

CLAIMS

I/We claim:

1. A method for monitoring neoplastic potential of a cell population following an engineered genomic editing event, comprising:

- (a) duplex sequencing one or more target double-stranded DNA molecules originating from the cell population following the engineered genomic editing event at a first time point;
- (b) identifying one or more variants among the target double-stranded DNA molecules;
- (c) determining a first variant frequency of the one or more variants among the target double-stranded DNA molecules originating from the cell population at the first time point;
- (d) duplex sequencing one or more target double-stranded DNA molecules originating from the cell population following the engineered genomic editing event at a second time point following the first time point;
- (e) determining a second variant frequency of the same one or more variants among the target double-stranded DNA molecules originating from the cell population at the second time point; and
- (f) comparing the first variant frequency for the one or more variants to the second variant frequency for the one or more variants.

2. The method of claim 1, wherein the step of duplex sequencing comprises:

- (i) preparing a sequencing library from double-stranded DNA molecules originating from the cell population, wherein preparing the sequence library comprises ligating adapter molecules to the plurality of double-stranded DNA fragments to generate a plurality of adapter-DNA molecules;

for at least some of the plurality of adapter-DNA molecules:

- (ii) sequencing first and second strands of the adapter-DNA molecules to provide a first strand sequence read and a second strand sequence read for the adapter-DNA molecules; and
- (iii) for the adapter-DNA molecules, comparing the first strand sequence read and the second strand sequence read to identify one or more correspondences between the first and second strand sequences reads.

3. The method of claim 1, wherein the cell population comprises genome-edited immune cells.

4. The method of claim 1, wherein the cell population is to be administered to a subject between the first and second time points.

5. The method of claim 1, wherein the engineered genomic editing event is a target endonuclease-mediated editing event; is a Cas-mediated editing event; is a Cas9-mediated editing event; is a CPF1-mediated editing event; is a modified CAS or CPF-1 mediated event; is carried out by a TALON, MEGATAL, Zinc-fingered nuclease, a homing endonuclease, or a restriction endonuclease; is a polynucleotide substrate-mediated homologous recombination event; is carried out by a retrovirus or another virus; or any combination of one or more of the foregoing.

6. The method of claim 1, wherein the second variant frequency for the one or more variants is a safe threshold variant frequency.

7. The method of claim 1, wherein following step (f), the method further comprises identifying a variant from the one or more variants identified in step (b) as a variant with a positive growth bias.

8. The method of claim 1, wherein the cell population is a cell-based therapy.

9. The method of claim 1, wherein the one or more variants identified in step (b) comprise a cell-cycle mutation.

10. The method of claim 1, wherein the one or more variants identified in step (b) are detected in one or more cancer driver.

11. The method of claim 10, wherein the one or more cancer drivers comprises ABL, ACC, BCR, BLCA, BRCA, CESC, CHOL, COAD, DLBC, DNMT3A, EGFR, ESCA, GBM, HNSC, RICH, KIRC, KIRP, LAML, LGG, LIHC, LUAD, LUSC, MESO, OV, PAAD, PCPG, PI3K, PIK3CA, BRAD, PTEN, READ, SARC, SKCM, STAD, TGCT, THCA, THUM, UCEC, UCS, and/or UVM.

12. The method of claim 10, wherein the cancer driver comprises *TP53*.

13. The method of claim 10, wherein the cancer driver comprises *HRAS*, *NRAS* or *KRAS*.

14. The method of claim 1, wherein the one or more variants are in a sequence of a tumor suppressor gene, an oncogene, and/or a proto-oncogene.

15. The method of claim 1, wherein the cell population is determined to have neoplastic potential if the second variant frequency is greater than the first variant frequency.

16. The method of claim 1, wherein the first and/or second variant frequency of the one or more variants is above a threshold variant frequency.

17. The method of claims 1, wherein the one or more variants are identified relative to a sequence from a reference cell population.

18. The method of claim 1, wherein both the first and second time points are within about 30 days, within about 45 days, within about 60 days, within about 75 days, or within about 90 days following the engineered genomic editing event.

19. The method of claim 1, wherein the second time point is 2 to about 90 days after the first time point.

20. The method of claim 1, wherein a clonal expansion of a cell in the cell population is determined to have occurred if the second variant frequency is greater than the first variant frequency.

21. The method of claim 1, further comprising comparing the first variant frequency and/or the second variant frequency for the one or more variants to an expected variant frequency.

22. The method of claim 21, wherein the expected variant frequency is determined from a comparable cell population that has not undergone the engineered genomic editing event.

23. The method of claim 21, wherein the cell population is determined to have neoplastic potential if the first and/or second variant frequency is greater than the expected variant frequency.

24. The method of claim 1, wherein the one or more variants are at one or more locations outside an intended locus for genomic editing.

25. The method of claim 1, wherein the variant frequency of the one or more variants is determined by calculating a number of duplex sequenced target double-stranded DNA molecules having a specified variant mapping to a genomic locus of interest per a total number of duplex sequenced target double-stranded DNA molecules mapping to the genomic locus of interest.

26. The method of claim 2, wherein prior to sequencing, the method further comprises selectively enriching the first and second strands for a subset of adapter-DNA molecules that map to one or more genomic loci of interest to provide an enriched sequencing library.