



<b>Title</b>	<b>Non-invasive prenatal assessment of trisomy 21 by multiplexed maternal plasma DNA sequencing: large scale validity study.</b>
<b>Author(s)</b>	<b>Chiu, RWK; Akolekar, R; Zheng, YWL; Leung, TY; Sun, H; Chan, KCA; Lun, FMF; Go, ATJI; Lau, ET; To, WWK; Leung, WC; Tang, RYK; Au-Yeung, SKC; Lam, H; Kung, YY; Zhang, X; Van Vugt, JMG; Minekawa, R; Tang, MHY; Wang, J; Oudejans, CBM; Lau, TK; Nicolaidis, KH; Lo, YMD</b>
<b>Citation</b>	<b>BMJ (Clinical Research Ed.), 2011, v. 342, article no. c7401</b>
<b>Issued Date</b>	<b>2011</b>
<b>URL</b>	<b><a href="http://hdl.handle.net/10722/180701">http://hdl.handle.net/10722/180701</a></b>
<b>Rights</b>	<b>Creative Commons: Attribution 3.0 Hong Kong License</b>

# Non-invasive prenatal assessment of trisomy 21 by multiplexed maternal plasma DNA sequencing: large scale validity study

Rossa W K Chiu, professor,<sup>1</sup> Ranjit Akolekar, clinical research fellow,<sup>3</sup> Yama W L Zheng, student,<sup>1</sup> Tak Y Leung, professor,<sup>2</sup> Hao Sun, assistant professor,<sup>1</sup> K C Allen Chan, associate professor,<sup>1</sup> Fiona M F Lun, postdoctoral fellow,<sup>1</sup> Attie T J I Go, professor,<sup>4</sup> Elizabeth T Lau, department manager and honorary assistant professor,<sup>5</sup> William W K To, consultant,<sup>6</sup> Wing C Leung, consultant,<sup>7</sup> Rebecca Y K Tang, consultant,<sup>8</sup> Sidney K C Au-Yeung, consultant,<sup>9</sup> Helena Lam, consultant,<sup>10</sup> Yu Y Kung, obstetrician,<sup>11</sup> Xiuqing Zhang, manager,<sup>12,13</sup> John M G van Vugt, professor,<sup>4</sup> Ryoko Minekawa, postdoctoral fellow,<sup>3</sup> Mary H Y Tang, consultant and honorary clinical associate professor,<sup>5</sup> Jun Wang, professor,<sup>12</sup> associate director,<sup>13</sup> Cees B M Oudejans, associate professor,<sup>4</sup> Tze K Lau, professor,<sup>2</sup> Kypros H Nicolaidis, professor,<sup>3</sup> Y M Dennis Lo, professor<sup>1,12</sup>

<sup>1</sup>Centre for Research into Circulating Fetal Nucleic Acids, Li Ka Shing Institute of Health Sciences, Department of Chemical Pathology, The Chinese University of Hong Kong, Hong Kong SAR, China

<sup>2</sup>Department of Obstetrics and Gynaecology, The Chinese University of Hong Kong

<sup>3</sup>Harris Birthright Research Centre for Foetal Medicine, King's College Hospital, London SE5 9RS, UK

<sup>4</sup>VU University Medical Center, 10081 HV Amsterdam, Netherlands

<sup>5</sup>Tsan Yuk Hospital, Department of Obstetrics and Gynaecology, University of Hong Kong, Hong Kong

<sup>6</sup>United Christian Hospital, Hospital Authority, Hong Kong

<sup>7</sup>Kwong Wah Hospital, Hospital Authority, Hong Kong

<sup>8</sup>Pamela Youde Nethersole Eastern Hospital, Hospital Authority, Hong Kong

<sup>9</sup>Tuen Mun Hospital, Hospital Authority, Hong Kong

<sup>10</sup>Princess Margaret Hospital, Hospital Authority, Hong Kong

<sup>11</sup>YY Kung Medical Centre, Hong Kong

<sup>12</sup>Joint Chinese University of Hong Kong-Beijing Genomics Institute Genome Research Centre, Hong Kong

<sup>13</sup>Beijing Genomics Institute at Shenzhen, Shenzhen, China

Correspondence to: Y M D Lo loym@cuhk.edu.hk

Cite this as: *BMJ* 2011;342:c7401 doi:10.1136/bmj.c7401

## ABSTRACT

**Objectives** To validate the clinical efficacy and practical feasibility of massively parallel maternal plasma DNA sequencing to screen for fetal trisomy 21 among high risk pregnancies clinically indicated for amniocentesis or chorionic villus sampling.

**Design** Diagnostic accuracy validated against full karyotyping, using prospectively collected or archived maternal plasma samples.

**Setting** Prenatal diagnostic units in Hong Kong, United Kingdom, and the Netherlands.

**Participants** 753 pregnant women at high risk for fetal trisomy 21 who underwent definitive diagnosis by full karyotyping, of whom 86 had a fetus with trisomy 21.

**Intervention** Multiplexed massively parallel sequencing of DNA molecules in maternal plasma according to two protocols with different levels of sample throughput: 2-plex and 8-plex sequencing.

**Main outcome measures** Proportion of DNA molecules that originated from chromosome 21. A trisomy 21 fetus was diagnosed when the z score for the proportion of chromosome 21 DNA molecules was  $>3$ . Diagnostic sensitivity, specificity, positive predictive value, and negative predictive value were calculated for trisomy 21 detection.

**Results** Results were available from 753 pregnancies with the 8-plex sequencing protocol and from 314 pregnancies with the 2-plex protocol. The performance of the 2-plex protocol was superior to that of the 8-plex protocol. With the 2-plex protocol, trisomy 21 fetuses were detected at 100% sensitivity and 97.9% specificity, which resulted in a positive predictive value of 96.6% and negative predictive value of 100%. The 8-plex protocol detected 79.1% of the trisomy 21 fetuses and 98.9% specificity, giving a positive predictive value of 91.9% and negative predictive value of 96.9%.

**Conclusion** Multiplexed maternal plasma DNA sequencing analysis could be used to rule out fetal trisomy 21 among high risk pregnancies. If referrals for amniocentesis or chorionic villus sampling were based on the sequencing test results, about 98% of the invasive diagnostic procedures could be avoided.

## INTRODUCTION

Trisomy 21, Down's syndrome, occurs in 1 in 800 live births.<sup>1</sup> Prenatal diagnosis of trisomy 21 requires invasive sampling of fetal genetic material through amniocentesis or chorionic villus sampling. However, these tests carry a risk of miscarriage of about 1%,<sup>2</sup> and they are therefore reserved for pregnancies considered to be at high risk of fetal trisomy 21. The traditional method of identifying the high risk group has been increased maternal age, but screening by this method would require invasive testing in about 5% of pregnant women and identify only 30% of affected fetuses.<sup>1</sup> In the past 20 years maternal age has been combined with ultrasonographic examination of the fetus and biochemical measurement of various proteins or hormones in the maternal circulation to improve identification of high risk pregnancies. This combined approach of screening can now identify more than 90% of affected fetuses, but there is still a need for invasive testing in 3–5% of the population.<sup>3–9</sup>

Cell-free DNA from the fetus has been found in the plasma of pregnant women, and this has been used successfully for non-invasive determination of the fetal sex and fetal RhD genotype in RhD negative women.<sup>10–13</sup> The basis of these tests is the detection of fetal-specific DNA sequences in maternal plasma.<sup>14</sup> The same approach of searching for fetal-specific nucleic acids, such as DNA methylation and mRNA

markers in maternal plasma, has been proposed for non-invasive detection of fetal aneuploidies.<sup>15-19</sup>

The advantages and disadvantages of these approaches have been reviewed in detail.<sup>20</sup> An alternative approach for non-invasive prenatal diagnosis of fetal trisomy 21 is to show the presence of an elevated amount of chromosome 21 sequences in maternal blood, because there are three rather than two copies of fetal chromosome 21.<sup>18</sup> However, fetal DNA molecules amount to just 10–20% of the total DNA circulating in the maternal plasma,<sup>21,22</sup> so any increment in the total amount (fetal and maternal) of chromosome 21 DNA molecules in the plasma of a trisomy 21 pregnancy would be substantially diluted by contributions from the mother.<sup>20</sup>

The difficulty in measuring such a small increment in chromosome 21 DNA concentration has recently been overcome with the use of massively parallel genomic sequencing.<sup>20</sup> This technique can identify and quantify millions of DNA fragments in biological samples in a span of days.<sup>23</sup> Three cohort studies have shown the feasibility of using the technique to identify fetuses with trisomy 21 by analysis of maternal plasma DNA.<sup>24-26</sup> The sample numbers studied were small because, typically, only a few samples could be analysed in each sequencing run.

The objective of this study is to validate the diagnostic performance and practical feasibility of massively parallel genomic sequencing for the non-invasive prenatal assessment of trisomy 21 in a large number of pregnancies that have undergone conventional screening and were clinically indicated for definitive testing by amniocentesis or chorionic villus sampling.

## METHODS

The primary research question of the study was to evaluate if maternal plasma DNA sequencing could accurately confirm or exclude fetal trisomy 21 compared with full karyotyping among pregnancies clinically indicated for amniocentesis or chorionic villus sampling. A secondary goal of the study was to develop laboratory protocols to improve the throughput of massively parallel genomic sequencing for handling large sample numbers.

### Recruitment criteria

We prospectively recruited pregnant women from eight obstetric units in Hong Kong, one unit from the Netherlands, and one prenatal diagnostic centre in the United Kingdom between October 2008 and May 2009. We also retrieved archived maternal plasma samples collected between October 2003 and September 2008 from trisomy 21 and non-trisomy 21 pregnancies matched for gestational ages in a ratio of approximately 1:5 from the participating sites in the Netherlands and UK. The inclusion criteria were singleton pregnancies with clinical indications for chorionic villus sampling or amniocentesis as per the existing obstetric practice of each recruitment unit. Women

were excluded if full karyotyping results were not available.

We recorded the maternal ages at the expected time of delivery, gestational ages at blood sampling, and indications for invasive testing.

### Sample processing

Peripheral venous blood samples (5–10 mL) were collected into tubes containing EDTA. A plasma sample would be accepted for analysis if it was collected before invasive obstetric procedures, was harvested within six hours of venepuncture, was at least 2 mL in volume, and was not haemolysed. The prospectively collected and retrieved archival plasma samples from the UK and the Netherlands were sent to Hong Kong in batches by overnight courier while kept frozen on dry ice. We performed all subsequent analyses prospectively at The Chinese University of Hong Kong.

### Maternal plasma DNA sequencing

We extracted and sequenced DNA molecules from 2.0–4.8 mL of maternal plasma using a protocol similar to that reported previously but with the introduction of multiplexing.<sup>24,27</sup> DNA molecules exist in maternal plasma as short fragments. The genomic identities of millions of DNA molecules per maternal plasma sample can be decoded in a sequencing run. The analytical goal of our test was to determine the proportion of sequenced plasma DNA molecules originating from chromosome 21 (percentage chromosome 21). This proportion is expected to be elevated in maternal plasma during a pregnancy with a trisomy 21 fetus. Multiplexing allows more than one plasma sample to be mixed and sequenced jointly in each discrete segment of a sequencing glass slide and thereby increases the number of samples that can be analysed in each sequencing run. We studied two levels of multiplexing, 2-plex and 8-plex, whereby DNA from two or eight maternal plasma samples were co-sequenced in each slide segment.

Briefly, a unique synthetic DNA “barcode” of six base pairs, referred to as an index, was introduced on to one end of each plasma DNA molecule. The index served as a signature for a sample, with one index used per maternal plasma sample. For example, for 8-plex sequencing, eight different indices were needed for each of the eight test or control samples that would be co-sequenced. We compiled the multiplexed sample mixtures by pooling eight or two maternal plasma DNA preparations for the 8-plex or 2-plex sequencing protocols.

We performed sequencing on the Genome Analyzer II (Illumina) for the 8-plex sample mixtures and on the Genome Analyzer IIx (Illumina) for the 2-plex sample mixtures. After the sequencing run, the actual DNA molecules that belonged to a specific sample could then be distinguished from those belonging to other samples by sorting the index sequences attached to the DNA molecules. We considered a sequencing result as valid and reportable only if the analysis of

the sample passed a set of quality control measures (listed in appendix 1).

### Sequencing data analysis

We processed the sequencing data with the use of a bioinformatics algorithm based on our previously reported strategy (appendix 2 on [bmj.com](http://bmj.com)).<sup>24</sup> In essence, the chromosomal origin of each sequenced “read” (the short fragments of DNA that are decoded) was identified by comparing with the reference human genome available from the Ensembl website ([ftp://ftp.ensembl.org/pub/release-48/fasta/homo\\_sapiens/dna/](http://ftp.ensembl.org/pub/release-48/fasta/homo_sapiens/dna/)). The percentage chromosome 21 within each maternal plasma sample was then calculated. The proportion of reads from chromosome Y (percentage chromosome Y) was similarly calculated and was used to estimate the fetal DNA concentrations (that is, the proportion of DNA molecules in the maternal plasma sample that originated from the fetus) in pregnancies with male fetuses. We derived a z score for chromosome 21 in a test sample by subtracting the mean percentage chromosome 21 of a reference set of euploid pregnancies (controls) from the percentage chromosome 21 of the test case and divided by the standard deviation of the value for percentage chromosome 21 among the reference sample set according to the equation:

$$Z \text{ score for percentage chromosome 21 in test case} = \frac{((\text{percentage chromosome 21 in test case}) - (\text{mean percentage chromosome 21 in reference controls}))}{(\text{standard deviation of percentage chromosome 21 in reference controls})}$$

We used a z score of >3 (representing a percentage chromosome 21 value greater than that of the 99.9th centile of the reference sample set for a one tailed distribution) as the cut-off value to determine if the percentage chromosome 21 was increased and hence fetal trisomy 21 was present.

All members of the sequencing and bioinformatics teams were blinded to the karyotype information. Disease classification based on the sequencing results and z scores was generated automatically by the bioinformatics algorithm. We then determined the diagnostic performance of the sequencing test by comparing its results with those of full karyotyping of the amniotic fluid or chorionic villus sample.

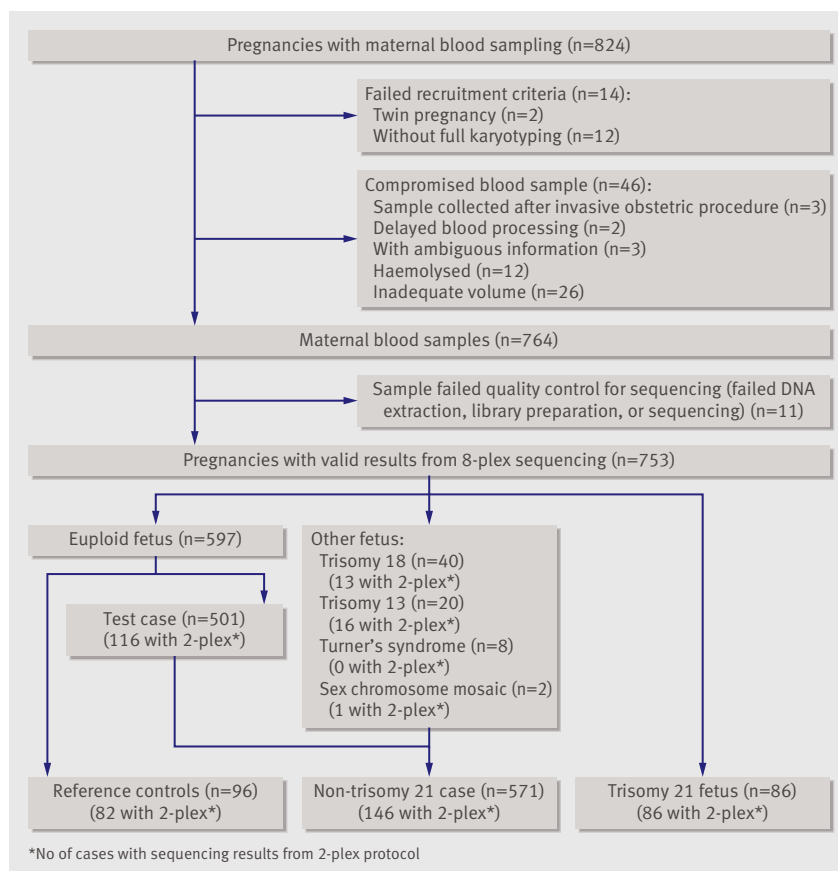
### Statistical analysis

We reported the observed diagnostic sensitivity and specificity values of the sequencing protocols based on the predefined cut-off point—namely, chromosome 21 z score of 3. We plotted the receiver operating characteristic (ROC) curves for the measurements of percentage chromosome 21 and determined the areas under the curves (AUC). ROC curve analysis was also performed to derive cut-off values for percentage chromosome Y to distinguish male and female fetuses. We measured the analytical imprecision of the sequencing protocols by calculating the coefficient of variation ( $SD/\text{mean} \times 100\%$ ) of the percentage chromosome 21 values among the control samples. The normality of quantitative variables was analysed with the use of the Kolmogorov-Smirnov test. We then used Student's *t* test or Mann-Whitney test and Pearson's correlation as appropriate. Two sided P values of <0.05 were considered to indicate statistical significance. ROC curve analyses were performed with the use of the MedCalc software (version 9.6.4.0). All other statistical calculations were performed with the use of the SigmaStat software (version 3.11).

## RESULTS

### Study participants

We recruited 576 pregnant women prospectively and retrieved archived maternal plasma samples from 248 women, which amounted to a total of 824 pregnancies (fig 1). Fourteen pregnant women (1.7%) did not meet the recruitment criteria, and 46 maternal plasma samples (5.6%) did not pass the specimen quality requirements. The quality control requirements were not met during the sequencing of 11 (1.4%) maternal plasma samples (appendix 1). Analytically valid sequencing results that met the criteria for reporting were achieved for 753 maternal plasma samples with the use of the 8-plex protocol. The characteristics of these pregnancies are shown in table 1. The median maternal age was



**Fig 1** Recruitment of participants. Numbers in parentheses are the number of cases with sequencing results from the 2-plex protocol



**Table 1** Characteristics of 753 pregnancies with maternal plasma DNA sequencing. Values are numbers (percentages) of subjects unless stated otherwise

Characteristic	Reference controls (n=96)*	Non-trisomy 21 fetuses (n=571)	Trisomy 21 fetuses (n=86)	All (n=753)
<b>Fetal sex:</b>				
Male	96 (100)	251 (44)	39 (45)	386 (51)
Female	0	318 (56)	47 (55)	365 (48)
Sex chromosome mosaic	0	2 (0.4)	0	2 (0.3)
<b>Maternal age at expected date of delivery (years):</b>				
<35	42 (44)	255 (45)	22 (26)	319 (42)
≥35	54 (56)	316 (55)	64 (74)	434 (58)
Median age	35.7	35.0	38.0	35.4
<b>Gestational age at maternal blood sampling (weeks+days):</b>				
≤13+6	80 (83)	410 (72)	67 (78)	557 (74)
>14+0 to 14+6	7 (7)	80 (14)	15 (17)	102 (14)
>15+0	9 (9)	81 (14)	4 (5)	94 (12)
Median gestational age	12+5	13+1	13+0	13+1
<b>Means of invasive testing:</b>				
Chorionic villus sampling	78 (81)	462 (81)	81 (94)	621 (82)
Amniocentesis	18 (19)	109 (19)	5 (6)	132 (18)
<b>Risk grouping:</b>				
High	64 (67)	437 (77)	81 (94)	582 (77)
Intermediate	6 (6)	33 (6)	0	39 (5)
Other	26 (27)	101 (18)	5 (6)	132 (18)
<b>Median fetal risk for trisomy 21:</b>				
In high risk group	1 in 50	1 in 51	1 in 2	1 in 43
In intermediate risk group	1 in 486	1 in 513	NA	1 in 502

\*Euploid male fetuses. NA=not applicable.

35.4 years, and the median gestational age at the time of maternal blood sampling was 13 weeks and 1 day.

Full karyotyping showed that there were 86 pregnancies with trisomy 21 fetuses (40 cases from the archival sample set) and 597 pregnancies with euploid fetuses (fig 1). The remainder comprised 40 fetuses with trisomy 18, 20 with trisomy 13, eight with Turner’s syndrome, and two sex chromosome mosaics, all of which were included as non-trisomy 21 cases in this study.

Among the sequenced samples, we randomly chose 96 pregnancies with euploid male fetuses as the reference control set for z score calculation. All 86 trisomy 21 cases, 82 reference controls, and 146 non-trisomy 21 cases were sequenced with the 2-plex protocol (fig 1).

Review of the clinical profile of the cases revealed there were three broad groups of clinical indications for referral for full karyotyping: (a) pregnancies with a risk higher than 1 in 300 for fetal trisomy 21 as estimated by conventional prenatal screening; (b) pregnancies with an intermediate risk (risk between 1 in 300 and 1 in 1000) for fetal trisomy 21; and (c) pregnancies with other risk indications, including a previous trisomy 21 pregnancy, ultrasound abnormalities, or risks for monogenic diseases. The numbers of pregnancies in the high, intermediate, and other risk groups were 582, 39, and 132, respectively (table 1). Among the high and intermediate risk groups, 570 (91.8%)

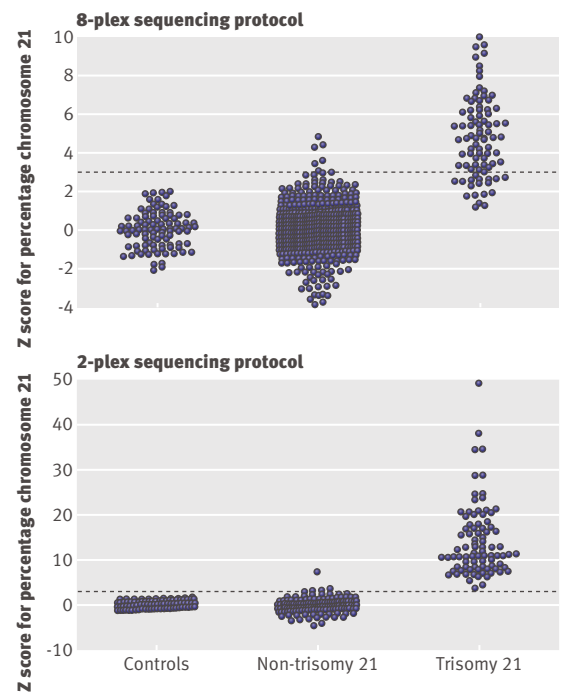
pregnant women underwent combined screening in the first trimester while the remaining 51 (8.2%) women received other forms of prenatal screening, including the integrated test and second trimester maternal serum screening.

**Diagnostic performance of maternal plasma DNA sequencing**

A mean of 0.3 million (SD 88 000) reads per sample from the 8-plex sequencing analysis and 2.3 million (SD 474 000) from 2-plex sequencing matched our bioinformatics criteria for calculation of percentage chromosome 21. As each sequencing glass slide can sequence only a finite number of DNA molecules, higher levels of multiplexing would lead to fewer DNA molecules to be sequenced per sample. Figure 2 shows the distribution of the chromosome 21 z scores among the controls, non-trisomy 21 group, and trisomy 21 group. Table 2 summarises the diagnostic performance of the sequencing test for detecting fetal trisomy 21.

With a chromosome 21 z score of 3 as the diagnostic cut-off point, the 2-plex protocol detected 100% of trisomy 21 fetuses at a 2.1% false positive rate—namely, 100% sensitivity and 97.9% specificity. This gave a positive predictive value of 96.6% and a negative predictive value of 100%.

The 8-plex protocol, however, detected 79.1% of the trisomy 21 fetuses at a false positive rate of 1.1%—that is, 79.1% sensitivity and 98.9% specificity. The positive predictive value was 91.9% and the negative predictive



**Fig 2** Z scores of percentage chromosome 21 (proportion of sequenced plasma DNA molecules originating from chromosome 21) determined by the 8-plex and 2-plex sequencing protocols. Broken lines indicate the z score cut-off value of 3

**Table 2** | Diagnostic performance of maternal plasma DNA sequencing for detecting fetal trisomy 21 and fetal sex

	8-plex sequencing protocol		2-plex sequencing protocol	
	True detection rate	False positive rate	True detection rate	False positive rate
<b>Trisomy 21 detection*</b>	Among 86 trisomy 21 cases	Among 571 non-trisomy 21 cases	Among 86 trisomy 21 cases	Among 146 non-trisomy 21 cases
	79.1% (68/86)	1.1% (6/571)	100% (86/86)	2.1% (3/146)
<b>Fetal sex (male) detection†</b>	Among 386 male fetuses	Among 365 female fetuses	Among 196 male fetuses	Among 117 female fetuses
	99.5% (384/386) (95% CI 98.1% to 99.9%)	0.8% (3/365) (95% CI 0.2% to 2.4%)	99.5% (195/196) (95% CI 97.2% to 99.9%)	0.8% (1/117) (95% CI 0.1% to 4.7%)

\*Z score for percentage chromosome 21 >3.

†Cut-off values for percentage chromosome Y identified by ROC analysis.

value was 96.9%. Eleven of the 18 false negative cases were from the archived samples. The distribution of the false negative and true positive cases among the prospectively collected and archived samples was not statistically significant ( $\chi^2$  test,  $P=0.258$ ). The median gestational ages for the true positive and false negative trisomy 21 cases from the 8-plex protocol were 13.0 weeks (interquartile range 12.3–13.5) and 12.9 weeks (12.3–13.5), respectively. The median risk level for trisomy 21 estimated by conventional screening for the true positive and false negative cases were 1 in 4 (interquartile range 1 in 5 to 1 in 2) and 1 in 5 (1 in 7 to 1 in 2), respectively. The true positive and false negative cases were not significantly different in terms of gestational ages (Mann-Whitney,  $P=0.538$ ) or risk levels (Mann-Whitney,  $P=0.466$ ). There were six false positive results in the 8-plex results and three in the 2-plex results. Because of the small number of false positive cases, no systematic features were identified. Profiles of these cases are shown in appendix 3 on [bmj.com](http://bmj.com).

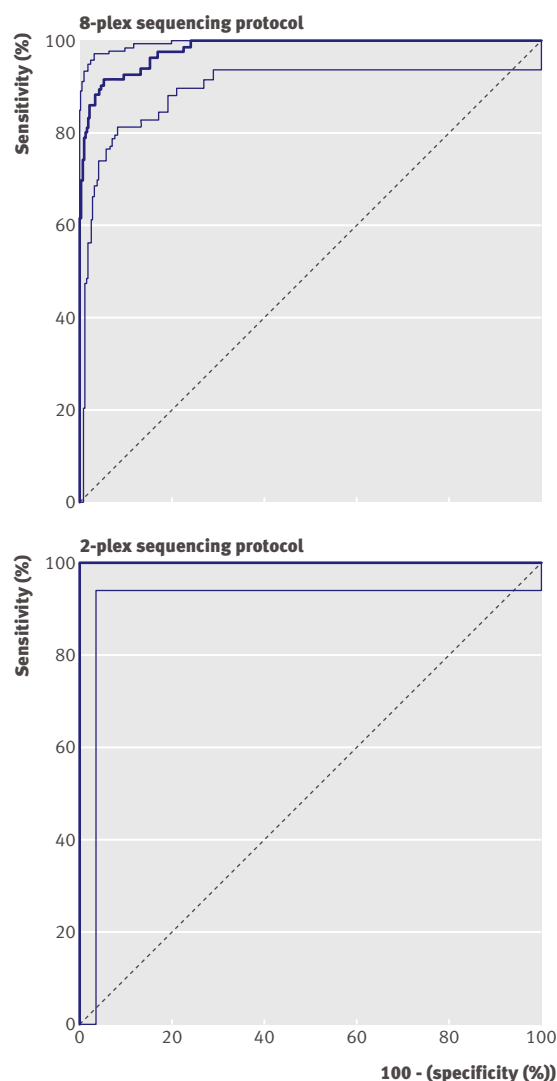
ROC curve analysis of the percentage chromosome 21 measurements for all test cases was done to estimate the diagnostic efficacy of the 2-plex and 8-plex sequencing protocols (fig 3). Area under the curve values for the 2-plex and 8-plex sequencing protocols were 1.00 (95% confidence interval 0.98 to 1.00) and 0.98 (0.97 to 0.99) respectively.

#### Utility of maternal plasma DNA sequencing for the prenatal assessment of trisomy 21

In this study we focused on pregnancies clinically indicated for invasive prenatal diagnostic testing because we required the availability of full karyotyping for comparison of the diagnostic performance of the sequencing test. However, this study design essentially limits the applicability of our data to pregnancies at increased risk for trisomy 21. To investigate if the sequencing test might be useful as a screening test for all pregnancies in general, we determined the post-test probabilities for women of different ages and hence with different prevalences of trisomy 21.<sup>28</sup> The data are shown in table 3. For a 20 year old woman at 12 weeks of gestation, her pretest probability for trisomy 21 by maternal age alone was 1 in 1068. If maternal plasma DNA sequencing with the 2-plex protocol was done, the post-test probability for a positive test result was 1 in 23 while the post-test probability for a negative test result was 1 in infinity. With the 8-plex

protocol, the post-test probabilities for a positive result and a negative result becomes 1 in 16 and 1 in 5082 respectively.

These data suggest that the sequencing test result, whether positive or negative, can substantially alter the probabilities for having a trisomy 21 fetus. The 2-



**Fig 3** | Diagnostic efficacy of diagnosis of fetal trisomy 21 by sequencing of maternal plasma DNA according to 8-plex and 2-plex protocols. Receiver operating characteristic (ROC) curves (faint lines showing 95% confidence intervals) for measurements of percentage chromosome 21

plex sequencing particularly highlights the key value of the sequencing test result for ruling out trisomy 21.

#### Factors affecting diagnostic performance

The diagnostic performance of the 2-plex protocol was superior to that of the 8-plex sequencing, with significantly higher z score values for the trisomy 21 cases (Mann-Whitney,  $P < 0.001$ ). The high z scores were due to a small standard deviation and, hence, better precision for estimating percentage chromosome 21. The coefficients of variation of percentage chromosome 21 values for the controls on each flow cell were calculated as a measure of precision (see appendix 4 on bmj.com). The mean coefficient of variation for measuring percentage chromosome 21 with the 8-plex protocol was 1.59% (SD 0.36%) and was higher than that for the 2-plex protocol, which was 0.66% (SD 0.25%) (*t* test,  $P < 0.001$ ).

The extent of increase in percentage chromosome 21 in maternal plasma for the trisomy 21 cases is governed by the fetal DNA concentration. We calculated the fetal DNA concentrations based on the percentage chromosome Y values among the 39 male fetuses with trisomy 21 and plotted them against the chromosome 21 z scores (shown in appendix 5 on bmj.com). The two variables are significantly correlated in both the 8-plex sequencing data (Pearson correlation coefficient 0.703,  $P < 0.0001$ ) and 2-plex data (Pearson correlation coefficient 0.462,  $P = 0.003$ ). The five pregnancies with male fetuses with trisomy 21 that we failed to detect using the 8-plex protocol (that is, with chromosome 21 z scores  $< 3$ ) had fetal DNA concentrations less than 10%.

Theoretically, at 10% fetal DNA concentration, percentage chromosome 21 is just 1.05 times higher in a plasma sample obtained from a trisomy 21 pregnancy than a euploid pregnancy.<sup>18</sup> To detect this increase in

percentage chromosome 21, the coefficient of variation of the test would need to be  $\leq 0.83\%$  (see appendix 6 on bmj.com). The mean coefficient of variation of the 8-plex protocol was greater than this value, whereas that of the 2-plex protocol was  $< 0.83\%$ . Hence, the 2-plex protocol performed better at detecting trisomy 21 among samples containing low fetal DNA concentrations.

#### Fetal sex determination by maternal plasma DNA sequencing

The presence of chromosome Y sequences among sequenced DNA molecules from maternal plasma should theoretically signify that the pregnancy involved a male fetus. However, we previously reported that the bioinformatics algorithm would incorrectly assign a small fraction of sequenced DNA reads to chromosome Y in pregnancies with female fetuses.<sup>24</sup> The percentage chromosome Y values of all studied pregnancies are shown in appendix 7 on bmj.com. Optimal cut-off values for percentage chromosome Y to distinguish male and female fetuses were identified by ROC curve analysis to be 0.0114% and 0.0095% for the 8-plex and 2-plex protocols, respectively. Table 2 shows the estimated diagnostic performance for fetal sex determination. We then determined the fetal DNA concentrations of 314 pregnancies with euploid male fetuses based on the percentage chromosome Y values obtained from the 8-plex protocol. The median fetal DNA concentration was 15.2% (interquartile range 10.6%–19.1%), and the distribution is shown in fig 4. Among the pregnancies with euploid male fetuses, the median fetal DNA concentration for the archival and prospectively collected samples were 14.7% and 15.4%, respectively (Mann-Whitney,  $P = 0.334$ ).

**Table 3** | Probabilities for a trisomy 21 fetus in women by age alone and according to result of maternal plasma DNA sequencing test

Maternal age (years)	Pretest probability*	Post-test probability			
		8-plex sequencing		2-plex sequencing	
		Positive test result†	Negative test result	Positive test result†	Negative test result
20	1 in 1068	1 in 16	1 in 5082	1 in 23	1 in infinity
25	1 in 946	1 in 14	1 in 4501	1 in 21	1 in infinity
30	1 in 626	1 in 10	1 in 2977	1 in 14	1 in infinity
31	1 in 543	1 in 9	1 in 2582	1 in 12	1 in infinity
32	1 in 461	1 in 7	1 in 2191	1 in 11	1 in infinity
33	1 in 383	1 in 6	1 in 1820	1 in 9	1 in infinity
34	1 in 312	1 in 5	1 in 1482	1 in 8	1 in infinity
35	1 in 249	1 in 4	1 in 1182	1 in 6	1 in infinity
36	1 in 196	1 in 4	1 in 930	1 in 5	1 in infinity
37	1 in 152	1 in 3	1 in 720	1 in 4	1 in infinity
38	1 in 117	1 in 3	1 in 553	1 in 3	1 in infinity
39	1 in 89	1 in 2	1 in 420	1 in 3	1 in infinity
40	1 in 68	1 in 2	1 in 320	1 in 2	1 in infinity
41	1 in 51	1 in 2	1 in 239	1 in 2	1 in infinity
42	1 in 38	1 in 2	1 in 177	1 in 2	1 in infinity

\*Pretest probabilities are based on prevalence of fetal trisomy 21 at 12th week of gestation.<sup>28</sup>

†A positive test result is a sample with a z score for percentage chromosome 21  $> 3$ .

## DISCUSSION

We show that fetal trisomy 21 could be detected in high risk pregnancies with high sensitivity and specificity by means of multiplexed sequencing of maternal plasma DNA. A key strength of the study is that the diagnostic performance was compared against full karyotyping, which is the gold standard for aneuploidy diagnosis. Another strength of this study is that the test performance was assessed based on a large number of trisomy 21 and unaffected pregnancies. In fact, this study reports the largest scale use of massively parallel genomic sequencing (so called next generation sequencing) for medical diagnostics to date and shows that it is practically feasible on a large scale.

Another strength of the study lies in the comparison of the diagnostic performance of sequencing protocols with two levels of sample throughput. The comparison allows us to deduce factors that are fundamental to the sensitive detection of fetal trisomy 21 and provide insight into ways for further improvements. Measurement of the percentage of plasma DNA that was from chromosome 21 was more precise with the use of the 2-plex protocol because seven times more plasma DNA molecules were analysed with this method than with the 8-plex protocol. All trisomy 21 fetuses, including those missed by the 8-plex protocol, were detected with the 2-plex protocol. In short, the diagnostic sensitivity of the sequencing approach improves with an increase in the number of plasma DNA molecules analysed per sample.<sup>18</sup>

There are some weaknesses in the study design. Our study was based on high risk pregnancies—those clinically indicated for amniocentesis or chorionic villus sampling—which ensured that full karyotyping results were available from each pregnancy. To investigate if the sequencing test might be useful for prenatal assessment of fetal trisomy 21 for all pregnancies in general, we determined the post-test probabilities for having a trisomy 21 fetus in women of different ages, and hence with different pretest risks. The data suggest that the sequencing test result can substantially alter the probabilities for having a trisomy 21 fetus even in women with low pretest risk.

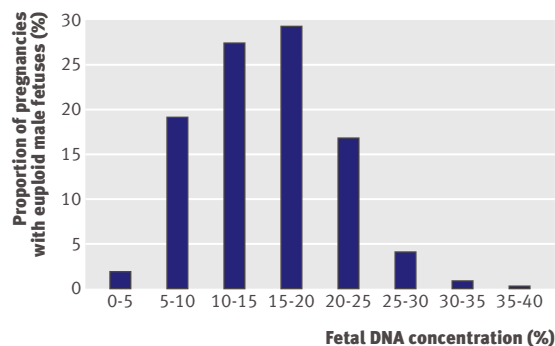
Another potential weakness of the study was that an archived sample set was included. However, both the

prospectively collected and archived maternal plasma samples were processed using the same protocols and analysed prospectively. There were also no significant differences in the fetal DNA concentrations and test performances between the two sample groups. In this study 5.6% of the collected maternal plasma samples were of compromised quality. Reportable sequencing results were not achieved for 11 samples. All these samples were identified by the predefined quality control steps. In routine practice, laboratory reports would not be issued for such samples, and another blood specimen would be requested.

We found other aneuploidies besides trisomy 21 among the recruited pregnant women. For this study, we focused on the diagnostic performance of the sequencing approach for fetal trisomy 21 because our previous data showed that the measurements of the proportion of DNA molecules from chromosomes 18 and 13 were much less precise.<sup>24,25</sup> More research is required to develop protocols to improve the precision for measuring amount of DNA molecules from chromosomes 18 and 13.<sup>29</sup>

Our data reveal that the main value of the maternal plasma DNA sequencing test is to rule out fetal trisomy 21. Hence, with the current diagnostic performance, it is more suitable as a screening test to stratify pregnancies whose risk for trisomy 21 warrants the consideration of amniocentesis or chorionic villus sampling. In this study, we performed the sequencing test after a pregnancy was deemed to be at increased risk for trisomy 21 according to the current prenatal screening programmes. The false positive rates of the current screening programmes are about 5%, and all of these pregnant women are offered the option of invasive testing. However, if we took into consideration the results of the sequencing test, trisomy 21 could be ruled out in 98% of those pregnancies. This would leave just 0.1% (that is,  $5\% \times (100\% - 98\%) \times 100\%$ ) of all pregnant women needing referrals for amniocentesis or chorionic villus sampling. Most of our studied pregnancies were in the first trimester, which suggests that it is possible to implement the test even in early pregnancy as a second tier screening test after the first trimester combined test, which is already in use in many parts of the world.

On the other hand, our post hoc analysis shows that the sequencing test result may alter clinical decisions even in women at low risk for trisomy 21. Hence, there is the potential to apply the sequencing test as a first tier screening test. However, massively parallel genomic sequencing is currently expensive. To determine its suitability as a first tier screening test would require formal investigation of both its diagnostic performance and cost benefit. We used the sequencing test in this study for detection of fetal trisomy 21, but, with some improvements in bioinformatics, it could be used to detect other aneuploidies, such as trisomy 18 and trisomy 13.<sup>29</sup> Furthermore, our group has recently shown that when more sophisticated sequencing protocols and bioinformatics algorithms are applied to the analysis of maternal plasma DNA, it may be possible to



**Fig 4** | Fetal DNA concentration found in maternal plasma among pregnancies with euploid male fetuses



### WHAT IS ALREADY KNOWN ON THIS TOPIC

Non-invasive prenatal detection of fetal trisomy 21 is achievable by massively parallel sequencing of maternal plasma DNA

Its diagnostic performance and practical feasibility in the clinical setting has not been tested on a large scale

### WHAT THIS STUDY ADDS

Among high risk pregnancies clinically indicated for invasive prenatal diagnosis, non-invasive detection of fetal trisomy 21 can be achieved with the use of multiplexed massively parallel sequencing of maternal plasma DNA at 100% sensitivity and 97.9% specificity, giving a 96.6% positive predictive value and 100% negative predictive value

The sequencing test could be used to rule out trisomy 21 among high risk pregnancies before proceeding to invasive diagnostic testing to reduce the number of cases requiring amniocentesis or chorionic villus sampling

perform a genetic and mutational scan across the whole genome of the fetus in a non-invasive manner.<sup>30</sup> Thus, it is likely that maternal plasma DNA sequencing will play an increasingly important role in the future developments of prenatal screening and diagnosis.

We thank members of the Lo laboratory for technical assistance in performing the sequencing runs.

**Contributors:** RWKC and YMDL designed and initiated the study, coordinated and managed the research process, interpreted the results and wrote the first draft of the report. RA, ATJIG, ETL, JMGvV, and RM contributed to the implementation of the clinical recruitment protocols and management of clinical data. TYL, MHYT, CBMO, TKL, and KHN contributed to the design of the clinical recruitment protocols, coordination of the recruitment of patients and management of clinical data. YWLZ, KCAC, and FMFL contributed to the development of the laboratory protocols and monitoring of the laboratory analyses. HS and KCAC contributed to the development of the bioinformatics analysis protocols. WWKT, WCL, RYKT, SKCA-Y, HL, and YYK contributed to the recruitment, collection and management of clinical data. XZ and JW contributed to the laboratory analysis. All authors provided critical comments on the drafts of the manuscript and approved the final version. RWKC and YMDL had full access to all the data in the study, drafted the manuscript, take responsibility for the integrity of the data and the accuracy of the data analysis, and are the guarantors.

**Competing interests:** All authors have completed the Unified Competing Interest form at [www.icmje.org/doi\\_disclosure.pdf](http://www.icmje.org/doi_disclosure.pdf) (available on request from the corresponding author) and declare that (1) RWKC, HS, ETL, MHYT, TKL, and YMDL had support from the University Grants Committee of the Government of the Hong Kong Special Administrative Region, China, under the Areas of Excellence Scheme (AoE/M-04/06) for the submitted work; RWKC and YMDL had support from Sequenom for the submitted work; YMDL has support from an endowed chair from the Li Ka Shing Foundation. (2) YMDL is a consultant to, and holds equities in, Sequenom; RWKC and YMDL have received travel grants from Life Technologies and Illumina; RWKC, YWLZ, HS, KCAC, FMFL, and YMDL have received travel grants from University Grants Committee of the Government of the Hong Kong Special Administrative Region, China, under the Areas of Excellence Scheme (AoE/M-04/06). (3) The authors' spouses, partners, and children have no financial relationships that may be relevant to the submitted work. (4) RWKC, YWLZ, KCAC, FMFL, and YMDL have filed patent applications on the detection of fetal nucleic acids in maternal plasma for non-invasive prenatal diagnosis. Part of this patent portfolio has been licensed to Sequenom and the Institut Jacques Boyd. The funders and sponsors have no role in the study design and the collection, analysis, and interpretation of data and the writing of the article and the decision to submit it for publication. The researchers are independent of the funders and sponsors.

**Ethical approval:** Approvals were obtained from the institutional review boards of each recruitment site: Joint Chinese University of Hong Kong-New Territories East Cluster Clinical Research Ethics Committee, Joint Institutional Review Board of the University of Hong Kong-Hospital Authority Hong Kong West Cluster, Clinical and Research Ethics

Committees of the Hospital Authority in the Kowloon Central/Kowloon East, Kowloon West, Hong Kong East, New Territories West Clusters, King's College Hospital Ethics Committee, and Ethics Committee of the VU University Medical Center. All participants gave informed written consent.

**Data sharing:** No additional data available.

- 1 Driscoll DA, Gross S. Clinical practice. Prenatal screening for aneuploidy. *N Engl J Med* 2009;360:2556-62.
- 2 Tabor A, Philip J, Madsen M, Bang J, Obel EB, Norgaard-Pedersen B. Randomised controlled trial of genetic amniocentesis in 4606 low-risk women. *Lancet* 1986;1:1287-93.
- 3 Snijders RJ, Noble P, Sebire N, Souka A, Nicolaides KH. UK multicentre project on assessment of risk of trisomy 21 by maternal age and fetal nuchal-translucency thickness at 10-14 weeks of gestation. Fetal Medicine Foundation First Trimester Screening Group. *Lancet* 1998;352:343-6.
- 4 Kagan KO, Wright D, Baker A, Sahota D, Nicolaides KH. Screening for trisomy 21 by maternal age, fetal nuchal translucency thickness, free beta-human chorionic gonadotropin and pregnancy-associated plasma protein-A. *Ultrasound Obstet Gynecol* 2008;31:618-24.
- 5 Malone FD, Canick JA, Ball RH, Nyberg DA, Comstock CH, Bukowski R, et al. First-trimester or second-trimester screening, or both, for Down's syndrome. *N Engl J Med* 2005;353:2001-11.
- 6 Wald NJ, Rodeck C, Hackshaw AK, Rudnicka A. SURUSS in perspective. *BJOG* 2004;111:521-31.
- 7 Wald NJ, Huttly WJ, Hackshaw AK. Antenatal screening for Down's syndrome with the quadruple test. *Lancet* 2003;361:835-6.
- 8 Kagan KO, Cicero S, Staboulidou I, Wright D, Nicolaides KH. Fetal nasal bone in screening for trisomies 21, 18 and 13 and Turner syndrome at 11-13 weeks of gestation. *Ultrasound Obstet Gynecol* 2009;33:259-64.
- 9 Kagan KO, Valencia C, Livanos P, Wright D, Nicolaides KH. Tricuspid regurgitation in screening for trisomies 21, 18 and 13 and Turner syndrome at 11+0 to 13+6 weeks of gestation. *Ultrasound Obstet Gynecol* 2009;33:18-22.
- 10 Lo YMD, Corbetta N, Chamberlain PF, Rai V, Sargent IL, Redman CW, et al. Presence of fetal DNA in maternal plasma and serum. *Lancet* 1997;350:485-7.
- 11 Lo YMD, Hjelm NM, Fidler C, Sargent IL, Murphy MF, Chamberlain PF, et al. Prenatal diagnosis of fetal Rhd status by molecular analysis of maternal plasma. *N Engl J Med* 1998;339:1734-8.
- 12 Wright CF, Chitty LS. Cell-free fetal DNA and RNA in maternal blood: implications for safer antenatal testing. *BMJ* 2009;339:b2451.
- 13 Finning K, Martin P, Summers J, Massey E, Poole G, Daniels G. Effect of high throughput RHD typing of fetal DNA in maternal plasma on use of anti-RhD immunoglobulin in RhD negative pregnant women: prospective feasibility study. *BMJ* 2008;336:816-8.
- 14 Lo YMD, Chiu RWK. Prenatal diagnosis: progress through plasma nucleic acids. *Nat Rev Genet* 2007;8:71-7.
- 15 Tong YK, Ding C, Chiu RWK, Gerovassili A, Chim SSC, Leung TY, et al. Noninvasive prenatal detection of fetal trisomy 18 by epigenetic allelic ratio analysis in maternal plasma: Theoretical and empirical considerations. *Clin Chem* 2006;52:2194-202.
- 16 Tong YK, Jin S, Chiu RWK, Ding C, Chan KCA, Leung TY, et al. Noninvasive prenatal detection of trisomy 21 by an epigenetic-genetic chromosome-dosage approach. *Clin Chem* 2010;56:90-8.
- 17 Lo YMD, Tsui NBY, Chiu RWK, Lau TK, Leung TN, Heung MM, et al. Plasma placental RNA allelic ratio permits noninvasive prenatal chromosomal aneuploidy detection. *Nat Med* 2007;13:218-23.
- 18 Lo YMD, Lun FMF, Chan KCA, Tsui NBY, Chong KC, Lau TK, et al. Digital PCR for the molecular detection of fetal chromosomal aneuploidy. *Proc Natl Acad Sci USA* 2007;104:13116-21.
- 19 Tsui NBY, Akolekar R, Chiu RWK, Chow KC, Leung TY, Lau TK, et al. Synergy of total PLAC4 RNA concentration and measurement of the RNA single-nucleotide polymorphism allelic ratio for the noninvasive prenatal detection of trisomy 21. *Clin Chem* 2010;56:73-81.
- 20 Chiu RWK, Cantor CR, Lo YMD. Non-invasive prenatal diagnosis by single molecule counting technologies. *Trends Genet* 2009;25:324-31.
- 21 Lo YMD, Tein MS, Lau TK, Haines CJ, Leung TN, Poon PM, et al. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. *Am J Hum Genet* 1998;62:768-75.
- 22 Lun FMF, Chiu RWK, Chan KCA, Leung TY, Lau TK, Lo YMD. Microfluidics digital PCR reveals a higher than expected fraction of fetal DNA in maternal plasma. *Clin Chem* 2008;54:1664-72.
- 23 Schuster SC. Next-generation sequencing transforms today's biology. *Nat Methods* 2008;5:16-8.
- 24 Chiu RWK, Chan KCA, Gao Y, Lau YYM, Zheng W, Leung TY, et al. Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc Natl Acad Sci USA* 2008;105:20458-63.

- 25 Chiu RWK, Sun H, Akolekar R, Clouser C, Lee C, McKernan K, et al. Maternal plasma DNA analysis with massively parallel sequencing by ligation for noninvasive prenatal diagnosis of trisomy 21. *Clin Chem* 2010;56:459-63.
- 26 Fan HC, Blumenfeld YJ, Chitkara U, Hudgins L, Quake SR. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proc Natl Acad Sci USA* 2008;105:16266-71.
- 27 Cronn R, Liston A, Parks M, Gernandt DS, Shen R, Mockler T. Multiplex sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis technology. *Nucleic Acids Res* 2008;36:e122.
- 28 Nicolaides KH. *The 11 to 13+6 weeks scan*. Fetal Medicine Foundation, 2004.
- 29 Fan HC, Quake SR. Sensitivity of noninvasive prenatal detection of fetal aneuploidy from maternal plasma using shotgun sequencing is limited only by counting statistics. *PLoS One* 2010;5:e10439.
- 30 Lo YMD, Chan KCA, Sun H, Chen EZ, Jiang P, Lun FMF, et al. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Sci Transl Med* 2010;2:61ra91.

**Accepted:** 14 December 2010