(12) UK Patent Application (19) GB (11) 2564848 (13) A

(43) Date of A Publication 30.01.2019

(21) Application No: 1711540.3

(22) Date of Filing: 18.07.2017

(71) Applicant(s):
Congenica Ltd.
Merlin Place, Milton Road, Cambridge, CB4 0DP, United Kingdom

(72) Inventor(s):
Nick Lench
Matt Hurles
John McGonigle
Alan Martin
Suzie Drury

(74) Agent and/or Address for Service:
Basck Ltd
16 Saxon Road, CAMBRIDGE, Cambridgeshire, CB5 8HS, United Kingdom

(51) INT CL:
C12Q 1/6855 (2018.01) ***C12Q 1/6809*** (2018.01)
C12Q 1/6827 (2018.01) ***C12Q 1/6869*** (2018.01)
C12Q 1/6883 (2018.01) ***G16B 30/00*** (2019.01)
G16B 30/10 (2019.01)

(56) Documents Cited:
WO 2018/049049 A1 **WO 2014/149134 A2**
US 20160251704 A1
PLoS One, Vol. 8, No. 6, Jun. 2013, Ge Huijuan, et al., "Noninvasive Prenatal Detection for Pathogenic CNVs: The Application in alpha-Thalassemia", Article No.: e67464.
Prenatal Diagnosis, Vol. 36, No. 1, Jan. 2016, Tynan J A, et al., "Application of risk score analysis to low-coverage whole genome sequencing data for the noninvasive detection of trisomy 21, trisomy 18, and trisomy 13.", pages 56-62.
Fetal Diagn Ther., Vol. 42, No. 3, epub Feb. 2017, F Crea et al., "The IONA® Test: Development of an Automated Cell-Free DNA-Based Screening Test for Fetal Trisomies 13, 18, and 21 That Employs the Ion Proton Semiconductor Sequencing Platform.", pages 218-224.

(58) Field of Search:
Other: **EPODOC, WPI, BIOSIS, MEDLINE, INTERNET**

(54) Title of the Invention: **Prenatal screening and diagnostic system and method**
Abstract Title: **Prenatal screening systems and methods comprising ligation of barcodes to cell-free DNA**

(57) Disclosed is prenatal screening and diagnostic system. The prenatal screening and diagnostic system includes a wet-laboratory arrangement for processing a blood sample to determine cell-free DNA readout data from the blood sample, and a data processing arrangement for processing the cell-free DNA readout data with reference to information stored in a database arrangement to generate a risk score indicative of whether or not there are genetic abnormalities in the blood sample that are indicative of fetal abnormalities or risk of developing abnormalities. The prenatal screening and diagnostic system is operable to ligate nucleic acid base molecular barcodes to fragments of the cell-free DNA present in the blood sample prior to amplifying the molecular barcode-ligated DNA fragments for sequencing the amplified molecular barcode-ligated fragments to generate the cell-free DNA readout data. Associated methods and computer programs are also claimed.

GB 2564848 A

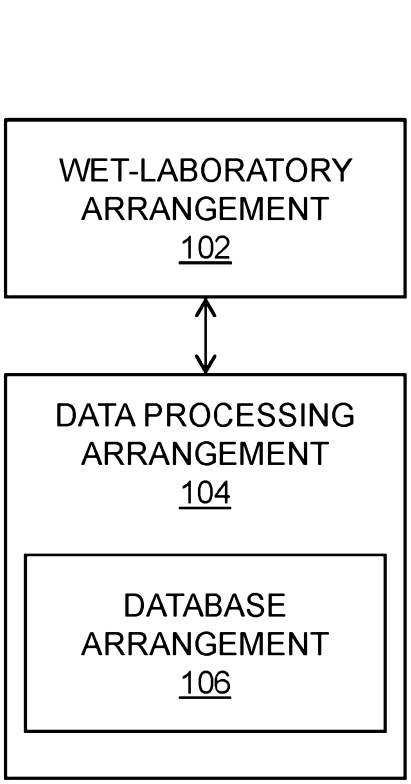
100
WET-LABORATORY ARRANGEMENT
102
DATA PROCESSING ARRANGEMENT
104
DATABASE ARRANGEMENT
106

FIG. 1

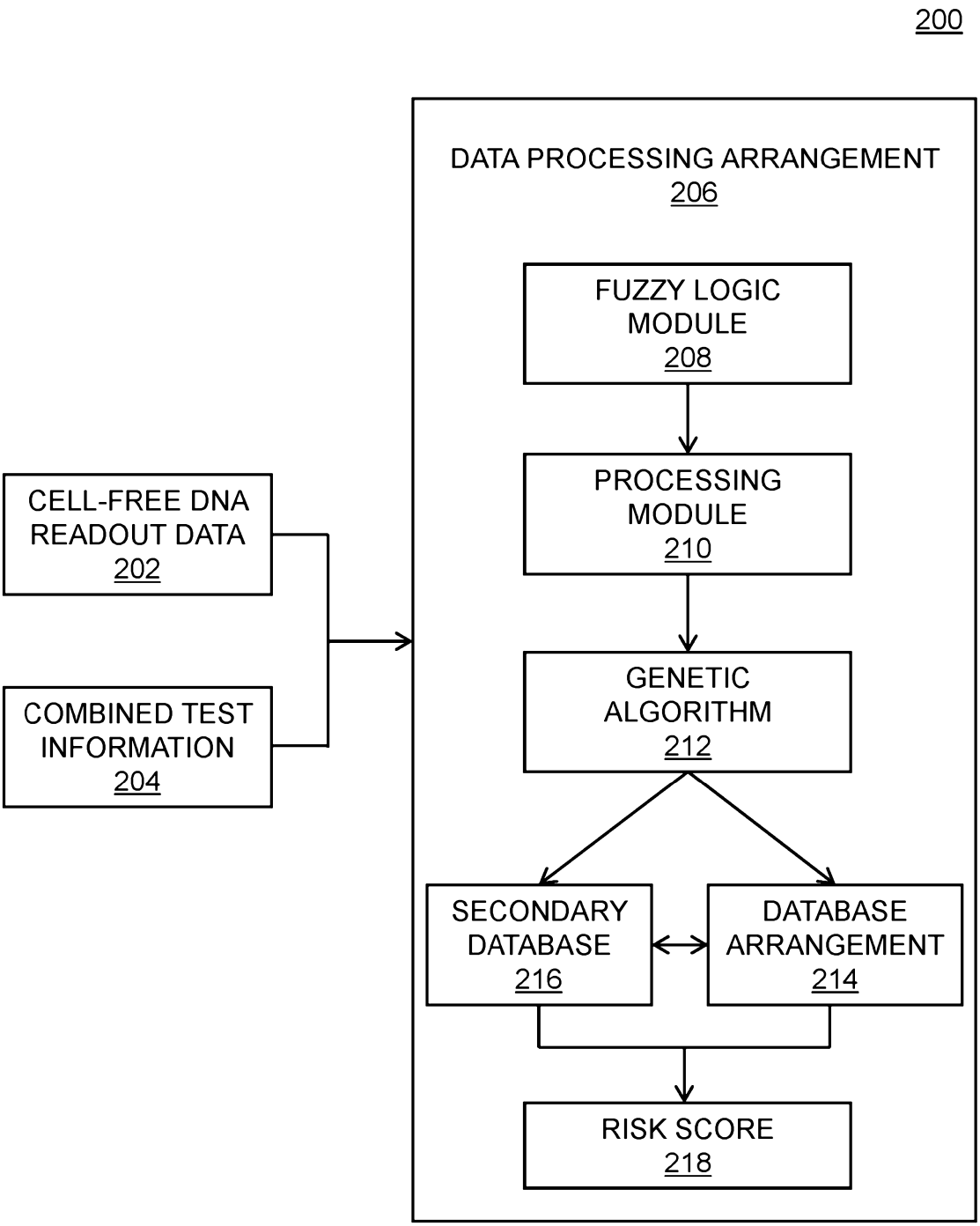
200
DATA PROCESSING ARRANGEMENT
206
FUZZY LOGIC MODULE
208
PROCESSING MODULE
210
GENETIC ALGORITHM
212
SECONDARY DATABASE
216
DATABASE ARRANGEMENT
214
RISK SCORE
218
CELL-FREE DNA READOUT DATA
202
COMBINED TEST INFORMATION
204

FIG. 2

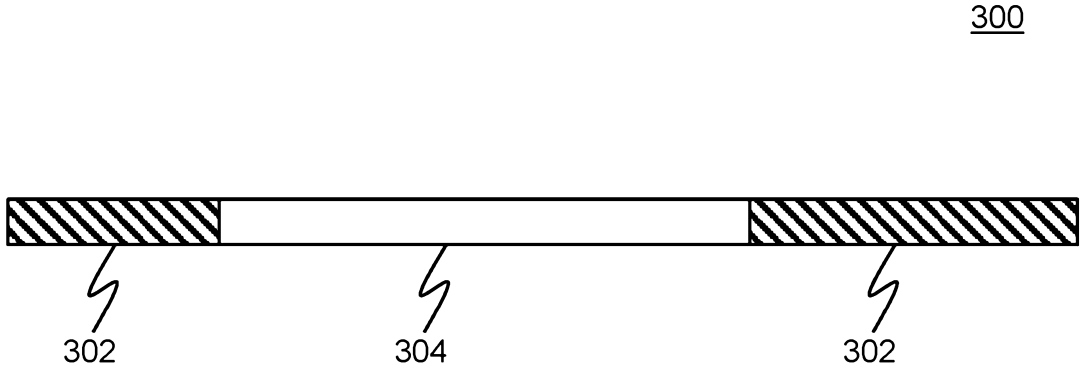

FIG. 3

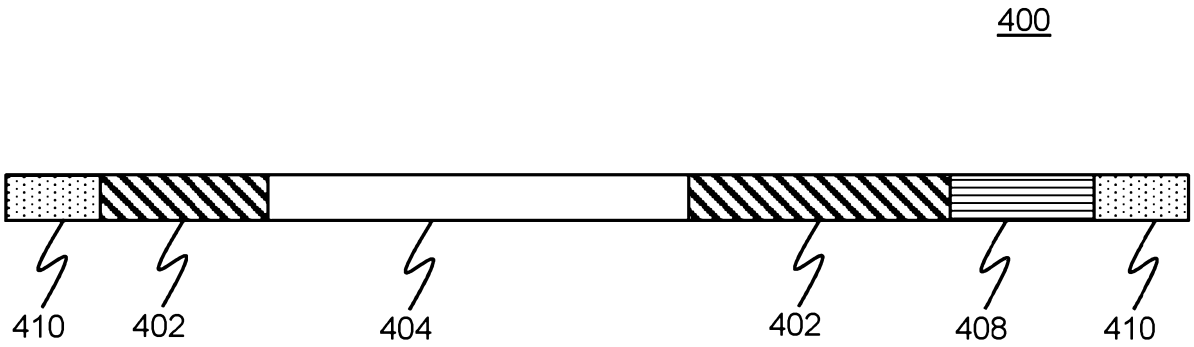

FIG. 4

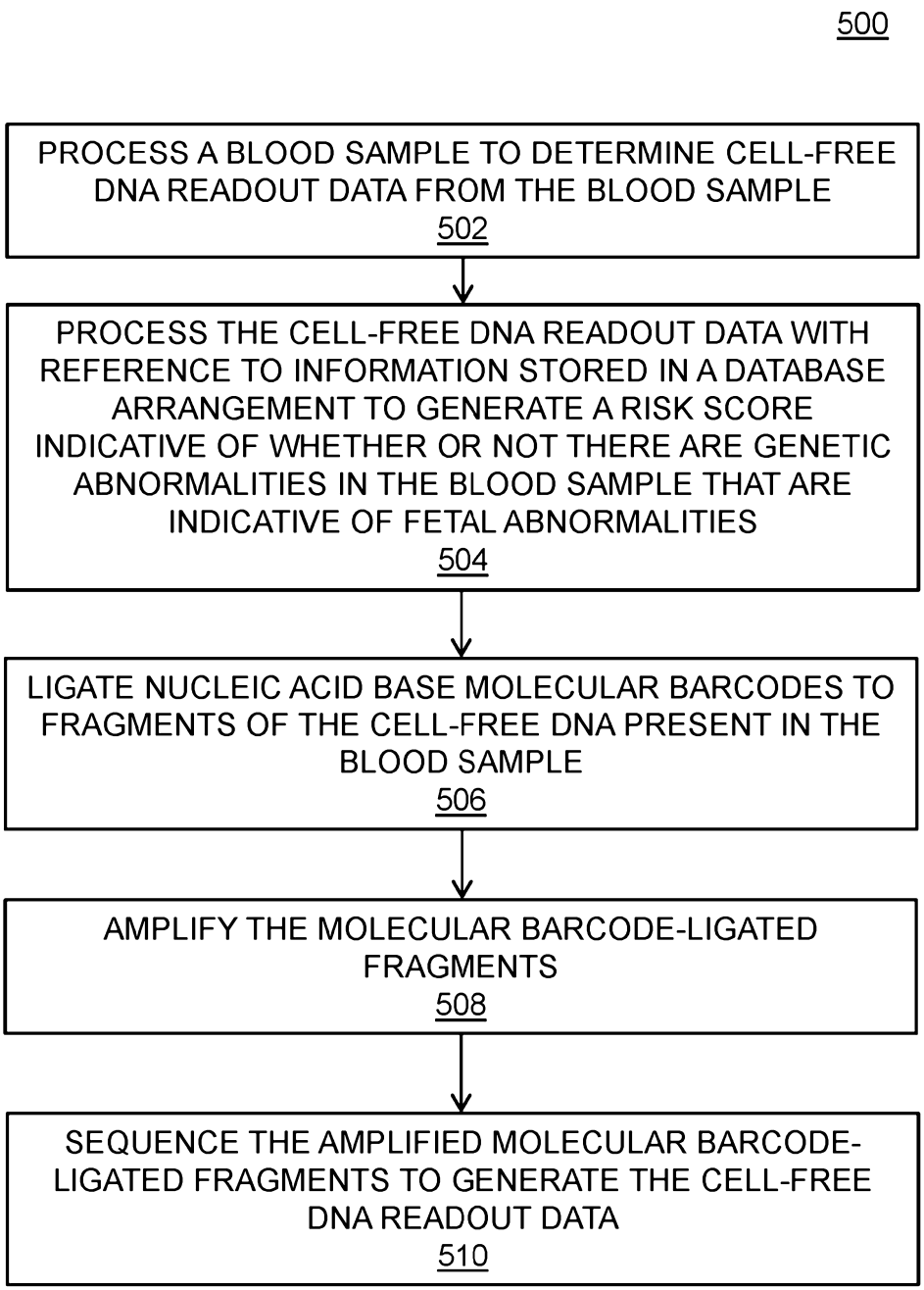
500
PROCESS A BLOOD SAMPLE TO DETERMINE CELL-FREE DNA READOUT DATA FROM THE BLOOD SAMPLE
502
PROCESS THE CELL-FREE DNA READOUT DATA WITH REFERENCE TO INFORMATION STORED IN A DATABASE ARRANGEMENT TO GENERATE A RISK SCORE INDICATIVE OF WHETHER OR NOT THERE ARE GENETIC ABNORMALITIES IN THE BLOOD SAMPLE THAT ARE INDICATIVE OF FETAL ABNORMALITIES
504
LIGATE NUCLEIC ACID BASE MOLECULAR BARCODES TO FRAGMENTS OF THE CELL-FREE DNA PRESENT IN THE BLOOD SAMPLE
506
AMPLIFY THE MOLECULAR BARCODE-LIGATED FRAGMENTS
508
SEQUENCE THE AMPLIFIED MOLECULAR BARCODE-LIGATED FRAGMENTS TO GENERATE THE CELL-FREE DNA READOUT DATA
510

FIG. 5

PRENATAL SCREENING AND DIAGNOSTIC SYSTEM AND METHOD

TECHNICAL FIELD

The present disclosure relates to prenatal screening and diagnostic systems that are operable to process maternal blood in order to determine fetal (foetal) characteristics. Moreover, the present disclosure concerns methods of using aforementioned prenatal screening systems, for example for processing maternal blood in order to determine fetal (foetal) characteristics. Additionally, the present disclosure is concerned with computer program products comprising a non-transitory computer-readable storage medium having computer-readable instructions stored thereon, the computer-readable instructions being executable by a computerized device comprising processing hardware to execute the aforesaid methods.

BACKGROUND

Zygote formation and associated subsequent fetal (foetal) development is a complex biological process that does not always occur without defects arising. It is of great societal benefit that such defects are detected reliably, for example as early as possible, during fetal growth.

Antenatal or prenatal screening is provided to pregnant women to prevent or treat potential health problems that may occur during their pregnancy. Such problems may affect both a given mother and/or a fetus of the mother. Moreover, the problems may be influenced by factors such as lifestyle, environmental or genetics. However, of particular importance are fetal (foetal) abnormalities that are genetic in origin. These abnormalities may be caused by mutations inherited from one or both parents, or may arise "*de novo*" (namely stochastically spontaneously). The mutations can range

extensively from changes in single nucleotides (namely, in 'rungs' of a given DNA molecule) to a presence of additional whole chromosomes; for example, the human genome usually includes 46 chromosomes (two pairs of 23 chromosomes), but abnormalities can result in a total of 45 or 47 chromosomes arising. Of particular clinical significance are chromosomal disorders known as '*aneuploidies*' that occur when there are an abnormal number of chromosomes (for example, as occurs in Down's syndrome). Many chromosomal disorders are incompatible with life or result in multiple congenital anomalies for a given newborn child.

Prenatal screening for fetal (foetal) chromosomal abnormalities during pregnancy is contemporarily widely available through public and private healthcare providers. This prenatal screening is normally carried out during a first trimester of pregnancy (namely 10 to 14 weeks after conception) and typically involves obtaining a maternal blood sample from a given mother combined with performing an ultrasound scan of a fetus of the given mother; such a procedure is known as a "*Combined Test*". When undertaking the "*Combined Test*", concentrations (namely, 'levels') of human chorionic gonadotrophin (hCG) and pregnancy-associated plasma protein (PAPP-A) are measured, together with executing a nuchal translucency (NT) scan; once other medical factors (for example, maternal age) have been taken into account, a risk-score is provided at a conclusion of the "*Combined Test*".

If a given mother's pregnancy is categorised as being of a high-risk, an invasive diagnostic procedure (for example, chorionic villus sampling, an amniocentesis, a cordocentesis) is offered to the given mother to confirm or rule out Down's syndrome (trisomy chromosome 21 - T21), Edwards's syndrome (trisomy chromosome 18 - T18) and Patau syndrome (trisomy chromosome 13 - T13). In such circumstances, pregnant women are also

offered a second ultrasound scan at 18 to 21 weeks (after conception) to check for structural fetal anomalies such as cardiac malformations, brain malformations and skeletal abnormalities. This second scan can be used to direct antenatal treatments, identify anomalies that require early intervention following delivery of a child or enable follow-on diagnostic testing and pregnancy management. Invasive tests such as chorionic villus sampling, amniocentesis and cordocentesis carry a 1% chance of miscarriage.

During recent years, non-invasive techniques (without an associated increased risk of miscarriage) have been developed for diagnosing fetal chromosomal anomalies that rely on a presence of circulating cell free fetal DNA (cffDNA) in the mother's blood. The testing of cell-free fetal DNA (cffDNA) has now entered routine clinical practice for non-invasive prenatal testing (NIPT) for aneuploidy (associated with chromosomes T21, T18, T13) and has a broad application as a replacement for the aforementioned 'Combined Test'. The anomalies that can be tested by NIPT are increasing as methods are developed for identifying sub-chromosomal rearrangements such as 22q11.2/DiGeorge syndrome and other nucleic acid based microdeletion syndromes. However, the false positive rate for these sub-chromosomal anomalies is considered to be too high to offer such testing on a screening basis, and is only offered if there is an accompanying clinical indication, such as a congenital heart defect. NIPT classified as 'testing' rather than 'diagnosis', as the cffDNA which is measured is derived from a placenta of a given mother rather than a fetus of the given mother, meaning that false positives can occur due to confined placental mosaicism. For this reason, it is recommended that positive NIPT results are confirmed by performing an invasive amniocentesis (often referred to colloquially as an 'amnio').

Non-invasive prenatal diagnosis (NIPD), namely a form of test, is classified as being a diagnostic assay with no need for performing a subsequent invasive assay to confirm results provided by the diagnostic assay. The use of NIPD is more limited than NIPT and is used for investigating fetuses at risk of single gene disorders (namely, affected by inherited *and de novo* mutations) or those that present with a suspicion of a monogenic disorder when performing a fetal ultrasound investigation.

Cell-free fetal DNA (cffDNA) circulates in maternal blood at approximately 10% of the total cell free component. Coupled with low concentrations of total cell free DNA, next generation sequencing library preparation methods for analysing this cell-free fetal DNA (cffDNA) material require Polymerase Chain Reaction (PCR) amplification to be employed. PCR amplification generally introduces errors, which can be at a rate higher than a fetal fraction; thus, a determination of a true variation in a given fetal genome may be obscured, or false positives obtained, as a result of such introduced errors (namely, stochastic noise). Even using PCR-free approaches, a final sequencing step employed requires copies of a given original DNA molecule to be made, which can also introduce errors. Such an introduction of errors is particularly relevant when trying to identify *de novo* variants which have occurred in a given fetus.

To address erroneous artefacts that PCR duplicates, analysis protocols typically remove duplicate reads based upon an assumption that two reads with a same given start and a same given end position have arisen from the PCR process, as opposed to being unique DNA molecules; such an assumption is essentially a form of correlation with a purpose of reducing stochastic noise when making measurements. The number of PCR duplicates tends to increase with lower (namely, smaller) starting amounts of DNA (as is the case with cfDNA). However, it has been determined that

unique molecules of cfDNA and cffDNA can have a same start and a same end point, due to preferences of where DNA is sheared, based upon open chromatin regions. This means that by applying a PCR duplicate removal step, 14% of genuine DNA fragments are being discarded (Chan *et al.*, 2016). Thus, such a 'correlation' approach is not without its own problems and inaccuracies.

A technical problem that the present disclosure seeks to address is how to identify unique DNA molecules, wherein the DNA molecules have a mutually same start position and a mutually same end position.

A solution that exists to identify unique DNA molecules is molecular barcoding (MBC); also known as Unique Molecular Indices (UMI). Molecular barcodes enable reads to be identified that have arisen as a result of sequencing error; only reads with a unique molecular barcode should be included in such an analysis. This means PCR artefacts can be removed and real variants kept.

Molecular barcoding of individual DNA molecules can be used to increase confidence in a given variant calling, when an expected frequency of mutant reads is at or below an error rate threshold of a DNA nucleic acid base sequencing method. Moreover, each original DNA fragment in a given sample, when implementing the method, is attached to a unique barcode. This barcode is typically a string of random nucleotides, degenerate nucleotides or defined nucleotides. Reads which contain a sequencing error can be removed from downstream analysis while processing various DNA fragment reads. Such a barcoding approach can account for PCR and sequencer errors, and may potentially improve detection of low allele frequency variants. Based upon this sequencing method using barcodes, the method is potentially susceptible to being used to identify which DNA fragments with mutually identical start and end points are genuine,

biological duplicates, and therefore which can be retained for use in further analysis.`

Known types of molecular barcoding include incorporation into sequencing adapters during library construction (duplex sequencing) (for example, Peng, Vijaya Satya, Lewis, Randad, & Wang, 2015); and smMIP (single molecule molecular inversion probes), wherein a method employs single-molecule tagging, optionally combined with multiplex targeted capture. Specifically, Hiatt, Pritchard, Salipante, O'Roak, & Shendure, (2013) first described this method, wherein 1312 smMIP oligos targeting coding sequences of 33 genes (approx. 125kb) were designed. Furthermore, publications describing smMIP include, a published US patent application US2016/0055293A describing such a method, systems implementing the method, and algorithms and software for MIP design associated with the method; a *BRCA* kit available using method; a published US patent application US2016/0055293A describing such a method, systems implementing the method, and algorithms and software for MIP design associated with the method. Additionally, smMIP for non-invasive prenatal diagnosis (NIPD) is being developed at Maastricht University and Radbound UMC. Such an approach appears presently to be on a single gene basis, rather than a panel of genes. Furthermore, it has been suggested that such an approach is a most favoured option for development as combined barcoding and enrichment, wherein there is focus upon a scalability of target regions.

Known commercially-available customizable methods for providing molecular barcodes include Agilent HaloPlex^{HS}; Agilent whitepaper on molecular barcoding; QiaSeq Targeted DNA Panel and ArcherDX Archer MBC Adapters; these names include trade marks (US: trademarks) ® TM.

Molecules are labelled with a unique sequence prior to performing PCR amplification. There is employed an adapter that contains a sample-specific

index of pre-defined sequences and a typically, but not exclusively, a random 8-mer molecular barcode. This random 8-mer molecular barcode is ligated to what is usually enzymatically fragmented gDNA before amplification. The random 8-mer, along with a random start site generated during the enzymatic shearing, is used to identify duplicates. The cfDNA samples that are of interest, in respect of technology described in the present disclosure, is not subjected to experimental enzymatic fragmentation, but by endogenous enzymatic processes.

Contemporary aforementioned methods that are currently available for molecular barcoding are restricted to a relatively small number of regions of interest in a DNA molecule via PCR amplicon approaches, meaning that associated DNA analysis has to be very targeted in order to achieve useful DNA sequence readout results.

Therefore, in light of the foregoing discussion, there exist problems associated with conventional pre-natal screening systems.

SUMMARY

The present disclosure seeks to provide an improved prenatal screening and diagnostic system that is capable of providing a lower occurrence of false-positives and false-negatives when the screening system is employed for providing a prenatal screening service.

Moreover, the present disclosure seeks to provide an improved method of using a prenatal screening and diagnostic system that is capable of providing a lower occurrence of false-positives and false-negatives when the screening system is employed for providing a prenatal screening service.

In a first aspect, embodiments of the present disclosure provide a prenatal screening and diagnostic system, wherein the prenatal screening and

diagnostic system includes a wet-laboratory arrangement for processing a blood sample to determine cell-free DNA readout data from the blood sample, and a data processing arrangement for processing the cell-free DNA readout data with reference to information stored in a database arrangement to generate a risk score indicative of whether or not there are genetic abnormalities in the blood sample that are indicative of fetal abnormalities, characterized in that the prenatal screening and diagnostic system is operable to ligate nucleic acid base molecular barcodes to fragments of the cell-free DNA present in the blood sample prior to amplifying the molecular barcode-ligated DNA fragments for sequencing the amplified molecular barcode-ligated fragments to generate the cell-free DNA readout data.

The present disclosure is of advantage in that it provides an improved personalized non-invasive system and method of identifying genetic abnormalities in a fetus.

Moreover the system disclosed herein is advantageous because it provides an improved method of identifying genetic abnormalities in a fetus, without an associated increased risk of miscarriage, as well as providing a higher accuracy with less risk of false negative and false positive results.

Embodiments of the disclosure are advantageous in terms of providing a rapid, simple, low cost, patient specific and highly efficient method and system of screening that can efficiently decrease an error that may arise during amplification of DNA sequences. Furthermore, the present disclosure provides a method of identifying unique DNA molecules, wherein the DNA molecules have a mutually same start position and a mutually same end position.

Optionally, the prenatal screening and diagnostic system is operable to implement the molecular barcode as an *n*-mer, wherein *n* is in a range of 3 to 100.

Optionally, *n* is in a range of 4 to 20.

Optionally, the *n* is 10.

Optionally, the molecular barcode includes a random sequence of nucleic acid bases.

More optionally, the fragments of cell-free DNA are generated in the prenatal screening and diagnostic system by endogenous enzymatic digestion.

Yet more optionally, start sites for ligating the molecular barcode are determined by the endogenous enzymatic digestion.

Optionally, the wet-laboratory arrangement is operable to incorporate the molecular barcode to a cell-free DNA library containing a fetal component, and to use the cell-free DNA library thereby obtained in hybridisation-based enrichment for identifying *de novo* variants when computing the risk score.

Optionally, the prenatal screening and diagnostic system is operable to generate the cell-free DNA fragments generated by endogenous enzymatic digestion, to ligate nucleic acid base molecular barcodes to the fragments to generate corresponding barcoded fragments, and to perform enrichment by, for example, hybridization using baits, or primers targeted at genes that are susceptible to causing genetic disease.

More optionally, the prenatal screening and diagnostic system is operable to perform non-invasive molecular diagnosis of a fetus which on ultrasound testing (for example, ultrasound imaging) presents with, for example, a skeletal abnormality and/or cardiac abnormality.

Optionally, the genetic disorder, such as skeletal abnormality and/or cardiac abnormality is caused by a *de novo* mutation.

In a second aspect, embodiments of the present disclosure provide a method of using a prenatal screening and diagnostic system, wherein the prenatal screening and diagnostic system includes a wet-laboratory arrangement for processing a blood sample to determine cell-free DNA readout data from the blood sample, and a data processing arrangement for processing the cell-free DNA readout data with reference to information stored in a database arrangement to generate a risk score indicative of whether or not there are genetic abnormalities in the blood sample that are indicative of genetic abnormalities, characterized in that the method includes:

(i) ligating nucleic acid base molecular barcodes to fragments of the cell-free DNA present in the blood sample;

(ii) amplifying the molecular barcode-ligated fragments; and

(iii) sequencing the amplified molecular barcode-ligated fragments to generate the cell-free DNA readout data.

Optionally, the method includes operating the prenatal screening and diagnostic system to implement the molecular barcode as an *n*-mer, wherein *n* is in a range 3 to 100.

Optionally, the method includes arranging for the molecular barcode to include a random sequence of nucleic acid bases.

Optionally, the method includes using fragments of cell-free DNA in the prenatal screening and diagnostic system generated by enzymatic digestion.

Optionally, the method includes determining start sites for ligating the molecular barcode by the endogenous enzymatic digestion.

More optionally, the method includes operating the wet-laboratory arrangement to incorporate the molecular barcode to a cell-free DNA library

containing a fetal component, and to use the cell-free DNA library thereby obtained in hybridisation-based enrichment for identifying *de novo* variants when computing the risk score.

Yet more optionally, the method includes operating the prenatal screening and diagnostic system utilising the cell-free DNA fragments generated by endogenous enzymatic digestion, to ligate nucleic acid base molecular barcodes to the fragments to generate corresponding barcoded fragments, and to perform enrichment by, for example, hybridization using baits or primers targeted at genes which for one or more diseases that are susceptible to causing genetic disease.

Optionally, the method includes operating the prenatal screening and diagnostic system to perform non-invasive molecular diagnosis of a fetus which on ultrasound testing (for example, ultrasound imaging) presents with an abnormality, for example skeletal and/or cardiac abnormality.

More optionally, abnormality is caused by a *de novo* mutation.

In a third aspect, embodiments of the present disclosure provide a computer program product comprising a non-transitory computer-readable storage medium having computer-readable instructions stored thereon, the computer-readable instructions being executable by a computerized device comprising processing hardware to execute the aforementioned method.

Additional aspects, advantages, features and objects of the present disclosure would be made apparent from the drawings and the detailed description of the illustrative embodiments construed in conjunction with the appended claims that follow.

It will be appreciated that features of the present disclosure are susceptible to being combined in various combinations without departing from the scope of the present disclosure as defined by the appended claims.

A better understanding of the present invention may be obtained through the following examples which are set forth to illustrate, but are not to be construed as limiting the present invention.

BRIEF DESCRIPTION OF DRAWINGS

Embodiments of the present disclosure will be more fully understood from examples described herein below and the accompanying drawings, which is given by way of illustration only, and thus are not limitative of the present invention, and wherein:

FIG. 1 is a block diagram of a prenatal screening and diagnostic system, in accordance with an embodiment of the present disclosure;

FIG. 2 is an illustration of a Kalman filter equivalent representation of the prenatal screening and diagnostic system of FIG. 1, in accordance with an embodiment of the present disclosure;

FIG. 3 is an illustration of molecular barcode-ligated DNA fragments, in accordance with an embodiment of the present disclosure;

FIG. 4 is an illustration of amplified molecular barcode-ligated fragment, in accordance with an embodiment of the present disclosure; and

FIG. 5 is an illustration of steps of a method of using the prenatal screening and diagnostic system of FIG. 1, in accordance with an embodiment of the present disclosure.

In the accompanying diagrams, an underlined number is employed to represent an item over which the underlined number is positioned or an item to which the underlined number is adjacent. A non-underlined number relates to an item identified by a line linking the non-underlined number to the item. When a number is non-underlined and accompanied by an associated arrow, the non-underlined number is used to identify a general item at which the arrow is pointing.

LIST OF ABBREVIATIONS

Abbreviation	Meaning
PCR	Polymerase Chain Reaction
cfDNA	Cell-Free DNA

DEFINITIONS

As used herein, the following terms shall have the following meanings:

As used herein, the term '*data processing arrangement*' refers to a process and/or system that can be embodied in software that determines a biological significance of acquired data (for example, an ultimate result of an assay). For example, a data processing arrangement can determine an amount of each nucleotide sequence species present, based upon the acquired data that is collected. A data processing arrangement may also control an instrument and/or a data collection system based upon results determined. A data processing arrangement and a data collection arrangement are often integrated and provide feedback to operate data acquisition performed by a given instrument, and hence provide assay-based judging methods as provided herein.

As used herein, the term '*database arrangement*' refers to a nucleic acid databases known in the art including, for example, GenBank®, dbEST®, dbSTS®, EMBL® (European Molecular Biology Laboratory) and DDBJ® (DNA Databank of Japan). BLAST® or similar tools can be used to search the identified sequences against a sequence database. ® denotes a registered trade mark (trademark).

As used herein, the term '*genetic information*' refers to information related to nucleic acids, altered nucleotide sequence, chromosomes, segments of chromosomes, polymorphic regions, translocated regions, the like or combinations of the foregoing. Furthermore, the nucleic acids may include, but are not limited to, DNA, cfDNA, cDNA, RNA, mRNA, t RNA and rRNA. Moreover, the genetic information may include information related to mutations, copy number variations, transversions, translocations, inversion, deletions, aneuploidy, partial aneuploidy, polyploidy, chromosomal instability, chromosomal structure alterations, gene fusions, chromosome fusions, gene truncations, gene amplification, gene duplications, chromosomal lesions, DNA lesions, abnormal changes in nucleic acid chemical modifications, abnormal changes in epigenetic patterns, abnormal changes in nucleic acid methylation infection or cancer.

As used herein, the term '*cell-free DNA*' refers to DNA that is not within a cell. In one embodiment, cell free DNA includes DNA circulating in blood. In another embodiment, cell free DNA includes DNA existing outside a cell. In yet another embodiment, cell free DNA includes DNA existing outside a cell as well as DNA present in a blood sample after such a blood sample has undergone partial or gentle cell lysing.

As used herein, the terms '*biological sample*' refers to a sample obtained from a female who is pregnant, wherein the sample may include, but is not limited to, plasma, serum, peripheral blood and urine. Typically, the sample is a maternal plasma sample, although other tissue sources that contain both maternal and fetal DNA are optionally used. Maternal plasma can be obtained from a peripheral whole blood sample from a pregnant woman and the plasma can be obtained by standard methods, for example by employing centrifuging processes. A volume in a range of 3 ml to 5 ml of plasma is typically sufficient to provide a suitable quantity of DNA material for

analysis. The cell free DNA can be extracted from the sample using standard techniques, wherein non-limiting examples of which include a QIASymphony® protocol (Qiagen) suitable for free fetal DNA isolation, or any other automated or manual extraction method suitable for cell free DNA isolation.

As used herein, the term '*biological characteristics*' refers to genetic variations, abnormalities, irregularities or mutations which range extensively from changes in single nucleotides to a presence of additional whole chromosomes or abnormal number of chromosomes. Such a chromosomal abnormality is a structural abnormality, including, but not limited to, copy number changes including microdeletions and microduplications, insertions, translocations, inversions and small-size mutations including point mutations and mutational signatures.

As used herein, the term '*wet-laboratory arrangement*' refers to a facility, clinic and/or a setup of: instruments, equipment and/or devices used for extraction, collection, processing and/or analysis of body fluid samples; instruments, equipment and/or devices used for extraction, collection, processing and/or analysis of genetic material; instruments, equipment and/or devices used for amplification, enrichment and/or processing of genetic material received from the body fluid samples; instruments, equipment and/or devices used for extraction and/or analysis of the genetic information received from the amplified genetic material. Herein the instruments, equipment and/or devices may include but not limited to centrifuge, ELISA®, spectrophotometer, PCR, RT- PCR, High-Throughput-Screening (HTS) system, Microarray system, ultrasound (scanning), genetic analyser, deoxyribonucleic acid (DNA) sequencer and SNP analyser. The wet-laboratory arrangement is operable to monitor and/or scan a given fetus. Herein, the wet-laboratory arrangement may include equipment,

instruments and/or devices for scanning the fetus; such examples include an ultrasound scan, presymptomatic genetic testing and/or combined tests.

As used herein, '*polymerase chain reaction (PCR)*' is a technique used in molecular biology to amplify a single copy, or a few copies, of a segment of DNA by several orders of magnitude, thereby generating potentially thousands of millions of copies of a particular given DNA sequence.

As used herein, '*bridge amplification*' or '*amplification*' is employed in massively parallel sequencing for DNA sequencing purposes using a concept of massively parallel processing, wherein use is made of miniaturized and parallelized platforms for sequencing of 1 million to 43 billion short reads (50 to 400 nucleic acid bases each) per instrument run.

As used herein, the term '*baits*' refers to a bioactive molecule which is used to detect other bioactive molecules such as genes of interest or target genes. This bait design, or primer design, will be in combination with the targeting of genes which are relevant to monogenic clinical disorders and for the enrichment of fetal DNA from a maternal plasma sample. The baits for example or primers are, prepared beforehand, and are optionally selected from a library of prepared baits or primers.

DETAILED DESCRIPTION

Practical implementation of the embodiments of the present disclosure are described in further detail below; these embodiments are operable to employ, unless otherwise indicated, conventional methods of diagnostics, molecular biology, cell biology, biochemistry and immunology within the skill of the art. Such techniques are explained fully in the literature, for example contemporary academic research literature pertaining to pregnancy and genetic material processing.

It will be appreciated that certain features of the present invention, which are for clarity described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the invention, which are for brevity, described in the context of a single embodiment, may also be provided separately and/or in any suitable sub-combination.

The following detailed description illustrates embodiments of the present disclosure and ways in which they can be implemented. Although some modes of carrying out the present disclosure has been disclosed, those skilled in the art would recognize that other embodiments for carrying out or practicing the present disclosure are also possible.

Referring to FIG. 1, there is shown a block diagram of a prenatal screening and diagnostic system **100**, in accordance with an embodiment of the present disclosure. The prenatal screening and diagnostic system **100** includes a wet-laboratory arrangement **102**, wherein the wet-laboratory arrangement **102** includes apparatus such as blood sample collection apparatus, centrifuges, PCR rapid gene sequencing apparatus and similar apparatuses. For example, the wet-laboratory arrangement **102** includes apparatus manufactured by Illumina Inc. for performing gene sequencing tasks. Furthermore, the prenatal screening and diagnostic system **100** is operable to process a blood sample in the wet-laboratory arrangement **102** to obtain cell-free DNA readout data therefrom.

In an embodiment, the prenatal screening and diagnostic system **100** is operable to perform non-invasive molecular diagnosis of a fetus which, on ultrasound testing (for example, ultrasound scanning), presents with (but not limited to) a skeletal abnormality and/or a cardiac abnormality. Optionally, the wet-laboratory arrangement **102** is operable to perform non-invasive molecular diagnosis of a fetus, such as ultrasound testing, to detect

fetal abnormalities, such as a skeletal abnormality and/or a cardiac abnormality. Furthermore, the prenatal screening and diagnostic system **100**, optionally, the wet-laboratory arrangement **102** is operable to generate an ultrasonic image or video of the fetus, to deduce the possibility of a fetal abnormality identified from the ultrasonic test.

Optionally, the wet-laboratory arrangement **102** may be operable to perform a combined test for prenatal screening of fetal chromosomal abnormalities. More optionally, the combined test may include, but is not limited to, a maternal blood test and an ultrasound scan of a fetus. Furthermore, the wet-laboratory arrangement **102** is operable to provide information representative of the combined test of the fetus.

In operation, the blood sample is obtained from a person, for example a pregnant mother. Optionally, with regard to the pregnant mother, the blood sample is a non-invasive sample, wherein collection of sample does not have an associated risk of miscarriage therewith. Furthermore, the blood sample includes plasma that includes, as a component part thereof, a mixture of cell-free DNA (cfDNA). Specifically, the cell-free DNA (cfDNA) may comprise a portion derived from the pregnant mother, from the placenta of the pregnant mother and/or from a fetus of the pregnant mother.

In an embodiment, a genetic abnormality may include genetic diseases that are present in the DNA sequences of a given mother. Specifically, such genetic diseases may or may not be inherited by a fetus of the given mother. Additionally, abnormalities may include diseases that may be inherited or that may arise *de novo* in the fetus.

Moreover, the prenatal screening and diagnostic system **100** further includes a data processing arrangement **104**, including a database arrangement **106**, for receiving cell-free DNA readout data from the wet-laboratory arrangement **102**. Optionally, the data processing arrangement

104 provides feedback data to the wet-laboratory arrangement **102** for controlling various tests performed thereat. Furthermore, the database arrangement **106** stores information comprising genomic mapping data and research data analysing structure, location and sequencing of human genes, and clinical effects of mutations and their co-relation with biological sequences and structures. Furthermore, the wet-laboratory arrangement **102** may be operable to amplify the fragments of DNA to provide amplified DNA for nucleic acid base sequencing or readout to generate the cell-free DNA readout data. In this exemplary embodiment, the wet-laboratory arrangement **102** may include a PCR or RT-PCR for amplifying the free fetal DNA fragments for providing a plurality of copies of the free fetal DNA to the data processing arrangement **104** for accessing genetic information in the database arrangement **106**. Additionally, the data processing arrangement **104** also includes data communication connections to networks such as the Internet®, for example for accessing various external databases associated with university research departments and hospitals.

Furthermore, the data processing arrangement **104** is operable to process the cell-free DNA readout data with reference to information stored in a database arrangement **106** to generate a risk score indicative of whether or not there are genetic abnormalities in the blood sample that are indicative, for example, of fetal abnormalities. Specifically, a risk score may be associated with a given fetus, wherein a higher risk score is indicative a higher possibility of a genetic abnormality. Furthermore, the risk score is generated after processing of cell-free DNA readout data with reference to information stored in the database arrangement **106**. Specifically, the cell-free DNA readout data may correspond to a given genomic information in the database arrangement **106**. Furthermore, such genomic information may be linked with a risk of a given genetic abnormality, as aforementioned. In an exemplary embodiment, the cell-free DNA readout data may comprise

a sequential arrangement of 'P–Q–R–S–P–Q–R–S' DNA base pairs with an anomaly 'P–R–Q–S'. In such an embodiment, the data processing arrangement **102** may compare the anomaly against sequential arrangements of DNA stored in the database arrangement **106**. Subsequently in the embodiment, the data processing arrangement **104** may assess if the anomaly may or may not cause a genetic disorder. Additionally, the data processing arrangement **104** may compare and provide the risk score representative of a risk to the fetus of inheriting or acquiring the genetic disorder. It will be appreciated that the DNA base pairs P, Q, R, S represent DNA base pairs adenine, thymine, guanine and cytosine for illustrative purposes only and do not represent the actual arrangement of the DNA base pairs which may be responsible for a specific disease.

In an embodiment, fragments of cell-free DNA are generated in the prenatal screening and diagnostic system **100** by endogenous enzymatic digestion. More specifically, strands of DNA may be fragmented (namely, cleaved) using enzymes. Furthermore, action sites of enzymes on the cell-free DNA may not be experimentally controlled. Examples of enzymes include, but are not limited to, cyanogen bromide (CNBr), N-bromosuccinimide.

Referring to FIG. 2, there is shown an illustration of a Kalman filter equivalent representation **200** of the prenatal screening and diagnostic system (such as the prenatal screening and diagnostic system **100** of FIG. 1), in accordance with an embodiment of the present disclosure. The Kalman filter equivalent representation **200** includes the cell-free DNA readout data **202** and the information representative of combined test of the fetus **204** to a data processing arrangement **206** (such as data processing arrangement **104** of FIG. 1). The data processing arrangement **206** may be operable to implement a Kalman filter on the cell-free DNA readout data **202**

information representative of combined test of the fetus **204**. The data processing arrangement **206** further includes a fuzzy logic module **208**, a processing module **210**, a genetic algorithm **212** for processing the cell-free DNA readout data with reference to the information stored in a database arrangement **214** (such as the database arrangement **106** of FIG. 1), a secondary database **216** (such as secondary database for storing the risk score **218** received from the processing module **210**. In this embodiment, the data processing system **206** may be operable to implement the Kalman filter on the genetic information received after prenatal screening tests performed by the wet-laboratory arrangement **102**. Furthermore, the genetic algorithm **212** may be operable to generate risk score by processing the cell-free DNA readout data with reference to the information stored in a database arrangement **214**.

In an exemplary embodiment, plasma sample derived from the aforementioned blood sample includes DNA sequences that are enriched for example by, but not limited to, hybridization or primer annealing. Specifically, the hybridization or primer-based enrichment is performed using baits or primers targeted at genes that are susceptible to causing genetic disease. In this embodiment, the processing module **210** is operable to validate target positions of genes that are susceptible to causing genetic diseases.

In an embodiment, the prenatal screening and diagnostic system **100** may be operable to differentiate maternal and fetal components of cell-free DNA. In this embodiment, such differentiation may be achieved by employing an assay design which enriches the fetal component and which aids in mapping of maternal and fetal reads.

In another exemplary embodiment, the prenatal screening and diagnostic system **100** may be operable to design baits or primers and to employ these

baits or primers at targeted positions on the genes that are susceptible to causing genetic diseases. Furthermore, the prenatal screening and diagnostic system **100** may be operable to avoid the maternal-specific regions in the blood sample.

Referring to FIG. 3, there is shown an illustration of molecular barcode-ligated DNA fragments **300**, in accordance with an embodiment of the present disclosure. The prenatal screening and diagnostic system **100** is operable to ligate nucleic acid base molecular barcodes **302** to fragments of the cell-free DNA **304** present in the blood sample. Optionally, the nucleic acid base molecular barcodes **302** may be ligated to fragments of cell-free DNA **304** and may be followed by subsequent enrichment using baits for hybridisation or primers targeted at genes that are susceptible to causing genetic disease.

In an embodiment, the prenatal screening and diagnostic system **100** is operable to implement the molecular barcode **302** as an *n*-mer. Optionally, *n* is in a range of 3 to 100. More optionally, *n* is in a range of 4 to 20. Yet more optionally, *n* is substantially 10. Specifically, the molecular barcode **302** may be implemented in a range of 3-mer to 100-mer. Furthermore, the molecular barcode **302** may be synthesized during the ligation process.

In an embodiment, the molecular barcode **302** includes a random sequence of nucleic acid bases. Specifically, the nucleic acid bases include adenine (A), cytosine (C), guanine (G), thymine (T). Furthermore, the molecular barcode **302** includes the random sequence of nucleic acid bases for identification of targeted genes that are susceptible to causing genetic disease.

In an embodiment, the molecular barcode **302** includes adapters (namely, linkers). Specifically, adapters are short, chemically synthesized, single-stranded or double-stranded oligonucleotide. More specifically, such

adapters may be comprised in the molecular barcode **302** and may facilitate ligation thereof.

According to an embodiment, start sites for ligating the molecular barcode **302** are determined by endogenous enzymatic digestion. As aforementioned, fragments of cell-free DNA are generated in the prenatal screening and diagnostic system **100** by endogenous enzymatic digestion. Furthermore, start sites may be generated on the fragments of cell-free DNA during enzymatic shearing of the cell-free DNA.

Referring to FIG. 4, there is shown an illustration of amplified molecular barcode-ligated fragment **400**, in accordance with an embodiment of the present disclosure. The prenatal screening and diagnostic system **100** is operable to amplify the molecular barcode-ligated DNA fragments **300** for sequencing the amplified molecular barcode-ligated fragments **400**. Specifically, the molecular barcode-ligated DNA fragments **300** are amplified by the prenatal screening arrangement **100**. Optionally, the amplification may include using a Polymerase Chain Reaction (PCR) technique. Specifically, such amplification techniques may amplify a single copy or a few copies of molecular barcode-ligated DNA fragments **300** by several orders of magnitude, thereby generating potentially thousands of millions of copies of the particular given DNA sequence. Furthermore, such amplification techniques may provide an error, such as duplication of a nucleic acid base, in such an amplification process, which may be incorrectly represented as indicative of a genetic abnormality. Furthermore, such error may be corrected during sequencing of the amplified molecular barcode-ligated fragments **400**. Consequently, the cell-free DNA readout data generated from sequencing process of the amplified molecular barcode-ligated fragments **400** may take into account the amplification error when generating the risk score.

In an embodiment, the amplified molecular barcode-ligated fragments **400** comprise the molecular barcodes **402** (such as the molecular barcodes **302**), fragments of the cell-free DNA **404** (such as the fragments of the cell-free DNA **304**). Furthermore, the amplified molecular barcode-ligated fragments **400** may comprise a sample-specific index **408**. Specifically, the sample-specific index **408** comprises a pre-defined sequence and a random 8-mer molecular barcode. Furthermore, if an error, such as the duplication of a nucleic acid base, is generated during the amplification process, the sample-specific index may be used during sequencing to identify the amplification error. Additionally, sites **410** may represent sites for amplification, wherein amplified DNA may be attached to the sites **410**. Alternatively or additionally, optionally, the sites **410** may comprise baits or primers used to perform enrichment of targeted genes that are susceptible to causing genetic disease.

The amplified molecular barcode-ligated fragments **400** are sequenced to generate cell-free DNA readout data. Optionally, sequencing process may account for errors generated during an amplification process or processes. More optionally, the cell-free DNA readout data is indicative of whether or not there are genetic abnormalities in the blood sample. Optionally, the skeletal abnormality and/or the cardiac abnormality is caused by a *de novo* mutation. Furthermore, the molecular barcode-ligated fragments are useful to employ for reducing stochastic noise (namely, stochastic error) generated during aforementioned sequencing process and/or during processing in the data processing arrangement **104**.

Additionally, the cell-free DNA readout data, generated subsequent to ligation, amplification and sequencing processes is used to generate a risk score which provides an accurate indication of whether or not there are genetic abnormalities in the blood sample.

In an embodiment, the wet-laboratory arrangement **102** is operable to incorporate the molecular barcode to a cell-free DNA library containing a fetal component, and to use the cell-free DNA library thereby obtained in hybridisation-based enrichment for identifying *de novo* variants when computing the risk score. Specifically, the wet-laboratory arrangement is operable to prepare a cell-free DNA library comprising information about cell-free DNA readout data and the molecular barcodes. Furthermore, the cell-free DNA library may be used in achieving a higher accuracy of identification of *de novo* variants in a given fetal DNA.

Referring to FIG. 5, there is shown an illustration of steps of a method **500** of using a prenatal screening and diagnostic system (such as the prenatal screening and diagnostic system **100** of FIG. 1), in accordance with an embodiment of the present disclosure. At a step **502**, a blood sample is processed to determine cell-free DNA readout data from the blood sample. At a step **504**, the cell-free DNA readout data is processed with reference to information stored in a database arrangement to generate a risk score indicative of whether or not there are genetic abnormalities in the blood sample that are indicative of fetal abnormalities. At a step **506**, nucleic acid base molecular barcodes are ligated to fragments of the cell-free DNA present in the blood sample. At a step **508**, the molecular barcode-ligated fragments are amplified. At a step **510**, the amplified molecular barcode-ligated fragments are sequenced to generate the cell-free DNA readout data.

The steps **502** to **510** are only illustrative and other alternatives can also be provided where one or more steps are added, one or more steps are removed, or one or more steps are provided in a different sequence without departing from the scope of the claims herein. Optionally, the method **500** includes operating the prenatal screening and diagnostic system to implement the molecular barcode as an *n*-mer, wherein *n* is in a range 3 to 100. Optionally, the method **500** includes arranging for the molecular

barcode to include a random sequence of nucleic acid bases. More optionally, the method **500** includes using fragments of cell-free DNA in the prenatal screening and diagnostic system generated by enzymatic digestion. Yet more optionally, the method **500** includes determining start sites for ligating the molecular barcode by the enzymatic digestion.

Optionally, the method **500** includes operating the wet-laboratory arrangement to incorporate the molecular barcode to a cell-free DNA library containing a fetal component, and to use the cell-free DNA library thereby obtained in hybridisation-based enrichment for identifying *de novo* variants when computing the risk score. Optionally, the method **500** includes operating the prenatal screening and diagnostic system utilising the cell-free DNA fragments generated by endogenous enzymatic digestion, to ligate nucleic acid base molecular barcodes to the fragments to generate corresponding barcoded fragments, and to perform enrichment by hybridization using baits targeted at genes which for one or more diseases that are susceptible to causing genetic disease. More optionally, the method **500** includes operating the prenatal screening and diagnostic system to perform non-invasive molecular diagnosis of a fetus which on ultrasound investigation (for example, non-invasive ultrasound imaging) presents with, for example, a skeletal abnormality and/or a cardiac abnormality. Yet more optionally, the genetic abnormality, for example, skeletal and/or cardiac abnormality is caused by a *de novo* mutation.

Optionally, the aforementioned method **500** of using the prenatal screening and diagnostic system is implemented by using a computer program product comprising a non-transitory computer-readable storage medium having computer-readable instructions stored thereon, the computer-readable instructions being executable by a computerized device comprising processing hardware.

Although use of the prenatal screening and diagnostic system **100** described in the foregoing, it will be appreciated that the prenatal screening and diagnostic system may be used for investigating other types of biological problems, and not merely restricted to prenatal screening tasks, for example: cancer risk determination; autistic risk determination; verification of organism performance after performing gene therapy; ionizing radiation damage identification to cell DNA; and/or diabetes risk determination.

Modifications to embodiments described in the foregoing are possible without departing from the scope of the invention as defined by the accompanying claims. Expressions such as "including", "comprising", "incorporating", "consisting of", "have", "is" used to describe and claim the present invention are intended to be construed in a non-exclusive manner, namely allowing for items, components or elements not explicitly described also to be present. Reference to the singular is also to be construed to relate to the plural. Numerals included within parentheses in the accompanying claims are intended to assist understanding of the claims and should not be construed in any way to limit subject matter claimed by these claims.

REFERENCES

[1] Chan, L.L., and Jiang, P. (2015). Bioinformatics analysis of circulating cell-free DNA sequencing data. Clin. Biochem. *48*, pp 962–975.

[2] Chan, K.C.A., Jiang, P., Sun, K., Cheng, Y.K.Y., Tong, Y.K., Cheng, S.H., Wong, A.I.C., Hudecova, I., Leung, T.Y., Chiu, R.W.K., *et al.* (2016). Second generation noninvasive fetal genome analysis reveals *de novo* mutations, single-base parental inheritance, and preferred DNA ends. Proc. Natl. Acad. Sci. U. S. A. 201615800.

[3] Chiu, R.W.K., Chan, K.C.A., Gao, Y., Lau, V.Y.M., Zheng, W., Leung, T.., Foo, C.H.F., Xie, B., Tsui, N.B.Y., Lun, F.M.F., et al. (2008). Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. Proc. Natl. Acad. Sci. U. S. A. *105*, pp 20458–20463.

[4] Kitzman, J.O., Snyder, M.W., Ventura, M., Lewis, A.P., Simmons, L.E., Gammill, H.S., Rubens, C.E., Santillan, D. a., Murray, J.C., Tabor, H.K., et al. (2012). Non-invasive Whole Genome Sequencing of Human Fetus. Sci. Transl. Med. *4*, pp 1–18.

[5] Lam, K.W.G., Jiang, P., Liao, G.J.W., Chan, K.C.A., Leung, T.Y., Chiu, R.W.K., and Lo, Y.M.D. (2012). Noninvasive prenatal diagnosis of monogenic diseases by targeted massively parallel sequencing of maternal plasma: Application to ??-thalassemia. Clin. Chem. *58*, pp 1467–1475.

[6] Lo, Y.M.D., Chan, K.C.A., Sun, H., Chen, E.Z., Jiang, P., Lun, F.M.F., Zheng, Y.W., Leung, T.Y., Lau, T.K., Cantor, C.R., et al. (2010). Maternal Plasma DNA Sequencing Reveals the Genome-Wide Genetic and Mutational Profile of the Fetus. Sci. Transl. Med. *2*, 61ra91-61ra91.

[7] New, M.I., Tong, Y.K., Yuen, T., Jiang, P., Pina, C., Chan, K.C.A., Khattab, A., Liao, G.J.W., Yau, M., Kim, S.-M., et al. (2014). Noninvasive prenatal diagnosis of congenital adrenal hyperplasia using cell-free fetal DNA in maternal plasma. J. Clin. Endocrinol. Metab. *99*, E1022-30.

[8] Xiong, L., Barrett, A.N., Hua, R., Tan, T.Z., Ho, S.S.Y., Chan, J.K.Y., Zhong, M., and Choolani, M. (2015). Non-invasive prenatal diagnostic testing for β-thalassaemia using cell-free fetal DNA and next generation sequencing. Prenat. Diagn. *35*, pp 258–265.

CLAIMS

We claim:

1. A prenatal screening and diagnostic system (100), wherein the prenatal screening and diagnostic system (100) includes a wet-laboratory arrangement (102) for processing a blood sample to determine cell-free DNA readout data from the blood sample, and a data processing arrangement (104) for processing the cell-free DNA readout data with reference to information stored in a database arrangement (106) to generate a risk score indicative of whether or not there are genetic abnormalities in the blood sample that are indicative of abnormalities, including fetal abnormalities, characterized in that the prenatal screening and diagnostic system (100) is operable to ligate nucleic acid base molecular barcodes to fragments of the cell-free DNA present in the blood sample prior to amplifying the molecular barcode-ligated DNA fragments for sequencing the amplified molecular barcode-ligated fragments to generate the cell-free DNA readout data.

2. A prenatal screening and diagnostic system (100) of claim 1, characterized in that the prenatal screening and diagnostic system (100) is operable to implement the molecular barcode as an *n*-mer, wherein *n* is in a range of 3 to 100.

3. A prenatal screening and diagnostic system (100) of claim 2, characterized in that *n* is in a range of 4 to 20.

4. A prenatal screening and diagnostic system (100) of claim 2, characterized in that the *n* is 10.

5. A prenatal screening and diagnostic system (100) of claim 1 or 2, characterized in that the molecular barcode includes a random sequence of nucleic acid bases.

6. A prenatal screening and diagnostic system (100) of any one of the preceding claims, characterized in that the fragments of cell-free DNA are generated in the prenatal screening and diagnostic system (100) by endogenous enzymatic digestion.

7. A prenatal screening and diagnostic system (100) of claim 6, characterized in that start sites for ligating the molecular barcode are determined by the enzymatic digestion.

8. A prenatal screening and diagnostic system (100) of any one of the preceding claims, characterized in that the wet-laboratory arrangement (102) is operable to incorporate the molecular barcode to a cell-free DNA library containing a fetal component, and to use the cell-free DNA library thereby obtained in hybridisation-based enrichment, or primer-based enrichment for identifying *de novo* variants when computing the risk score.

9. A prenatal screening and diagnostic system (100) of any one of the preceding claims, characterized in that the prenatal screening and diagnostic system (100) is operable to generate the cell-free DNA fragments generated by endogenous enzymatic digestion, to ligate nucleic acid base molecular barcodes to the fragments to generate corresponding barcoded fragments, and to perform enrichment by hybridization using baits or primer-based enrichment targeted at genes that are susceptible to causing genetic diseases.

10. A prenatal screening and diagnostic system (100) of any one of the preceding claims, characterized in that the prenatal screening and diagnostic system (100) is operable to perform non-invasive molecular diagnosis of a fetus which may on ultrasound investigation present with an abnormality, such as a skeletal and/or a cardiac abnormality.

11. A prenatal screening and diagnostic system (100) of claim 8, characterized in that the abnormality is caused by a *de novo* mutation, or that the de novo mutation could cause an abnormality.

12. A method of using a prenatal screening and diagnostic system (100), wherein the prenatal screening and diagnostic system (100) includes a wet-laboratory arrangement (102) for processing a blood sample to determine cell-free DNA readout data from the blood sample, and a data processing arrangement (104) for processing the cell-free DNA readout data with reference to information stored in a database arrangement (106) to generate a risk score indicative of whether or not there are genetic abnormalities in the blood sample that are indicative of genetic disorders, characterized in that the method includes:

(i) ligating nucleic acid base molecular barcodes to fragments of the cell-free DNA present in the blood sample;

(ii) amplifying the molecular barcode-ligated fragments; and

(iii) sequencing the amplified molecular barcode-ligated fragments to generate the cell-free DNA readout data.

13. A method of claim 12, characterized in that the method includes operating the prenatal screening and diagnostic system (100) to implement the molecular barcode as an *n*-mer, wherein *n* is in a range 3 to 100.

14. A method of claim 12 or 13, characterized in that the method includes arranging for the molecular barcode to include a random sequence of nucleic acid bases.

15. A method of any one of claims 12 to 14, characterized in that the method includes using fragments of cell-free DNA in the prenatal screening and diagnostic system (100) generated by endogenous enzymatic digestion.

16. A method of claim 15, characterized in that the method includes determining start sites for ligating the molecular barcode by the enzymatic digestion.

17. A method of any one of claims 12 to 16, characterized in that the method includes operating the wet-laboratory arrangement (102) to incorporate the molecular barcode to a cell-free DNA library containing a fetal component, and to use the cell-free DNA library thereby obtained in hybridisation-based, or primer-based enrichment for identifying *de novo* variants when computing the risk score.

18. A method of any one of claims 12 to 17, characterized in that the method includes operating the prenatal screening and diagnostic system (100) utilising the cell-free DNA fragments generated by endogenous enzymatic digestion, to ligate nucleic acid base molecular barcodes to the fragments to generate corresponding barcoded fragments, and to perform enrichment by hybridization using baits targeted at genes, or primers targeted at genes for that are susceptible to causing one or more diseases

19. A method of any one of claims 12 to 18, characterized in that the method includes operating the prenatal screening and diagnostic system (100) to perform non-invasive molecular diagnosis of a fetus which may on ultrasound present with an abnormality, such as skeletal or cardiac abnormality, or which may be at risk of developing an abnormality associated with genetic disease

20. A method of claim 19, characterized in that the abnormality, for example, skeletal and/or cardiac abnormality is caused by a *de novo* mutation.

21. A computer program product comprising a non-transitory computer-readable storage medium having computer-readable instructions stored

thereon, the computer-readable instructions being executable by a computerized device comprising processing hardware to execute a method as claimed in any one of claims 12 to 20.

Intellectual Property Office

Application No:	GB1711540.3	**Examiner:**	Dr Philip Mountjoy
Claims searched:	1-21	**Date of search:**	30 April 2018

Patents Act 1977: Search Report under Section 17

Documents considered to be relevant:

Category	Relevant to claims	Identity of document and passage or figure of particular relevance
X	1-21	PLoS One, Vol. 8, No. 6, Jun. 2013, Ge Huijuan, et al., "Noninvasive Prenatal Detection for Pathogenic CNVs: The Application in alpha-Thalassemia", Article No.: e67464. See entire document, particularly the abstract, page 3 of 7 (column 2), page 5 of 7, column 1, to page 6, column 1.
X	1-21	Fetal Diagn Ther., Vol. 42, No. 3, epub Feb. 2017, F Crea et al., "The IONA® Test: Development of an Automated Cell-Free DNA-Based Screening Test for Fetal Trisomies 13, 18, and 21 That Employs the Ion Proton Semiconductor Sequencing Platform.", pages 218-224. See entire document, particularly the abstract, page 220 (column 1), page 22 (column 1), and page 223 (column 1).
X	1-21	WO2014/149134 A2 (GUARDANT HEALTH INC) - See entire document, particularly the abstract, paragraphs [0013], [0084], [0094]-[0096], [00225], [00232], [00237]-[00239], and the claims.
X	1-21	US2016/251704 A1 (TALASAZ et al.) - See entire document, particularly paragraph [0095], [0175], [0208], [0237]-[0238], [0321], and the claims.
X	1-21	Prenatal Diagnosis, Vol. 36, No. 1, Jan. 2016, Tynan J A, et al., "Application of risk score analysis to low-coverage whole genome sequencing data for the noninvasive detection of trisomy 21, trisomy 18, and trisomy 13.", pages 56-62. See entire document, particularly the abstract, and page 57, column 2.
X,E	1-20	WO2018/049049 A1 (BAYLOR COLLEGE OF MEDICINE) - See entire document, particularly the abstract, paragraphs [0006]-[0008], [0020], [0045]-[0047], [0049], [0055], and the claims.

Categories:

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

Field of Search:

Search of GB, EP, WO & US patent documents classified in the following areas of the UKC^{X} :

Worldwide search of patent documents classified in the following areas of the IPC

The following online and other databases have been used in the preparation of this search report

EPODOC, WPI, BIOSIS, MEDLINE, INTERNET

International Classification:

Subclass	Subgroup	Valid From
C12Q	0001/6855	01/01/2018
C12Q	0001/6809	01/01/2018
C12Q	0001/6827	01/01/2018
C12Q	0001/6869	01/01/2018
C12Q	0001/6883	01/01/2018
G06F	0019/22	01/01/2011