Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US2010/001293

International filing date: 30 April 2010 (30.04.2010)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 61/174,470
Filing date: 30 April 2009 (30.04.2009)

Date of receipt at the International Bureau: 20 May 2010 (20.05.2010)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

260605

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

May 19, 2010

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *61/174,470*
FILING DATE: *April 30, 2009*
RELATED PCT APPLICATION NUMBER: *PCT/US10/01293*

THE COUNTRY CODE AND NUMBER OF YOUR PRIORITY APPLICATION, TO BE USED FOR FILING ABROAD UNDER THE PARIS CONVENTION, IS *US61/174,470*

Certified by

David W. Kappos

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

PTO/SB/16 (12-08)
Approved for use through 06/30/2010. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET – Page 1 of 2

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

VIA EFS-web e-Filed: 4/30/2009

INVENTOR(S)		
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Greg	Porreca	Ocean City, New Jersey 08226
Uri	Laserson	Boston, Massachusetts 02215
Jin Billy	Li	Medford, Massachusetts 02155

Additional inventors are being named on the _____ separately numbered sheets attached hereto.

TITLE OF THE INVENTION (500 characters max):

METHOD FOR REDUCING REPRESENTATIONAL BIAS IN MULTIPLEX AMPLIFICATION OF NUCLEIC ACIDS

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

[X] The address corresponding to Customer Number: 23628

OR

[ ] Firm or Individual Name

Address

City	State	Zip
Country	Telephone	Email

ENCLOSED APPLICATION PARTS (*check all that apply*)

[X] Application Data Sheet. See 37 CFR 1.76 [ ] CD(s), Number of CDs _____

[X] Drawing(s) *Number of Sheets* 3 [ ] Other (specify) _____

[X] Specification (e.g. description of the invention) *Number of Pages* 28

Fees Due: Filing Fee of $220 ($110 for small entity). If the specification and drawings exceed 100 sheets of paper, an application size fee is also due, which is $270 ($135 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

METHOD OF PAYMENT OF THE FILING FEE AND APPLICATION SIZE FEE FOR THIS PROVISIONAL APPLICATION FOR PATENT

[X] Applicant claims small entity status. See 37 CFR 1.27.

[ ] A check or money order made payable to the *Director of the United States Patent and Trademark Office* is enclosed to cover the filing fee and application size fee (if applicable).

[X] Payment by credit card. ~~Form PTO-2038 is attached.~~

110.00

TOTAL FEE AMOUNT ($)

[X] The Director is hereby authorized to charge the filing fee and application size fee (if applicable) or credit any overpayment to Deposit Account Number: 23/2825 .

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

Petitioner Illumina, Inc., EX1006, p. 3

PTO/SB/16 (12-08)
Approved for use through 06/30/2010. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

[X] No.

[ ] Yes, the name of the U.S. Government agency and the Government contract number are:

_____

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

SIGNATURE /Patrick R.H. Waller/ Date April 30, 2009

TYPED or PRINTED NAME Patrick R.H. Waller REGISTRATION NO. *(if appropriate)* 41,418

TELEPHONE 617.646.8000 Docket Number: G0814.70000US00

Petitioner Illumina, Inc., EX1006, p. 4

VIA EFS-web
e-Filed: 04/30/2009

Application Data Sheet

Application Information

Application Type::	Provisional
Subject Matter::	Utility
Suggested Group Art Unit::	N/A
CD-ROM or CD-R::	None
Sequence submission::	None
Computer Readable Form (CRF)::	No
Title::	METHOD FOR REDUCING REPRESENTATIONAL BIAS IN MULTIPLEX AMPLIFICATION OF NUCLEIC ACIDS
Attorney Docket Number::	G0814.70000US00
Request for Early Publication::	No
Request for Non-Publication::	No
Small Entity::	Yes
Petition included::	No
Secrecy Order in Parent Appl.::	No

Applicant Information

Applicant Authority Type::	Inventor
Primary Citizenship Country::	United States
Status::	Full Capacity
Given Name::	Greg
Family Name::	Porreca
City of Residence::	Ocean City
State or Province of Residence::	NJ
Country of Residence::	US
Street of Mailing Address::	911 Gardens Parkway

City of Mailing Address::	Ocean City
State or Province of Mailing Address::	NJ
Postal or Zip Code of Mailing Address::	08226

Applicant Authority Type::	Inventor
Primary Citizenship Country::	United States
Status::	Full Capacity
Given Name::	Uri
Family Name::	Laserson
City of Residence::	Boston
State or Province of Residence::	MA
Country of Residence::	United States
Street of Mailing Address::	645 Beacon Street, #215
City of Mailing Address::	Boston
State or Province of Mailing Address::	MA
Postal or Zip Code of Mailing Address::	02215

Applicant Authority Type::	Inventor
Primary Citizenship Country::	People's Republic of China
Status::	Full Capacity
Given Name::	Jin
Middle Name::	Billy
Family Name::	Li
City of Residence::	Medford
State or Province of Residence::	Massachusetts
Country of Residence::	United States
Street of Mailing Address::	100 Station Landing, Unit 204

City of Mailing Address::	Medford
State or Province of Mailing Address::	Massachusetts
Postal or Zip Code of Mailing Address::	02155

Correspondence Information

Correspondence Customer Number::	23628

Representative Information

Representative Customer Number::	23628

Domestic Priority Information

Foreign Priority Information

Assignee Information

Assignee Name::	Good Start Genetics, Inc.
Street of Mailing Address::	312 Marlborough St., #4
City of Mailing Address::	Boston
State or Province of Mailing Address::	Massachusetts
Country of Mailing Address::	United States
Postal or Zip Code of Mailing Address::	02116

METHOD FOR REDUCING REPRESENTATIONAL BIAS IN MULTIPLEX AMPLIFICATION OF NUCLEIC ACIDS

Field of Invention

The invention relates to methods that reduce bias and increase reproducibility in multiplex detection of genetic loci, e.g., for diagnostic purposes.

Background of the Invention

Molecular inversion probe technology is used to detect or amplify particular nucleic acid sequences in potentially complex mixtures. Use of molecular inversion probes has been demonstrated for detection of single nucleotide polymorphisms (Hardenbol et al. 2005 Genome Res 15:269-75) and for preparative amplification of large sets of exons (Porreca et al. 2007 Nat Methods 4:931-6, Krishnakumar et al. 2008 Proc Natl Acad Sci USA 105:9296-301). One of the main benefits of the method is in its capacity for a high degree of multiplexing, because generally thousands of targets may be captured in a single reaction containing thousands of probes. However, challenges associated with, for example, amplification efficiency (See, e.g., Turner EH, et al., Nat Methods. 2009 Apr 6:1-2.) have limited the practical utility of the method in research and diagnostic settings.

Summary of the Invention

Aspects of the disclosure are based, in part, on the discovery of effective methods for overcoming challenges associated with systematic errors (bias) in multiplex genomic capture and sequencing methods, namely high variability in target nucleic acid representation and unequal sampling of heterozygous alleles in pools of captured target nucleic acids. Accordingly, in some embodiments, the disclosure provides methods that reduce variability in the detection of target nucleic acids in multiplex capture methods. In other embodiments, methods improve allelic representation in a capture pool and, thus, improve variant detection outcomes. In certain embodiments, the disclosure provides preparative methods for capturing target nucleic acids (e.g., genetic loci) that involve the use of different sets of multiple probes (e.g., molecular inversion probes MIPs) that capture overlapping regions of a target nucleic acid to achieve a more uniform representation of the target nucleic acids in a capture pool compared with methods of the

Petitioner Illumina, Inc., EX1006, p. 8

prior art. In other embodiments, methods reduce bias, or the risk of bias, associated with large scale parallel capture of genetic loci, e.g., for diagnostic purposes. In other embodiments, methods are provided for increasing reproducibility (e.g., by reducing the effect of polymorphisms on target nucleic acid capture) in the detection of a plurality of genetic loci in parallel. In further embodiments, methods are provided for reducing the effect of probe synthesis and/or probe amplification variability on the analysis of a plurality of genetic loci in parallel.

According to some aspects, the disclosure relates to multiplex diagnostic methods. In some embodiments, multiplex diagnostic methods comprise capturing a plurality of genetic loci in parallel (e.g., a genetic locus of Table 1). 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 provides methods for analyzing multiple genetic loci (e.g., a plurality of target nucleic acids selected from Table 1) from a patient sample, such as a blood, pre-implantation embryo, chorionic villus or amniotic fluid sample.

According to some aspects, methods of analyzing a plurality of genetic loci are provided. In some embodiments, the methods comprise contacting each of a plurality of target nucleic acids with a probe set, wherein each probe set comprises a plurality of different probes, each probe having a central region flanked by a 5' region and a 3' region that are complementary to nucleic acids flanking the same strand of one of a plurality of subregions of the target nucleic acid, wherein the subregions of the target nucleic acid are different, and wherein each subregion overlaps with at least one other subregion, isolating a plurality of nucleic acids each having a nucleic acid sequence of a different subregion for each of the plurality of target nucleic acids, and analyzing the isolated nucleic acids.

In other embodiments, methods comprise contacting each of a plurality of target nucleic acids with a probe set, wherein each probe set comprises a plurality of different probes, each probe having a central region flanked by a 5' region and a 3' region that are

Petitioner Illumina, Inc., EX1006, p. 9

complementary to nucleic acids flanking the same strand of one of a plurality of subregions of the target nucleic acid, wherein the subregions of the target nucleic acid are different, and wherein a portion of the 5' region and a portion of the 3' region of a probe have, respectively, the sequence of the 5' region and the sequence of the 3' region of a different probe, isolating a plurality of nucleic acids each having a nucleic acid sequence of a different subregion for each of the plurality of target nucleic acids, and analyzing the isolated nucleic acids.

In some embodiments of the methods disclosed herein the target nucleic acids are selected from Table 1.

According to some aspects of the methods disclosed herein, the analyzing comprises genotyping a polymorphism of a target nucleic acid. In some embodiments, the genotyping may comprise determining that alleles of the target nucleic acid are heterozygous or homozygous. In further embodiments, the genotyping may comprise determining the sequence of a polymorphism and comparing that sequence to a control sequence that is indicative of a disease risk. In some embodiments, the polymorphism is selected from Table 2.

In some embodiments, a diagnosis, prognosis, or disease risk assessment is provided to a subject based on a genotype determined for that subject at one or more genetic loci (e.g., based on the analysis of a biological sample obtained from that subject). In some embodiments, an assessment is provided to a couple, based on their respective genotypes at one or more genetic loci, of the risk of their having one or more children having a genotype associated with a disease or condition (e.g., a homozygous or heterozygous genotype associated with a disease or condition). In some embodiments, a subject or a couple may seek genetic or reproductive counseling in connection with a genotype determined according to embodiments of the invention.

BRIEF DESCRIPTION OF DRAWINGS

Fig. 1 depicts a tiled probe layout, showing twelve captured sub-targets (also referred to herein as subregions of a target nucleic acid) tiled across a single target. Each position in the target is covered by three sub-targets such that MIP capture performance per basepair is averaged across three probes.

Fig. 2 depicts a staggered probe layout, showing the targets captured by a set of three MIPs. Each MIP captures the full target, shown in black, plus (in some cases)

Petitioner Illumina, Inc., EX1006, p. 10

additional extra-target sequence, shown in gray, such that the targeting arms of each MIP fall on different sequence. Each position in the target is covered by three sub-targets such that MIP performance per basepair is averaged across three probes.

Fig. 3 depicts an alternating staggered probe layout, showing the targets captured by a set of three MIPs. Each MIP captures the full target, shown in black, plus (in some cases) additional extra-target sequence, shown in gray, such that the targeting arms of each MIP fall on different sequence. Each position in the target is covered by three sub-targets such that MIP performance per basepair is averaged across three probes.

DETAILED DESCRIPTION

Aspects of the disclosure are based, in part, on the discovery of effective methods for overcoming challenges associated with systematic errors (bias) in multiplex genomic capture and sequencing methods, namely high variability in target nucleic acid representation and unequal sampling of heterozygous alleles in pools of captured target nucleic acids (e.g., isolated from a biological sample). Accordingly, in some embodiments, the disclosure provides methods that reduce variability in the detection of target nucleic acids in multiplex capture methods. In other embodiments, methods improve allelic representation in a capture pool and, thus, improve variant detection outcomes. In certain embodiments, the disclosure provides preparative methods for capturing target nucleic acids (e.g., genetic loci) that involve the use of different sets of multiple probes (e.g., molecular inversion probes MIPs) that capture overlapping regions of a target nucleic acid to achieve a more uniform representation of the target nucleic acids in a capture pool compared with methods of the prior art. In other embodiments, methods reduce bias, or the risk of bias, associated with large scale parallel capture of genetic loci, e.g., for diagnostic purposes. In other embodiments, methods are provided for increasing reproducibility (e.g., by reducing the effect of polymorphisms on target nucleic acid capture) in the detection of a plurality of genetic loci in parallel. In further embodiments, methods are provided for reducing the effect of probe synthesis and/or probe amplification variability on the analysis of a plurality of genetic loci in parallel.

According to some aspects, the disclosure relates to multiplex diagnostic methods. In some embodiments, multiplex diagnostic methods comprise capturing a plurality of genetic loci in parallel (e.g., a genetic locus of Table 1). In some

Petitioner Illumina, Inc., EX1006, p. 11

embodiments, 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 provides methods for analyzing multiple genetic loci (e.g., a plurality of target nucleic acids selected from Table 1) from a patient sample, such as a blood, pre-implantation embryo, chorionic villus or amniotic fluid sample. A patient or subject may be a human. However, aspects of the invention are not limited to humans and may be applied to other species (e.g., mammals, birds, reptiles, other vertebrates or invertebrates) as aspects of the invention are not limited in this respect. A subject or patient may be male or female. In some embodiments, in connection with reproductive genetic counseling, samples from a male and female member of a couple may be analyzed. In some embodiments, for example, in connection with an animal breeding program, samples from a plurality of male and female subjects may be analyzed to determine compatible or optimal breeding partners or strategies for particular traits or to avoid one or more diseases or conditions.

In some aspects, the disclosure provides probe sets that comprise a plurality of different probes. As used herein, a 'probe' is a nucleic acid having a central region flanked by a 5' region and a 3' region that are complementary to nucleic acids flanking the same strand of a target nucleic acid or subregion thereof. An exemplary probe is a molecular inversion probe (MIP). A 'target nucleic acid' may be a genetic locus. Exemplary genetic loci are disclosed herein in Table 1 (RefSeqGene Column).

While probes have been typically designed to meet certain constraints (e.g. melting temperature, G/C content, etc.) known to partially affect capture/amplification efficiency (Ball et al (2009) Nat Biotech 27:361-8 AND Deng et al (2009) Nat Biotech 27:353-60), a set of constraints which is sufficient to ensure either largely uniform or highly reproducible capture/amplification efficiency has not previously been achieved. As disclosed herein, uniformity and reproducibility can be increased by designing multiple probes per target, such that each base in the target is captured by more than one probe. In some embodiments, the disclosure provides multiple MIPs per target to be captured, where each MIP in a set designed for a given target nucleic acid has a central

Petitioner Illumina, Inc., EX1006, p. 12

region and a 5' region and 3' region ('targeting arms') which hybridize to (at least partially) different nucleic acids in the target nucleic acid (immediately flanking a subregion of the target nucleic acid). Thus, differences in efficiency between different targeting arms and fill-in sequences may be averaged across multiple MIPs for a single target, which results in more uniform and reproducible capture efficiency.

In some embodiments, the methods involve designing a single probe for each target (a target can be as small as a single base or as large as a kilobase or more of contiguous sequence).

It may be preferable, in some cases, to design probes to capture molecules (e.g., target nucleic acids or subregions thereof) having lengths in the range of 1 - 200 bp (as used herein, a bp refers to a base pair on a double-stranded nucleic acid – however, where lengths are indicated in bps, it should be appreciated that single-stranded nucleic acids having the same number of bases, as opposed to base pairs, in length also are contemplated by the invention). However, probe design is not so limited. For example, probes can be designed to capture targets having lengths in the range of up to 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 1000, or more bps, in some cases.

It is to be appreciated that the length of a capture molecule (e.g., a target nucleic acid or subregion thereof) is selected based upon multiple considerations. For example, where analysis of a target involves sequencing, e.g., with a next-generation sequencer, the target length should typically match the sequencing read-length so that shotgun library construction is not necessary. However, it should be appreciated that captured nucleic acids may be sequenced using any suitable sequencing technique as aspects of the invention are not limited in this respect.

It is also to be appreciated that some target nucleic acids are too large to be captured with one probe. Consequently, it may be necessary to capture multiple subregions of a target nucleic acid in order to analyze the full target.

In some embodiments, a subregion of a target nucleic acid is at least 1 bp. In other embodiments, a subregion of a target nucleic acid is at least 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 bp or more. In other embodiments, a subregion of a target nucleic acid has a length that is up to 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or more percent of a target nucleic acid length.

Petitioner Illumina, Inc., EX1006, p. 13

The skilled artisan will also appreciate that consideration is made, in the design of MIPs, for the relationship between probe length and target length. In some embodiments, MIPs are designed such that they are several hundred basepairs (e.g., up to 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 bp or more) longer than corresponding target (e.g., subregion of a target nucleic acid, target nucleic acid).

In some embodiments, lengths of subregions of a target nucleic acid may differ. For example, if a target nucleic acid contains regions for which probe hybridization is not possible or inefficient, it may be necessary to use probes that capture subregions of one or more different lengths in order to avoid hybridization with problematic nucleic acids and capture nucleic acids that encompass a complete target nucleic acid.

Aspects of the invention involve using multiple probes, e.g., MIPs, to amplify each target nucleic acid. In some embodiments, the set of probes for a given target can be designed to 'tile' across the target, capturing the target as a series of shorter sub-targets. Alternately, the set can be designed to 'stagger' the exact positions of the hybridization regions flanking the target, capturing the full target (and in some cases capturing flanking non-target sequence) with multiple probes having different targeting arms, obviating the need for tiling. The particular approach chosen will depend on the nature of the target set. For example, if small regions are to be captured, a staggered-end approach might be appropriate, whereas if longer regions are desired, tiling might be chosen. In all cases, the amount of bias-tolerance for probes targeting pathological loci can be adjusted ('dialed in') by changing the number of different MIPs used to capture a given molecule.

In some embodiments, the 'coverage factor', or number of probes used to capture a basepair in a molecule, is an important parameter to specify. Different numbers of probes per target are indicated depending on whether one is using the tiling approach (Figure 1) or one of the staggered approaches (Figures 2, 3). In certain embodiments, a coverage factor of about 3 to about 10 is used. However, the methods are not so limited and coverage factors of up to 2, 3, 4, 5, 6, 7, 8, 9, 10, 20 or more may be used. It is to be appreciated that the coverage factor selected may depend the probe layout being employed. For example, in the tiling approach, for a desired coverage factor, the number of probes per target is typically a function of target length, sub-target length, and spacing between adjacent sub-target start locations (step size). For example, for a

Petitioner Illumina, Inc., EX1006, p. 14

desired coverage factor of 3, a 200 bp target with a start-site separation of 20 bp and sub-target length of 60 bp may be encompassed with 12 MIPs (Figure 1). Thus, a specific coverage factor may be achieved by varying the number of probes per target nucleic acid and the length of the molecules captured. In the staggered approach, a fixed-length target nucleic acid is captured as several subregions or as 'super-targets', which are molecules comprising the target nucleic acid and additional flanking nucleic acids, which may be of varying lengths. For example, a target of 50 bp can be captured at a coverage factor of 3 with 3 probes in either a 'staggered' (Figure 2) or 'alternating staggered' configuration (Figure 3).

The coverage factor will be driven by the extent to which detection bias is tolerable. In some cases, where the bias tolerance is small, it may be desirable to target more subregions of target nucleic acid with, perhaps, higher coverage factors. In some embodiments, the coverage factor is up to 2, 3, 4, 5, 6, 7, 8, 9, 10 or more.

In some embodiments, when a tiled probe layout is used, when the target length is greater than 1 bp and when a step size (distance between the 5'-end of a target and the 5' end of its adjacent target) is less than the length of a target or subregion thereof, it is possible to compute probe number for a particular target based on target length (T), sub-target length (S), and coverage factor (C), such that probe number = T/(S/C) + (C-1).

In some aspects, the disclosure provides methods to increase the uniformity of amplification efficiency when multiple molecules are amplified in parallel; methods to increase the reproducibility of amplification efficiency; methods to reduce the contribution of targeting probe variability to amplification efficiency; methods to reduce the effect on a given target nucleic acid of polymorphisms in probe hybridization regions; and/or methods to simplify downstream workflows when multiplex amplification by MIPs is used as a preparative step for analysis by nucleic acid sequencing.

Polymorphisms in the target nucleic acid under the regions flanking a target can interfere with hybridization, polymerase fill-in, and/or ligation. Furthermore, this may occur for only one allele, resulting in allelic drop-out, which ultimately decreases downstream sequencing accuracy. In some embodiments, using a set of MIPs having multiple hybridization sites for the capture of any given target, the probability of loss

Petitioner Illumina, Inc., EX1006, p. 15

from polymorphism is substantially decreased because not all targeting arms in the set of MIPs will cover the location of the mutation.

Probes for MIP capture reactions may be synthesized on programmable microarrays because of the large number of sequences required. Because of the low synthesis yields of these methods, a subsequent amplification step is required to produce sufficient probe for the MIP amplification reaction. The combination of multiplex oligonucleotide synthesis and pooled amplification results in uneven synthesis error rates and representational biases. By synthesizing multiple probes for each target, variation from these sources may be averaged out because not all probes for a given target will have the same error rates and biases.

Multiplex amplification strategies disclosed herein may be used analytically, as in detection of SNPs, or preparatively, often for next-generation sequencing or other sequencing techniques. In the preparative setting, the output of an amplification reaction is generally the input to a shotgun library protocol, which then becomes the input to the sequencing platform. The shotgun library is necessary in part because next-generation sequencing yields reads significantly shorter than amplicons such as exons. In addition to the bias-reduction afforded by the multi-tiled approach described here, tiling also obviates the need for shotgun library preparation. Since the length of the capture molecule can be specified when the probes, e.g., MIPs, are designed, it can be chosen to match the readlength of the sequencer. In this way, reads can 'walk' across an exon by virtue of the start position of each capture molecule in the probe set for that exon.

In some aspects, the disclosure provides methods for analyzing a plurality 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

Petitioner Illumina, Inc., EX1006, p. 16

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 & 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.

Petitioner Illumina, Inc., EX1006, p. 17

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; 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

Petitioner Illumina, Inc., EX1006, p. 18

Gene name	Gene ID	Description	Gene aliases	OMIM	RefSeqGene	Chromosome map position
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
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., EX1006, p. 19

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
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

Petitioner Illumina, Inc., EX1006, p. 20

Gene name	Gene ID	Description	Gene aliases	OMIM	RefSeqGene	Chromosome map position
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; DXS1272E; KDM5C	314690	NG_008085.1	Xp11.22-p11.21
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

Petitioner Illumina, Inc., EX1006, p. 21

Gene name	Gene ID	Description	Gene aliases	OMIM	RefSeqGene	Chromosome map position
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

Table 2: Polymorphisms

Gene name	Gene ID	SNP ID	Mutation
CFTR	1080	rs63500661	TCACATCACCAAGTTAAAAAAAAAAA[A/G]GGGGCGGGGGGGCAGAATGAAAATT
CFTR	1080	rs63107760	AAACAAGGATGAATTAAGTTTTTTTT[-/T]AAAAAAGAAACATTTGGTAAGGGGA
CFTR	1080	rs62469443	ATCACCAAGTTAAAAAAAAAAAAGGG[A/G]CGGGGGGGCAGAATGAAAATTGCAT
CFTR	1080	rs62469442	CTATTGAACCAGAACCAAACAGGAAT[A/G]CCATAGCATTTTGTAAACTAAACTG
CFTR	1080	rs62469441	CAGGAGTTCAAGACCAGCCTACTAAA[A/C]CACACACACACACACACACACAC
CFTR	1080	rs62469439	GATTAAATAATAGTGTTTATGTACCC[C/G]GCTTATAGGAGAAGAGGGTGTGTGT
CFTR	1080	rs62469438	ATTGTTATCTTTTCATATAAGGTAAC[A/T]GAGGCCCAGAGAGATTAAATAACAT
CFTR	1080	rs62469437	TAATTTTAATTAAGTAAATTTAATTG[A/G]TAGATAAATAAGTAGATAAAAAATA
CFTR	1080	rs62469436	GTATAAAAAAAAAAAAAAAAAAAGTT[A/T]GAATGTTTTCTTGCATTCAGAGCCT
CFTR	1080	rs62469435	ATACTAAAAATTTAAAGTTCTCTTGC[A/G]ATATATTTTCTTAATATCTTAC

Petitioner Illumina, Inc., EX1006, p. 22

Gene name	Gene ID	SNP ID	Mutation
			ATC
CFTR	1080	rs62469434	TGCTGGGATTACAGGCGTGAGCCACC[A/G]CGCCTGGCCTGATGGGACATATTTT
CFTR	1080	rs62469433	CTACAATATAAGTATAGTATTGCAAA[A/C]CCATCAGGAAGGGTGTTAACTATTT
CFTR	1080	rs61763210	GTTGTCTCCAAACTTTTTTTCAGGTG[-/AGA]AGGTGGCCAACCGAGCTTCGGAAAG
CFTR	1080	rs61720488	TTTTTTCATAAAAGATTATATAAAGG[A/C]TATTGCTTTTGAATCACAAACACTA
CFTR	1080	rs61481156	ATCTAGTGAGCAGTCAGGAAAGAGAA[C/T]TTCCAGATCCTGGAAATCAGGGTTA
CFTR	1080	rs61443875	TAGAGTATAAAAAAAAAAAAAAAAAA[-/A]GTTTGAATGTTTTCTTGCATTCAGA
CFTR	1080	rs61312222	TGCAAATGCCAACTATCAAAGATATT[C/G]GAGTATACTGTCAATAAACTTCATA
CFTR	1080	rs61159372	TCCTCAACAGTTAGAAACAATATTTT[C/G]AGTGATTTCCCATGCCAACTTTACT
CFTR	1080	rs61094145	TTTTTGGTATTGTTGTTAAATAAGTG[A/G]GAATTCAATACAGTATAATGTCTGT
CFTR	1080	rs61086387	CTTGAAATCGGATATATATATATATA[-/TGTATATATATATATATATATATATATATACATATATATATATA]GTATTATCCCTGTTTTCACAGTTTT
CFTR	1080	rs60996744	AGAGGGGCTGTGAAGGACACCAAGGA[A/G]GAGACTAAGAGCCAGGAGGGAAAAC
CFTR	1080	rs60960860	TAGAGTTTATTAGCTTTTACTACTCT[A/G]CTTAGTTACTTTGTGTTACAGAATA
CFTR	1080	rs60923902	ACTAGTGATGATGAGCTTCTTTTCAT[-/AT]GTTTGTTGGCTGCATAAATGTCTTC
CFTR	1080	rs60912824	GCAGAGAAAAGAGGGGCTGTGAAGGA[C/G]ACCAAGGAGGAGACTAAGAGCCAGG
CFTR	1080	rs60887846	TTCAGAGGTCTACCACTGGTGCATAC[G/T]CTAATCACAGTGTCGAAAATTTTAC
CFTR	1080	rs60793174	AAGAAAGAGCAAAAGAGGGCAAACTT[C/T]TCATACATTTTTGATGTCGAAACCA
CFTR	1080	rs60788575	CCTAAAGTTTAAAAAGAAAAAAAAAA[-/A]GGAAGAAGGAATTAAAAATCCAAAG
CFTR	1080	rs60760741	GTGTGTGTGTGTATATATATATATAT[A/T]TATATATTTTTTTTTTCCTGAGCCA
CFTR	1080	rs60456599	AAACTGTTGATGTTTTCATTTATTTA[C/G]ATCATTGGAAAACTTTAGATTCTAG
CFTR	1080	rs60363249	TTTATCCATTCTTAACCAGAACAGAC[A/G]TTTTTTCAGAGCTGGTCCAGGAAAA
CFTR	1080	rs60355115	TTGAAATCGGATATATATATATATAT[A/G]TATATATATATATATATATATATAT
CFTR	1080	rs60308689	TAGTTTTTTATTTCCTCATATTATTT[-/T]CAGTGGCTTTTTCTTCCACATCTTT
CFTR	1080	rs60271242	ACATAGTTCTCAGTGGTACAACTACA[A/G]GTGATTTCTCTTTTCTTATTTCTGG
CFTR	1080	rs60010318	AGAGCAATGGCATCCCTTGTCTTGTG[C/T]TATACAGGATGCAGCAATTTATAGG
CFTR	1080	rs59961323	TTCTGTCTACATAAGATGTCATACTA[A/G]ATTATCTTTTCCAGCATGCATTCAG
CFTR	1080	rs59961270	CAGGGTGGCATGTTAGGCAGTGCTTA[A/G]AATAAATGAGTTGGTTATACAAGTA

Petitioner Illumina, Inc., EX1006, p. 23

Gene name	Gene ID	SNP ID	Mutation
CFTR	1080	rs59837506	AGGACACACACACACACACACACACA[-/CA]TGCACACACATTTAAATAGATGCAT
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[-/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]CAGACCGAGAGAACAGGAA

Petitioner Illumina, Inc., EX1006, p. 24

Gene name	Gene ID	SNP ID	Mutation
			GCCTGA
GBA	2629	rs28559737	AAGGACAAAGGCAAAGAGACAAAGGC[G/T]CAACACTGGGGGTCCCCAG AGAGTG
GBA	2629	rs28373017	TACCTAGTCACTTCCTGCCTCCATGG[C/T]GCAAAAGGGGATGGGTGTGCC TCTT
GBA	2629	rs12752133	CTCTTCCGAGGTTCCACCCTGAACAC[C/T]TTCCTGCTCCCTCGTGGTGTAG AGT
GBA	2629	rs12747811	TTCTGACTGGCAACCAGCCCCACTCT[C/T]TGGGAGCCCTCAGGAATGAAC TTGC
GBA	2629	rs12743554	gctcagcctcccaggctggagtgcag[A/T]ggcgcgatctcggctcaccgcaacc
GBA	2629	rs12041778	CATGAACCACATCAAATGAGATTTAG[C/T]GGGAGTGGCACACACAGTCA TGACC
GBA	2629	rs12034326	AAGCAGCCCTGGGGAGTCGGGGCGGG[A/G]CCTGGATTGGAAAAGAGAC GGTCAC
GBA	2629	rs11558184	CTCCAAGTTCTGGGAGCAGAGTGTGC[A/G]GCTAGGCTCCTGGGATCGAG GGATG
GBA	2629	rs11430678	GTTCCTCCAGTAAttttttttttttt[-/G/T]gttttgagacagagtcttgccctgt
GBA	2629	rs11264345	CTAGTACCTTACTTCCCTCAAGTTCA[A/T]TCATCTCACAGATATTTCCTGA GCA
GBA	2629	rs10908459	aattagccgtgcgtggtggcgggtgc[C/T]tgtaatcccacgtacttgggaggct
GBA	2629	rs10796940	CCATGGCCAGCCGGGGAGGGGACGGG[A/C]ACACACAGACCCACACAGA GACTCA
GBA	2629	rs10668496	agcgagagagagagagagagagagag[-/AG]gagCCGGCGCGAGAACTACGCATGC
GBA	2629	rs7416991	CGTAGCAGTTAGCAGATGATAGGCGG[C/G/T]GAAATCTTATTTCACAGGG CATTAA
GBA	2629	rs4024049	CTGGCCCTGGTGACAGTGGGGCTGTG[C/T]GTGGGGCCAGAGCCTTCTCAG AGGT
GBA	2629	rs4024048	CAGATACTGGCCCTGGTGACAGTGGG[A/G]CTGTGCGTGGGGCCAGAGCC TTCTC
GBA	2629	rs4024047	GACAGATACTGGCCCTGGTGACAGTG[G/T]GGCTGTGCGTGGGGCCAGAG CCTTC
GBA	2629	rs3841430	GGCTCctctctctctctctctctc[-/TC]gctcgctctctcgctctctcgctct
GBA	2629	rs3754485	GTTTCAGACCAGCCTGGCCAACATAG[C/T]GAAACCCCATCTCTACTAAAA ATAA
GBA	2629	rs3205619	AGTGGGCGATTGGATGGAGCTGAGTA[C/T]GGGGCCCATCCAGGCTAATC ACACC
GBA	2629	rs2990227	CCGGGCTCCGTGAATGTTTGTCACAT[C/G]TCTGAAGAACGTATGAATTAC ATAA
GBA	2629	rs2990226	GAATCCCAACCCCGACGCTCGTCGCC[C/G]GGCTCCGTGAATGTTTGTCAC ATGT
GBA	2629	rs2990225	GCGAATCCCAACCCCGACGCTCGTCG[C/T]CGGGCTCCGTGAATGTTTGTC ACAT
GBA	2629	rs2990224	TGGGCAGAAGTCAGGGTCCAAAGAAA[G/T]GGCAAAGAAAAGTGTcagtggct ca
ASPA	443	rs63751297	TAAGAAAGACGTTTTTGATTTTTTTC[A/G]GACTTCTCTGGCTCCACTACCC TGC
ASPA	443	rs62071301	CTGATTCCTGGCCAGGAGCGGTGGCT[C/T]ACGCCTGTAATCCCAGCGCTT TGGG
ASPA	443	rs62071300	TAAAAATGCTGATTCCTGGCCAGGAG[C/T]GGTGGCTCACGCCTGTAATCC CAGC
ASPA	443	rs62071299	TTTAAAAATGCTGATTCCTGGCCAGG[A/C]GCGGTGGCTCACGCCTGTAAT CCCA
ASPA	443	rs62071297	CAAGACCTGTCAAAGATCTGAGAAAT[A/T]TTACCCGACTTACAAGCTAAC

Petitioner Illumina, Inc., EX1006, p. 25

Gene name	Gene ID	SNP ID	Mutation
			CATT
ASPA	443	rs61697033	ACTGTAATAAGTGCTGTAAAAGAAAT[A/G]CACAAAATAATATAGCAGAG GGTAT
ASPA	443	rs60743592	CTTGAGGTCAGGAGTTCAAGACCAGT[C/T]TGGGCAACATGGGGAAAACC TTGTC
ASPA	443	rs60666840	AGGTTGCAGTGAGCCGAGATCATGCC[A/G]TTGCACTCCAGCCGGGGCAA CAAAA
ASPA	443	rs60147514	ACAAGTGTCTTGAAATTATCTGTGAT[C/T]TGCTATAGAGCAATACTTTTGT AAA
ASPA	443	rs59930743	GTGGGTATATGCAGCTCTATGCACTA[C/T]CTGCTCATTTATTTGGTAAATC TAA
ASPA	443	rs59690349	TGTGTGTGTGTGCGTGTGTGTGTGTG[- /TGTGTGTG]ATCATAAGAGTGGCTGCAGCAAACT
ASPA	443	rs59676360	AGTCTGGAGTGCAATGGTGCAATCTC[A/G]GCTCACTGCAGCCTCCACCTC CGGG
ASPA	443	rs59335404	CTCCTAATGGATATTTCCTAAATTTT[G/T]CTGAACAGAATTTAACTTGAGC TGG
ASPA	443	rs58879097	ATTTAAAAATGGATTTCTAGAAAAAC[A/G]ATCACATACTTGAATATTTTA GCAA
ASPA	443	rs58686774	CTATAAATGGGTAGCATGAGGGATTC[A/G]AGGAGGTGGCTGAAAGAAGC ACGTA
ASPA	443	rs57511162	AAGAAACCAAGCATAGTAGAGTGTTA[A/G]AAAACCAAAGCAACTAAACA ACTGT
ASPA	443	rs55859596	CGGGGCTCAGAACTTGTAACAGAAAA[A/T]TAAAATATACTCCACTCAAG GGAAT
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	CTTTTTCCAGATTTTTTTTTTTTTT[- /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[- /A]TTCTCCTGCCTCAGCCTCCCAAGTA
ASPA	443	rs35002210	CCTCCCTGTGATCCGAAGTAGCAGAC[A/G]TACTTAACTTCCATGGTGGAT TGTT
ASPA	443	rs34744839	AAAACATTATTATATCTAGAAAAAAA[- /A]TGTATCTTAACCATTGTGGGAAGTG

Petitioner Illumina, Inc., EX1006, p. 26

Gene name	Gene ID	SNP ID	Mutation
ASPA	443	rs34680506	TTGAAGGTAAAATCATAGGGAGTTGG[-/G]AGCTGTCCTCTTGCGCTGAATCAGT
ASPA	443	rs34365618	ACTTGTGGCCTTTTTGGAGAGGTTAG[-/CA]ACTCTGAAAACTCTGTCCCTGGACC
ASPA	443	rs34275920	GAAGGAGAAAAAGAGAGGAAATAAGT[-/T]AAAATAATAAACACAATTAATAAAG
ASPA	443	rs34109510	TGTATACATGTGCCATGTTGGTGTGC[C/T]GCACCCATTAACTCGTCATTTAGCA
ASPA	443	rs34054576	TCACCTGTCACCTCCTATAGAACTTT[-/C]CCCTGACCCTCCTCTATAGCATTAA
ASPA	443	rs34015272	ATAAATGATCATCATTCACAGTAGGG[-/G]TTTTGTTTTGTTTTTTTTCTGGAAA
ASPA	443	rs34002091	ACAGACATATCTACAAACACACTTTT[-/T]CACATATTTGTGTAAGTCATTTATG
ASPA	443	rs28940574	AAAGACAACTAAACTAACGCTCAATG[A/C]AAAAAGTATTCGCTGCTGTTTACAT
ASPA	443	rs28940279	TACCGTGTACCCCGTGTTTGTGAATG[A/C]GGCCGCATATTACGAAAAGAAAGAA
ASPA	443	rs17850703	CAGGGCTGGAGGTAAAACCATTTATT[A/G]CTAACCCCAGAGCAGTGAAGAAGTG
ASPA	443	rs17222495	TTCTTCATTGCCTATTGAAGAGAGAG[C/T]GGAATGCTTTGGTTGCCAGATATGG
ASPA	443	rs17175228	CACAAGATCTCATTACTCAGGAGCTG[C/T]CCAAGTGTCTAATGTACTTAGTTAA
ASPA	443	rs16953074	TTCTGTGTAACATTTCATTTAAGCAA[A/G]GGATTCGGCAAATCAAAAATTGTCA
ASPA	443	rs16953070	TAAAACGTATTGAAGGTATTATTGAC[G/T]CTGTTGAAGCAAAGAGAACAAAACA
HEXA	3073	rs62022858	ATCTGCTCTTCCAGTTGGATGACAAG[C/T]CTTGCTGTCTAACACCTGCTGCAGA
HEXA	3073	rs62022857	CCATTTTTTGTTGTATTTTTTTTTTC[C/T]TGAATACTTTTTATCGCAGTTGGTT
HEXA	3073	rs62017872	CCCTGTCTCTAAAAGAAAAAAAAAAA[A/G]AAAAAAAAAAGAAAACAAAACCCAA
HEXA	3073	rs62017871	AGTGGCTCCAAAAAGGTCATGGAACC[C/T]CTTGAGGATGATGCAAATTGACTCT
HEXA	3073	rs61662730	TAAAGTTACTTTTCTTTTATTGACTT[C/T]CCCTTATTTTTTAACCTTATGCTTT
HEXA	3073	rs61329913	CAGAGTTAAAAAAAAAAAAAAAAAAA[-/A]GGAAGTAGCAGCAACAGCTTGGAAA
HEXA	3073	rs60920713	GTTGCCCAGGGTTGAGTGCAGAGGCA[C/T]ATCTTGGCTCACAGCAACCTCTGCC
HEXA	3073	rs60783213	AAGGCTTTTTTTTTTTTTTTTTTTT[-/TTTT]GAGACAGAGTCTTGCTGTGTCACCC
HEXA	3073	rs60644867	GCCTACATTCTGCAAAGAGGAGGGAA[C/G]ATTCACAGCTCCATACTTGAACCCT
HEXA	3073	rs60288568	CCAAAGGAGAATAGCTCTAGGGGAGG[C/G]AGGTGGATGAGTATGCATGGGGGAG
HEXA	3073	rs59888548	GACTCCATCTCAAAAAAAAAAAAAAA[-/A]TGCAGTCTAATGGCAGAATTAGACT
HEXA	3073	rs59733856	TTATTTATTTATTTATTTATTTTTGA[A/G]ACAGGGTCTCTGTTGTCCAGGCTGG
HEXA	3073	rs59427837	TTTTGAGGCAGGGTCTCACTCTGTTG[C/T]CCAGGGTTGAGTGCAGAGGCACATC

Petitioner Illumina, Inc., EX1006, p. 27

Gene name	Gene ID	SNP ID	Mutation
HEXA	3073	rs59171976	CGCCTTGCGAAGGCCCCACAGCTTGC[C/T]TGTGACAAACGTTCATAGGCAAATG
HEXA	3073	rs58706602	GGAGGTCTGTACAAAGCACCACCTAC[C/T]TCATGGGTCAGTTTCCACAGCAGAA
HEXA	3073	rs58696963	GAATCTTATAATTCACTGTGTACCTC[-/CCTC]TGTTTCATATTTTCGCAATTGAACT
HEXA	3073	rs58610850	AACATAGTATCTAATATAGCTTTACA[C/T]CCAAAGCCAAAATATGAATACACTG
HEXA	3073	rs58016062	TTGTTTTGTTTTGTTTGGGGGGGGGG[-/G]TTGTTTTTCTGAGAGGGAGTCTTGC
HEXA	3073	rs57733983	CATACCAAAGGGCAGCTGGAGGGATA[C/T]AGACGGAAGTCATGTGGAGAGTGAA
HEXA	3073	rs57476645	CAGGTGTGAGCCACCACGACCACCAA[A/T]TTAGCTCTTTTTACTCCTTCCCTTC
HEXA	3073	rs56870003	AGTGGTAGCTGATTTTGCTTCTGGAT[A/C]CTTTGCCACCTTCCCACTCTTTAAT
HEXA	3073	rs56338339	AAAGACCTGTTTCTTAAAAAAAAAAA[-/AGAAAAAAAAAAA]GAAAGAAAAGAAAAGAAAAAAACAG
HEXA	3073	rs55995352	TAAAAAATCTTTCAATGAGGAGATGT[C/T]CCCAGAGCAAGACAGCTGTAGGATG
HEXA	3073	rs55860138	AAAAGAAAAAAAAAAAAAAAAAAAA[-/A]GAAAACAAAACCCAAACCCATAAAG
HEXA	3073	rs55743646	CCTGTCTCTAAAAGAAAAAAAAAAAA[A/G]AAAAAAAAAGAAAACAAAACCCAAA
HEXA	3073	rs55665666	GTTATCATAGAAAAATATCACACTCT[-/GT]CTGTATCCCCACTTCCAGAAACTGT
HEXA	3073	rs36106892	CAGGAGCTCATAGAATTACATACAAT[-/C]TTTTTTTTTTTTTTTTTGAGACAGCG
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[-

Petitioner Illumina, Inc., EX1006, p. 28

Gene name	Gene ID	SNP ID	Mutation
			/G]GGAAATAAACATCTTATTGAATTTA
HEXA	3073	rs34093438	GGACCCCTGAAAGGCACAAGACACCC[-/T]TTCAGGTTCACACTTCCTGAAAGCT
HEXA	3073	rs34085965	CCACCAATCACCAGAGCCTTCTGCTC[A/G]GGGGTACCTGAGGGAAAACAAGCAA
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

The following examples illustrate non-limiting embodiments of the invention.

Design a set of capture probes for a human target exon

All targets are captured as a set of partially-overlapping subtargets. For example, in the tiling approach, a 200 bp target exon might be captured as a set of 12 subtargets, each 60bp in length (Figure 1). Each subtarget is chosen such that it partially overlaps two or three other targets.

All probes are composed of three regions: 1) a 20 bp 'targeting arm' comprised of sequence which hybridizes immediately upstream from the sub-target, 2) a 30 bp 'constant region' comprised of sequence used as a pair of amplification priming sites, and 3) a second 20 bp 'targeting arm' comprised of sequence which hybridizes immediately downstream from the sub-target. Targeting arm sequences will be different for each capture probe in a set, while constant region sequence will be the same for all probes in the set, allowing all captured targets to be amplified with a single set of primers. Targeting arm sequences should be designed such that any given pair of 20 bp sequences is unique in the target genome (to prevent spurious capture of undesired sites). Additionally, melting temperatures should be matched for all probes in the set such that hybridization efficiency is uniform for all probes at a constant temperature (e.g. 60C). Targeting arm sequences should be computationally screened to ensure they do not form

Petitioner Illumina, Inc., EX1006, p. 29

strong secondary structure that would impair their ability to basepair with the genomic target.

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)
- 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)

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

Petitioner Illumina, Inc., EX1006, p. 30

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')
- 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 length.
- Ligate adapters required for amplification by the sequencing platform used. If necessary, purify ligated product from unligated product and adapters.

Petitioner Illumina, Inc., EX1006, p. 31

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.

What is claimed is:

Petitioner Illumina, Inc., EX1006, p. 32

Claims

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

contacting each of a plurality of target nucleic acids with a probe set,

wherein each probe set comprises a plurality of different probes, each probe having a central region flanked by a 5' region and a 3' region that are complementary to nucleic acids flanking the same strand of one of a plurality of subregions of the target nucleic acid, wherein the subregions of the target nucleic acid are different, and wherein each subregion overlaps with at least one other subregion,

isolating a plurality of nucleic acids each having a nucleic acid sequence of a different subregion for each of the plurality of target nucleic acids, and

analyzing the isolated nucleic acids.

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

contacting each of a plurality of target nucleic acids with a probe set,

wherein each probe set comprises a plurality of different probes, each probe having a central region flanked by a 5' region and a 3' region that are complementary to nucleic acids flanking the same strand of one of a plurality of subregions of the target nucleic acid, wherein the subregions of the target nucleic acid are different, and wherein a portion of the 5' region and a portion of the 3' region of a probe have, respectively, the sequence of the 5' region and the 3' region of a different probe,

isolating a plurality of nucleic acids each having a nucleic acid sequence of a different subregion for each of the plurality of target nucleic acids, and

analyzing the isolated nucleic acids.

3. The method of claim 1 or 2, wherein the target nucleic acids are selected from Table 1.

4. The method of claim 1 or 2, wherein the analyzing comprises genotyping a polymorphism of the target nucleic acid.

Petitioner Illumina, Inc., EX1006, p. 33

5. The method of claim 4, wherein the polymorphism is selected from Table 2.

Petitioner Illumina, Inc., EX1006, p. 34

Abstract of the Disclosure

The invention relates to methods that reduce bias and increase reproducibility in multiplex detection of genetic loci.

Petitioner Illumina, Inc., EX1006, p. 35

Electronic Acknowledgement Receipt	
EFS ID:	5256599
Application Number:	61174470
International Application Number:	
Confirmation Number:	9948
Title of Invention:	METHOD FOR REDUCING REPRESENTATIONAL BIAS IN MULTIPLEX AMPLIFICATION OF NUCLEIC ACIDS
First Named Inventor/Applicant Name:	Greg Porreca
Customer Number:	23628
Filer:	Patrick R.H. Waller/Trish McDonald
Filer Authorized By:	Patrick R.H. Waller
Attorney Docket Number:	G0814.70000US00
Receipt Date:	30-APR-2009
Filing Date:	
Time Stamp:	22:08:00
Application Type:	Provisional

Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	$110
RAM confirmation Number	5756
Deposit Account	232825
Authorized User	WALLER,PATRICK R.H.
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Provisional Cover Sheet (SB16)	G081470000US00-TRN-PRW.pdf	45766 /99f/dda40/5aae/4b61345afeb24939ed454dcf	no	2
Warnings:					
This is not a USPTO supplied Provisional Cover Sheet SB16 form.					
Information:					
2	Application Data Sheet	G081470000US00-ADS-PRW.pdf	21051 3ff643d5b782464c0efc876893fbd9073a52b873	no	3
Warnings:					
Information:					
This is not an USPTO supplied ADS fillable form					
3	Fee Worksheet (PTO-875)	G081470000US00-FEE-PRW.pdf	44552 44488229c858c4aebb0639279770da5a6399f070	no	1
Warnings:					
Information:					
4		G081470000US00-PROAPL-PRW.pdf	152979 eeb1275495597b76eda1921d075f086c9e19b6440	yes	28

Multipart Description/PDF files in .zip description

Document Description	Start	End
Specification	1	25
Claims	26	27
Abstract	28	28

Warnings:					
Information:					
5	Drawings-only black and white line drawings	G081470000US00-DRW-PRW.pdf	602374 1c1bc60f2f4d40735e0ea98e3b041cb2841cf82a	no	3
Warnings:					
Information:					
6	Fee Worksheet (PTO-875)	fee-info.pdf	29731 5fe0719ef53f9849eaee0bec67eee060cce91b40	no	2
Warnings:					
Information:					

Total Files Size (in bytes):	896453

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

Figure 1

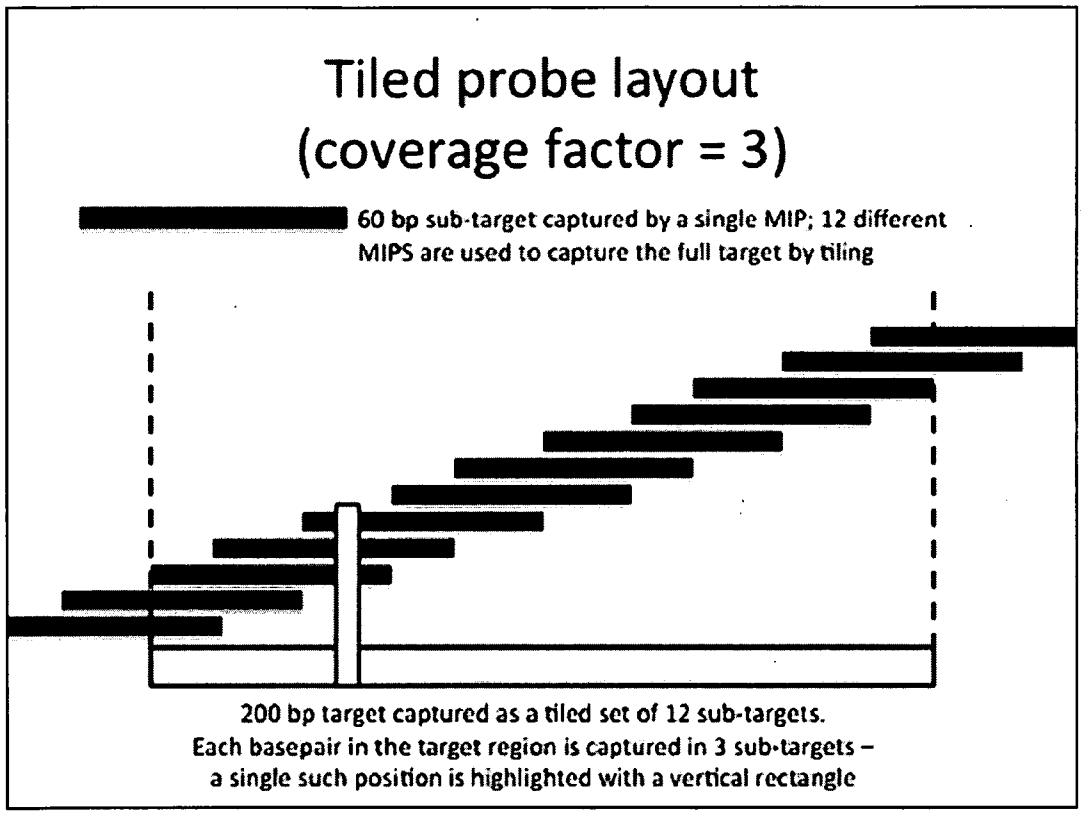


Figure 1. Tiled probe layout, showing ten captured sub-targets tiled across a single target. Each position in the target is covered by three sub-targets such that MIP performance per basepair is averaged across three probes.

Figure 2

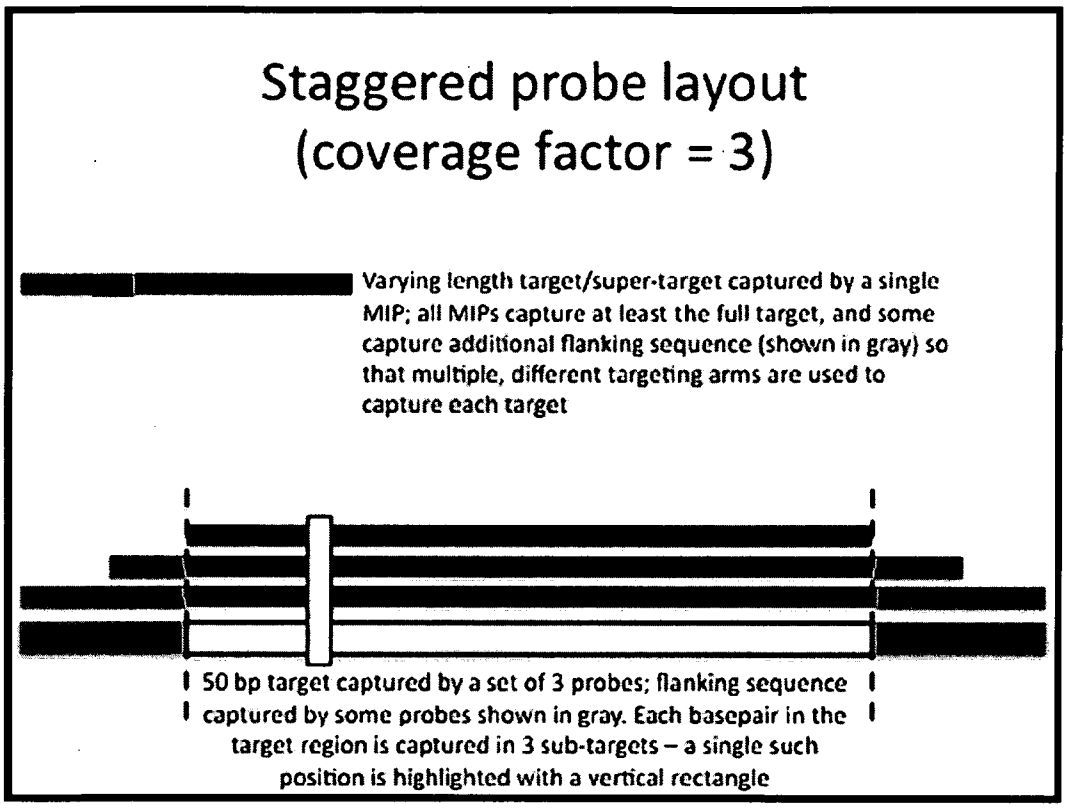


Figure 2. Staggered probe layout, showing the targets captured by a set of three MIPs. Each MIP captures the full target, shown in black, plus (in some cases) additional extra-target sequence, shown in gray, such that the targeting arms of each MIP fall on different sequence. Each position in the target is covered by three sub-targets such that MIP performance per basepair is averaged across three probes.

Figure 3

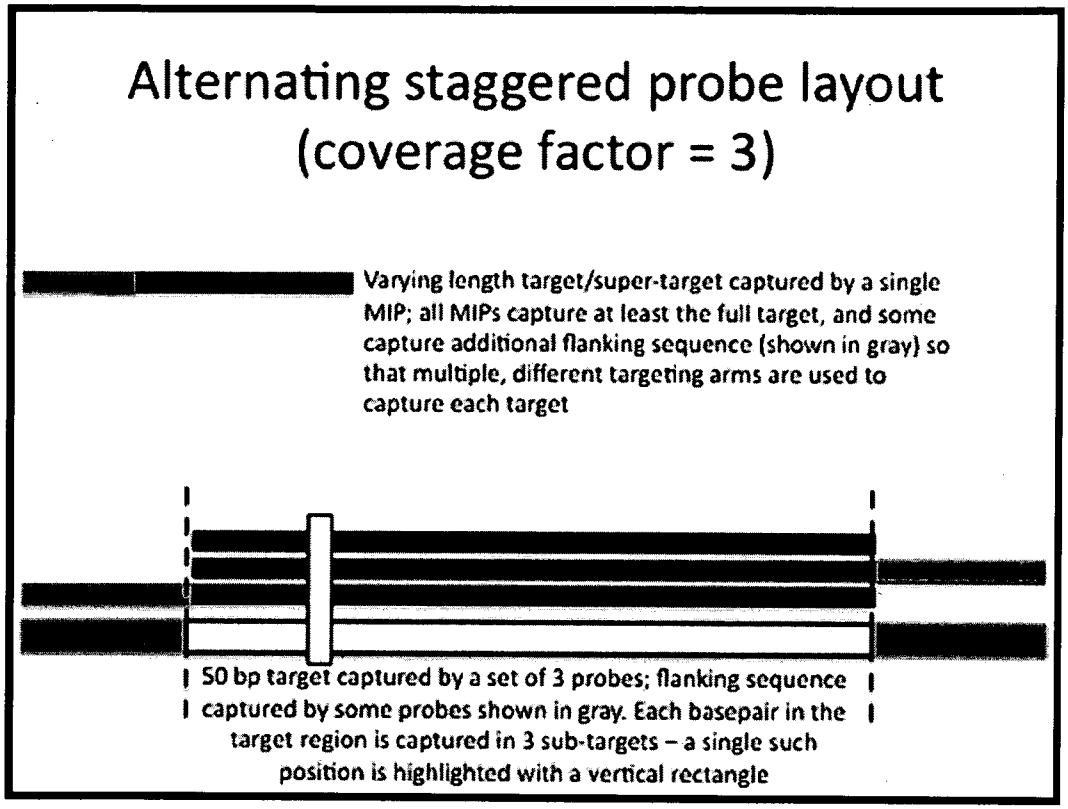


Figure 3. Alternating staggered probe layout, showing the targets captured by a set of three MIPs. Each MIP captures the full target, shown in black, plus (in some cases) additional extra-target sequence, shown in gray, such that the targeting arms of each MIP fall on different sequence. Each position in the target is covered by three sub-targets such that MIP performance per basepair is averaged across three probes.