

What is claimed is:

1. A method of analyzing expression of a protein-coding region of DNA in a cell, the method comprising:
 - replacing a protein-coding region of the DNA in the cell with a donor vector comprising a variant of the protein-coding region and a first barcode identifying that variant, wherein the cell generates mRNA comprising an expression of the variant and an expression of the first barcode;
 - coupling, to the mRNA, a second barcode corresponding to the cell;
 - reverse transcribing the mRNA, having the second barcode coupled thereto, into cDNA;
 - sequencing the cDNA;
 - sequencing the donor vector or cDNA using amplicon sequencing; and
 - correlating the donor vector sequence and the cDNA sequence to identify the variant and the cell's expression of the variant.
2. The method of claim 1, wherein the donor vector comprises
 - (a) a promoter region,
 - (b) right and left homology arms, the variant and the first barcode being between the right and left homology arms, or
 - (c) both (a) and (b).
3. The method of claim 2, wherein the barcode is located between the promoter region and the variant.
4. The method of claim 2, wherein the promoter region comprises a reverse promoter region.
5. The method of claim 4, wherein
 - (a) the reverse promoter region is disposed between the first barcode and the variant,
 - (b) the expression of the variant of the protein-coding region is in the forward direction, and wherein the expression of the first barcode is in the reverse direction, or
 - (c) both (a) and (b).

6. The method of claim 2, further comprising:
 - using a first polymerase chain reaction (PCR) process to generate a first amplicon of the donor sequence that includes the variant, the first barcode, and the right homology arm and substantially excludes the left homology arm; and
 - using a second PCR process to generate a second amplicon of the first amplicon that includes the variant and the first barcode and substantially excludes the right and left homology arms.
7. The method of claim 6, wherein sequencing the donor vector comprises sequencing the second amplicon.
8. The method of claim 6, wherein the second amplicon has a length of about 1000 bases or fewer.
9. The method of claim 1, wherein the mRNA comprises:
 - a first mRNA molecule comprising the expression of the variant, and
 - a second mRNA molecule comprising the expression of the first barcode.
10. The method of claim 9, wherein coupling the second barcode to the mRNA comprises:
 - coupling a first molecule of the second barcode to the first mRNA molecule; and
 - coupling a second molecule of the second barcode to the second mRNA molecule.
11. The method of claim 10, wherein the cDNA comprises a first cDNA molecule comprising a reverse transcription of the variant and the second barcode, and a second cDNA molecule comprising a reverse transcription of the protein coding region and the second barcode, and sequencing the cDNA comprises sequencing the first and second cDNA molecules.
12. The method of claim 1, wherein replacing the initial protein-coding region comprises:
 - using a CRISPR-associated protein guide RNA ribonucleoprotein (Cas-gRNA RNP) to cut the DNA in the cell; and
 - using homology-directed repair (HDR) to repair the cut in the DNA using the donor vector.

13. The method of claim 12, further comprising inserting first and second plasmids into the cell,
wherein the donor vector is located on the first plasmid; and
wherein the cell expresses the Cas-gRNA RNP using the second plasmid.
14. The method of claim 1, wherein the donor vector comprises a lentiviral vector.
15. The method of claim 1, wherein the donor vector further comprises a puromycin resistance gene, the method further comprising contacting the cell with puromycin to enrich for the cell.
16. The method of claim 15, wherein the first barcode is located on a UTR terminus of the puromycin resistance gene.
17. The method of claim 1, further comprising cleaving the first barcode from the variant in the cell.
18. A method of analyzing expression of a protein-coding region of DNA in a collection of cells, the method comprising:
replacing the initial protein coding-region of the DNA in each of the cells with a donor vector comprising a variant of the protein-coding region and a first barcode identifying that variant, wherein the cells receive different variants than one another;
obtaining mRNA from the cells, the mRNA from each cell comprising an expression of the variant of the protein-coding region in that cell and an expression of the first barcode;
coupling, to the mRNA from each cell, a second barcode corresponding to that cell;
reverse transcribing the mRNA, having the second barcode coupled thereto, into cDNA;
sequencing the cDNA;
sequencing the donor vector; and
correlating the donor vector sequence and the cDNA sequence to identify the variant in each of the cells and that cell's expression of that variant.

19. The method of claim 18, wherein the different variants are saturationally mutagenized.
20. A collection of polynucleotides from a collection of cells, the polynucleotides comprising first and second mRNA molecules from each of the cells, wherein, for each cell:
the first mRNA molecule comprises a first molecule of a barcode corresponding to that cell and an expression of a variant in that cell, and
the second mRNA molecule comprises the barcode corresponding to that cell and an expression of a first barcode corresponding to the variant.
21. The collection of polynucleotides of claim 20, wherein the different variants are saturationally mutagenized.
22. A method, comprising:
providing a barcoded homology donor vector comprising a semi-random barcode on termini of a foreign transcript, the donor vector including homology arms and mutations;
knocking-in the barcoded homology donor vector to the vicinity of an exon to be edited to create a variant on the exon; and
cleaving the variant using a CRISPR-associated protein guide RNA ribonucleoprotein (Cas-gRNA RNP).
23. The method of claim 22, wherein the barcode is placed on UTR termini of the donor vector so that it may be expressed and detectable in scRNA-seq.
24. The method of claim 22 or claim 23, wherein the donor vector comprises a puromycin resistance gene.
25. The method of claim 22, wherein providing the barcoded homology donor vector comprises:
using a first polymerase chain reaction (PCR) to specifically amplify the knocked-in region with a genomically edited allele;
using a second PCR, using the product of the first PCR as a template, to link the barcode with variants in an amplicon; and
performing amplicon sequencing using the product from the second PCR.

26. The method of claim 25, wherein the amplicon sequencing covers both the barcode and the variants.
27. A method, comprising:
adding semi-random variant barcodes to UTR regions of a saturationally mutagenized variant library;
coupling cell barcodes to the variant barcodes;
reading the variant barcodes out in scRNA-seq; and
linking the variant barcodes to the variants of the library using a separate sequencing operation.
28. The method of claim 27, wherein the semi-random variant barcode may be placed downstream of promoters or upstream of terminators of the variant library.
29. The method of claim 27 or claim 28, wherein linking the variant barcodes to the variants of the library comprises generating tiled polymerase chain reaction (PCR) amplicons by using one set of primers to amplify the barcode on one side, and another set of primers to amplify the variants on the other side, such that each amplicon links a respective segment of the variant to the barcode.