Claims

What is claimed is:

Claim 1. A method for constructing and using a chimeric reference standard specific to an individual subject for improving accuracy of analysis of a plurality of portions of a genome of the subject, the method comprising:

comparing a first portion of the genome of the subject to each reference standard of a plurality of reference standards stored in a database, wherein the genome of the subject is based on a biological sample obtained from the subject;

selecting a portion of a first reference standard that satisfies a predetermined matching threshold to a first portion of the genome of the subject based on the comparison of the first portion to the plurality of reference standards;

comparing a second portion of the genome of the subject to each reference standard of the plurality of reference standards;

selecting a portion of a second reference standard that satisfies a predetermined matching threshold to the second portion of the genome of the subject based on the comparison of the second portion to the plurality of references standards;

creating a chimeric reference standard, whereby the creating comprises inserting the selected portions of the first reference standard and the second reference standard into respective portions of a hash table maintained in one or more memory devices; and

using the hash table and the genome of the subject in order to perform one or more of a mapping, aligning, and/or variant calling procedure.

Claim 2. The method of claim 1, wherein the first reference standard is a linear reference standard, and wherein the second reference standard is a cultural reference standard.

Claim 3. The method of claim 1 or 2, wherein the first reference standard is a linear reference standard, and wherein the second reference standard is a different linear reference standard.

Claim 4. The method of any one of claims 1 through 3, wherein the first reference standard is cultural reference standard, and wherein the second reference standard is a different cultural reference standard.

Claim 5. The method of any one of claims 1 through 4, wherein at least one of the first reference standard and the second reference standard is a previously generated chimeric reference standard.

Claim 6. A non-transitory computer-readable medium storing software comprising instructions executable by one or more computers which, upon such execution, cause the one or more computers to perform operations for constructing and using a chimeric reference standard specific to an individual subject for improving accuracy of analysis of a plurality of portions of a genome of the subject, the operations comprising:

comparing a first portion of the genome of the subject to each reference standard of a plurality of reference standards stored in a database, wherein said genome is based on a biological sample obtained from said subject;

selecting a portion of a first reference standard that satisfies a predetermined matching threshold to a first portion of the genome of the subject based on the comparison of the first portion to the plurality of reference standards;

comparing a second portion of the genome of the subject to each reference standard of the plurality of reference standards;

selecting a portion of a second reference standard that satisfies a predetermined matching threshold to the second portion of the genome of the subject based on the comparison of the second portion to the plurality of references standards;

creating a chimeric reference standard, whereby the creating comprises inserting the selected portions of the first reference standard and the second reference standard into respective portions of a hash table maintained in one or more memory devices; and

using the hash table and the genome of the subject in order to perform one or more of a mapping, aligning, and/or variant calling procedure.

Claim 7. The computer-readable medium of claim 6, wherein the first reference standard is a linear reference standard, and wherein the second reference standard is a cultural reference standard.

Claim 8. The computer-readable medium of claim 6 or 7, wherein the first reference standard is a linear reference standard, and wherein the second reference standard is a different linear reference standard.

Claim 9. The computer-readable medium of any one of claim 6 through 8, wherein the first reference standard is cultural reference standard, and wherein the second reference standard is a different cultural reference standard.

Claim 10. The computer-readable medium of any one of claims 6 through 9, wherein at least one of the first reference standard and the second reference standard is a previously generated chimeric reference standard.

Claim 11. A system for constructing and using a chimeric reference standard specific to an individual for improving accuracy of nucleic acid sequence analysis of nucleic acid sequences from the subject, the system comprising:

one or more

computers and one or more storage devices storing instructions that are operable, when executed by the one or more computers, to cause the one or more computers to perform operations, the operations comprising:

comparing a first portion of a genome of the subject to each reference standard of a plurality of reference standards stored in a database, wherein the genome of the subject is based on a biological sample obtained from the subject;

selecting a portion of a first reference standard that satisfies a predetermined matching threshold to a first portion of the genome of the subject based on the comparison of the first portion to the plurality of reference standards;

comparing a second portion of the genome of the subject to each reference standard of the plurality of reference standards;

selecting a portion of a second reference standard that satisfies a predetermined matching threshold to the second portion of the genome of the subject based on the comparison of the second portion to the plurality of references standards;

creating a chimeric reference standard, whereby the creating comprises inserting the selected portions of the first reference standard and the second reference standard into respective portions of a hash table maintained in one or more memory devices; and

using the hash table and the genome of the subject in order to perform one or more of a mapping, aligning, and/or variant calling procedure.

Claim 12. The system of claim 11, wherein the first reference standard is a linear reference standard, and wherein the second reference standard is a cultural reference standard.

Claim 13. The system of claim 11 or 12, wherein the second reference standard is a different reference standard.

Claim 14. The system of any one of claims 11 through 13, wherein the first reference standard is cultural reference standard, and wherein the second reference standard is a

different cultural reference standard.

Claim 15. The system of any one of claims 11 throgh 14, wherein the first first reference standard and the second reference standard is a previously generated chimeric reference standard.