

Noninvasive detection of fetal trisomy 21 by sequencing of DNA in maternal blood: a study in a clinical setting

Mathias Ehrich, MD; Cosmin Deciu, MSc; Tricia Zwiefelhofer; John A. Tynan, DPhil; Lesley Cagasan, MSc; Roger Tim, DPhil; Vivian Lu; Ron McCullough, DPhil; Erin McCarthy; Anders O. H. Nygren, DPhil; Jarrod Dean; Lin Tang, DPhil; Don Hutchison, MSc; Tim Lu, DPhil; Huiquan Wang, DPhil; Vach Angkachatchai, DPhil; Paul Oeth, MSc; Charles R. Cantor, DPhil; Allan Bombard, MD; Dirk van den Boom, DPhil

OBJECTIVE: We sought to evaluate a multiplexed massively parallel shotgun sequencing assay for noninvasive trisomy 21 detection using circulating cell-free fetal DNA.

STUDY DESIGN: Sample multiplexing and cost-optimized reagents were evaluated as improvements to a noninvasive fetal trisomy 21 detection assay. A total of 480 plasma samples from high-risk pregnant women were employed.

RESULTS: In all, 480 prospectively collected samples were obtained from our third-party storage site; 13 of these were removed due to insufficient quantity or quality. Eighteen samples failed prespecified assay quality control parameters. In all, 449 samples remained: 39 tri-

somy 21 samples were correctly classified; 1 sample was misclassified as trisomy 21. The overall classification showed 100% sensitivity (95% confidence interval, 89–100%) and 99.7% specificity (95% confidence interval, 98.5–99.9%).

CONCLUSION: Extending the scope of previous reports, this study demonstrates that plasma DNA sequencing is a viable method for noninvasive detection of fetal trisomy 21 and warrants clinical validation in a larger multicenter study.

Key words: circulating cell-free fetal DNA, massively parallel shotgun sequencing, maternal blood, NIPD, noninvasive prenatal diagnosis

Cite this article as: Ehrich M, Deciu C, Zwiefelhofer T, et al. Noninvasive detection of fetal trisomy 21 by sequencing of DNA in maternal blood: a study in a clinical setting. *Am J Obstet Gynecol* 2011;204:205.e1-11.

Trisomy 21 is the most common chromosomal aneuploidy in live born infants. The overall incidence of trisomy 21 is approximately 1 in 800 births in the general population,¹ but this risk increases to 1 in 35 term births for women 45 years of age.²⁻⁴ Advanced maternal age is only one factor contributing to increased risk. When other factors, such as positive serum screening results, fetal ultrasound abnormality, or family history are included, the odds of being affected given a positive result of

★ EDITORS' CHOICE ★

Down syndrome can be as high as 1 in 9 using the integrated test.⁵ For women in this high-risk group, an invasive diagnostic procedure is currently the only way to confirm the diagnosis of trisomy 21, commonly by means of a fetal karyotype. Although the safety of the invasive procedures, specifically genetic amniocentesis and chorionic villus sampling (CVS), has improved greatly since their introduction, there remains a well-rec-

ognized risk of iatrogenic fetal loss. Tests that could better identify those women who would most benefit from confirmatory invasive diagnostic tests are of great public health interest. Since the initial seminal work by Merkatz et al,⁶ continuous efforts have focused on increasing the specificity of primary screening methods; eg, by including more serum protein markers or through the addition of ultrasound findings suggestive of fetal aneuploidy. Consequently, screening tests have greatly improved in their clinical sensitivity and specificity over the last 2 decades. These developments have also led to contemporary testing programs involving a variety of potential screening algorithms.⁷ Accurate gestational dating through ultrasound is a critical element to achieve high accuracy, but may not be readily available to all pregnant women.

A new approach to detect fetal aneuploidy analyzes fetal DNA itself rather than the surrogate biochemical or ultrasound markers in current maternal serum screening protocols. In 1997, Lo et al⁸ reported that circulating cell-free (ccf) fetal (ccff) DNA is present in the plasma of pregnant women. DNA of fetal

From Sequenom Inc (Drs Ehrich, Cantor, Bombard, and van den Boom), and Sequenom Center for Molecular Medicine LLC (Mr Oeth, Mr Deciu, Ms Zwiefelhofer, Ms Cagasan, Ms V. Lu, Ms McCarthy, Mr Dean, Mr Hutchison, and Drs Tynan, Tim, McCullough, Nygren, Tang, Lu, Wang, and Angkachatchai), San Diego, CA.

Received Sept. 30, 2010; revised Dec. 28, 2010; accepted Dec. 28, 2010.

Reprints: Dirk van den Boom, DPhil, Sequenom Inc, 3595 John Hopkins Ct., San Diego, CA 92121. dvandenboom@sequenom.com.

DISCLOSURE: All authors of this article are employees and shareholders of Sequenom, Inc. or its subsidiaries, and therefore, a potential conflict of interest exists.

0002-9378/free • © 2011 Mosby, Inc. All rights reserved. • doi: 10.1016/j.ajog.2010.12.060



For Editors' Commentary, see Table of Contents



See related editorial, page 183

origin ranges between 2% and 40% with a mean around 10% of the total ccf DNA across varying gestational ages.⁹⁻¹² The ccf DNA is cleared from the maternal bloodstream within hours after birth; thus, misdiagnosis from carryover contamination from a previous pregnancy is unlikely.¹³ A noninvasive ccf method for prenatal Rhesus D testing in Europe has already been widely adopted.¹⁴

In comparison to Rhesus D testing, aneuploidy detection from ccf DNA is far more challenging. In principle, aneuploidy detection could be enabled through a variety of methods including the analysis of single nucleotide polymorphisms,¹⁵ DNA methylation,^{16,17} or fetally expressed RNA transcripts.^{18,19} The most convincing data to date for a generally applicable test, however, have been generated through massively parallel shotgun sequencing (MPSS) of ccf DNA. Two groups have independently shown that MPSS can unambiguously identify plasma samples from women carrying a trisomy 21 fetus²⁰⁻²³ compared to samples from women with euploid fetuses. These studies were performed with small numbers of clinical samples and, while these preliminary results are very promising, the true clinical performance remains to be established. As originally described in 2008, the overall cost of a sequencing-based test was prohibitive in terms of potential deployment in clinical practice. However, next-generation sequencing methods such as MPSS are rapidly evolving with concomitant declines in reagent and instrument costs.

We have implemented several process improvements in MPSS for noninvasive aneuploidy detection using ccf DNA. These modifications provide an affordable testing procedure with the potential for widespread utilization. Because such a test, first and foremost, has to be safe and efficacious we designed a blinded study that tested a total of 480 plasma samples collected from pregnant woman at high risk for fetal aneuploidy.

MATERIALS AND METHODS

Study design

The study was set up to include at least 40 trisomy 21 samples, a design chosen to achieve a lower 95% confidence bound

of 91% when all trisomy 21 cases are correctly identified. We matched trisomy 21 samples with euploid samples at a ~1:11 ratio, slightly higher than the more typical prevalence in a high-risk group of 1 in 15.

Patients at increased risk for fetal Down syndrome and other chromosomal aneuploidies were asked to participate in this prospective study. Risks included a positive serum biochemical screening test; advanced maternal age (≥ 35 years at the estimated date of delivery); a fetal ultrasound finding suggestive of Down syndrome; or a personal/family history of Down syndrome. Patient informed consent was obtained for peripheral blood sampling and for the inclusion of karyotype results from an already scheduled, subsequent invasive diagnostic procedure. Fetal karyotypes or quantitative fluorescent PCR results were obtained as part of regular clinical care on either CVS or genetic amniocentesis samples. These data were unknown to the investigators prior to unblinding. The sample demographics were representative for pregnant women at high risk for fetal trisomy 21 (Table 1).

Samples were blinded to the investigators and prospectively collected, processed, and stored at an independent, contracted, third-party location (Biostorage Technologies Inc [BST], Indianapolis, IN). All information was kept within an independent, third-party database (Pharmaceutical Research Associates Inc [PRA], Raleigh, NC). A total of 480 samples were requested from PRA and provided by BST for analysis at Sequenom Center for Molecular Medicine, San Diego, CA. Karyotype results were unknown to the investigators and data analysts until after completion of all sample testing and submission for review. The MPSS results were sent to an independent, third-party biostatistician who had all clinical information including confirmatory karyotypes. The data were matched and unblinded by this third-party biostatistician and the concordance of the results was reported.

Sample collection

For the study presented here, samples were collected at clinical practices active

in the treatment of patients undergoing invasive prenatal diagnosis by CVS (first trimester) and genetic amniocentesis (second trimester) and, for some of the cases, from pregnancy termination centers. Eight samples were collected for research purposes under Food and Drug Administration approval (FDA Establishment Identifier no. 3005208435). All remaining samples were collected under institutional review board (IRB) approval (Western Institutional Review Board [WIRB] no. 20091396, WIRB no. 20080757, Compass IRB no. 00351). All samples, demographics, and karyotype results were completely blinded to the laboratory investigators by the third-party clinical research organization (PRA) and the BST facility. Patients were approached during their genetic counseling sessions and, if they gave their informed consent, the study protocol dictated that phlebotomy was to be performed prior to their invasive procedure. The vast majority of samples were collected after August 2009 and none were collected before May 2009; therefore, the oldest samples in the study were ~10 months old. Samples were all collected at specifically contracted processing centers operating under study-specific protocols. None of the samples were obtained and analyzed as fresh samples; ie, all were processed and frozen before shipment to the central, independent biostorage facility (<http://biostorage.com/>; a full description of the independent nature of this widely used biostorage company is detailed on their World Wide Web site).

All samples were collected and processed under the same protocol: 10 mL of maternal whole blood was drawn into an EDTA-K2 spray-dried Vacutainer (Becton Dickinson, Franklin Lakes, NJ), stored, and transported to the processing laboratory on wet ice. Within 6 hours of the blood draw, the maternal whole blood was centrifuged (Eppendorf 5810R plus swing out rotor) chilled (4°C) at 2500g for 10 minutes and the plasma was collected. The plasma was centrifuged a second time (Eppendorf 5810R plus fixed angle rotor) at 4°C at 15,000g for 10 minutes. After the second spin, the plasma was removed from the

pellet that formed at the bottom of the tube and distributed into 4-mL plasma bar-coded aliquots. In this study, only a single 4-mL plasma aliquot from each patient was used for DNA isolation.

MPSS aneuploidy detection

DNA was prepared from 4 mL of maternal plasma. The short lengths of ccfDNA afford direct use in preparing the libraries of DNA fragments that were sequenced. In practice, 4 different libraries each containing a synthetic oligonucleotide sequence as a bar code were mixed and analyzed together (multiplexing). The bar code revealed which library each sequence read represented. Eight separate mixtures of 4 libraries were analyzed in parallel. One MPSS process required about 2 days and yielded 36 bases of sequence from each DNA fragment. Approximately 5 million 36-base fragments were sequenced from each library. These represented about 6% of the human genome in each sample. As is standard in MPSS, the 36-base reads were processed to exclude poor-quality data and then matched to a reference human genome to determine their chromosome origin. The fraction of reads is proportionate to chromosome size. Thus, typically 8.5% of all reads are from chromosome 1, while only about 1.2% are from chromosome 22.^{20,21}

A fetus with trisomy 21 contributes additional genetic material to the total pool of ccfDNA. Consequently, in comparison to women carrying a euploid fetus, a slightly larger contribution of sequence reads mapping to chromosome 21 is observed in a plasma sample of a woman carrying a fetus with Down syndrome. Ccf DNA in plasma from a pregnant woman with a euploid fetus shows an average 1.35% of all aligned sequence reads located on chromosome 21. A variety of analytical methods have been published to detect an overabundance of genetic material from chromosome 21 in trisomic pregnancies.^{20,21,23} These use some form of normalization to calibrate the results against a known set of euploid reference samples. Contributions greater than the reference range are then indicative of additional genetic material from chromosome 21 and in many cases can

be interpreted as a fetal trisomy 21. In this study, a modification of a method used by Chiu et al²⁰ was used for classification. Prior to the main study a set of known euploid reference samples was used to calculate the mean and standard deviation (SD) of the representation of chromosome 21 (percentage of reads obtained from chromosome 21). Then, for every test sample, the distance, measured in SD, from the mean in the euploid reference set was calculated. A fixed cutoff of 2.5 SD was used to identify samples with an overrepresentation of chromosome 21 material.

Assay design

Compared to previously published studies,^{20,21} 3 important modifications were made to the sequencing protocol. We used custom purified enzymes in the library generation process to achieve a reduction in assay cost. We employed the latest sequencing biochemistry available for the GAIIX sequencer (Illumina Inc, San Diego, CA) in combination with the manufacturer's analysis software CASAVA version 1.6. These changes increased the number of sequence reads from approximately 13 to 20 million per lane. We also used indexing primers during library amplification to allow analysis of multiple samples in a single sequencing reaction ("multiplexing" vs "monoplexing"). In this study, 4 samples were analyzed per lane ("4-plex" or "tetraplexing"), which equates to approximately 3 to 5 million available sequence reads per sample. The combination of these modifications enabled 4 times higher throughput at about 4 times lower cost.

DNA extraction

The Qiagen ccf nucleic acid kit (Qiagen, Hilden, Germany) was used according to the manufacturer's specifications. The resulting DNA was eluted in 55 μ L of buffer AVE (part of the Qiagen kit).

Quality control of extracted DNA

The quantity of the extracted DNA was determined with an assay that uses simultaneous quantification of fetal and total ccfDNA. This fetal quantifier assay (FQA) was recently published¹⁰ and uses methylation-sensitive restriction en-

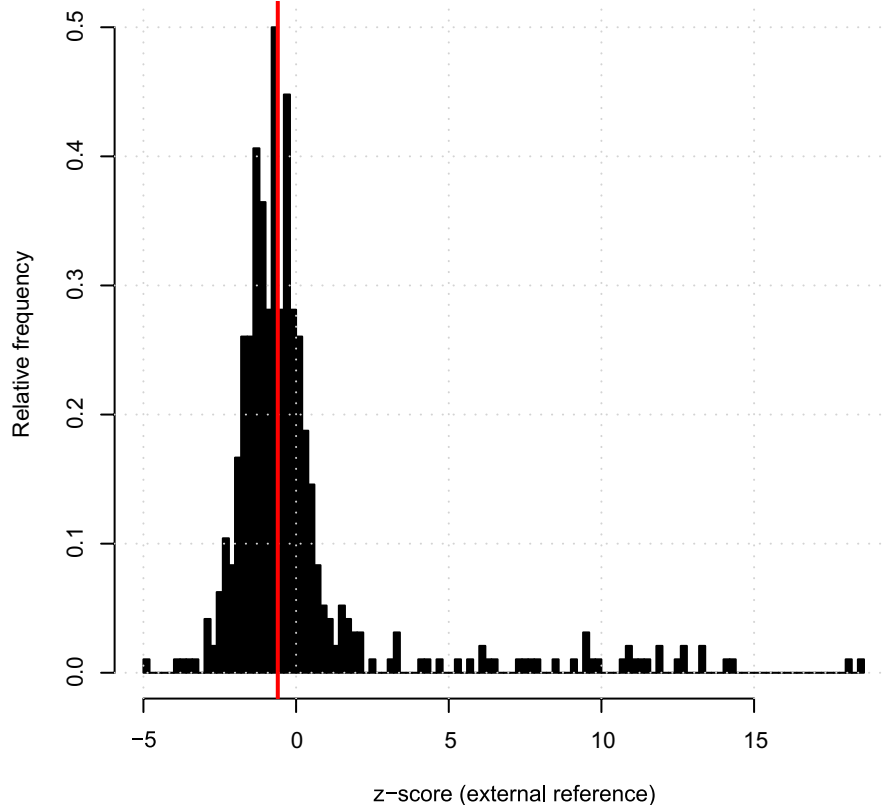
zymes to eliminate the maternal contribution of genomic regions that are methylated in fetal DNA and unmethylated in maternal DNA. The remaining nondigested fetal DNA is coamplified in the presence of a known amount of synthetic oligonucleotide to permit competitive polymerase chain reaction (PCR). This synthetic oligonucleotide has an identical sequence to the target genomic DNA, apart from 1 nucleotide that can be targeted by single-base extension and quantitative matrix-assisted laser desorption/ionization time-of-flight mass spectrometric analysis. To ensure accurate results, the assay comprises multiple markers in 4 different categories. Three markers are used to measure total DNA amounts. Three markers are used to measure chromosome Y copy numbers; 2 markers interrogate the efficiency of the methylation-specific digestion reaction, and 5 markers are used to measure fetal DNA amounts.

Methylation-based DNA discrimination was performed using 10 μ L of eluted DNA per reaction. All reagents and apparatus were obtained from Sequenom Inc, San Diego, CA, unless stated otherwise. Digestion of plasma DNA was performed for 30 minutes at 41°C by adding 25 μ L of a mixture containing 3.5X PCR buffer, 2.22 mmol/L MgCl₂, 10 U HhaI (New England Biolabs, Ipswich, MA), 10 U HpaII (New England Biolabs), and 10 U ExoI (New England Biolabs). Exonuclease was added to eliminate single-stranded DNA that would escape digestion and overestimate the fetal fraction. After the restriction was complete, the enzymes were inactivated and the DNA denatured by heating the mixture for 10 minutes at 98°C. All steps following the restriction reaction were performed according to Nygren et al.¹⁰

Library preparation

The extracted ccf DNA was used for library preparation without further fragmentation or size selection, because ccf DNA is already naturally fragmented, having an average length of approximately 160 base pairs. Low binding Eppendorf tubes were used to store 55 μ L of DNA eluent at 4°C following extraction until the library preparation had started.

FIGURE 1
Histogram of z-scores calculated using 24 reference samples



A red vertical line is drawn at $z = -0.6$ and corresponds to the median z-score for all samples; its expected position, under the assumption that reference samples have same distribution of chromosome 21 representation as all euploid samples in this study, is at $z = 0$.

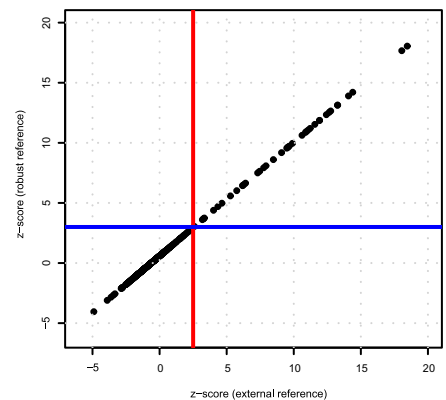
Ehrich. Noninvasive detection of fetal trisomy 21. *Am J Obstet Gynecol* 2011.

Storage times ranged from 24-72 hours. The library preparation was carried out according to the manufacturer's specifications (Illumina Inc) with some modifications. Enzymes and buffers were sourced from Enzymatics (End Repair Mix -LC; dNTP Mix [25 mmol/L each]; Exo(-) Klenow polymerase; 10X Blue Buffer; 100 mmol/L dATP; T4 DNA Ligase; 2X Rapid Ligation Buffer) and New England Biolabs (Phusion MM). Adapter oligonucleotides, indexing oligonucleotides, and PCR primers were obtained from Illumina Inc.

Library preparation was initiated by taking 40 μL of ccf DNA for end repair, retaining 15 μL for QC by FQA. End repair was performed with a final concentration of 1X End Repair buffer, 24.5 $\mu\text{mol/L}$ each dNTPs, and 1 μL of End Repair enzyme mix. The end repair reaction was carried out at room tempera-

ture for 30 minutes and the products were cleaned with Qiagen Qiaquick columns, eluting in 36 μL of elution buffer (EB). 3' mono-adenylation of the end-repaired sample was performed by mixing it with a final concentration of 1X Blue Buffer, 192 $\mu\text{mol/L}$ dATP, and 5 U of Exo(-) Klenow Polymerase. The reaction was incubated at 37°C for 30 minutes and cleaned up with Qiagen MinElute columns, eluting the products in 14 μL of EB. Adapters were ligated to the fragments by incubating for 10 minutes at room temperature with 1X Rapid Ligation buffer, 48.3 nmol/L Index PE Adapter Oligos, and 600 U T4 DNA Ligase. The ligation reaction was cleaned up with QiaQuick columns, and the sample eluted in 23 μL of EB. The adapter-modified sample was enriched by amplifying with a high-fidelity polymerase. The entire 23 μL eluent of each sample

FIGURE 2
Robust versus externally referenced z-scores



Relationship between z-scores calculated based on 24 external samples using mean and SD (x-axis) and using median and median absolute deviation (y-axis). The blue line indicates the empirically derived z-score cutoff of 2.5, which is comparable to the cutoff of 3 (red line) using the robust method.

Ehrich. Noninvasive detection of fetal trisomy 21. *Am J Obstet Gynecol* 2011.

was mixed with 1X Phusion MM, Illumina Inc PE 1.0 and 2.0 primers, and 1 of 12 index primers for a total PCR reaction volume of 50 μL . The sample was amplified in a 0.65-mL PCR tube using a MJ Research (Bio-Rad, Hercules, CA) Model PTC-200 thermal cycler. The PCR conditions were an initial denaturation at 98°C for 30 seconds, 15 cycles of denaturation at 98°C for 10 seconds, annealing at 65°C for 30 seconds, and extension at 72°C for 30 seconds. A final extension at 72°C for 5 minutes was followed by a 4°C hold. The PCR products were cleaned with MinElute columns and the libraries eluted in 17 μL of EB.

Quality control of generated sequencing library

The libraries were quantified via SYBR Green quantitative PCR (qPCR) analysis as outlined by Meyer et al.²⁴ Each library was diluted 1:10⁸ and quantified against a library standard using Power SYBR Green PCR Master Mix (ABI, Foster City, CA).

Each sample or standard was assayed in triplicate, including triplicate non-template control reactions. The sample was gently inverted or pipetted up and

down to mix, then spun down. In the reaction 2 μL of the 1:10⁸ dilution are added to a reaction mix containing 9 μL of Ultrapure Water, 12.5 μL 2x Power Mix, 0.5 μL of each forward (GAT ACG GCG ACC ACC GAG AT) and reverse (CAA GCA GAA GAC GGC ATA CGA G) primer at 10 $\mu\text{mol/L}$, and 0.5 μL of 1 U/ μL uracil-N-glycosylase. Amplification was performed on an ABI 7500 (Applied Biosystem, Foster City, CA). The cycling protocol began with a 2-minute uracil-N-glycosylase decontamination step at 50°C, this was followed by Hot-Start activation 95°C for 10 minutes. The program cycling was commenced and continued through 46 cycles of 15 seconds denaturation at 95°C followed by 1 minute of annealing/extension at 60°C. The final step was a 15-second denaturation at 95°C.

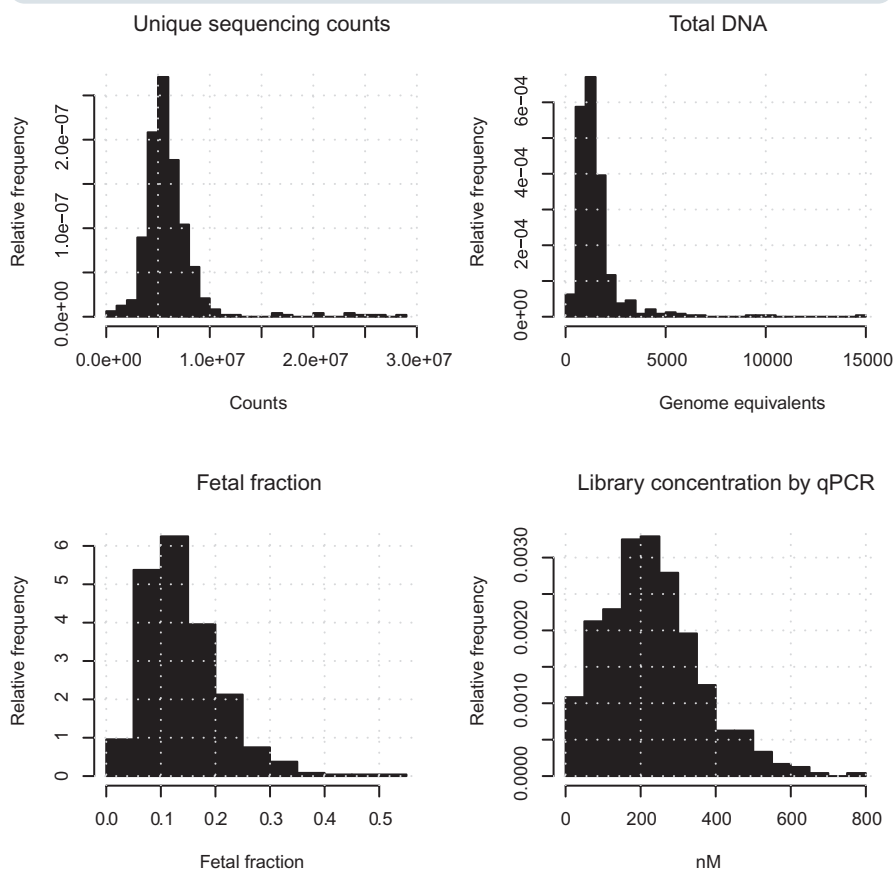
Clustering and sequencing

Clustering and sequencing were performed according to standard Illumina Inc protocols. Individual libraries were normalized to a 5-nmol/L concentration and then clustered in 4-plex format to a final flow cell loading concentration of 1.75 pmol/L per sample or 7 pmol/L per flow cell lane. The cBOT instrument and v4 Single-Read cBOT reagent kit (Illumina Inc) were used. Thirty-six cycles of single-read multiplexed sequencing were performed on the Genome Analyzer IIx with Paired-End module using v4 SBS reagent kits and supplemental Multiplex Sequencing Primer kits (Illumina Inc). Image analysis and base calling were performed with RTA1.6/SCS2.6 software (Illumina Inc). Sequences were aligned to the UCSC hg19 human reference genome (nonrepeat-masked) using CASAVA version 1.6 (Illumina Inc).

Data analysis

Sequence reads unique to a chromosome were counted, up to 1 mismatch (UI counts), and the chromosome 21-specific genomic representation was calculated based on these unique sequence reads. The fractional genomic representation of chromosome 21 (also referred to as the percentage of chromosome 21) was determined by dividing the number of sequence reads from chromosome 21

FIGURE 3
Distributions of quality control parameters



Histograms showing distribution of quality control parameters (each histogram has optimal number of bins, as calculated with Scott's rule). Very high values of unique sequence counts (>10 million) are mainly obtained from samples that were analyzed in monoplex.

Ehrich. Noninvasive detection of fetal trisomy 21. *Am J Obstet Gynecol* 2011.

by all sequence reads excluding sequence reads from chromosomes X and Y. The fractional genomic representation was then standardized by subtracting the mean of a control group and dividing by the standard deviation (SD) of that same control group. Using a set of known euploid samples as a control group, this method determines the distance in SD of the tested sample to the mean of the euploid reference group. This metric, standardized fractional genomic representation (the so-called z-scores), is the metric used to classify samples as euploid or trisomy 21. Details of this procedure are outlined in Chiu et al.²⁰

Ideally, the standardization process would be based on the true mean and true SD as calculated from a very large set of euploid samples. In the absence

of such a large set, a control group of 24 euploid samples from a previous experiment was used. Given both the limited sample size and the latent differences between the 2 different experiments, this control group may produce biased estimates of the true mean and true SD. In the design of the current study, similar to an expected clinical setting, it was anticipated that the majority of the samples would be euploid. Therefore the distribution of the z-scores should have a large normally distributed component centered on 0 and with SD close to 1. Any significant departure from this situation would be an indicator of improper standardization. Alternatively, a robust standardization can be employed given the data from the current experiment, by using

TABLE 1
Demographics of 449 analyzed samples

Demographic	Median	Range
Maternal age, y (n = 448)	37	(18–47)
Gestational age, wk (n = 448)	16	(8–36)
Maternal weight, lb (n = 425)	153	(96–314)
Variable	Percent	(N _{affected} /N _{total})
Indication for testing^a		
Positive serum screening	30.2	(133/441)
Advanced maternal age	68.3	(306/448)
Ultrasound abnormality	12.9	(57/441)
Positive family history	5.2	(23/441)
Not specified	10.2	(45/441)
Procedure		
CVS	19	(84/442)
Genetic amniocentesis	81	(358/442)
Confirmation		
Karyotype	59.9	(269/449)
FISH	2.9	(13/449)
Both	35.6	(160/449)
QF-PCR	1.6	(7/449)

For some patients not all information was available. Number of patients used to calculate statistics is indicated for each parameter.

CVS, chorionic villus sampling; FISH, fluorescent in situ hybridization; QF-PCR, quantitative fluorescent polymerase chain reaction.

^a Some patients had >1 indication.

Ehrlich. Noninvasive detection of fetal trisomy 21. *Am J Obstet Gynecol* 2011.

the median and median absolute deviation for the calculation of z-scores.

Twenty-four known euploid samples from a previous study were used to determine mean and SD of the percent of chromosome 21 representation needed for calculating the z-scores for the set of 480 samples. The determination of the mean and SD of the distribution of z-scores was performed by applying an iterative censoring approach. In each iteration we excluded the most extreme values (outside of 3 SD) and recalculated mean and SD. The values for mean and SD approached a stable value after 10 iterations. Using this method, we estimate the true mean to be -0.6 and the SD to be 1.03. Based on these values the empirically derived z-score cutoff was set to 2.5 (z-score cutoff = mean + 3 SD). This distribution of the z-scores is displayed in Figure 1. The z-score cutoff

was derived and applied to the data before unblinding.

We used the same 24 reference samples and sequenced them in monoplex format. The resulting data were used to calculate mean and SD for this monoplex reference dataset. For the set of 10 samples that were run in monoplex we did not have enough data available for an equivalent bias estimate. Consequently, we applied the same z-score cutoff to the entire dataset.

The empirically derived threshold of $z = 2.5$ correlates well with the cutoff of $z = 3$, when z-scores are calculated using the robust standardization procedure described earlier. In effect, we are still applying the same rule of 3 SD but our approach compensates for a biased control group. Figure 2 depicts this correlation.

In future analysis, the results from studies like the one presented here can be

used to derive a more robust mean and SD for the calculations of z-scores. Enrolling more patients will clearly benefit the analytical procedure and over time may make empirically derived cutoffs unnecessary.

Quality control criteria

All incoming plasma samples were inspected to assure each was intact and had sufficient volume.

Assay-specific QC criteria were derived from the results of the pilot studies. Figure 3 depicts the distribution of values of the QC assays obtained for the entire set of 467 samples that underwent QC. A subset of 449 samples had sufficient quality in all control assays to be enrolled into the analysis.

The following cutoffs were used: for samples that were analyzed in 4-plex:

- Minimum fetal fraction, as estimated with the FQA: 3.9%.
- Minimum total DNA per sample, as estimated with the FQA: 556 copies.
- Minimum library concentration, as measured with a qPCR assay: 32.3 nmol.
- Minimum number of total unique, postfilter counts (U1 counts): 3,044,976.

The following cutoffs were used for samples that were analyzed in monoplex:

- Minimum total DNA per sample, as estimated with the FQA: 556 copies.
- Minimum library concentration, as measured with a qPCR assay: 32.3 nmol.
- Minimum number of total unique, postfilter counts (U1 counts): 12,179,904.

The confidence intervals (CIs) for sensitivity and specificity were calculated using Wilson score method incorporating continuity correction.²⁵⁻²⁷

All calculations were done using the R environment for statistical computing.²⁸

Raw data for the analyzed patient samples will be made available upon request.

Pilot studies and implementation of calling rules

We carried out 3 pilot studies to assess the general feasibility of the multiplexing, to test the processing infrastructure, and to test the procedure for requesting

and receiving blinded samples from a third party. The data obtained in these pilot studies were used to establish the quality criteria and calling rules for the main study reported in this article.

In the first pilot study, we analyzed 86 samples from women with euploid fetuses and 10 from women with a trisomy 21 fetus: 84 were correctly identified as euploid and 8 were correctly identified as trisomy 21 and 4 samples were misclassified, including 2 false positives and 2 false negatives. The source documentation for all samples was reviewed and revealed that the 2 false-negative samples were incorrectly annotated by the third-party providing the sample, resulting in the selection of 2 euploid samples incorrectly classified as trisomy 21. Thus, in retrospect, these samples were correctly identified as euploid. The 2 false positives remained discrepant. The sample provider revised the sample selection and receiving procedure manuals to mitigate against future transcription errors.

Next, we tested the established internal sample processing infrastructure and the quality control (QC) documentation. In this second pilot study, we processed 96 samples from women with euploid fetuses according to standard operating procedures. We correctly identified 95 samples as euploid and 1 sample was misclassified as a trisomy 21 sample.

Finally, in the third pilot study, we verified the modified protocol to request, review, and process a set of blinded samples and to send the results to an external biostatistician. Forty euploid samples and 4 trisomy 21 samples were requested from PRA and BST. The results documented that 1 sample had to be excluded from the analysis because of a failed sequencing reaction. The remaining 43 samples were classified correctly (40 euploid and 3 trisomy 21).

RESULTS

The study comprised 480 plasma samples from pregnant women who were a priori at high risk for fetal trisomy 21. Details of sample demographics are pro-

TABLE 2
Reasons for sample exclusion

Variable	Criteria	Samples excluded
Incoming inspection	Plasma volume <3.5 mL	9
Processing error	Sample dropped during DNA extraction	1
	Samples mixed into each other	2
	Tube broke during centrifugation	1
		Subtotal: 13
Quality control ^a	Fetal percentage	7
	Total DNA	7
	Library concentration	15
	Unique counts	11
		Subtotal: 18 ^a

Quality control was performed with 2 sets of acceptance criteria. One set was applied to all 457 samples that were sequenced in 4-plex and second set was applied to all 10 samples that were sequenced in monoplex. Cutoff values are depicted in Figure 1. Set of 31 samples excluded from analysis contained 28 euploids and 3 trisomy 21 samples.

^a Some samples failed >1 criterion.

Ehrich. Noninvasive detection of fetal trisomy 21. *Am J Obstet Gynecol* 2011.

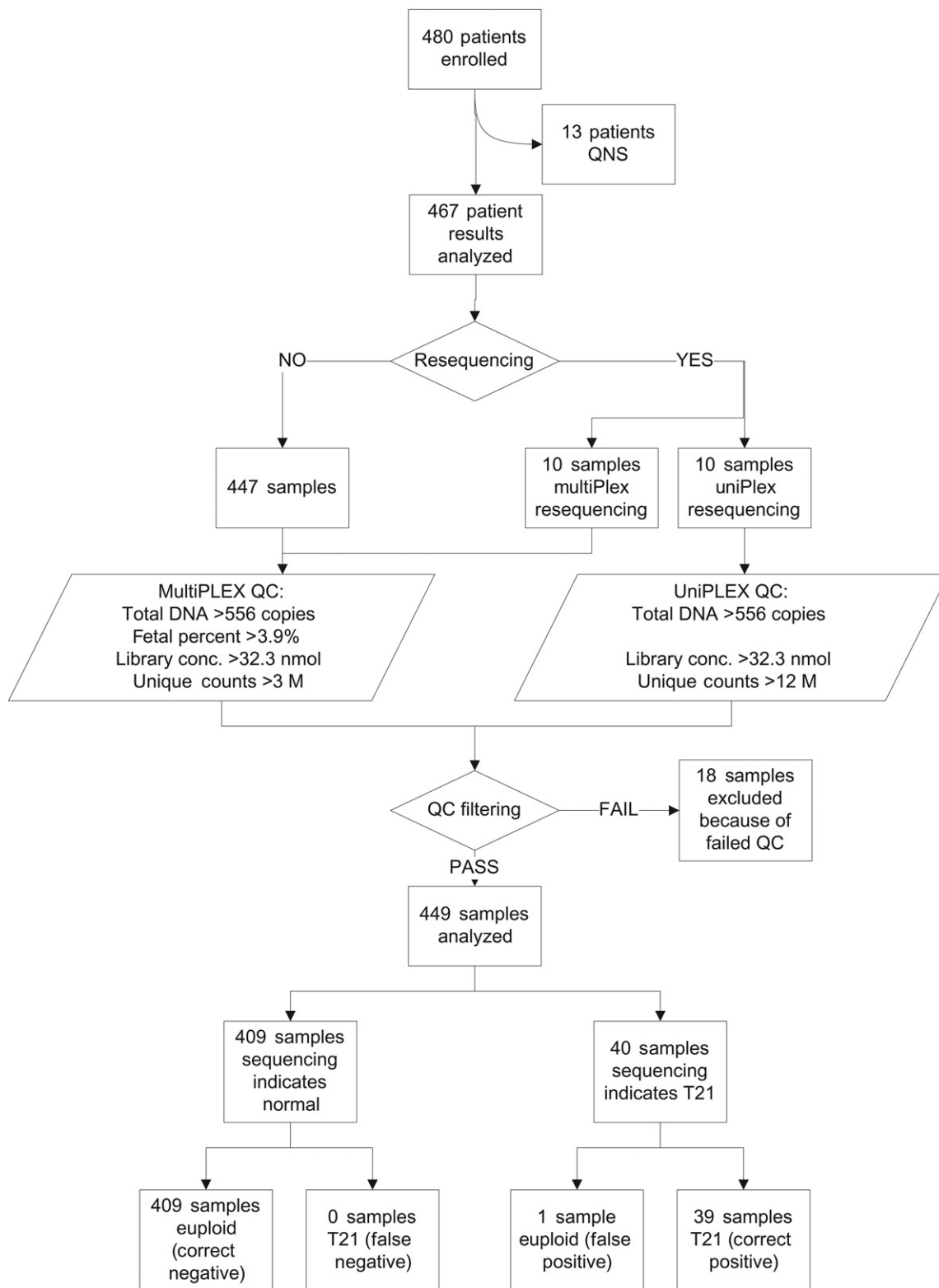
vided in Table 1. The samples were processed in 5 batches of 96 each. Each batch required approximately 10 days from DNA extraction to the final sequencing result. Since batches were processed in parallel, the entire study (including data analysis) was completed in July and August 2010.

Thirteen samples of the 480 were excluded as preanalytic failures because of insufficient quality (see Table 2 for details and quality-not-sufficient definitions). Twenty samples of the remaining 467 failed the initial sequencing QC. These showed sequence counts below the predetermined cutoff (<3 million) but high-quality libraries (concentration >32.3 nmol). Libraries from samples that had >3.9% fetal DNA were resequenced in their original tetraplex format (n = 10), while samples with lower fetal DNA percentages were resequenced in monoplex (n = 10). Only the resequencing results were considered. Application of assay QC cutoffs to the entire set of 467 samples thus excluded 18 samples (4%) from analysis. Figure 4 and Table 2 provide an overview of the QC process. Three excluded samples were subsequently identified as trisomy 21: 1 tube broke during centrifugation, 1 failed library and sequencing QC, and 1 failed sequencing QC after it had been resequenced.

Of the 449 samples that passed QC and were analyzed, 410 were euploid and 39 were trisomy 21. Of the trisomy 21 samples, all 39 were correctly identified. Of the euploid samples, 409 of the 410 were correctly classified while 1 was misclassified as trisomy 21 (false positive). Thus, the overall classification showed 100% sensitivity (95% CI, 89–100%) and 99.7% specificity (95% CI, 98.5–99.9%). Figure 5 shows the z-scores for euploid and trisomy 21 samples. Additionally, we plotted the nonnormalized fractional representation of chromosome 21 stratified by sample type and multiplexing level (Figure 6). By use of the appropriate reference set for each of the mono- and tetraplex samples, a single z-score based classification cutoff can be applied.

To assess the value of the QC methods, we also performed a post hoc analysis of the classification accuracy of all 467 samples before applying our QC metrics (excluding only the 13 samples with preanalytic failures). The accuracy was comparable to the 449 samples described above. However, the full set of 467 had 1 trisomy 21 incorrectly called euploid (false negative). The sensitivity in this sample group was 97% (95% CI, 86–100%) and the specificity of 99.8% was marginally higher due to the increased sample number (95% CI, 98.5–100%).

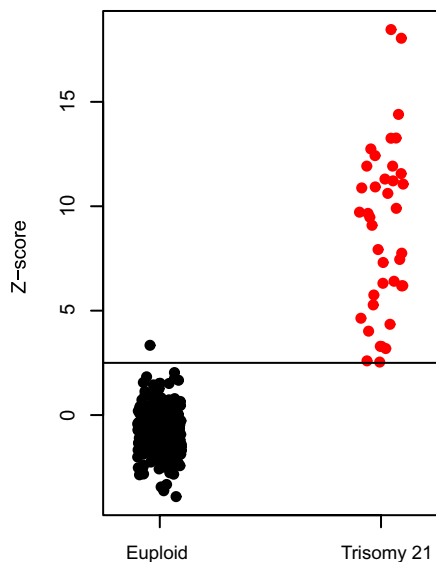
FIGURE 4
Flowchart representing decision matrix for sample analysis



QC, quality control; QNS, quality not sufficient.

Ehrich. Noninvasive detection of fetal trisomy 21. *Am J Obstet Gynecol* 2011.

FIGURE 5
The z-scores for euploid (*black*)
and trisomy 21 (*red*) samples



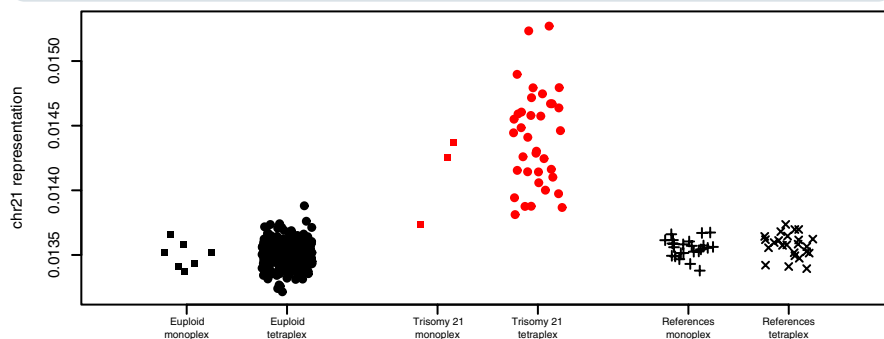
Depicted are the z-scores for the set of 449 analyzed samples, including 410 euploid (*black*) and 39 trisomy 21 (*red*) samples. To avoid overplotting, data points are separated in horizontal direction by some small random "jitter." The z-score cutoff of 2.5 is shown as *solid line*. Because a set of 24 euploid samples from an external reference set was used to calculate the mean and SD, the distribution of z-scores for euploid samples centers on -0.6 instead of 0, and hence a cutoff of $z = 2.5$ is more appropriate than the usual cutoff of 3.0.

Ehrich. *Noninvasive detection of fetal trisomy 21. Am J Obstet Gynecol* 2011.

COMMENT

Noninvasive prenatal aneuploidy detection has evolved and steadily improved over the last 20 years. Now, prenatal screening is on the verge of another major transformation. Rapidly advancing sequencing technologies are reaching a point where an individual's personal genome becomes available as a standard source for medically relevant genetic information. We took advantage of recent advances in MPSS, not to determine the full genomic sequence, but rather to use sequence-specific tags of known chromosomal location as a quantitative representation of individual chromosomes in maternal plasma. In this test, a sufficiently disproportionate representation of chromosome 21 is indicative of fetal

FIGURE 6
Distribution of chromosome 21 fractional representation



Non-normalized fractional representation of chromosome 21 stratified by multiplexing level and sample type. The data for reference sets are slightly different. In particular, the estimated variance of the reference set for monoplex samples is lower than for tetraplex samples (both $n = 24$). Therefore, after standardization to the appropriate reference set, the trisomy 21 (T21) sample with the lowest chromosome 21 representation in monoplex can still be correctly classified.

Ehrich. *Noninvasive detection of fetal trisomy 21. Am J Obstet Gynecol* 2011.

trisomy 21.²⁰ The present blinded study comprised 480 patient samples of which 42 samples had a fetus with trisomy 21. Based on predefined, pretesting QC criteria, 31 samples including 3 trisomy 21 samples had insufficient quality and were excluded from further processing. The overall rate of successful completion of sample analysis in this study is $>94\%$. If confirmed in subsequent studies, this rate would be clinically acceptable for such a high complexity assay. Thirteen samples were excluded because of insufficient sample volume or processing errors, both unrelated to the sequencing assay but likely to occur in clinical practice (Table 2). A clinical laboratory will resort to a second plasma aliquot of the initial patient sample if the standard collection procedure uses more than one 10-mL tube of blood. Eighteen additional samples were excluded through the application of our QC processes. For some of these patients the second plasma aliquot might yield useable results.

In the remaining 449 samples, the test achieved a sensitivity of 100% and $>99\%$ specificity. While these numbers are encouraging, they must be viewed objectively. The number of samples analyzed in this feasibility study (39 trisomy 21 samples) is still relatively limited, and this small sample cohort may not span the entire spectrum of cases that presents in clinical practice. Future clinical vali-

dation studies with larger numbers of trisomy 21 samples will reveal the true sensitivity of this noninvasive aneuploidy test. The specificity can be assessed much more accurately from our dataset. The lower 95% CI for the specificity is $>98\%$. Given the large number of euploid samples tested, the specificity is less likely to change significantly when more euploid samples from high-risk pregnancies are tested.

Further improvements will be necessary over the next years for this test to find more widespread adoption as a screening tool. The complexity of the assay has to be reduced to allow decentralized processing in less specialized clinical laboratories. Lower failure rates and quicker turnaround times are desirable to meet patient expectations in prenatal screening. Ideally, a screening test will also include trisomy 18, as done in current practice, and potentially trisomy 13. Finally, the clinical sensitivity and specificity achieved in the high-risk population needs to be validated in a general population screen.

Today trisomy 21 detection by MPSS in high-risk women is complicated but manageable if implemented with appropriate procedures and quality metrics. This assay holds considerable promise for prenatal aneuploidy detection in pregnant women at high risk for a trisomy 21 fetus. If the current results are confirmed in a larger

multicenter trial, this method could have a substantial impact on future clinical practice. The number of pregnant women who are candidates to undergo invasive testing owing to advances in prenatal fetal ultrasound scanning and serum biochemical screening for fetal chromosome abnormalities, including Down syndrome, has increased with the advent of modern, more complex screening algorithms. There remains the traditional risk group of women who are aged ≥ 35 years at term, and this group too has been growing over the last 3 decades. In 1980 in the United States, approximately 4.5% of all pregnant women were of advanced maternal age; by 2007, that figure had increased to 14%.²⁹ Consequently, being at high risk for a trisomy 21 pregnancy is becoming more relevant to a larger group of women. Although the merits of current screening technologies cannot be understated, ultimately these women are faced with the decision to undergo an invasive procedure, which to this day continues to add risk to the pregnancy. Our data show that a non-invasive prenatal trisomy 21 test from ccf DNA might be used in concert with other clinical assessments, such as ultrasound, and become an option to better identify those women who would, or would not, benefit from confirmatory invasive diagnostic tests.

In this study, we have taken the next step in evaluating the use of ccf DNA sequencing and have shown that it has the potential to be highly accurate. We have also implemented technical improvements to overcome the previous barrier to implementing this new technology by increasing the throughput and reducing the cost. These technical assay improvements are essential to enable an adoption into clinical practice once a planned clinical validation of the test in a larger multicenter study has been completed. ■

ACKNOWLEDGMENTS

We thank Dennis Lo and Rossa Chiu for advice on the sequencing and data analysis protocols and the provision of the standard for DNA library quantification. We thank Yong Chen, DPhil, Dave Henderson, and Ryan Biltz for their excellent work on establishing early versions of the analytical scripts and the necessary information technology infrastructure. We also thank Haijing Lu, Rochelle Jean-Jacques, and Stephanie

Lin for their outstanding work on sample processing and DNA extractions. We are also grateful to Sonia Minassian, DPhil, for acting as the independent third party within the unblinding process as well as Aruna Bansal, DPhil, for independent verification of the sample datasets submitted to the third party for unblinding. We wish to gratefully acknowledge the assistance of the following individuals in the recruitment of patients for this study: Erol Amon, MD, St. Louis University; Keith Aqua, MD and Cynthia Goldberg, MD, Visions Clinical Research; Susan Baker, MD, University of South Alabama; Richard Beyerlein, MD, Clinical Trials of America; Richard Bronsteen, MD, William Beaumont Hospital; Jay Cohen, MD, Discovery Clinical Research; Larry Cousins, MD, San Diego Perinatal Center; Micah Harris, MD, Women's Health Research; Greigh Hirata, MD, Fetal Diagnostic Institute of the Pacific; Heather Holmes, MD, Texas Tech University Health Services Center; Meena Khandelwal, MD, Cooper University Hospital; Helena Kilpatrick, MD, Magnolia Ob/Gyn Research Center; Brian Kirshon, MD, Houston Perinatal Associates; Robert Klein, MD, Global Ob/Gyn Center; Steven Lenowitz, MD; Jorge Lodeiro, MD, Kingsport Perinatology; B. Kerry Lowder, MD, Woman's Clinic; Elaine Lyon, PhD, ARUP Laboratories; Fred-erico Mariona, MD, Maternal Fetal Medicine of MI; Dibe Martin, MD, South Florida Perinatal Medicine; Rovena Reagan, MD, Women's Health Care of San Diego; Jaime Rodriguez, MD, Sheridan Clinical Research; Christopher Robinson, MD, Medical University of South Carolina; Carl Rose, MD, Mayo Clinic; Milroy Samuel, MD, Complete Healthcare for Women; Asad Sheikh, MD, Gioi Smith-Nguyen, MD, Sharp Grossmont Center for Clinical Research; Spectrum Health Maternal Fetal Medicine; Advanced Biosciences Resources.

REFERENCES

1. Centers for Disease Control and Prevention. Improved national prevalence estimates for 18 selected major birth defects-United States, 1999-2001. *Morbidity and Mortality Weekly* 2006;54:1301-5.
2. Savva GM, Morris JK, Mutton DE, Alberman E. Maternal age-specific fetal loss rates in Down syndrome pregnancies. *Prenat Diagn* 2006; 26:499-504.
3. Morris JK, Mutton DE, Alberman E. Revised estimates of the maternal age specific live birth prevalence of Down's syndrome. *J Med Screen* 2002;9:2-6.
4. Haddow JE, Palomaki GE, Canick JA, Knight GJ. Prenatal screening for open neural tube defects and Down's syndrome. In: Rodeck CH, Whittle MJ, eds. *Fetal medicine: basic science and clinical practice*. London (United Kingdom): Churchill Livingstone Elsevier; 2009.
5. Wald NJ, Watt HC, Hackshaw AK. Integrated screening for Down's syndrome on the basis of tests performed during the first and second trimesters. *N Engl J Med* 1999;341:461-7.

6. Merkatz IR, Nitowsky HM, Macri JN, Johnson WE. An association between low maternal serum alpha-fetoprotein and fetal chromosome abnormalities. *Am J Obstet Gynecol* 1984; 148:886-94.
7. American College of Obstetricians and Gynecologists. ACOG practice bulletin no. 77. *Obstet Gynecol* 2007;109:217-25.
8. Lo YMD, Corbetta N, Chamberlain PF, et al. Presence of fetal DNA in maternal plasma and serum. *Lancet* 1997;350:485-7.
9. Sikora A, Zimmermann BG, Rusterholz C, et al. Detection of increased amounts of cell-free fetal DNA with short PCR amplicons. *Clin Chem* 2010;56:136-8.
10. Nygren AOH, Dean J, Jensen TJ, et al. Quantification of fetal DNA by use of methylation-based DNA discrimination. *Clin Chem* 2010;56:1627-35.
11. Fan HC, Blumenfeld YJ, Chitkara U, Hudgins L, Quake SR. Analysis of the size distributions of fetal and maternal cell-free DNA by paired-end sequencing. *Clin Chem* 2010;56: 1279-86.
12. Lun FM, Chiu RW, Allen Chan KC, Yeung Leung T, Kin Lau T, Dennis Lo YM. Microfluidics digital PCR reveals a higher than expected fraction of fetal DNA in maternal plasma. *Clin Chem* 2008;54:1664-72.
13. Lo YMD, Zhang J, Leung TN, Lau TK, Chang AMZ, Hjelm NM. Rapid clearance of fetal DNA from maternal plasma. *Am J Hum Genet* 1999;64:218-24.
14. Chitty LS, van der Schoot CE, Hahn S, Avent ND. SAFE-the special non-invasive advances in fetal and neonatal evaluation network: aims and achievements. *Prenat Diagn* 2008; 28:83-8.
15. Dhallan R, Au WC, Mattagajasingh S, et al. Methods to increase the percentage of free fetal DNA recovered from the maternal circulation. *JAMA* 2004;291:1114.
16. Tong YK, Jin S, Chiu RWK, et al. Noninvasive prenatal detection of trisomy 21 by an epigenetic-genetic chromosome-dosage approach. *Clin Chem* 2010;56:90-8.
17. Chu T, Burke B, Bunce K, Surti U, Allen Hogge W, Peters DG. A microarray-based approach for the identification of epigenetic biomarkers for the noninvasive diagnosis of fetal disease. *Prenat Diagn* 2009;29:1020-30.
18. Lo YM, Tsui NB, Chiu RW, et al. Plasma placental RNA allelic ratio permits noninvasive prenatal chromosomal aneuploidy detection. *Nat Med* 2007;13:218-23.
19. Tsui NBY, Akolekar R, Chiu RWK, 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 RW, Chan KC, Gao Y, et al. Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc Natl Acad Sci U S A* 2008;105:20458-63.

- 21.** 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 U S A* 2008;105:16266-71.
- 22.** Chiu RW, Sun H, Akolekar R, 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.
- 23.** Chu T, Bunce K, Hogge WA, Peters DG. Statistical model for whole genome sequencing and its application to minimally invasive diagnosis of fetal genetic disease. *Bioinformatics* 2009;25:1244-50.
- 24.** Meyer M, Briggs AW, Maricic T, et al. From micrograms to picograms: quantitative PCR reduces the material demands of high-throughput sequencing. *Nucleic Acids Res* 2008;36:e5.
- 25.** Blythe CR, Still HA. Binomial confidence intervals. *J Am Stat Assoc* 1983;78:108-16.
- 26.** Fleiss JL. *Statistical methods for rates and proportions*. New York: Wiley; 1981.
- 27.** Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* 1998;17:857-72.
- 28.** R Development Core Team. *R: a language and environment for statistical computing*. Vienna (Austria): R Foundation for Statistical Computing; 2008.
- 29.** Martin JA, Hamilton BE, Sutton PD, et al. Births: final data for 2007. *Natl Vital Stat Rep* 2010;58:1-86.