

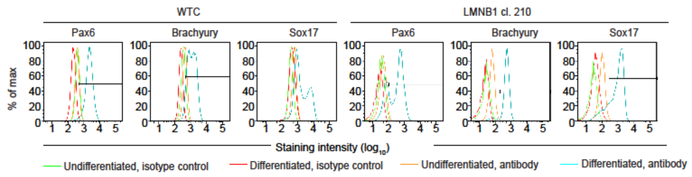
Morphology, pluripotency marker expression and directed differentiation into germ layers and cardiomyocytes are used to vet and rank clones

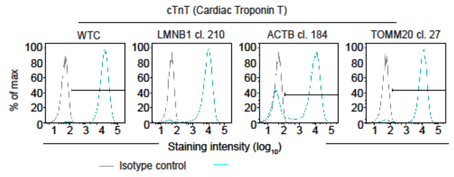


-Clones expanded after verification of correct editing retain stem cell morphology
-Final clones obustly express pluripotency markers and differentiate into multiple fates
-Minor quantifiable differences in pluripotency and differentiation are used to rank clones
-Cardiomyocytes are structurally mature









Clones with correct editing are validated by quality control criteria at a high rate

Cell Line/ Final clone	Pluripotency markers, % positive (n) Oct3/4	Nanog	Sox2	SSEA-3	TRA-1-60	SSEA-1
Unedited	>95 (7)	>97 (7)	>97 (7)	>97 (7)	>97 (7)	<2 (7)
PXN cl. 50	>97 (4)	>99 (4)	>98 (4)	100 (4)	>98 (4)	<4 (4)
TUBA1B cl. 105	>98 (2)	100 (2)	>99 (2)	100 (2)	>98 (2)	<6 (2)
LMNB1 cl. 210	>95 (2)	>99 (2)	100 (2)	>98 (2)	>86 (2)	<4 (2)
TOMM20 cl. 27	>98 (3)	100 (3)	>99 (3)	>98 (3)	>91 (3)	<9 (3)
DSP cl. 65	>98 (3)	>99 (3)	>99 (3)	>97 (3)	>86 (3)	<5 (3)
ACTB cl. 184	95	100	100	100	94	8
SEC61B cl. 55	97	100	100	95	97	1
FBL cl. 6	97	100	100	98	91	2
MYH10 cl.80	96	98	98	100	99	4
TJP1 cl. 20	100	99	100	100	94	4

Cell Line/ Final clone	Germ layer markers, % positive (n) Brachyury (mesoderm)	Sox17 (endoderm)	Pax6 (ectoderm)	Cardiomyocyte differentiation % cTnT+ (n)	% of expts with beating cells (n)	Range in day of beating initiation
Unedited	84-97 (3)	47-70 (3)	95-99 (3)	70-98 (4)	89 (38)	d8-d15
PXN cl. 50	96	58	65	66-98 (4)	75 (16)	d7-d17
TUBA1B cl. 105	92	80	89	87-96 (4)	75 (12)	d8-d20
LMNB1 cl. 210	93	81	72	89-97 (4)	80 (10)	d7-d12
TOMM20 cl. 27	94	82	93	80-95 (2)	75 (12)	d7-d13
DSP cl. 65	91	57	78	92-95 (2)	50 (10)	d7-d15
ACTB cl. 184	>94 (2)	48 (2)	>98 (2)	72	100 (4)	d7-d8
SEC61B cl. 55	>97 (2)	76-79 (2)	797 (2)	95	80 (5)	d7-d8
FBL cl. 6	>96 (2)	57-75 (2)	>98 (2)	64	75 (4)	d7-d9
MYH10 cl.80	>94 (2)	52-64 (2)	98 (2)	75-99 (2)	100 (4)	d7-d16
TJP1 cl. 20	>94 (2)	53-66 (2)	>99 (2)	75-93 (3)	100 (8)	d7-d17

Experiment	Clones isolated (n)	Clones analyzed (n)	Clones passed all other QC (passed/analyzed)	Final clone chosen for distribution (crRNA used)
SEC61B	118	58	3/3	cl. 55 (Cr2)
FBL	58	20	2/2	cl. 6 (Cr1)
ACTB	192	181	5/5	cl. 184 (Cr2)
TJP1	93	91	4/4	cl. 20 (Cr3)
PXN	79	68	4/5	cl. 50 (Cr2)
LMNB1	202	93	3/3	cl. 210 (Cr1)
TOMM20	96	83	4/4	cl. 27 (Cr1)
DSP	173	164	2/3	cl. 65 (Cr2)
TUBA1B	144	138	4/5	cl. 105 (Cr1)
MYH10	102	93	5/5	cl. 80 (Cr4)
Total	1257	989	36/39	

36/39 clones validated according to:
-normal karyotype
-normal morphology
-robust pluripotency
-cardiomyocyte differentiation
-correct HDR
-absence of NHEJ in high-risk off-target sites

Resources

Poster T-2034 Thursday 6/15: CREATING A CARDIOMYOCYTE PIPELINE FOR GENE EDITED HUMAN IPSCS
Allen Cell Explorer: http://www.allencell.org/
Allen Cell Collection: https://catalog.coriell.org/1/AllenCellCollection
Allen Plasmid Collection: https://www.addgene.org/allen-institute-cell-science/

We wish to thank the Allen Institute for Cell Science founder, Paul G. Allen, for his vision, encouragement, and support.