From: Young, Daniel W.
To: "Greg Porreca"
Cc: Waller, Patrick
Subject: RE: new app
Date: Thursday, May 14, 2009 10:29:52 PM
Attachments: G0814.70002US00 - PROVISIONAL PT APP.DOC
G0814.70002US00 - FIGURES.DOC

Thanks, Greg.

Attached is a draft of Uri's application. We are still thinking about the term "source sequence", and may change it. But please review this draft and let us know your comments. I have included some specific questions in bold.

Thanks
Dan

-----Original Message-----
From: Greg Porreca [mailto:gporreca@gsgenetics.com]
Sent: Thursday, May 14, 2009 5:13 PM
To: Young, Daniel W.
Cc: Waller, Patrick
Subject: Re: new app

Bob's address:

Robert Wassman
90 West Shore Drive
Marblehead, MA 01945

On Thu, May 14, 2009 at 4:27 PM, Daniel Young <Daniel.Young@wolfgreenfield.com> wrote:

> Hi Greg,
>
> We will call you in ten minutes to discuss this new application.
>
> Thanks
> Dan

From: Greg Porreca [mailto:gporreca@gsgenetics.com]
Sent: Thursday, May 14, 2009 1:30 PM
To: Waller, Patrick; Young, Daniel W.
Subject: new app

Gregory Porreca, Ph.D.
Director of Technology
Good Start Genetics



8 St Mary's Street, Suite 621
Boston, MA 02215
P: 617-358-6004
F: 617-358-5677

Petitioner Illumina, Inc., Ex. 1061, p. 2
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

TITLE OF THE INVENTION

METHODS FOR REDUCING BIAS IN PREPARATIONS OF NUCLEIC ACIDS

Field of the Invention

The invention relates to preparative steps in DNA sequencing-related technologies that reduce bias and increase the reliability and accuracy of downstream quantitative applications.

Background of the Invention

There are currently many genomics assays that utilize next-generation (polony-based) sequencing to generate data, including genome resequencing, RNA-seq for gene expression, bisulphite sequencing for methylation, and Immune-seq, among others. In order to make quantitative measurements (including genotype calling), these methods utilize the counts of sequencing reads of a given genomic locus as a proxy for the representation of that sequence in the original sample of nucleic acids. The majority of these techniques require a preparative step to construct a high-complexity library of DNA molecules that is representative of a sample of interest. This may include chemical or biochemical treatment of the DNA (e.g., bisulphite treatment), capture of a specific subset of the genome (e.g., padlock probe capture, solution hybridization), and a variety of amplification techniques (e.g., polymerase chain reaction, whole genome amplification, rolling circle amplification).

Systematic errors are common problems associated with genome amplification and sequencing library construction techniques. For example, genomic sequencing library may contain an over- or under-representation of particular sequences from a source genome as a result of systematic errors (bias) in the library construction process. Such bias can be a particularly problematic when it results in target sequences from a genome being absent or undetectable in the sequencing libraries. For example, an under-representation of particular allelic sequences (e.g., heterozygotic alleles) from a genome in a sequencing library can result in an apparent homozygous representation in a sequencing library. As most downstream sequencing library quantification techniques depend on stochastic counting processes, these problems have typically been addressed by sampling enough (over-sampling) to obtain a minimum number of observations necessary to make statistically significant decisions. However, the strategy of oversampling is generally limited to

Petitioner Illumina, Inc., Ex. 1061, p. 3
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

elimination of low-count Poisson noise, and the approach wastes resources and increases the expense required to perform such experiments. Moreover, oversampling can result in a reduced statistical confidence in certain conclusions (e.g., diagnostic calls) based on the data. Accordingly, new approaches are needed for overcoming bias in sequencing library preparatory methods.

Brief Summary of the Invention

Aspects of the disclosure are based, in part, on the discovery of methods for overcoming problems associated with systematic errors (bias) in genome capture, amplification and sequencing methods, namely high variability in the capture and amplification of nucleic acids and disproportionate representation of heterozygous alleles in sequencing libraries. Accordingly, in some embodiments, the disclosure provides methods that reduce variability in the capture and amplification of nucleic acids. In other embodiments, the methods improve allelic representation in sequencing libraries and, thus, improve variant detection outcomes. In certain embodiments, the disclosure provides preparative methods for capturing source nucleic acids (e.g., genetic loci) that involve the use of source sequences to uniquely tag individual nucleic acid molecules. In some embodiments, the source sequence permit the detection of bias based on the occurrence of combinations of source and target sequences observed in a sequencing reaction. In other embodiments, the methods reduce bias, or the risk of bias, associated with the capture, amplification and sequencing of genetic loci, e.g., for diagnostic purposes.

According to some aspects, methods of genotyping a plurality of genetic loci are provided. In some embodiments, the methods involve obtaining a plurality of molecular inversion probe (MIP) capture products, each comprising a molecular inversion probe sequence, a target sequence and a source sequence, in which the target sequence is a captured genomic locus, amplifying the plurality of molecular inversion probe capture products, and genotyping target sequences by determining, for each target sequence, the sequence of at least a threshold number of unique combinations of target sequence and source sequence of molecular inversion probe capture products in the amplified plurality.

In some embodiments, a plurality of molecular inversion probe capture products is obtained by capturing target sequences from a genomic sample with a plurality of molecular inversion probes, each comprising a source sequence.

In some embodiments of the methods disclosed herein, capturing is performed under conditions wherein the likelihood of obtaining two or more molecular inversion probe capture

Petitioner Illumina, Inc., Ex. 1061, p. 4
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

products with identical combinations of target and source sequences is equal to or less than a predetermined value, optionally wherein the predetermined value is about 0.05.

In some embodiments of the methods disclosed herein, the threshold number for a specific target sequence is selected based on a desired statistical confidence for the genotype of the specific target sequence.

In some embodiments, the methods disclosed herein comprise determining a statistical confidence for the genotype of a specific target sequence based on the number of unique combinations of target sequence and source sequence of molecular inversion probe capture products comprising the specific target sequence.

According to some aspects, methods of analyzing a plurality of genetic loci are provided. In some embodiments, the methods comprise obtaining a plurality of molecular inversion probe capture products, each comprising a molecular inversion probe sequence, a target sequence and a source sequence, amplifying the plurality of molecular inversion probe capture products, determining numbers of occurrence of combinations of target sequence and source sequence of molecular inversion probe capture products in the amplified plurality, and if the number of occurrence of a specific combination of target sequence and source sequence exceeds a predetermined value, detecting bias in the amplification of the molecular inversion probe comprising the specific combination.

In some embodiments, the methods disclosed herein comprise genotyping target sequences in a plurality of molecular inversion probe capture products, in which the genotyping comprises correcting for bias, if detected.

In some embodiments of the methods disclosed herein, the target sequence is a sequence of a gene selected from Table 1.

Brief Description of the Drawings

Figure 1a, b, and c depict various methods for combining source sequence and target sequences (NNNN depicts a source sequence).

Figure 2 depicts a method for genotyping based on target and source sequences. Sequencing reads of target and sources sequences are collapsed to make diploid genotype calls.

Figure 3 depicts results of a simulation of a MIP capture reaction in which MIP probes, each having a source sequence of 15 nucleotides, are combined with 10000 target sequence copies (e.g., genome equivalents). In this simulated reaction, the probability of

Petitioner Illumina, Inc., Ex. 1061, p. 5
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

capturing one or more copies of a target sequence having the same source sequence is 0.05. The Y axis reflects the number of observations. The X axis reflects the number of independent occurrences of target:source combinations.

Figure 4 depicts a graph of sequencing coverage, which is needed to ensure that alleles are sampled to sufficient depth (either 10x or 20x minimum sampling per allele, assuming 1000 targets). X axis is total per-target coverage required, and Y axis is the probability that a given total coverage will result in at least 10x or 20x coverage for each allele.

Detailed Description of the Invention

Aspects of the disclosure are based, in part, on the discovery of methods for overcoming problems associated with systematic errors (bias) in genome capture, amplification and sequencing methods, namely high variability in the capture and amplification of nucleic acids and disproportionate representation of heterozygous alleles in sequencing libraries. Accordingly, in some embodiments, the disclosure provides methods that reduce variability in the capture and amplification of nucleic acids. In other embodiments, the methods improve allelic representation in sequencing libraries and, thus, improve variant detection outcomes. In certain embodiments, the disclosure provides preparative methods for capturing source nucleic acids (e.g., genetic loci) that involve the use of source sequences to uniquely tag individual nucleic acid molecules. In some embodiments, the source sequence permit the detection of bias based on the frequency with which pairs of source and target sequences are observed in a sequencing reaction. In other embodiments, the methods reduce bias, or the risk of bias, associated with the capture, amplification and sequencing of genetic loci, e.g., for diagnostic purposes.

According to some aspects, the disclosure relates to multiplex diagnostic methods. In some embodiments, the multiplex diagnostic methods comprise capturing a plurality of target nucleic acids (genetic loci) in parallel (e.g., a genetic locus of Table 1) using capture probes having unique source sequences. In some embodiments, the genetic loci possess one or more polymorphisms (e.g., a polymorphism of Table 2) the genotypes of which correspond to disease causing alleles. Accordingly, in some embodiments, the disclosure provides methods for assessing multiple heritable disorders in parallel. In some embodiments, methods are provided for diagnosing multiple heritable disorders in parallel at a pre-implantation, prenatal, perinatal, or postnatal stage. In some embodiments, the disclosure

Petitioner Illumina, Inc., Ex. 1061, p. 6
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

provides methods for analyzing multiple target nucleic acids (e.g., a plurality of genetic loci selected from Table 1) from a patient sample, such as a blood, pre-implantation embryo, chorionic villus or amniotic fluid sample.

In some embodiments, the invention relates to compositions comprising, and methods for using, a large library of identifier nucleic acids consisting of "source sequences" in preparative steps for reducing the effects of the bias and for facilitating genotyping of target sequences. As used herein, a "source sequence" is a sequence of a nucleic acid of a predefined length, which has been synthesized by random nucleotide addition, and which is incorporated in a molecular inversion probe to identify the origin of molecular inversion probe capture products.

During a library preparation process, each nucleic acid molecule may be tagged with a unique source sequence in a configuration that permits the source sequence to be sequenced along with the target nucleic acid sequence of interest (the nucleic acid sequence for which the library is being prepared, e.g., a polymorphic sequence). The incorporation of the nucleic acid comprising a source sequence at a particular step allows the detection and correction of biases in subsequent steps of the protocol.

A large library of unique source sequences may be created by using degenerate, random-sequence polynucleotides of defined length. The source sequences of the polynucleotides may be read at the final stage of the sequencing. The observations of the source sequences may be used to detect and correct biases in the final sequencing read-out of the library. For example, the total possible number of source sequences, which may be produced, e.g., randomly, is 4^N, where N is the length of the source sequence. Thus, it is to be understood that the length of the source sequence may be adjusted such that the size of the population of MIPs having unique source sequences is sufficient to produce a library of MIP capture products in which identical independent target sequence and source sequence combinations (target:source sequences) are rare.

In the final readout of a sequencing process, each read may have an additional unique source sequence. When source sequence are distributed randomly in a library, all the unique source sequences will be observed about an equal number of times. The number of occurrences of a source sequence may follow a Poisson distribution. *[Right? Assuming we had a uniform distribution of barcodes that were amplified, what happens to their distribution post-PCR if the efficiencies are unequal? I suppose this also depends on the distribution of the efficiencies.]* ***GREG: please confirm if the Poission distribution statement is accurate.).***

Petitioner Illumina, Inc., Ex. 1061, p. 7
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Overrepresentation of target:source sequence combinations in a pool of MIP capture products (e.g., amplified MIP capture products) is indicative of bias in the amplification protocol (e.g., bias in the amplification process). Target:source sequence combinations that are statistically overrepresented are indicative of bias in the protocol at one or more steps between the incorporation of the source sequences into MIPs and the actual sequencing of the MIP capture products.

The number of reads of a given target:source sequence may be indicative (may serve as a proxy) of the amount of that target sequence present in the originating sample. In some embodiments, the numbers of occurrence of sequences in the originating sample is the quantity of interest. Using the methods disclosed herein, the occurrence of source sequences in a pool of MIPs may be predetermined (e.g., may be the same for all source sequences). Accordingly, changes in the occurrence of source sequences after amplification and sequencing may be indicative of bias in the protocol. Bias may be corrected to provide an accurate representation of the composition of the original MIP pool, e.g., for diagnostic purposes.

A MIP library (each MIP in the library having a unique source sequence) may be constructed such that the number of MIP molecules in the library is significantly larger than the number prospective target molecules to be captured using the library. This may ensure that the capture products include only unique target:source sequences, because the capture reaction undersamples the total population of unique source sequences in the MIP library. For example, an experiment utilizing 1 μg of genomic DNA will contain about ~150,000 copies of a diploid genome. For a MIP library, each MIP in the library comprising a randomly produced 12-mer source sequence (~1.6 million possible unique source sequences), there would be more than 100 unique source sequences per genomic copy. For a MIP library, each MIP in the library comprising a randomly produced 15-mer source sequence (~1 billion possible unique source sequences), there would be more than 7000 unique source sequences per genomic copy. Therefore, the probability of the same source sequence being incorporated multiple times is incredibly small. Thus, it is to be appreciated that the length of the source sequence is to be selected based on the amount of target sequence in a MIP capture reaction and the desired probability for having multiple, independent occurrences of target:source sequence combinations. The source sequence may be, up to 5, up to 6, up to 7 up to 8, up to 9, up to 10, up to 11, up to 12, up to 13, up to 14, up to 15, up to 16, up to 17,

Petitioner Illumina, Inc., Ex. 1061, p. 8
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

up to 18, up to 19, up to 20, up to 21, up to 22, up to 23, up to 24, up to 25, or more nucleotides in length.

The skilled artisan will appreciate that as part of a MIP library preparation process, adapters may be ligated onto the ends of the molecules of interest. Adapters often contain PCR primer sites (for amplification or emulsion PCR) and/or sequencing primer sites. In addition, barcodes may be included, for example, to uniquely identify individual samples (e.g., patient samples) that may be mixed together. (See, e.g., USPTO Publication Number US 2007/0020640 A1 (McCloskey et al.)

The actual incorporation of the random source sequences can be performed through various methods known in the art. For example, nucleic acids comprising source sequences may be incorporated by ligation. This is a flexible method, because molecules having source sequence can be ligated to any blunt-ended nucleic acids. The sequencing primers must be incorporated subsequently such that they sequence both the source sequence and the target sequence. Alternatively, the sequencing adaptors can be synthesized with the random source sequences at their 3' end (as degenerate bases), so that only one ligation must be performed. Another method is to incorporate the source sequence into a PCR primer, such that the primer structure is arranged with the common adaptor sequence followed by the random source sequence followed by the PCR priming sequence (in 5' to 3' order). A source sequence and adaptor sequence (which may contain the sequencing primer site) are incorporated as tags. Another method to incorporate the source sequences is to synthesize them into a padlock probe prior to performing a gene capture reaction. The source sequence is incorporated 3' to the targeting arm but 5' to the amplification primer that will be used downstream in the protocol. Another method to incorporate the source sequences is as a tag on a gene-specific or poly-dT reverse-transcription primer. This allows the source sequence to be incorporated directly at the cDNA level.

In some embodiments, at the incorporation step, the distribution of source sequences can be assumed to be uniform. In this case, bias in any part of the protocol would change the uniformity of this distribution, which can be observed after sequencing. This allows the source sequence to be used in any preparative process where the ultimate output is sequencing of many molecules in parallel.

Source sequences may be incorporated into probes (e.g., MIPs) of a plurality when they are synthesized on-chip in parallel, such that degeneracy of the incorporated nucleotides is sufficient to ensure near-uniform distribution in the plurality of probes. It is to be appreciated that amplification of pool of unique source sequences may itself introduce bias in

Petitioner Illumina, Inc., Ex. 1061, p. 9
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

the initial pool. However, in most practical cases, the scale of synthesis (e.g., by column synthesis, chip based synthesis, etc.) is large enough that amplification of an initial pool of source sequences is not necessary. By avoiding amplification or selection step on the pool of unique source sequences, potential bias may be minimized.

One example of the use of the source sequences is in genome re-sequencing. Considering that the raw accuracy of most next-generation sequencing instruments is relatively low, it is crucial to oversample the genomic loci of interest. Furthermore, since there are two alleles at every locus, it is important to sample enough to ensure that both alleles have been observed a sufficient number of times to determine with a sufficient degree of statistical confidence whether the sample is homozygous or heterozygous. Indeed, the sequencing is performed to sample the composition of molecules in the originating sample. However, after multiple reads have been collected for a given locus, it is possible that due to bias (e.g., caused by PCR amplification steps), a large fraction of the reads are derived from a single originating molecule. This would skew the population of target sequences observed, and would affect the outcome of the genotype call. For example, it is possible that a locus that is heterozygous is called as homozygous, because there are only a few observations of the second allele out of many observations of that locus. However, if information is available on source sequences, this situation could be averted, because the over-represented allele would be seen to also have an over-represented source sequence (i.e., the sequences with the overrepresented source sequence all originated from the same single molecule). Therefore, the sequences and corresponding distribution of source sequences can be used as an additional input to the genotype-calling algorithm to significantly improve the accuracy and confidence of the genotype calls.

Another embodiment of the invention is in the quantitative analysis of gene expression. When RNA is sequenced, the relative numbers of counts for different transcripts is used as a proxy for the relative abundance of the transcript in the original sample of RNA. As described above, the library preparation protocols may introduce biases that change the relative amounts of the RNA species in the sample that actually gets sequenced (post-processing). However, if each RNA/cDNA molecule is coupled to a unique random source sequence, then any repeated source sequences in the sequencing reactions can simply be collapsed into each other. By collapsing similar clones in such a way that the sequencing data only includes a single representative for each source sequence (or target:source sequence combination), the biases that were introduced in preparation are essentially eliminated. **[IS THIS GSG WORK??]**

Petitioner Illumina, Inc., Ex. 1061, p. 10
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Random source sequences methodology may be used retrospectively on samples that do not explicitly incorporate unique source sequences. Most oligonucleotide synthesis methods are error-prone. The fraction of oligos with errors depends on the length of the oligonucleotide and the synthesis method employed. For example, using traditional column synthesis of 25-mers, we would expect that 1 in 40 PCR primers would contain a synthesis error (assuming one error every 1,000 bases). These error-containing primers are also incorporated in the PCR reaction. Since the errors are distributed in a random fashion, they are unique and rare; therefore, the error-containing PCR primers can be considered as unique random source sequences as well. Since these primers typically get sequenced as well, the errors can be detected and used as random source sequences to look for biases in the same way described above. **[IS THIS GSG WORK??]**

In some aspects, the disclosure provides methods for analyzing a plurality of target sequences which are genetic loci or portions of genetic loci (e.g., a genetic locus of Table 1). The genetic loci may be analyzed by sequencing to obtain a genotype at one or more polymorphisms (e.g., SNPs). Exemplary polymorphisms are disclosed in Table 2. The skilled artisan will appreciate that other polymorphisms are known in the art and may be identified, for example, by querying the Entrez Single Nucleotide Polymorphism database, for example, by searching with a GeneID from Table 1.

The mutations listed in Table 2 are documented polymorphisms in several disease-associated genes (CFTR is mutated in cystic fibrosis, GBA is mutated in Gaucher disease, ASPA is mutated in Canavan disease, HEXA is mutated in Tay Sachs disease). The polymorphisms are of several types: insertion/deletion polymorphisms which will cause frameshifts (and thus generally interrupt protein function) unless the insertion/deletion length is a multiple of 3bp, and substitutions which can alter the amino acid sequence of the protein and in some cases cause complete inactivation by introduction of a stop codon.

The methods disclosed here are also useful for diagnosis of one or more heritable disorders. In some embodiments, a heritable disorder that may be diagnosed with the methods disclosed herein is a genetic disorder that is prevalent in the Ashkenazi Jewish population. In some embodiments, the heritable disorders are selected from: 21-Hydroxylase-Defiocient Congenital Adrenal Hyperplasia; ABCC8-Related Hyperinsulinism; Alpha-Thalassemia, includes Constant Spring, & MR associated; Arylsulfatase A Deficiency-Metyachromatic Leukodystrophy; Biotinidase Deficiency-Holocarboxylase Synthetase Deficiency; Bloom's Syndrome; Canavan Disease; CFTR-Related Disorders-cystic fibrosis; Citrullinemia Type I; Combined MMA & Homocystinuria-dblC; Dystrophinopathies (DMD

Petitioner Illumina, Inc., Ex. 1061, p. 11
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

& BMD); Familial Dysautonomia; Fanconi Anemia-FANCC; Galactosemia-Classical: Galactokinase Defiency & Galactose Epimerase Deficiency; Gaucher Disease; GJB2-Related DFNB 1 Nonsyndromic Hearing Loss and Deafness; Glutaric acidemia Type 1; Hemoglobinopathies beta-chain disorders; Glycogen Storage Disease Type 1A; Maple Syrup Urine Disease; Types 1A, 1B, 2, 3; Medium Chain Acyl-Coenzyme A; Dehydrogenase Deficiency-MCADD; Methylmalonic Acidemia; Mucolipidosis IV; Nemaline Myopathy; Nieman-Pick Type A-Acid Sphingomyelinase Deficiency; Non-Ketotic Hyperglycinemia-Glycine Encephalopathy; Ornithine Transcarbamylase Deficiency; PKU Phenylalanine Hydroxylase Deficiency; Propionic Acidemia; Short Chain Acyl-CoA Dehydrogenase Deficiency-SCADD; Smith-Lemli-Opitz Syndrome; Spinal Muscular Atrophy (SMN1)-SMA; Tay Sachs-HexA Deficiency; Usher Synbdrome-Type I (Type IB, Type IC, Type ID, Type IF, Type IG) ; X-Linked Mental Retardation ARX-Related Disorders; X-Linked Mental Retardation with Cerebellar Cypoplasia and sistinctive Facial Appearance; X-Linked Mental Retardation; includes 9, 21, 30, 46, 58, 63, 88, 89; X-linked mental retardation: FM1-Related Disorders-FRXA, Fragile X MR; X-linked SMR: Renpenning Syndrome 1; Zellweger Spectrum disorders - Peroxisomal Bifunctional Enzyme Deficiencies including Zellweger, NALD, infantile Refsums.

Table 1: Target Nucleic Acids

Gene name	Gene ID	Description	Gene aliases	OMIM	RefSeqGene	Chromosome map position
CYP21A2	1589	cytochrome P450, family 21, subfamily A, polypeptide 2	CAH1; CPS1; CA21H; CYP21; CYP21B; P450c21B; MGC150536; MGC150537; CYP21A2	201910	NG_008337.1	6p21.3
ABCC8	6833	ATP-binding cassette, sub-family C (CFTR/MRP), member 8	HI; SUR; HHF1; MRP8; PHHI; SUR1; ABC36; HRINS; TNDM2; ABCC8	600509	NG_008867.1	11p15.1
ATRX	546	alpha thalassemia/mental retardation syndrome X-linked (RAD54 homolog, S. cerevisiae)	SHS; XH2; XNP; ATR2; SFM1; RAD54; MRXHF1; RAD54L; ZNF-HX; MGC2094; ATRX	300032	NG_008838.1	Xq13.1-q21.1
ARSA	410	arylsulfatase A	MLD; ARSA	607574	NG_009260.1	22q13.31-qter;

Petitioner Illumina, Inc., Ex. 1061, p. 12
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	Description	Gene aliases	OMIM	RefSeqGene	Chromosome map position
						22q13.33
PSAP	5660	prosaposin	GLBA; SAP1; FLJ00245; MGC110993; PSAP	176801	NG_008835.1	10q21-q22
BTD	686	biotinidase	BTD	609019	NG_008019.1	3p25
HLCS	3141	holocarboxylase synthetase (biotin-(proprionyl-Coenzyme A-carboxylase (ATP-hydrolysing)) ligase)	HCS; HLCS	609018	NC_000021.7	21q22.1; 21q22.13
BLM	641	Bloom syndrome, RecQ helicase-like	BS; RECQ2; RECQL2; RECQL3; MGC126616; MGC131618; MGC131620; BLM	604610	NG_007272.1	15q26.1
ASPA	443	aspartoacylase (Canavan disease)	ASP; ACY2; ASPA	608034	NG_008399.1	17pter-p13
CFTR	1080	cystic fibrosis transmembrane conductance regulator (ATP-binding cassette sub-family C, member 7)	CF; MRP7; ABC35; ABCC7; CFTR/MRP; TNR-CFTR; dJ760C5.1; CFTR	602421	NC_000007.12	7q31.2
ASS1	445	argininosuccinate synthetase 1	ASS; CTLN1; ASS1	603470	NG_011542.1	9q34.1
MMACHC	25974	methylmalonic aciduria (cobalamin deficiency) cblC type, with homocystinuria	cblC; FLJ25671; DKFZp564I122; RP11-291L19.3; MMACHC	609831	NC_000001.9	1p34.1
IKBKAP	8518	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase complex-associated protein	FD; DYS; ELP1; IKAP; IKI3; TOT1; FLJ12497; DKFZp781H1425; IKBKAP	603722	NG_008788.1	9q31
FANCC	2176	Fanconi anemia, complementation group C	FA3; FAC; FACC; FLJ14675; FANCC	227645	NG_011707.1	9q22.3
GALK1	2584	galactokinase 1	GK1; GALK; GALK1	604313	NG_008079.1	17q24
GALT	2592	galactose-1-phosphate uridylyltransferase	GALT	606999	NC_000009.10	9p13
GALE	2582	UDP-galactose-4-epimerase	SDR1E1; FLJ95174; FLJ97302; GALE	606953	NG_007068.1	1p36-p35

Petitioner Illumina, Inc., Ex. 1061, p. 13
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	Description	Gene aliases	OMIM	RefSeqGene	Chromosome map position
GBA	2629	glucosidase, beta; acid (includes glucosylceramidase)	GCB; GBA1; GLUC; GBA	606463	NG_009783.1	1q21
GJB2	2706	gap junction protein, beta 2, 26kDa	HID; KID; PPK; CX26; DFNA3; DFNB1; NSRD1; DFNA3A; DFNB1A; GJB2	121011	NG_008358.1	13q11-q12
GCDH	2639	glutaryl-Coenzyme A dehydrogenase	GCD; ACAD5; GCDH	608801	NG_009292.1	19p13.2
G6PC	2538	glucose-6-phosphatase, catalytic subunit	G6PT; GSD1; GSD1a; MGC163350; G6PC	232200	NG_011808.1	17q21
HBB	3043	hemoglobin, beta	CD113t-C; beta-globin; HBB	141900	NG_000007.3	11p15.5
BCKDHA	593	branched chain keto acid dehydrogenase E1, alpha polypeptide	MSU; MSUD1; OVD1A; BCKDE1A; FLJ45695; BCKDHA	608348	NC_000019.8	19q13.1-q13.2
BCKDHB	594	branched chain keto acid dehydrogenase E1, beta polypeptide	E1B; FLJ17880; dJ279A18.1; BCKDHB	248611	NG_009775.1	6q13-q15
DBT	1629	dihydrolipoamide branched chain transacylase E2	E2; E2B; BCATE2; MGC9061; DBT	248610	NG_011852.1	1p31
DLD	1738	dihydrolipoamide dehydrogenase	E3; LAD; DLDH; GCSL; PHE3; DLD	238331	NG_008045.1	7q31-q32
ACADM	34	acyl-Coenzyme A dehydrogenase, C-4 to C-12 straight chain	MCAD; ACAD1; MCADH; FLJ18227; FLJ93013; FLJ99884; ACADM	607008	NG_007045.1	1p31
MMAA	166785	methylmalonic aciduria (cobalamin deficiency) cblA type	cblA; MGC120010; MGC120011; MGC120012; MGC120013; MMAA	607481	NG_007536.1	4q31.22
MMAB	326625	methylmalonic aciduria (cobalamin deficiency) cblB type	ATR; cblB; MGC20496; MMAB	607568	NG_007096.1	12q24
MUT	4594	methylmalonyl Coenzyme A mutase	MCM; MUT	609058	NG_007100.1	6p12.3
MCOLN1	57192	mucolipin 1	ML4; MLIV; MST080; TRPML1; MSTP080; TRP-ML1; TRPM-L1; MCOLN1	605248	NC_000019.8	19p13.3-p13.2

Petitioner Illumina, Inc., Ex. 1061, p. 14
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	Description	Gene aliases	OMIM	RefSeqGene	Chromosome map position
ACTA1	58	actin, alpha 1, skeletal muscle	ACTA; ASMA; CFTD; MPFD; NEM1; NEM2; NEM3; CFTD1; CFTDM; ACTA1	102610	NG_006672.1	1q42.13
TPM3	7170	tropomyosin 3	TM3; TRK; NEM1; TM-5; TM30; TM30nm; TPMsk3; hscp30; MGC3261; FLJ41118; MGC14582; MGC72094; OK/SW-cl.5; TPM3	191030	NG_008621.1	1q21.2
TNNT1	7138	troponin T type 1 (skeletal, slow)	ANM; TNT; STNT; TNTS; FLJ98147; MGC104241; TNNT1	191041	NG_011829.1	19q13.4
NEB	4703	nebulin	NEM2; NEB177D; FLJ11505; FLJ36536; FLJ39568; FLJ39584; DKFZp686C1456; NEB	161650	NG_009382.1	2q22
SMPD1	6609	sphingomyelin phosphodiesterase 1, acid lysosomal	ASM; NPD; SMPD1	607608	NG_011780.1	11p15.4-p15.1
GLDC	2731	glycine dehydrogenase (decarboxylating)	GCE; NKH; GCSP; HYGN1; MGC138198; MGC138200; GLDC	238300	NC_000009.10	9p22
GCSH	2653	glycine cleavage system protein H (aminomethyl carrier)	GCE; NKH; GCSH	238330	NC_000016.8	16q23.2
AMT	275	aminomethyltransferase	GCE; NKH; GCST; AMT	238310	NC_000003.10	3p21.2-p21.1
OTC	5009	ornithine carbamoyltransferase	OCTD; MGC129967; MGC129968; MGC138856; OTC	300461	NG_008471.1	Xp21.1
PAH	5053	phenylalanine hydroxylase	PH; PKU; PKU1; PAH	612349	NG_008690.1	12q22-q24.2
DHPR	5860	quinoid dihydropteridine reductase	DHPR; PKU2; SDR33C1; FLJ42391; QDPR	612676	NG_008763.1	4p15.31
PTS	5805	6-pyruvoyltetrahydropterin synthase	PTPS; FLJ97081; PTS	261640	NG_008743.1	11q22.3-q23.3
PCCA	5095	propionyl Coenzyme A carboxylase, alpha polypeptide	PCCA	232000	NG_008768.1	13q32

Petitioner Illumina, Inc., Ex. 1061, p. 15
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	Description	Gene aliases	OMIM	RefSeqGene	Chromosome map position
PCCB	5096	propionyl Coenzyme A carboxylase, beta polypeptide	DKFZp451E113; PCCB	232050	NG_008939.1	3q21-q22
ACADS	35	acyl-Coenzyme A dehydrogenase, C-2 to C-3 short chain	SCAD; ACAD3; ACADS	606885	NG_007991.1	12q22-qter
DHCR7	1717	7-dehydrocholesterol reductase	SLOS; DHCR7	602858	NC_000011.8	11q13.2-q13.5
SMN1	6606	survival of motor neuron 1, telomeric	SMA; SMN; SMA1; SMA2; SMA3; SMA4; SMA@; SMNT; BCD541; T-BCD541; SMN1	600354	NG_008691.1	5q13
HEXA	3073	hexosaminidase A (alpha polypeptide)	TSD; MGC99608; HEXA	606869	NG_009017.1	15q23-q24
MYO7A	4647	myosin VIIA	DFNB2; MYU7A; NSRD2; USH1B; DFNA11; MYOVIIA; MYO7A	276903	NG_009086.1	11q13.5
USH1C	10083	Usher syndrome 1C (autosomal recessive, severe)	PDZ73; AIE-75; DFNB18; PDZ-45; PDZ-73; NY-CO-37; NY-CO-38; ush1cpst; PDZ-73/NY-CO-38; USH1C	605242	NC_000011.8	11p15.1-p14
CDH23	64072	cadherin-like 23	USH1D; DFNB12; FLJ00233; FLJ36499; KIAA1774; KIAA1812; MGC102761; DKFZp434P2350; CDH23	605516	NG_008835.1	10q21-q22
PCDH15	65217	protocadherin 15	USH1F; DFNB23; DKFZp667A1711; PCDH15	605514	NG_009191.1	10q21.1
SANS	124590	Usher syndrome 1G (autosomal recessive)	SANS; ANKS4A; FLJ33924; USH1G	607696	NG_007882.1	17q25.1
ARX	170302	aristaless related homeobox	ISSX; PRTS; MRX29; MRX32; MRX33; MRX36; MRX38; MRX43; MRX54; MRX76; MRX87; MRXS1; ARX	300382	NG_008281.1	Xp21
OPHN1	4983	oligophrenin 1	OPN1; MRX60; OPHN1	300127	NG_008960.1	Xq12
JARID1C	8242	lysine (K)-specific demethylase 5C	MRXJ; SMCX; MRXSJ; XE169; JARID1C;	314690	NG_008085.1	Xp11.22-p11.21

Petitioner Illumina, Inc., Ex. 1061, p. 16
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	Description	Gene aliases	OMIM	RefSeqGene	Chromosome map position
			DXS1272E; KDM5C			
FTSJ1	24140	FtsJ homolog 1 (E. coli)	JM23; MRX9; SPB1; TRM7; CDLIV; MRX44; FTSJ1	300499	NG_008879.1	Xp11.23
SLC6A8	6535	solute carrier family 6 (neurotransmitter transporter, creatine), member 8	CRT; CT1; CRTR; MGC87396; SLC6A8	300036	NC_000023.9	Xq28
DLG3	1741	discs, large homolog 3 (Drosophila)	MRX; MRX90; NEDLG; NE-Dlg; SAP102; SAP-102; KIAA1232; DLG3	300189	NC_000023.9	Xq13.1
TM4SF2	7102	tetraspanin 7	A15; MXS1; CD231; MRX58; CCG-B7; TM4SF2; TALLA-1; TM4SF2b; DXS1692E; TSPAN7	300096	NG_009160.1	Xp11.4
ZNF41	7592	zinc finger protein 41	MRX89; MGC8941; ZNF41	314995	NG_008238.1	Xp11.23
FACL4	2182	acyl-CoA synthetase long-chain family member 4	ACS4; FACL4; LACS4; MRX63; MRX68; ACSL4	300157	NG_008053.1	Xq22.3-q23
PQBP1	10084	polyglutamine binding protein 1	SHS; MRX55; MRXS3; MRXS8; NPW38; RENS1; PQBP1	300463	NC_000023.9	Xp11.23
PEX1	5189	peroxisomal biogenesis factor 1	ZWS1; PEX1	602136	NG_008341.1	7q21.2
PXMP3	5828	peroxisomal membrane protein 3, 35kDa	PAF1; PEX2; PMP3; PAF-1; PMP35; RNF72; PXMP3	170993	NG_008371.1	8q21.1
PEX6	5190	peroxisomal biogenesis factor 6	PAF2; PAF-2; PXAAA1; PEX6	601498	NG_008370.1	6p21.1
PEX10	5192	peroxisomal biogenesis factor 10	NALD; RNF69; MGC1998; PEX10	602859	NG_008342.1	1p36.32
PEX12	5193	peroxisomal biogenesis factor 12	PAF-3; PEX12	601758	NG_008447.1	17q12
PEX5	5830	peroxisomal biogenesis factor 5	PXR1; PTS1R; PTS1-BP; FLJ50634; FLJ50721; FLJ51948; PEX5	600414	NG_008448.1	12p13.31
PEX26	55670	peroxisomal biogenesis factor 26	FLJ20695; PEX26M1T; Pex26pM1T; PEX26	608666	NG_008339.1	22q11.21

Petitioner Illumina, Inc., Ex. 1061, p. 17
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Table 2: Polymorphisms

Gene name	Gene ID	SNP ID	Mutation
CFTR	1080	rs63500661	TCACATCACCAAGTTAAAAAAAAAAA[A/G]GGGGCGGGGGGGCAGAAT GAAAATT
CFTR	1080	rs63107760	AAACAAGGATGAATTAAGTTTTTTTT[- /T]AAAAAAGAAACATTTGGTAAGGGGA
CFTR	1080	rs62469443	ATCACCAAGTTAAAAAAAAAAAAGGG[A/G]CGGGGGGGCAGAATGAAA ATTGCAT
CFTR	1080	rs62469442	CTATTGAACCAGAACCAAACAGGAAT[A/G]CCATAGCATTTTGTAAACT AAACTG
CFTR	1080	rs62469441	CAGGAGTTCAAGACCAGCCTACTAAA[A/C]CACACACACACACACACAC ACACAC
CFTR	1080	rs62469439	GATTAAATAATAGTGTTTATGTACCC[C/G]GCTTATAGGAGAAGAGGGT GTGTGT
CFTR	1080	rs62469438	ATTGTTATCTTTTCATATAAGGTAAC[A/T]GAGGCCCAGAGAGATTAAAT AACAT
CFTR	1080	rs62469437	TAATTTTAATTAAGTAAATTTAATTG[A/G]TAGATAAATAAGTAGATAAA AAATA
CFTR	1080	rs62469436	GTATAAAAAAAAAAAAAAAAAAAAGTT[A/T]GAATGTTTTCTTGCATTCA GAGCCT
CFTR	1080	rs62469435	ATACTAAAAATTTAAAGTTCTCTTGC[A/G]ATATATTTTCTTAATATCTTA CATC
CFTR	1080	rs62469434	TGCTGGGATTACAGGCGTGAGCCACC[A/G]CGCCTGGCCTGATGGGACA TATTTT
CFTR	1080	rs62469433	CTACAATATAAGTATAGTATTGCAAA[A/C]CCATCAGGAAGGGTGTTAA CTATTT
CFTR	1080	rs61763210	GTTGTCTCCAAACTTTTTTTCAGGTG[- /AGA]AGGTGGCCAACCGAGCTTCGGAAAG
CFTR	1080	rs61720488	TTTTTTCATAAAAGATTATATAAAGG[A/C]TATTGCTTTTGAATCACAAA CACTA
CFTR	1080	rs61481156	ATCTAGTGAGCAGTCAGGAAAGAGAA[C/T]TTCCAGATCCTGGAAATCA GGGTTA
CFTR	1080	rs61443875	TAGAGTATAAAAAAAAAAAAAAAAAA[- /A]GTTTGAATGTTTTCTTGCATTCAGA
CFTR	1080	rs61312222	TGCAAATGCCAACTATCAAAGATATT[C/G]GAGTATACTGTCAATAAACT TCATA
CFTR	1080	rs61159372	TCCTCAACAGTTAGAAACAATATTTT[C/G]AGTGATTTCCCATGCCAACT TTACT

Petitioner Illumina, Inc., Ex. 1061, p. 18
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	SNP ID	Mutation
CFTR	1080	rs61094145	TTTTTGGTATTGTTGTTAAATAAGTG[A/G]GAATTCAATACAGTATAATG TCTGT
CFTR	1080	rs61086387	CTTGAAATCGGATATATATATATATA[- /TGTATATATATATATATATATATATATATATACATATATATATATA]GTA TTATCCCTGTTTTCACAGTTTT
CFTR	1080	rs60996744	AGAGGGGCTGTGAAGGACACCAAGGA[A/G]GAGACTAAGAGCCAGGAG GGAAAAC
CFTR	1080	rs60960860	TAGAGTTTATTAGCTTTTACTACTCT[A/G]CTTAGTTACTTTGTGTTACAG AATA
CFTR	1080	rs60923902	ACTAGTGATGATGAGCTTCTTTTCAT[- /AT]GTTTGTTGGCTGCATAAATGTCTTC
CFTR	1080	rs60912824	GCAGAGAAAAGAGGGGCTGTGAAGGA[C/G]ACCAAGGAGGAGACTAAG AGCCAGG
CFTR	1080	rs60887846	TTCAGAGGTCTACCACTGGTGCATAC[G/T]CTAATCACAGTGTCGAAAAT TTTAC
CFTR	1080	rs60793174	AAGAAAGAGCAAAAGAGGGCAAACTT[C/T]TCATACATTTTTGATGTCG AAACCA
CFTR	1080	rs60788575	CCTAAAGTTTAAAAAGAAAAAAAAAA[- /A]GGAAGAAGGAATTAAAAATCCAAAG
CFTR	1080	rs60760741	GTGTGTGTGTGTATATATATATATAT[A/T]TATATATTTTTTTTTTCCTGA GCCA
CFTR	1080	rs60456599	AAACTGTTGATGTTTTCATTTATTTA[C/G]ATCATTGGAAAACTTTAGATT CTAG
CFTR	1080	rs60363249	TTTATCCATTCTTAACCAGAACAGAC[A/G]TTTTTTCAGAGCTGGTCCAG GAAAA
CFTR	1080	rs60355115	TTGAAATCGGATATATATATATATAT[A/G]TATATATATATATATATATA TATAT
CFTR	1080	rs60308689	TAGTTTTTTATTTCCTCATATTATTT[- /T]CAGTGGCTTTTTCTTCCACATCTTT
CFTR	1080	rs60271242	ACATAGTTCTCAGTGGTACAACTACA[A/G]GTGATTTCTCTTTTCTTATTT CTGG
CFTR	1080	rs60010318	AGAGCAATGGCATCCCTTGTCTTGTG[C/T]TATACAGGATGCAGCAATTT ATAGG
CFTR	1080	rs59961323	TTCTGTCTACATAAGATGTCATACTA[A/G]ATTATCTTTTCCAGCATGCAT TCAG
CFTR	1080	rs59961270	CAGGGTGGCATGTTAGGCAGTGCTTA[A/G]AATAAATGAGTTGGTTATA CAAGTA
CFTR	1080	rs59837506	AGGACACACACACACACACACACACA[- /CA]TGCACACACATTTAAATAGATGCAT

Petitioner Illumina, Inc., Ex. 1061, p. 19
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	SNP ID	Mutation
CFTR	1080	rs59572090	TAAAAAATTGGTATAATGAAATTGCA[C/T]TTGTAGTCTTTGGACATTTAAATCC
CFTR	1080	rs59548252	TTTCAATACTTAAGAGGTACGCAGAG[A/G]AAAGAGGGGCTGTGAAGGACACCAA
CFTR	1080	rs59519859	CAGCAATGAATATTTTGAGGCTGAGG[C/T]GCTGAGGGGTAAAATTGCAGCCTGG
CFTR	1080	rs59509837	TTATGGTTTATATTTTTGTGTCTTCT[-/CTTT]AACACATCTTTTCTAGCAGAATTCA
CFTR	1080	rs59417037	GTATTTTAGTTTTTTTTTTTTGTTTG[-/T]TTTGTTTTGTTTTGTTTTGTTTTTG
CFTR	1080	rs59159458	TGGGTGACTCCATTTTTACTTTTAGT[C/T]TGGTCTGTTGAGGCCTCGTGAGAGA
CFTR	1080	rs59048119	TATTTTCATGTATTTTAGTTTTTTTT[-/TTTT]GTTTGTTTTGTTTTGTTTTGTTTTG
CFTR	1080	rs58970500	GTGTGTGTGTATATATATATATATAT[A/T]TATATTTTTTTTTTCCTGAGCCAAA
CFTR	1080	rs58942292	AACCTATTAGCATGTCTGGCAGAAAA[-/A]TAGATACTTAATAAATTTCTTAAAT
CFTR	1080	rs58917054	GAGGCTTAGACAGTTTAAGTAACTCA[A/G]GCATGGTTACACAACTAGCTAGGGC
CFTR	1080	rs58837484	GTGTGAGTATTATGAGACCATATGTT[A/G]GGAGATTTTATTTGGTATTGAGGAT
CFTR	1080	rs58829491	GAAACCCCACCCCTTCTATAGTTTTC[C/T]CTTTAATATTTACAATGGAACCATT
CFTR	1080	rs58805195	CATATATATATAGTGTGTGTGTGTGT[A/G]TATATATATATATATATATATTTTT
GBA	2629	rs60866785	CGAGCGAGAGAGAGAGAGAGAGAGAG[-/AG]GAGCCGGCGCGAGAACTACGCATGC
GBA	2629	rs60239603	GGCAGGTAATATCTAGTACCTTACTT[A/T]TATTTCCTGAGCACATTCTACATTT
GBA	2629	rs56310840	GGCCAGGAATGGGAGTGCTTAGGTGC[A/G]GAGGTGGCACTGTTCCCGCAGCTGC
GBA	2629	rs41264927	TGAAAACTCCATCCCCTCAGGGTCAT[C/T]AGATGAAGAGAAGACCACAGGGGTT
GBA	2629	rs41264925	TGTAGGTAAGGGTCACATGTGGGAGA[C/G]GCAGCTGTGGGTAGGTCAGCCCTGT
GBA	2629	rs36024691	CCAAGAAGGCGCCATTACACTCCAGC[-/C]TGGGCGACAGGGCGAGACTCCCTCA
GBA	2629	rs36024092	TGCCACACCCAGCTAATTTGTGTGTG[-

Petitioner Illumina, Inc., Ex. 1061, p. 20
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	SNP ID	Mutation
			/G]TATGTGTGTGTATGTATGTGTGTGT
GBA	2629	rs35682967	GTTCCTCCAGTAATTTTTTTTTTTTT[-/T]GGTTTTGAGACAGAGTCTTGCCCTG
GBA	2629	rs35033592	ATCATGCCCAGATAATTTTTTTTTTT[-/T]GTATTTTAGTAGACACAGGGTTTCA
GBA	2629	rs34732744	CGAGCGAGAGAGAGAGAGAGAGAGAG[-/AG]GAGCCGGCGCGAGAACTACGCATGC
GBA	2629	rs34620635	CCTGTGAGGGGCACATTCCTTAGTAG[-/C]TAAGGAGTTGGGGGTGTGAAGATCC
GBA	2629	rs34302637	ACAGGCTACTGGCTGGGCCCAGGCAA[-/A]GGGGGCCTTGGCAGGAAAAGTTCCT
GBA	2629	rs33949225	GCGAGAGAGAGAGAGAGAGAGAGAGG[-/AG]AGCCGGCGCGAGAACTACGCATGCG
GBA	2629	rs28678003	AAGAAGAAAAATAAAAAGAAAGTGGG[C/T]CAGACCGAGAGAACAGGAAGCCTGA
GBA	2629	rs28559737	AAGGACAAAGGCAAAGAGACAAAGGC[G/T]CAACACTGGGGGTCCCCAGAGAGTG
GBA	2629	rs28373017	TACCTAGTCACTTCCTGCCTCCATGG[C/T]GCAAAAGGGGATGGGTGTGCCTCTT
GBA	2629	rs12752133	CTCTTCCGAGGTTCCACCCTGAACAC[C/T]TTCCTGCTCCCTCGTGGTGTAGAGT
GBA	2629	rs12747811	TTCTGACTGGCAACCAGCCCCACTCT[C/T]TGGGAGCCCTCAGGAATGAACTTGC
GBA	2629	rs12743554	gctcagcctcccaggctggagtgcag[A/T]ggcgcgatctcggctcaccgcaacc
GBA	2629	rs12041778	CATGAACCACATCAAATGAGATTTAG[C/T]GGGAGTGGCACACACAGTCATGACC
GBA	2629	rs12034326	AAGCAGCCCTGGGGAGTCGGGGCGGG[A/G]CCTGGATTGGAAAAGAGACGGTCAC
GBA	2629	rs11558184	CTCCAAGTTCTGGGAGCAGAGTGTGC[A/G]GCTAGGCTCCTGGGATCGAGGGATG
GBA	2629	rs11430678	GTTCCTCCAGTAAttttttttttttt[-/G/T]gttttgagacagagtcttgccctgt
GBA	2629	rs11264345	CTAGTACCTTACTTCCCTCAAGTTCA[A/T]TCATCTCACAGATATTTCCTGAGCA
GBA	2629	rs10908459	aattagccgtgcgtggtggcgggtgc[C/T]tgtaatcccacgtacttgggaggct
GBA	2629	rs10796940	CCATGGCCAGCCGGGGAGGGGACGGG[A/C]ACACACAGACCCACACAGAGACTCA
GBA	2629	rs10668496	agcgagagagagagagagagagagag[-/AG]gagCCGGCGCGAGAACTACGCATGC

Petitioner Illumina, Inc., Ex. 1061, p. 21
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	SNP ID	Mutation
GBA	2629	rs7416991	CGTAGCAGTTAGCAGATGATAGGCGG[C/G/T]GAAATCTTATTTCACAGG GCATTAA
GBA	2629	rs4024049	CTGGCCCTGGTGACAGTGGGGCTGTG[C/T]GTGGGGCCAGAGCCTTCTCA GAGGT
GBA	2629	rs4024048	CAGATACTGGCCCTGGTGACAGTGGG[A/G]CTGTGCGTGGGGCCAGAGC CTTCTC
GBA	2629	rs4024047	GACAGATACTGGCCCTGGTGACAGTG[G/T]GGCTGTGCGTGGGGCCAGA GCCTTC
GBA	2629	rs3841430	GGCTCctctctctctctctctctc[-/TC]gctcgctctctcgctctctcgctct
GBA	2629	rs3754485	GTTTCAGACCAGCCTGGCCAACATAG[C/T]GAAACCCCATCTCTACTAAA AATAA
GBA	2629	rs3205619	AGTGGGCGATTGGATGGAGCTGAGTA[C/T]GGGGCCCATCCAGGCTAAT CACACC
GBA	2629	rs2990227	CCGGGCTCCGTGAATGTTTGTCACAT[C/G]TCTGAAGAACGTATGAATTA CATAA
GBA	2629	rs2990226	GAATCCCAACCCCGACGCTCGTCGCC[C/G]GGCTCCGTGAATGTTTGTCA CATGT
GBA	2629	rs2990225	GCGAATCCCAACCCCGACGCTCGTCG[C/T]CGGGCTCCGTGAATGTTTGT CACAT
GBA	2629	rs2990224	TGGGCAGAAGTCAGGGTCCAAAGAAA[G/T]GGCAAAGAAAAGTGTcagtgg ctca
ASPA	443	rs63751297	TAAGAAAGACGTTTTTGATTTTTTTC[A/G]GACTTCTCTGGCTCCACTACC CTGC
ASPA	443	rs62071301	CTGATTCCTGGCCAGGAGCGGTGGCT[C/T]ACGCCTGTAATCCCAGCGCT TTGGG
ASPA	443	rs62071300	TAAAAATGCTGATTCCTGGCCAGGAG[C/T]GGTGGCTCACGCCTGTAATC CCAGC
ASPA	443	rs62071299	TTTAAAAATGCTGATTCCTGGCCAGG[A/C]GCGGTGGCTCACGCCTGTAA TCCCA
ASPA	443	rs62071297	CAAGACCTGTCAAAGATCTGAGAAAT[A/T]TTACCCGACTTACAAGCTA ACCATT
ASPA	443	rs61697033	ACTGTAATAAGTGCTGTAAAAGAAAT[A/G]CACAAAATAATATAGCAGA GGGTAT
ASPA	443	rs60743592	CTTGAGGTCAGGAGTTCAAGACCAGT[C/T]TGGGCAACATGGGGAAAAC CTTGTC
ASPA	443	rs60666840	AGGTTGCAGTGAGCCGAGATCATGCC[A/G]TTGCACTCCAGCCGGGGCA ACAAAA
ASPA	443	rs60147514	ACAAGTGTCTTGAAATTATCTGTGAT[C/T]TGCTATAGAGCAATACTTTT GTAAA

Petitioner Illumina, Inc., Ex. 1061, p. 22
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	SNP ID	Mutation
ASPA	443	rs59930743	GTGGGTATATGCAGCTCTATGCACTA[C/T]CTGCTCATTTATTTGGTAAATCTAA
ASPA	443	rs59690349	TGTGTGTGTGTGCGTGTGTGTGTGTG[-/TGTGTGTG]ATCATAAGAGTGGCTGCAGCAAACT
ASPA	443	rs59676360	AGTCTGGAGTGCAATGGTGCAATCTC[A/G]GCTCACTGCAGCCTCCACCTCCGGG
ASPA	443	rs59335404	CTCCTAATGGATATTTCCTAAATTTT[G/T]CTGAACAGAATTTAACTTGAGCTGG
ASPA	443	rs58879097	ATTTAAAAATGGATTTCTAGAAAAAC[A/G]ATCACATACTTGAATATTTTAGCAA
ASPA	443	rs58686774	CTATAAATGGGTAGCATGAGGGATTC[A/G]AGGAGGTGGCTGAAAGAAGCACGTA
ASPA	443	rs57511162	AAGAAACCAAGCATAGTAGAGTGTTA[A/G]AAAACCAAAGCAACTAAACAACTGT
ASPA	443	rs55859596	CGGGGCTCAGAACTTGTAACAGAAAA[A/T]TAAAATATACTCCACTCAAGGGAAT
ASPA	443	rs55742972	TACTACACTTCACGGATACTGTACTT[-/GTACTT]TTTTTCCAAATTGAAGGTTTTTGGC
ASPA	443	rs55640436	TTGTTTTTGTTTTTGTTTTTGTTTTT[-/GTTTTTGTTTTT]TGAGATGGAGTCTCGCTCTGTCGCC
ASPA	443	rs36225687	TTTGCCTTACTACACTTCACGGATAC[-/TGTACT]TGTACTTTTTTTCCAAATTGAAGGT
ASPA	443	rs36051310	GAGGTGGCTGAAAGAAGCACGTATCC[-/C]TGATGGCATGGTTGCGGGTTATATG
ASPA	443	rs36034906	GAGAAAAGCAGTTCCTGGAACACCCC[-/C]ACCCCTTAACCCCTTATCTCTGCTT
ASPA	443	rs36033666	TTACATATGTATACATGTGCCATGTT[-/T]GGTGTGCCGCACCCATTAACTCGTC
ASPA	443	rs35730123	CTTTTTCCAGATTTTTTTTTTTTTTT[-/T]GAGACAGAGTTTCACTCTTGTTGCC
ASPA	443	rs35629100	TTTGGAAATCTTAAGCTTTTATTTGG[-/G]TGTCACAGAGAAACAGGATCTGTAT
ASPA	443	rs35614631	TACTTTAAGTTTTAGGGTACATGTGC[-/A]CCATGTGCAGGTTTGTTACATATGT
ASPA	443	rs35225782	ATTCATGACCAGCCACATAAATGCAC[-/A]GTATTACTTCGCAAGCATGCCAATG
ASPA	443	rs35178659	GTGCACTAGAATTAGCTAAAGTGGGG[-/G]AAAAAAAGATGCATTTGATGGTCTA
ASPA	443	rs35095578	AACCTCCACCTCCCAGGTTCAAGAGA[-

Petitioner Illumina, Inc., Ex. 1061, p. 23
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	SNP ID	Mutation
			/A]TTCTCCTGCCTCAGCCTCCCAAGTA
ASPA	443	rs35002210	CCTCCCTGTGATCCGAAGTAGCAGAC[A/G]TACTTAACTTCCATGGTGGA TTGTT
ASPA	443	rs34744839	AAAACATTATTATATCTAGAAAAAAA[- /A]TGTATCTTAACCATTGTGGGAAGTG
ASPA	443	rs34680506	TTGAAGGTAAAATCATAGGGAGTTGG[- /G]AGCTGTCCTCTTGCGCTGAATCAGT
ASPA	443	rs34365618	ACTTGTGGCCTTTTTGGAGAGGTTAG[- /CA]ACTCTGAAAACTCTGTCCCTGGACC
ASPA	443	rs34275920	GAAGGAGAAAAAGAGAGGAAATAAGT[- /T]AAAATAATAAACACAATTAATAAAG
ASPA	443	rs34109510	TGTATACATGTGCCATGTTGGTGTGC[C/T]GCACCCATTAACTCGTCATTT AGCA
ASPA	443	rs34054576	TCACCTGTCACCTCCTATAGAACTTT[- /C]CCCTGACCCTCCTCTATAGCATTAA
ASPA	443	rs34015272	ATAAATGATCATCATTCACAGTAGGG[- /G]TTTTGTTTTGTTTTTTTTCTGGAAA
ASPA	443	rs34002091	ACAGACATATCTACAAACACACTTTT[- /T]CACATATTTGTGTAAGTCATTTATG
ASPA	443	rs28940574	AAAGACAACTAAACTAACGCTCAATG[A/C]AAAAAGTATTCGCTGCTGT TTACAT
ASPA	443	rs28940279	TACCGTGTACCCCGTGTTTGTGAATG[A/C]GGCCGCATATTACGAAAAGA AAGAA
ASPA	443	rs17850703	CAGGGCTGGAGGTAAAACCATTTATT[A/G]CTAACCCCAGAGCAGTGAA GAAGTG
ASPA	443	rs17222495	TTCTTCATTGCCTATTGAAGAGAGAG[C/T]GGAATGCTTTGGTTGCCAGA TATGG
ASPA	443	rs17175228	CACAAGATCTCATTACTCAGGAGCTG[C/T]CCAAGTGTCTAATGTACTTA GTTAA
ASPA	443	rs16953074	TTCTGTGTAACATTTCATTTAAGCAA[A/G]GGATTCGGCAAATCAAAAAT TGTCA
ASPA	443	rs16953070	TAAAACGTATTGAAGGTATTATTGAC[G/T]CTGTTGAAGCAAAGAGAAC AAAACA
HEXA	3073	rs62022858	ATCTGCTCTTCCAGTTGGATGACAAG[C/T]CTTGCTGTCTAACACCTGCT GCAGA
HEXA	3073	rs62022857	CCATTTTTTGTTGTATTTTTTTTTTC[C/T]TGAATACTTTTTATCGCAGTTG GTT
HEXA	3073	rs62017872	CCCTGTCTCTAAAAGAAAAAAAAAAA[A/G]AAAAAAAAAAGAAAACAA AACCCAA

Petitioner Illumina, Inc., Ex. 1061, p. 24
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	SNP ID	Mutation
HEXA	3073	rs62017871	AGTGGCTCCAAAAAGGTCATGGAACC[C/T]CTTGAGGATGATGCAAATT GACTCT
HEXA	3073	rs61662730	TAAAGTTACTTTTCTTTTATTGACTT[C/T]CCCTTATTTTTTAACCTTATGC TTT
HEXA	3073	rs61329913	CAGAGTTAAAAAAAAAAAAAAAAAAA[-/A]GGAAGTAGCAGCAACAGCTTGGAAA
HEXA	3073	rs60920713	GTTGCCCAGGGTTGAGTGCAGAGGCA[C/T]ATCTTGGCTCACAGCAACCT CTGCC
HEXA	3073	rs60783213	AAGGCTTTTTTTTTTTTTTTTTTT[-/TTTT]GAGACAGAGTCTTGCTGTGTCACCC
HEXA	3073	rs60644867	GCCTACATTCTGCAAAGAGGAGGGAA[C/G]ATTCACAGCTCCATACTTG AACCCT
HEXA	3073	rs60288568	CCAAAGGAGAATAGCTCTAGGGGAGG[C/G]AGGTGGATGAGTATGCATG GGGGAG
HEXA	3073	rs59888548	GACTCCATCTCAAAAAAAAAAAAAAA[-/A]TGCAGTCTAATGGCAGAATTAGACT
HEXA	3073	rs59733856	TTATTTATTTATTTATTTATTTTTGA[A/G]ACAGGGTCTCTGTTGTCCAGG CTGG
HEXA	3073	rs59427837	TTTTGAGGCAGGGTCTCACTCTGTTG[C/T]CCAGGGTTGAGTGCAGAGGC ACATC
HEXA	3073	rs59171976	CGCCTTGCGAAGGCCCCACAGCTTGC[C/T]TGTGACAAACGTTCATAGGC AAATG
HEXA	3073	rs58706602	GGAGGTCTGTACAAAGCACCACCTAC[C/T]TCATGGGTCAGTTTCCACAG CAGAA
HEXA	3073	rs58696963	GAATCTTATAATTCACTGTGTACCTC[-/CCTC]TGTTTCATATTTTCGCAATTGAACT
HEXA	3073	rs58610850	AACATAGTATCTAATATAGCTTTACA[C/T]CCAAAGCCAAAATATGAATA CACTG
HEXA	3073	rs58016062	TTGTTTTGTTTTGTTTGGGGGGGGGG[-/G]TTGTTTTTCTGAGAGGGAGTCTTGC
HEXA	3073	rs57733983	CATACCAAAGGGCAGCTGGAGGGATA[C/T]AGACGGAAGTCATGTGGAG AGTGAA
HEXA	3073	rs57476645	CAGGTGTGAGCCACCACGACCACCAA[A/T]TTAGCTCTTTTTACTCCTTC CCTTC
HEXA	3073	rs56870003	AGTGGTAGCTGATTTTGCTTCTGGAT[A/C]CTTTGCCACCTTCCCACTCTT TAAT
HEXA	3073	rs56338339	AAAGACCTGTTTCTTAAAAAAAAAAA[-/AGAAAAAAAAAAA]GAAAGAAAAGAAAAGAAAAAAACAG
HEXA	3073	rs55995352	TAAAAAATCTTTCAATGAGGAGATGT[C/T]CCCAGAGCAAGACAGCTGT

Petitioner Illumina, Inc., Ex. 1061, p. 25
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	SNP ID	Mutation
			AGGATG
HEXA	3073	rs55860138	AAAAGAAAAAAAAAAAAAAAAAAAA[-/A]GAAAACAAAACCCAAACCCATAAAG
HEXA	3073	rs55743646	CCTGTCTCTAAAAGAAAAAAAAAAAA[A/G]AAAAAAAAAGAAAACAAAACCCAAA
HEXA	3073	rs55665666	GTTATCATAGAAAAATATCACACTCT[-/GT]CTGTATCCCCACTTCCAGAAACTGT
HEXA	3073	rs36106892	CAGGAGCTCATAGAATTACATACAAT[-/C]TTTTTTTTTTTTTTTTGAGACAGCG
HEXA	3073	rs36091525	TTGAGAATCTTATAATTCACTGTGTA[-/CCTC]CCTCTGTTTCATATTTTCGCAATTG
HEXA	3073	rs35949555	CCACTACCACAGTGCCTAGAGAACAA[C/T]ATGTGTTTAATAATATTTAAATAAT
HEXA	3073	rs35827424	CCCTGTCTCTAAAAGAAAAAAAAAAA[-/A]AAAAAAAAAAGAAAACAAAACCCAA
HEXA	3073	rs35729578	CCATTATATCATTCATTTCCCACTCA[-/T]TTTCTTCATTCCAACCAAGATATAT
HEXA	3073	rs35649102	TCCGTCTCAAAAAAAAAAAAAAAAAG[-/A]GAAAGGAATTATTCTCATGTATACA
HEXA	3073	rs35118677	CTGGGGCAGTTAAAAAGAAAAACAAA[-/C]CCCTGGTCCCTGCCCTTGAGGAGAT
HEXA	3073	rs35005352	CTCCAGGGTCCCATTCCAGGACCACA[-/C]GCCTGCTACCTCTGCAGCTCACTCA
HEXA	3073	rs34736306	GGATTGACATATACCAGTTAGACGGA[-/T]TTTTTTTTTCCATAAACCAGGCTCA
HEXA	3073	rs34607939	ACAAATAATTACTACATATCTACAAC[A/G]TTCCAGATACAGAAGAAATGGCCAA
HEXA	3073	rs34496117	TAAACACACTTGAAACATCATATAAA[-/ATG]ATATTACTACAAGACTTAACCGTAA
HEXA	3073	rs34300017	ACACAGGTAATCCATGTTTATTATAG[-/A]AAAATGCCACATTACTCTTTATTGA
HEXA	3073	rs34206496	AGTTATCATAGAAAAATATCACACTC[-/TG]TCTGTATCCCCACTTCCAGAAACTG
HEXA	3073	rs34110830	AATGAACTTACAGGAAGGTAATATAT[-/G]GGAAATAAACATCTTATTGAATTTA
HEXA	3073	rs34093438	GGACCCCTGAAAGGCACAAGACACCC[-/T]TTCAGGTTCACACTTCCTGAAAGCT
HEXA	3073	rs34085965	CCACCAATCACCAGAGCCTTCTGCTC[A/G]GGGGTACCTGAGGGAAAACAAGCAA

Petitioner Illumina, Inc., Ex. 1061, p. 26
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Gene name	Gene ID	SNP ID	Mutation
HEXA	3073	rs34004907	AAAGACTGAAAAAACATTCATAACTA[-/T]TTTTCTTGTTATCCTCGGAAATGTC
HEXA	3073	rs28942072	TATCTTCATCTTGGAGGAGATGAGGT[C/T]GATTTCACCTGCTGGAAGTCCAACC
HEXA	3073	rs28942071	TTGCCTATGAACGTTTGTCACACTTC[C/T]GCTGTGAGTTGCTGAGGCGAGGTGT
HEXA	3073	rs28941771	GCTTGCTGTTGGATACATCTCGCCAT[C/T]ACCTGCCACTCTCTAGCATCCTGGA
HEXA	3073	rs28941770	CCGGGGCTTGCTGTTGGATACATCTC[G/T]CCATTACCTGCCACTCTCTAGCATC

Examples

Example 1: Use of source sequences to detect and correct bias in a MIP-capture reaction of a set of exon targets

The first step in performing the detection/correction is to determine how many source sequences are necessary for the given sample. In this example, 1000 genomic targets corresponding to 1000 exons were captured. Since the source sequence is part of the probe, it will measure/report biases that occur from the earliest protocol steps. Also, being located in the backbone, the source sequence can easily be sequenced from a separate priming site, and therefore not impact the total achievable read-length for the target sequence. MIP probes are synthesized using standard column-based oligonucleotide synthesis by any number of vendors (e.g. IDT), and source sequences are introduced as 'degenerate' positions in the backbone. Each degenerate position increases the total number of source sequences synthesized by a factor of 4, so a 10 nt degenerate region implies a source sequence complexity of ~1e6 species.

Hybridize capture probes to human genomic sample

Assemble hybridization reaction:

- 1.0 ul capture probe mix (~2.5 pmol)
- 2.0 ul 10X Ampligase buffer (Epicentre)
- 6.0 ul 500 ng/ul human genomic DNA (~16.7 fmol)

Petitioner Illumina, Inc., Ex. 1061, p. 27
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

- 11 ul dH2O

In a thermal cycler, heat reaction to 95C for 5 min to denature genomic DNA, then cool to 60C. Allow to incubate at 60C for 40 hours.

Convert hybridized probes into covalently-closed circular products containing subtargets

Prepare fill-in / ligation reaction mixture:

- 0.25 ul 2 mM dNTP mix (Invitrogen)
- 2.5 ul 10X Ampligase buffer (Epicentre)
- 5.0 ul 5 U/ul Taq Stoffel fragment (Applied Biosystems)
- 12.5 ul 5 U/ul Ampligase (Epicentre)
- 4.75 ul dH2O

Add 1.0 ul of this mix to the hybridized probe reaction, and incubate at 60C for 10 hours.

Purify circularized probe / subtarget products from un-reacted probes and genomic DNA

Prepare exonuclease reaction mixture:

- 21 ul fill-in / ligation reaction product
- 2.0 ul 10X exonuclease I buffer (New England Biolabs)
- 2.0 ul 20 U/ul exonuclease I (New England Biolabs)
- 2.0 ul 100 U/ul exonuclease III (New England Biolabs)

Incubate at 37C for 60 min, then heat-inactivate by incubating at 80C for 15 min. Immediately cool to 4C for storage.

Amplify circular material by PCR using primers specific to the 'constant region' of the probes

Prepare PCR mixture:

- 5.0 ul 10X Accuprime reaction buffer (Invitrogen)
- 1.5 ul 10 uM CP-2-FA (5'-GCACGATCCGACGGTAGTGT-3')
- 1.5 ul 10 uM CP-2-RA (5'-CCGTAATCGGGAAGCTGAAG-3')

Petitioner Illumina, Inc., Ex. 1061, p. 28
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

- 0.4 ul 25 mM dNTP mix (Invitrogen)
- 2.0 ul heat-inactivated exonuclease reaction mix
- 1.5 ul 10X SybrGreen (Invitrogen)
- 0.4 ul 2.5 U/ul Accuprime Pfx polymerase (Invitrogen)
- 37.7 ul dH2O

Thermal cycle in real-time thermal cycler according to the following protocol, but stop cycling before amplification yield plateaus (generally 8-12 cycles):

1. 95C for 5 min
2. 95C for 30 sec
3. 58C for 60 sec
4. 72C for 60 sec
5. goto 2, N more times

Prepare a shotgun next-generation sequencing library for analysis

- Purify desired amplicon population from non-specific amplification products by gel extraction.
- Concatemerize amplicons into high-molecular weight products suitable for shearing
- Mechanically shear, using either a nebulizer, BioRuptor, Hydroshear, Covaris, or similar instrument. DNA should be sheared into fragments several hundred basepairs in lenth.
- Ligate adapters required for amplification by the sequencing platform used. If necessary, purify ligated product from unligated product and adapters.

Perform sequencing of library according to manufacturer's directions (e.g. Illumina,ABI, etc), reading both the target sequence and the source sequence.

Analyze data by correcting for any biases detected by quantitation of source sequence abundance

Construct a table of target:source abundances from the read data, e.g.:

Petitioner Illumina, Inc., Ex. 1061, p. 29
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Target ID	Source Sequence ID	Count
1	3547	1
2	4762	1
1	9637	1
1	1078	5
3	4762	1
1	2984	1

All 'count' entries should be '1', since any particular target:source mapping will not occur more than once by chance, and therefore will only be observed if bias was present somewhere in the sample preparation process. For any target:source combination observed more than once, all such reads are 'collapsed' into a single read before consensus basecalls are determined. This will cancel the effect of bias on consensus basecall accuracy. Figure 2 depicts a method for making diploid genotype calls in which repeat target:source combination are collapsed.

Example 2: Source sequence design for MIP capture reactions

For a set of targets, the number of source sequences necessary to be confident (within some statistical bounds) that a certain source sequence will not be observed more than once by chance in combination with a certain target sequence was determined. The total number of unique source sequences for a certain source sequence length is determined as $4^{(\text{Length in nucleotides of the source sequence})}$. For a molecular inversion probe capture reaction that uses MIP probes having source sequences, the probability of performing the capture reaction and capturing one or more copies of a target sequence having the same source sequence is calculated as: $p = 1 - [N! / (N-M)!] / [N^M]$, wherein N is the total number of possible unique source sequences and M is the number of target sequence copies in the capture reaction. Thus, by varying the source sequence length it is possible to perform a MIP capture reaction in which the probability of capturing one or more copies of a target sequence having the same source sequence is set at a predetermined probability value.

For example, for a source sequence of 15 nucleotides in length, there are 1,073,741,824 possible source sequences. A MIP capture reaction in which MIP probes, each

Petitioner Illumina, Inc., Ex. 1061, p. 30
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

having a source sequence of 15 nucleotides, are combined with 10000 target sequence copies (e.g., genome equivalents), the probability of capturing one or more copies of a target sequence having the same source sequence is 0.05. In this example, the MIP reaction will produce very few (usually 0, but occasionally 1 or more) targets where multiple copies are tagged with the same source sequence. Figure 3 depicts results of a simulation for 100000 capture reactions having 15 nucleotide source sequences and 10000 target sequences.

Example 3: Assessment of the probability for obtaining enough sequencing reads to make accurate base-calls at multiple independent loci, as a function of sequencing coverage.

Monte Carlo simulations were performed to determine sequencing coverage requirements. The simulations assume 10000 genomic copies of a given locus (target) half mom alleles and half dad alleles. The simulations further assume 1% efficiency of capture for the MIP reaction. The simulation samples from a capture mix 100 times without replacement to create a set of 100 capture products. The simulation then samples from the set of 100 capture products with replacement (assuming unbiased amplification) to generate 'reads' from either mom or dad. The number of reads sampled depends on the coverage. The number of independent reads from both mom and dad necessary to make a high-quality base-call (assumed to be 10 or 20 reads) were then determined. The process was repeated 1000 times for each coverage level, and the fraction of times that enough reads from both parents were successfully obtained was determined. This fraction was raised to the power 1000, assuming we have 1000 independent loci that must obtain successful base-calls, plotted (See Figure 4). Result show that roughly 50x coverage is required to capture each allele >= 10x with > 0.95 probability.

It will be clear that the methods may be practiced other than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present disclosure are possible in light of the above teachings and, therefore, are within the scope of the claims. Preferred features of each aspect of the disclosure are as for each of the other aspects mutatis mutandis. The documents including patents, patent applications, journal articles, or other disclosures mentioned herein are hereby incorporated by reference in their entirety. In the event of conflict, the disclosure of present application controls, other than in the event of clear error.

Petitioner Illumina, Inc., Ex. 1061, p. 31
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

What is claimed is:

Petitioner Illumina, Inc., Ex. 1061, p. 32
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Claims

1. A method of genotyping a plurality of genetic loci, the method comprising:

obtaining a plurality of molecular inversion probe capture products, each comprising a molecular inversion probe sequence, a target sequence and a source sequence, wherein the target sequence is a captured genomic locus,

amplifying the plurality of molecular inversion probe capture products, and

genotyping target sequences by determining, for each target sequence, the sequence of at least a threshold number of unique combinations of target sequence and source sequence of molecular inversion probe capture products in the amplified plurality.

2. The method of claim 1, wherein the obtaining comprises capturing target sequences from a genomic sample with a plurality of molecular inversion probes, each comprising a source sequence.

3. The method of claim 1 or 2, wherein the capturing is performed under conditions wherein the likelihood of obtaining two or more molecular inversion probe capture products with identical combinations of target and source sequences is equal to or less than a predetermined value, optionally wherein the predetermined value is about 0.05.

4. The method of any one of claims 1-3, wherein the threshold number for a specific target sequence is selected based on a desired statistical confidence for the genotype of the specific target sequence

5. The method of any one of claims 1-3, further comprising determining a statistical confidence for the genotype of a specific target sequence based on the number of unique combinations comprising the specific target sequence.

6. A method of analyzing a plurality of genetic loci, the method comprising:

obtaining a plurality of molecular inversion probe capture products, each comprising a molecular inversion probe sequence, a target sequence and a source sequence,

amplifying the plurality of molecular inversion probe capture products,

Petitioner Illumina, Inc., Ex. 1061, p. 33
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

determining numbers of occurrence of combinations of target sequence and source sequence of molecular inversion probe capture products in the amplified plurality, and

if the number of occurrence of a specific combination of target sequence and source sequence exceeds a predetermined value, detecting bias in the amplification of the molecular inversion probe comprising the specific combination.

7. The method of claim 6, further comprising genotyping target sequences in the plurality, wherein the genotyping comprises correcting for bias, if detected.

8. The method of any of claims 1-7, wherein the target sequence is a sequence of a gene selected from Table 1.

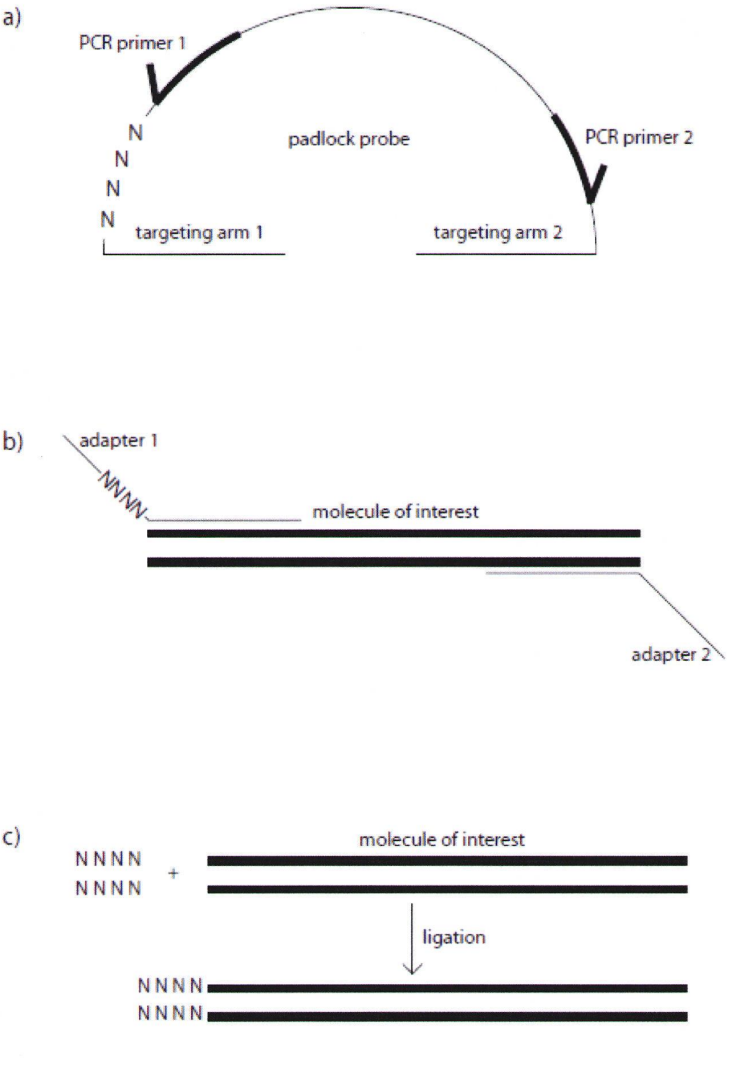
Petitioner Illumina, Inc., Ex. 1061, p. 34
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

Abstract

The invention relates to preparative steps in DNA sequencing-related technologies that reduce bias and increase the reliability and accuracy of downstream quantitative applications.

Petitioner Illumina, Inc., Ex. 1061, p. 35
Illumina Inc. v. Molecular Loop Biosciences, Inc.
IPR2024-01419

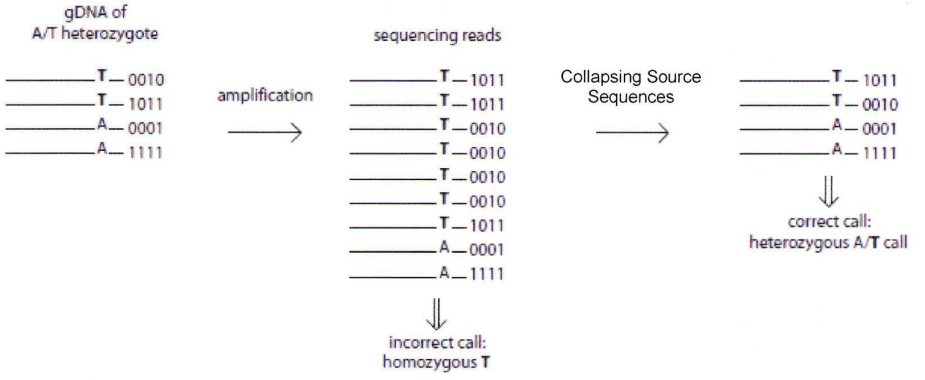
FIGURE 1



1661347.1

ML-IPR-000096

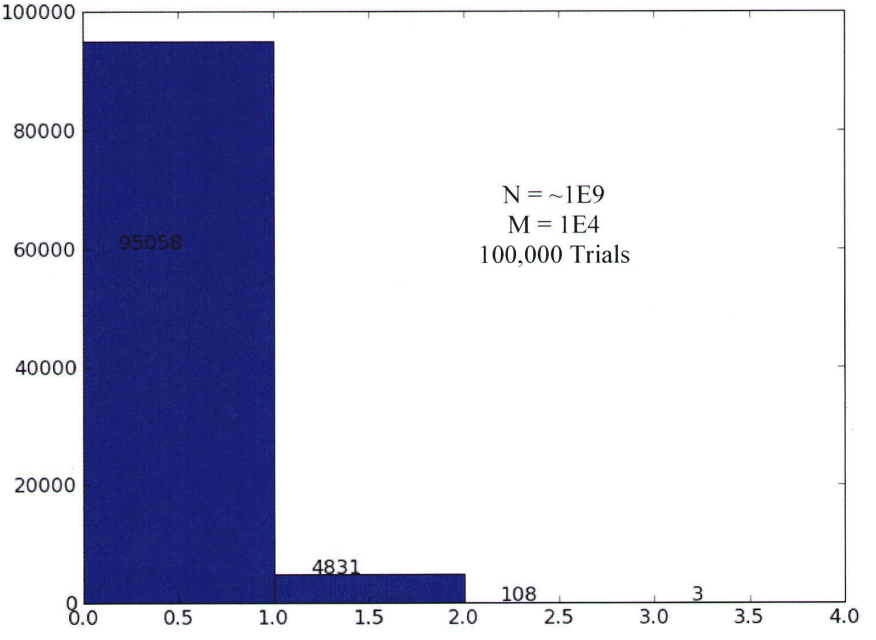
FIGURE 2



1661347.1

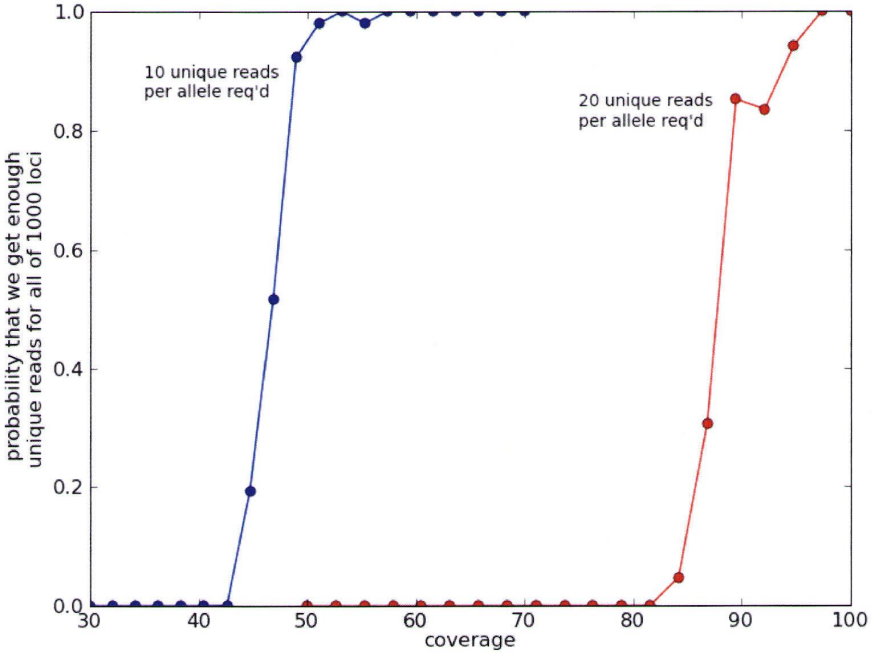
ML-IPR-000097

FIGURE 3



100,000 Trials

FIGURE 4


1.0
0.8
0.6
0.4
0.2
0.0
probability that we get enough unique reads for all of 1000 loci
10 unique reads per allele req'd
20 unique reads per allele req'd
30
40
50
60
70
80
90
100
coverage