CLAIMS

I/We claim:

1. A method for detecting and/or quantifying a donor source of nucleic acid from a mixture, comprising:

providing the mixture comprising target double-stranded DNA molecules from one or more donor sources, wherein the target double-stranded DNA molecules contain one or more genetic polymorphisms;

generating an error-corrected sequence read for each of a plurality of the target double-stranded DNA molecules in the mixture, comprising:

ligating adapter molecules to the plurality of target double-stranded DNA molecules to generate a plurality of adapter-DNA molecules;

generating a set of copies of an original first strand of the adapter-DNA molecule and a set of copies of an original second strand of the adapter-DNA molecule;

sequencing one or more copies of the original first and second strands to provide a first strand sequence and a second strand sequence; and

comparing the first strand sequence and the second strand sequence to identify one or more correspondences between the first and second strand sequences; and

identifying a donor source of nucleic acid present in the mixture of nucleic acid by deconvolving the error-corrected sequence reads into individual genotypes.

2. A method for detecting and/or quantifying a donor source of nucleic acid from a mixture, comprising:

generating error-corrected sequence reads from raw sequencing data, wherein:

the raw sequencing data is generated from a mixture comprising target double-stranded DNA molecules from one or more donor sources, and wherein the target double-stranded DNA molecules contain one or more genetic polymorphisms; and

generating the error-corrected sequence reads comprises:

comparing a first strand sequence read from the raw sequencing data and a second strand sequence read from the raw sequencing data, wherein the first strand sequence read and the second strand sequence read are derived from a same target double-stranded DNA molecule, and

identifying one or more correspondences between the first and second strand sequence reads; and

identifying a donor source of nucleic acid present in the mixture of nucleic acid by deconvolving the error-corrected sequence reads into individual genotypes.

3. The method of claim 1 or claim 2, wherein one or more of the donor sources have known genotypes.

4. The method of claim 1 or claim 2, wherein one or more of the donor sources have unknown genotypes.

5. The method of claim 1 or claim 2, wherein the mixture comprises one or more unknown individual genotypes, and wherein deconvolving the error-corrected sequence reads into individual genotypes comprises:

identifying microhaplotype allele combinations present within individual target double-stranded DNA molecules that map to one or more genetic loci in a reference sequence;

evaluating all possible mixing proportions against all possible genotypes present at each genetic locus within the one or more genetic loci; and

determining a list of all possible individual genotypes that adequately fit the identified microhaplotype allele combinations and all possible mixing proportions evaluated.

6. The method of claim 1 or claim 2, wherein the mixture comprises one or more known individual genotypes, and wherein deconvolving the error-corrected sequence reads into individual genotypes comprises:

identifying microhaplotype allele combinations present within individual target double-stranded DNA molecules in the mixture;

summing total counts of each allele donated from each known individual genotype; and

determining a mixing proportion of each known genotype present in the mixture.

7. The method of any of claims 1-6, further comprising comparing one or more individual genotypes to a database comprising a plurality of known genotypes to identify the one or more donor sources.

8. The method of any one of claims 1-7, wherein the mixture comprises more than one donor source, and wherein the method further comprises determining the proportion of each donor source from the more than one donor sources present in the mixture by calculating the proportion of each genetic polymorphism or the proportion of a substantially unique combination of genetic polymorphisms present in the error-corrected sequence reads.

9. The method of any one of claims 1-3 and 6-8, wherein the target double-stranded DNA molecules were extracted from one or more cord blood samples.

10. The method of any one of claims 1-8, wherein the target double-stranded DNA molecules were extracted from a forensic sample.

11. The method of any one of claims 1-3 and 6-8, wherein the target double-stranded DNA molecules were extracted from a patient with a stem cell or organ transplant.

12. The method of any one of claims 1-8, wherein the target double-stranded DNA molecules were extracted from a patient, and wherein identifying the one or more donor sources present in the mixture includes measuring a level of microchimerism in the patient.

13. The method of any one of claims 1-8, wherein the target double-stranded DNA molecules were extracted from a tumor sample.

14. The method of any one of claims 1-13, further comprising quantifying a relative abundance of each individual genotype present in the mixture.

15. The method of any one of claims 1-14, wherein the one or more genetic polymorphisms comprise a microhaplotype.

16. The method of any one of claims 1 and 3-15, wherein generating an error-corrected sequence read for each of a plurality of the target double-stranded DNA molecules in the mixture further comprises selectively enriching one or more targeted genomic regions prior to sequencing.

17. The method of claim 16, wherein the one or more targeted genomic regions comprises a microhaplotype site in the genome.

18. The method of any one of claims 2-15, wherein the target double-stranded DNA molecules in the mixture are selectively enriched for one or more targeted genomic regions prior to generating raw sequencing data.

19. The method of claim 18, wherein the one or more targeted genomic regions comprises a microhaplotype site in the genome.

20. A non-transitory computer-readable storage medium comprising instructions that, when executed by one or more processors, performs a method of any one of claims 2-15 and 18-19.

21. The non-transitory computer-readable storage medium of claim 20, further comprising instructions for computing a mixing proportion of each identified donor source.

22. A computer system for performing a method of any one of claims 2-15 and 18-19 for detecting and/or quantifying a donor source of nucleic acid from a mixture, the system comprising: at least one computer with a processor, memory, database, and a non-transitory computer readable storage medium comprising instructions for the processor(s), wherein said processor(s) are configured to execute said instructions to perform operations comprising the methods of any one of claims 2-15 and 18-19.