The
CRISPR
Journal

RESEARCH ARTICLE

Clonally Selected Lines After CRISPR-Cas Editing Are Not Isogenic

Arijit Panda,[1,†] Milovan Suvakov,[1,†] Jessica Mariani,[2,3,†] Kristen L. Drucker,[4,†] Yohan Park,[5,†] Yeongjun Jang,[1] Thomas M. Kollmeyer,[4] Gobinda Sarkar,[4] Taejeong Bae,[1] Jean J. Kim,[6] Wan Hee Yoon,[5,*] Robert B. Jenkins,[4,*] Flora M. Vaccarino,[2,3,*] and Alexej Abyzov[1,*]

Abstract

The CRISPR-Cas9 system has enabled researchers to precisely modify/edit the sequence of a genome. A typical editing experiment consists of two steps: (1) editing cultured cells; (2) cell cloning and selection of clones with and without intended edit, presumed to be isogenic. The application of CRISPR-Cas9 system may result in off-target edits, whereas cloning will reveal culture-acquired mutations. We analyzed the extent of the former and the latter by whole genome sequencing in three experiments involving separate genomic loci and conducted by three independent laboratories. In all experiments we hardly found any off-target edits, whereas detecting hundreds to thousands of single nucleotide mutations unique to each clone after relatively short culture of 10–20 passages. Notably, clones also differed in copy number alterations (CNAs) that were several kb to several mb in size and represented the largest source of genomic divergence among clones. We suggest that screening of clones for mutations and CNAs acquired in culture is a necessary step to allow correct interpretation of DNA editing experiments. Furthermore, since culture associated mutations are inevitable, we propose that experiments involving derivation of clonal lines should compare a mix of multiple unedited lines and a mix of multiple edited lines.

Introduction

The Clustered Regularly Interspaced Short Palindromic Repeats CRISPR-Cas9 system has enabled researchers to precisely modify the sequence of a genome. The idea is borrowed from the defense mechanism of bacteria fighting against a virus attack.[1,2] The system uses an RNA-protein complex consisting of two main components: a Cas9 effector protein and a single guide RNA (sgRNA) targeting a sequence of about 20 nucleotides. Once the RNA-protein complex enters the cell, the sgRNA recognizes a complementary target sequence in the nuclear DNA and Cas9 makes double-stranded break (DSB) in the DNA. The repair of the break may result in the sequence variation at the target site, such as an indel creating a loss of function (LOF) mutation in a particular locus.

If provided, a template DNA can be incorporated into the targeted site, to insert a desired variation or sequence (such a selectable marker) in a particular position of the genome. Through this process, any DNA sequence can be precisely edited with an efficiency far superior to that of other endogenous genome editing methods such as ZINC finger[3] or TALEN nucleases.[4] Currently, CRISPR-Cas9 system is used extensively to study the effect of a genetic variation within the human genome.[5] Modified CRISPR-Cas9 systems can also be applied to induce single-strand breaks, activate gene expression, and activate or repress noncoding elements.[6,7]

In a typical genome editing experiment, cultured cells are transfected with plasmids or transduced with a viral library (or libraries) encoding sgRNAs and the Cas9 nuclease. From the bulk cell culture, clonal subpopulations (e.g., clones with and without the intended edits) are isolated and expanded for further study. Of particular interest are studies that utilize human induced pluripotent

[1]Department of Quantitative Health Sciences, Center for Individualized Medicine, Mayo Clinic, Rochester, Minnesota, USA; [2]Child Study Center, Yale University, New Haven, Connecticut, USA; [3]Department of Neuroscience, Yale University, New Haven, Connecticut, USA; [4]Department of Lab Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA; [5]Aging and Metabolism Research Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, USA; and [6]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas, USA.
[†]All these authors contributed equally to this study.

*Address correspondence to: Alexej Abyzov, Department of Quantitative Health Sciences, Center for Individualized Medicine, Mayo Clinic, Rochester, MN 55905, USA, E-mail: abyzov.alexej@mayo.edu; Wan Hee Yoon, Aging and Metabolism Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA; Robert B. Jenkins, Aging and Metabolism Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA; Flora M. Vaccarino, Child Study Cemter and Department of Neuroscience, Yale University, New Haven, CT 06520, USA.

stem cell (iPSC) lines. In this case, the lines are derived from somatic cells of an individual, and the CRISPR-Cas9 system is often used to study the phenotypic effect of a particular genome edit (such as a deletion, substitution, or an indel creating a LOF mutation in a particular locus) on that individual's genetic background. Typically, cells of the edited iPSC line are cloned, and clones with or without a given edit are compared and referred to as "isogenic" except for the induced mutation. This approach is often highlighted as particularly rigorous and powerful, in that it allows comparing the effect of single mutations in an otherwise uniform genetic background.

However, the genomes of selected clones can be different. Differences can be introduced because of unintended edits of the Cas9 enzyme, so-called off-target effects.[8–10] More importantly, genomic differences can arise from the natural accumulation of mutations in cells of the parental (iPSC) line.[11,12] Although these mutations are present in a small percentage of cells in a line, they can be disregarded as a potential source of phenotypic variation. However, upon cloning a cell from the lines those mutations will be manifested in all cells of each clonal derivative, becoming a potential source of phenotypic variation between clones. Off-target effects have been studied in multiple reports,[13,14] but less emphasis has been given to understand the effect of clonal selection and manifestation of pre-existing somatic mutations in the clones.[15]

Here we studied the effect of clonal selection in CRISPR-Cas9 editing experiments by analyzing data from three independent studies that were done in three different institutions (Yale University, Mayo Clinic, and Oklahoma Medical Research Foundation [OMRF]/Baylor College of Medicine [BCM]). In each experiment, multiple clones were expanded, and single nucleotide variations (SNVs) and copy number alterations (CNAs) in each of the clones were analyzed.

Materials and Methods
Experiment performed at Yale University
Multiple iPSC lines were generated from skin fibroblasts of a 52-year-old male, S1123-01 (Fig. 1A and Supplementary Fig. S1A). iPSC line number 3 (S1123-01-i3 at passage number 11) was selected to generate a LOF mutation in the *FOXG1* gene by CRISPR-Cas9 genome editing. In brief, four different guide RNAs (gRNAs) targeting the beginning of the first and only exon of the *FOXG1* gene were tested for their abilities to mediate cleavage and indel formation, and the gRNA 5′ CAGGCTGTTGATGCTGAACG 3′ followed by the PAM seq 5′ AGG 3′ was chosen. Cas9/gRNA mediated 1-bp insertion was achieved by DSB of DNA by Cas9 at the targeted site and error-prone repair of the break by nonhomologous end joining (NHEJ).

Postediting, multiple single cells were cloned and expanded. Upon genotyping at the target site by Sanger sequencing, one control (B10_CTR) and two edited clones (A4 and A8) were picked for whole genome sequencing (WGS). Selected edited clones had a 1-bp insertion (dT) into the *FOXG1* exon (at chr14:29,236,536 GTT->GTTT in GRCh37 reference genome coordinates). The insertion, heterozygous in A4 and homozygous in A8, resulted in a frameshift and premature stop codon, giving rise to a FOXG1 truncated protein. The passage numbers were 19, 19, and 21 for B10_CTR, A4, and A8, respectively. All iPSC lines were grown in mTESR1 media (StemCell Technologies) on Matrigel-coated dishes (Corning Matrigel Matrix Basement Membrane Growth Factor Reduced) and propagate using Dispase (StemCell Technologies).

Experiment performed at the Mayo Clinic
The AA7 (parental) line was subcloned from Normal Human Astrocyte (NHA) cell line expressing the E6 and E7 subunits of the human papillomavirus (HPV).[16] Cells were grown in Dulbecco's Modified Eagle Medium, with glutamax, 10% fetal bovine serum, and 1% penicillin/streptomycin. Both haplotypes in parental line AA7 have an A allele at chr8:130,645,692 (GRCh37). The line was edited with the use of a CRISPR nickases and the experiment relied on the homology directed repair to incorporate a G allele at this position. Two individual transfections were done using two different pairs of gRNAs (NC1 and NC2) and the same donor sequence (Fig. 1A and Supplementary Fig. S1B).

The donor DNA contained the sequence of the region with altered allele and the sequence of a hygromycin resistance gene flanked by LoxP sequences for removal after editing. After transfections, cells were selected with antibiotic, subcloned and sequenced by Sanger sequencing to determine the incorporation of the G allele in the desired location. Two clones that were homozygous for the G allele at the location of interest were selected, as well as a clone that remained homozygous for A at the location of interest. All the clones that were selected were within two passages of each other at the time of editing.

Experiment performed at OMRF/BCM
We obtained skin fibroblasts carrying the heterozygous constitutional variant (NM_001170535.1; c.1582C>T: p.Arg528Trp) in the *ATPase family, AAA domain containing 3A* (*ATAD3A*) gene from a child in a family.[17] Using Sendai virus-mediated reprogramming, we derived three iPSC lines (iPSC Nos. 5, 7, and 10) from the patient
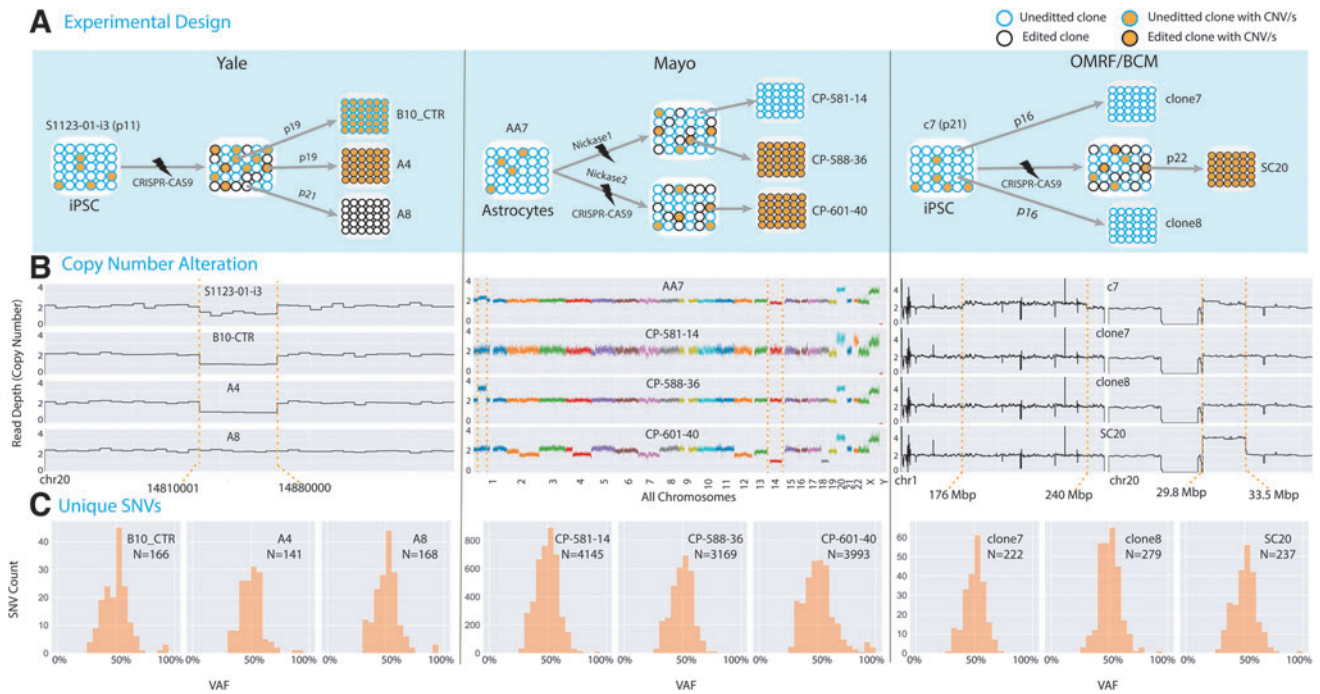
FIG. 1. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR-Cas9) edited clones are not isogenic. **(A)** In three different experiments conducted by three different groups, CRISPR-Cas9 unedited (in blue circles) and edited (in black circles) single cell clones were analyzed. Orange-filled circles represent clones with CNA. **(B)** Copy number profiles for the analyzed samples. CNA inherited from unedited parental lines are shown between orange broken lines. In the experiment conducted at Yale, the heterozygous deletion on chromosome 20 was mosaic in the parental unedited iPSC line (90% of cells) [see **(A)**], and two derived clones (unedited and edited) inherited it. Similarly, in the experiment done at Mayo, multiple mosaic CNA in the parental line (such as duplication on chromosome 1 and deletion of chromosome 14) were inherited by only one out of three clones. In addition, one edited line, CP-601-40, was genomically unstable. In the experiment by OMRF/BCM, a mosaic duplication on chromosome 20 inherited from the unedited parental iPSC line was present in only one of the clones. Clones were selected to avoid mosaic duplication on chromosome 1 present in the parental line. **(C)** Distribution of private SNVs by allele frequency in each clone (Supplementary Table S1). Clones have hundreds and thousands of point mutations that arose before cloning during culturing of parental lines (see Supplementary Fig. S5 for variant calling). OMRF, Oklahoma Medical Research Foundation; BCM, Baylor College of Medicine; CNA, copy number alteration; iPSC, induced pluripotent stem cell; SNV, single nucleotide variation.

fibroblasts (Fig. 1A and Supplementary Fig. S1C). Karyotyping G-banding showed normal genomic integrity for iPSC numbers 7 and 10, and iPSC number 7 was selected for further studies (c7 in Fig. 1). Two unedited control clones (clone7 and clone8) from the c7 line were selected and sequenced.

The iPSC number 7 was subjected to editing to correct the pathogenic variant allele (c.1582C>T) in the *ATAD3A* gene by employing CRISPR-Cas9-mediated site-specific DSB. We used nucleofection of ribonucleoprotein (RNP) method consisting of sgRNAs and SpCas9 protein[18] instead of vector-mediated Cas9 method. The variant specific 5′ GACGGAGGGCATGTCGGGCT 3′ gRNA and GGG PAM sequence were used. One clone SC20 with

corrected pathogenic allele was picked for the analysis. The passage numbers were 21, 16, 16, and 22 for c7, clone7, clone8, and SC20, respectively.

Sequencing data processing
Reads were aligned to the GRCh37 reference genome using BWA-Mem2.[19] As a result, for each sample an alignment file in BAM format was generated. GATK-HaplotypeCaller was used to call for germline mutations.

Checking for off-target edits
The Cas-OFFinder[20] was used to predict off-target sites using experiment specific gRNA and PAM sequence. We considered off-target site of up to four mismatch

bases to each gRNA-PAM sequence. For the experiment performed at Yale, we investigated 288 possible off-target sites and found no mutations around them. For experiment performed at Mayo Clinic, two pairs of nickase gRNAs, NC1.A/NC1.B and NC2.A/NC2.B, were used (Supplementary Fig. S2). Out of a total 1229 possible off-target sites for all nickases, only 2 with 4 mismatches had nearby mutations.

The chr14:51726108:A>T mutation was nearby the genomic site similar to NC1.B gRNA (AAAa-GagTTTCTTcAACAAATGG; mismatched bases are in lowercase). Similarly, mutation chr7:9805446:C>T was nearby the site similar to NC2.B gRNA (ACGGCTTCT-GAtaAAaTCtTAGG; mismatched bases are in lowercase). Given high count of possible off-target sites and higher mutation burden in the lines from that experiment (Supplementary Fig. S5), we think that likely those mutations are the results of cloning rather than editing and are coincidently next to the off-target sites. For experiment performed at OMRF/BCM, 94 possible off-target sites were predicted and none of them had mutations.

Checking for off-target insertion of template DNA

For the Mayo Clinic experiment, reads for each clone were realigned after adding a new contig to the GRCh37 reference genome. The contig contained the 150 bp of upstream homology arm, the sequence of a hygromycin resistance gene flanked by LoxP sequences, and 150 bp of downstream homology arm. We checked the reads that were aligned to the new contig and had a mate-pair read aligned to other chromosomes/contigs. As expected, most mates of the reads at the flanking region of contig mapped to chromosome 8 near the gRNA sites.

Overall, 98.6–100% (AA7: 126/127, CP-581-14: 634/638, CP-588-36: 297/297, and CP-601-40: 378/383) of reads matched to the new contig and chromosome 8. The rest of the reads (AA7: 1/127, CP-581-14: 4/638, and CP-601-40: 5/383) aligned randomly across the genome and did not form any cluster. Hence, we concluded that no significant evidence of incorrect insertion was found. In addition, the reads aligned to the flanking regions of the donor DNA template were checked and we found no evidence that the whole DNA template was inserted somewhere else in the genome.

Copy number analysis

CNVpytor was used to check for CNAs.[21] The mean-shift caller was used, and the read depth information was collected from the alignment files. Samples were processed with multiple bin sizes (10k, 100k). The commands were as follows:

```
cnvpytor -root file.pytor -rd [sample bam]
cnvpytor -root file.pytor -his 10000 100000
cnvpytor -root file.pytor -partition 10000 100000
cnvpytor -root file.pytor -call 10000 100000
cnvpytor -root file.pytor -snp file.vcf -sample sample_name
cnvpytor -root file.pytor -mask_snps
cnvpytor -root file.pytor -baf 10000 100000
```

We extracted read depth from alignment file for bins of 10 and 100 kbp. After GC correction was performed, we used the mean-shift approach to segment read depth signal. Furthermore, we used B-allele frequency (BAF) of SNVs within called region to confirm existence of CNAs. Selected regions with high confidence difference from other regions are called. For each segment two p-values were calculated reflecting significance of coverage deviation from an average and deviation of BAF from 0.5. Regions with at least one of p-values smaller than 10^{-20} and with coverage difference by at least 25% from average were analyzed.

Calling somatic point mutations

GATK-Mutect2[22] was used to call for somatic mutations. The following command was used:

```
GATK4—java-options "-Xmx12G -Djava.io.tmpdir=$OUTDIR/
tmp -XX:-UseParallelGC" Mutect2 \
-R $REF_GENOME -I $TUMOR_BAM -I $NORMAL_BAM -normal
$NORMAL_SAMPLE \
-O $OUTDIR/${NORMAL_SAMPLE}_${TUMOR_SAMPLE}.mutect
.vcf.gz

GATK4—java-options "-Xmx12G -Djava.io.tmpdir=$OUTDIR/
tmp -XX:-UseParallelGC" \
FilterMutectCalls -R $REF_GENOME \
-V $OUTDIR/${NORMAL_SAMPLE}_${TUMOR_SAMPLE}.mutect
.vcf.gz \
-O $OUTDIR/${NORMAL_SAMPLE}_${TUMOR_SAMPLE}.mutect
.filtered.vcf.gz

GATK version: 4.1.8.1; Reference Genome: GRCh37
```

In each clone, mutations were called by comparing against other clones and against bulk sample from the parental line. We then considered the overlapping set of calls from the three comparisons as culture-derived somatic mutations in the parental line transmitted to each of the clones (Supplementary Fig. S5). Such mutations must have variant allele frequency (VAF) distribution looking similar to a symmetrical bell-shaped curve centered around 50%, that is, almost all mutations are on one haplotype of autosomes. Indeed, we observed such a shape (Fig. 1C). To eliminate mutation that may have arisen during culturing of clones, we selected only mutations with at least 30% VAF.

Results

Experimental designs

Three independent experiments were conducted in three different institutions (Fig. 1A and Supplementary Fig. S1). In the experiment performed at Yale University, an iPSC line (i3) generated from skin fibroblasts of an adult male (S1123-01) was subjected to editing. A locus in the *FOXG1* gene was targeted to produce a truncated, LOF allele of the *FOXG1* gene. A double-stand DNA break in the *FOXG1* gene was generated by Cas9 using a gRNA binding to the sequence corresponding to N-terminal of the produced protein. When repaired by NHEJ, an error prone DSB repair pathway, the sequence at the break may contain small deletion or insertions (indels). Postediting, multiple clones were derived, expanded, and upon checking the target site with Sanger sequencing, one control (B10_CTR) and two clones with 1-bp insertion (A4 and A8) were picked for WGS.

The experiment performed at the Mayo Clinic used an NHA cell line (AA7) expressing the E6 and E7 subunits of the HPV. The line was edited to replace a fragment of DNA on chromosome 8 with a designed template sequence consisting of the same DNA fragment introducing SNP rs55705857 and a hygromycin resistance gene (for selection of cells with successful editing) (Fig. 1A and Supplementary Fig. S1B). Multiple clones were expanded. Upon checking the target allele with Sanger sequencing, one control (CP-581-14) and two edited clones (CP-588-36 and CP-601-40) were picked for WGS.

The OMRF/BCM experiment used an iPSC line derived from fibroblasts of a 9-year-old female patient (line c7) carrying a constitutional heterozygous variant (chr1:1,464,679 C>T; GRCh37) in exon 15 of the *ATAD3A* gene. The parental line was edited by introducing a DSB at the variant's allele to correct the variant by homology directed repair (Fig. 1A and Supplementary Fig. S1C). Two unedited control clones (clone7 and clone8) and one edited clone (SC20) were selected for WGS. In every experiment, we also performed WGS of the primary unedited lines.

Off-target effects

In each experiment, we determined off-target sites across the genome of edited lines considering all genomic sites that had up to four mismatched bases with respect to gRNAs. Only in the experiment performed by the Mayo Clinic, we identified two possible off-target mutations in proximity to those off-target sites in picked clones (see Checking for Off-Target Edits section). Particularly, in the experiment by OMRF/BCM, there were two genomic sites homologous to the targeted site, represented by the paralog genes *ATAD3B* and *ATAD3C*; but

no off-target editing was detected. In addition, in the experiment by Mayo Clinic, we searched for reads supporting an insertion of the designed sequence in off-target sites and found no significant support for such an event(s) in any of the analyzed clones. We, therefore, concluded that likely no off-target editing occurred in any of the experiments.

Copy number alterations

In the experiment carried out at Yale, we discovered only one CNA discordant between clones (Fig. 1B). That was a 70 kbp heterozygous deletion in clones B10_CTR and A4 and a normal copy number region with two different haplotypes in clone A8 (Supplementary Fig. S3). Analysis of the parental iPSC line (S1123-01-i3) revealed that the heterozygous deletion was already present in it at ∼90% cell frequency. From these observations we inferred that the CNAs was mosaic in the parental unedited line and that clones B10_CTR and A4 arose from cells with the deletion, whereas clone A8 from a cell without the deletion. This example demonstrates that, independently of the Cas9-induced edits, clonal lines are not necessarily isogenic due to genomic heterogeneity of cells in the parental line.

The same conclusion can be drawn from the other two experiments. In the Mayo experiment, a mosaic duplication on chromosome 1 was present in the parental line AA7 in about 10% of cells, was inherited by one of the edited clones (CP-588-36) but not by other clones. Similarly, a mosaic deletion of chromosome 14 in the parental line (AA7) was only transmitted to one of the edited clones CP-601-40 (Fig. 1B and Supplementary Fig. S4). These CNAs were totally independent from the intended Cas9-induced edits. All samples from this experiment have duplication on chromosome 20 and q arm of chromosome X, as they were constitutional to the parental line.

In the OMRF/BCM experiment, clones were selected to avoid a mosaic duplication on chromosome 1 in the parental line. But similar to the aforementioned experiments, a mosaic duplication on chromosome 20 present in the unedited line was transmitted only to clone SC20. Again, this was independent of the intended Cas9-induced edit in *ATAD3A* gene.

Point mutations in individual clones

Genomic heterogeneity of parental cell line could also manifest itself in differences between clones in other variant types, so we investigated unique SNVs in each clone. To discover unique SNVs we compared clones with each other and with the parental line. In each clone SNVs inherited from the founder cell must typically have a VAF of 50%, as the majority are present on one

haplotype of autosomes. Indeed, the VAF distributions of unique SNVs in each clone were bell-shaped and centered at 50% (Fig. 1C), indicating that most of them represent mutations from the founder cell of the clone. Applying a cutoff of 30% VAF to select SNVs likely inherited from founder cell (see Materials and Methods section), we detected hundreds to thousands of SNVs in each clone (Fig. 1C and Supplementary Fig. S5 and Supplementary Table S1).

Compared with iPSC clones from both Yale and OMRF/BCM cohort, the astrocyte lines (samples from Mayo) had a 10-fold higher number of SNVs, likely reflecting that the parental astrocyte line is genomically unstable. This was also reflected by the fact that clone CP-601-40 had a high fraction ($\sim$16.5%) of its genome affected by CNAs, which was outstanding even when compared with other clones of the same line. More SNVs in iPSC lines of the OMRF/BCM experiment as compared with iPSC lines of the Yale experiment may reflect earlier divergence of clones in the former case, because control clones were selected before editing and associated culture.

Discussion/Conclusions

Our study demonstrates that culture clones selected after DNA editing experiments are not isogenic and, consequently, may diverge phenotypically, independently from the intended edits. Such differences among clones arise as a consequence of the edited lines being mosaic, that is, harboring genomic variants accumulated *in vitro*. Mosaicism in iPSC lines naturally occurs as a result of culture and could potentially be influenced by the genetic background or culture conditions favoring genomic instability.[23] It was estimated that during culture, individual iPSC, intestinal, and liver stem cell lines accumulate, respectively, 3.5 ± 0.5, 7.2 ± 1.1, and 8.3 ± 3.6 base substitutions per population doubling.[24]

By the time a typical experiment selects a line to be used in a CRISPR-Cas study, hundreds of point mutations are present in each cell of the line, which will be manifested, upon cloning the line, as genomic difference between clones. Most notably, clones are different in CNAs that were several kb to several mb in size. This represents, by far, the largest source of genomic divergence among clones compared with SNVs variations, as exemplified by the duplications on chromosome 1 in the OMRF/BCM and Mayo experiments. In fact, such CNAs could be comparable or even larger than the total number of genomic bases different between two random individuals or lines derived from them. Thus, clonal selection during DNA editing experiment is a large and underappreciated source of genomic variability.

We suggest that for a meaningful comparison between edited lines, each derived clone should be subjected to WGS to screen for unintended unique genomic variations that can have potential functional consequences. At a minimum, each clone should be screened for large CNAs to ascertain the largest source of genomic divergence between clonal lines. The functional consequences of variants outside coding regions are particularly hard to predict and could confound the interpretation of genome editing experiments. To mitigate this issue, we also propose that genome editing experiments that involve derivation of clones compare a mix of multiple unedited clones with that of edited clones.

Ethics Statement

Data Availability

Authors' Contributions

Author Disclosure Statement

Funding Information

Supplementary Material

Supplementary Table S1
Supplementary Figure S1
Supplementary Figure S2
Supplementary Figure S3
Supplementary Figure S4
Supplementary Figure S5

References

1. Jinek M, Chylinski K, Fonfara I, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 2012;337(6096):816–821; doi: 10.1126/science.1225829
2. Garneau JE, Dupuis ME, Villion M, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 2010;468(7320):67–71; doi: 10.1038/nature09523
3. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A 1996;93(3):1156–1160; doi: 10.1073/pnas.93.3.1156
4. Gaj T, Gersbach CA, Barbas CF, 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol 2013;31(7):397–405; doi: 10.1016/j.tibtech.2013.04.004
5. Ran FA, Hsu PD, Wright J, et al. Genome engineering using the CRISPR-Cas9 system. Nat Protoc 2013;8(11):2281–2308; doi: 10.1038/nprot.2013.143
6. Ma E, Harrington LB, O'Connell MR, et al. Single-stranded DNA cleavage by divergent CRISPR-Cas9 enzymes. Mol Cell 2015;60(3):398–407; doi: 10.1016/j.molcel.2015.10.030
7. Hanna RE, Doench JG. Design and analysis of CRISPR-Cas experiments. Nat Biotechnol 2020;38(7):813–823; doi: 10.1038/s41587-020-0490-7
8. Pattanayak V, Lin S, Guilinger JP, et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat Biotechnol 2013;31(9):839–843; doi: 10.1038/nbt.2673
9. Fu Y, Foden JA, Khayter C, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 2013;31(9):822–826; doi: 10.1038/nbt.2623
10. Boyle EA, Andreasson JOL, Chircus LM, et al. High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding. Proc Natl Acad Sci U S A 2017;114(21):5461–5466; doi: 10.1073/pnas.1700557114
11. Kwon EM, Connelly JP, Hansen NF, et al. iPSCs and fibroblast subclones from the same fibroblast population contain comparable levels of sequence variations. Proc Natl Acad Sci U S A 2017;114(8):1964–1969; doi: 10.1073/pnas.1616035114
12. Laurent LC, Ulitsky I, Slavin I, et al. Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. Cell Stem Cell 2011;8(1):106–118; doi: 10.1016/j.stem.2010.12.003
13. Kim D, Bae S, Park J, et al. Digenome-seq: Genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. Nat Methods 2015;12(3):237–243, 1 p following 243; doi: 10.1038/nmeth.3284
14. Tsai SQ, Nguyen NT, Malagon-Lopez J, et al. CIRCLE-seq: A highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. Nat Methods 2017;14(6):607–614; doi: 10.1038/nmeth.4278
15. Assou S, Bouckenheimer J, De Vos J. Concise review: Assessing the genome integrity of human induced pluripotent stem cells: What quality control metrics? Stem Cells 2018;36(6):814–821; doi: 10.1002/stem.2797
16. Ohba S, Mukherjee J, See WL, et al. Mutant IDH1-driven cellular transformation increases RAD51-mediated homologous recombination and temozolomide resistance. Cancer Res 2014;74(17):4836–4844; doi: 10.1158/0008-5472.CAN-14-0924
17. Harel T, Yoon WH, Garone C, et al. Recurrent de novo and biallelic variation of ATAD3A, encoding a mitochondrial membrane protein, results in distinct neurological syndromes. Am J Hum Genet 2016;99(4):831–845; doi: 10.1016/j.ajhg.2016.08.007
18. Liang X, Potter J, Kumar S, et al. Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. J Biotechnol 2015;208:44–53; doi: 10.1016/j.jbiotec.2015.04.024
19. Md V, Misra S, Li H, et al. Efficient architecture-aware acceleration of BWA-MEM for multicore systems. 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS). IEEE Conf Proc 2019;2019:314–324; doi: 10.1109/ipdps.2019.00041
20. Bae S, Park J, Kim J-S. Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 2014;30(10):1473–1475; doi: 10.1093/bioinformatics/btu048
21. Suvakov M, Panda A, Diesh C, et al. CNVpytor: A tool for copy number variation detection and analysis from read depth and allele imbalance in whole-genome sequencing. Gigascience 2021;10(11); doi: 10.1093/gigascience/giab074
22. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 2013;31(3):213–219; doi: 10.1038/nbt.2514
23. International Stem Cell Initiative; Amps K, Andrews PW, et al. Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. Nat Biotechnol 2011;29(12):1132–1144; doi: 10.1038/nbt.2051
24. Thompson O, von Meyenn F, Hewitt Z, et al. Low rates of mutation in clinical grade human pluripotent stem cells under different culture conditions. Nat Commun 2020;11(1):1528; doi: 10.1038/s41467-020-15271-3