Noninvasive prenatal diagnosis of monogenic diseases by digital size selection and relative mutation dosage on DNA in maternal plasma

Fiona M. F. Lun[a,b,1], Nancy B. Y. Tsui[a,b,1], K. C. Allen Chan[a,b], Tak Y. Leung[c], Tze K. Lau[c], Pimlak Charoenkwan[d], Katherine C. K. Chow[a,b], Wyatt Y. W. Lo[a,b], Chanane Wanapirak[d], Torpong Sanguansermsri[d], Charles R. Cantor[e,2], Rossa W. K. Chiu[a,b], and Y. M. Dennis Lo[a,b,2]

[a]Centre for Research into Circulating Fetal Nucleic Acids, Li Ka Shing Institute of Health Sciences, Departments of [b]Chemical Pathology and [c]Obstetrics and Gynaecology, Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR; [d]Faculty of Medicine, Chiang Mai University, Thailand; and [e]Sequenom, San Diego, CA 92121

Contributed by Charles R. Cantor, October 15, 2008 (sent for review September 18, 2008)

Prenatal diagnosis of monogenic diseases, such as cystic fibrosis and β-thalassemia, is currently offered as part of public health programs. However, current methods based on chorionic villus sampling and amniocentesis for obtaining fetal genetic material pose a risk to the fetus. Since the discovery of cell-free fetal DNA in maternal plasma, the noninvasive prenatal assessment of paternally inherited traits or mutations has been achieved. Due to the presence of background maternal DNA, which interferes with the analysis of fetal DNA in maternal plasma, noninvasive prenatal diagnosis of maternally inherited mutations has not been possible. Here we describe a digital relative mutation dosage (RMD) approach that determines if the dosages of the mutant and wild-type alleles of a disease-causing gene are balanced or unbalanced in maternal plasma. When applied to the testing of women heterozygous for the CD41/42 (–CTTT) and hemoglobin E mutations on *HBB*, digital RMD allows the fetal genotype to be deduced. The diagnostic performance of digital RMD is dependent on interplay between the fractional fetal DNA concentration and number of DNA molecules in maternal plasma. To achieve fetal genotype diagnosis at lower volumes of maternal plasma, fetal DNA enrichment is desired. We thus developed a digital nucleic acid size selection (NASS) strategy that effectively enriches the fetal DNA without additional plasma sampling or experimental time. We show that digital NASS can work in concert with digital RMD to increase the proportion of cases with classifiable fetal genotypes and to bring noninvasive prenatal diagnosis of monogenic diseases closer to reality.

digital PCR | fetal DNA | mass spectrometry | microfluidics | thalassemia

Noninvasive prenatal diagnosis of fetal genetic diseases would obviate the risk of fetal miscarriage associated with current invasive procedures such as chorionic villus sampling. The discovery of cell-free fetal DNA in maternal plasma has offered new avenues for noninvasive prenatal diagnosis (1). Reliable noninvasive assessment of paternally inherited traits such as fetal sex and *RHD* status for the management of sex-linked disorders and RhD incompatibility, respectively, by circulating fetal DNA analysis has been achieved (2) and implemented for clinical use (3). Recently, noninvasive prenatal detection of fetal trisomy 21 has also been achieved (4, 5). Monogenic diseases, such as β-thalassemia and cystic fibrosis, are the other main conditions for which prenatal diagnosis is considered (6). However, a number of issues have hindered development in this area.

First, fetal DNA circulates in maternal plasma within a high background of maternal DNA (7). The maternally inherited fetal alleles present in maternal plasma are therefore difficult to discern from the background DNA of the mother. Thus, investigators have so far focused on the plasma detection of paternally inherited fetal alleles that are not present in the maternal genome (8). By detecting the presence of fetal-specific paternally inherited mutant alleles in maternal plasma, diagnosis of autosomal dominant diseases transmitted by the father could be made noninvasively (9–11), whereas the absence of such alleles could be used to exclude the fetal inheritance of autosomal recessive diseases (12, 13). These strategies have been applied to achondroplasia, myotonic dystrophy, Huntington chorea, and β-thalassemia (9–13). However, those approaches could not be applied to situations where the mother has an autosomal dominant mutation or when the mother and father are both carriers for the same autosomal recessive mutation (13).

Another issue that hindered research on circulating fetal DNA is its low concentration in maternal plasma. Although we recently showed that cell-free fetal DNA is present at higher concentrations than previously thought, it still amounts to only ≈10% to 20% of all DNA in maternal plasma (7, 14). Low fetal DNA concentration in maternal plasma has led to false-negative results and wrong diagnoses (15). Quantitative analysis of circulating fetal DNA is also less precise at low concentrations (5). Hence, researchers have been investigating methods for circulating fetal DNA enrichment. DNA molecules in maternal plasma are fragmented, with the fetal ones shorter than the background maternal ones (16). Li *et al.* (17) used gel electrophoresis to select for short DNA molecules in maternal plasma for enriching fetal DNA and reported improved sensitivities in detecting paternally inherited fetal β-thalassemia point mutations. Researchers have also attempted to suppress the amount of maternal background DNA (18). However, the gel electrophoresis technique may be prone to DNA contamination, and the suppression technique has not been universally reproducible (19–21).

Although the background maternal DNA interferes with the analysis of fetal DNA in maternal plasma, we recently developed methods for the noninvasive detection of fetal trisomy 21 (5). We applied digital PCR (22) and developed two approaches, namely digital RNA-SNP and digital relative chromosome dosage (RCD) for fetal aneuploidy detection. Both methods exploit the high analytical precision of digital PCR to detect the presence of an overrepresentation of chromosome 21 sequences in maternal

Author contributions: R.W.K.C. and Y.M.D.L. designed research; F.M.F.L., N.B.Y.T., K.C.A.C., T.Y.L., T.K.L., P.C., K.C.K.C., W.Y.W.L., C.W., and T.S. performed research; F.M.F.L., N.B.Y.T., K.C.A.C., C.R.C., R.W.K.C., and Y.M.D.L. analyzed data; and R.W.K.C. and Y.M.D.L. wrote the paper.

Conflict of interest statement: F.M.F.L., N.B.Y.T., K.C.A.C., C.R.C., R.W.K.C., and Y.M.D.L. have filed patent applications on the detection of fetal nucleic acids in maternal plasma for noninvasive prenatal diagnosis. Part of this patent portfolio has been licensed to Sequenom. C.R.C. is chief scientific officer of and holds equities in Sequenom. Y.M.D.L is a consultant to and holds equities in Sequenom.

Freely available online through the PNAS open access option.

[1]F.M.F.L. and N.B.Y.T. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: ccantor@sequenom.com or loym@cuhk.edu.hk.

This article contains supporting information online at www.pnas.org/cgi/content/full/0810373105/DCSupplemental.

plasma for pregnancies involving a trisomy 21 fetus. Digital RNA-SNP determines if an imbalance between heterozygous alleles of a fetal-derived placentally expressed RNA transcript originating from chromosome 21 exists in maternal plasma. Digital RCD determines if there exists an overrepresentation of the total (maternal + fetal) amount of DNA sequences from a chromosome 21 locus with reference to one on another chromosome. Digital RCD is feasible in theory, but fetal DNA enrichment would be needed to enhance its practicality.

Here we propose to adopt the principles of both digital RNA-SNP and digital RCD to develop a digital relative mutation dosage (RMD) approach for the noninvasive prenatal diagnosis of monogenic diseases. Digital RMD allows one to determine if the fetus has inherited the maternal mutant allele despite the coexistence of the background maternal DNA in maternal plasma. To enhance the efficiency of digital RMD and digital RCD, we developed a molecular enrichment strategy, termed digital nucleic acid size selection (NASS), which increases the fractional DNA concentration of the fetal loci of interest and thus helps to overcome the challenge posed by low fetal-DNA concentrations.

Results

Digital RMD. The principles of digital RMD are schematically illustrated in Fig. 1*A*. Clinically, digital RMD is required when a female heterozygous for the locus of interest is pregnant. Similar to digital RNA-SNP (5), digital real-time PCR assays were developed to discriminate the mutant from wild-type alleles of the gene locus. We determine if the amounts of mutant and wild-type alleles in a maternal plasma sample are in allelic balance or imbalance. Allelic balance is expected when the fetal genotype is identical to that of the mother (i.e., heterozygous). Allelic imbalance occurs if there is an under- or overrepresentation of the mutant allele with respect to the wild-type allele—underrepresentation would suggest a homozygous normal fetus, while overrepresentation would indicate a fetus homozygous for the mutant allele.

Digital RMD is similar to digital RCD (5) in that the allelic ratio between the mutant and wild-type alleles is determined by counting both the maternal and fetal contributions. Thus, the fractional fetal DNA concentration directly influences the expected extent of allelic imbalance when the fetus is homozygous for either allele. For example, in a maternal plasma sample containing 100 genome equivalents (GE) of DNA with 10% fetal DNA, the maternal compartment would contribute 90 GE of DNA. When the pregnancy involves a heterozygous mother and a homozygously affected fetus, the mother contributes 90 copies of each of the mutant and wild-type alleles; the fetus contributes 20 (2×10) copies of the mutant allele. Hence, there would be a total of 110 (90 + 20) copies of the mutant allele and 90 copies of the wild-type allele. The mutant allele is overrepresented by a ratio of 11:9 with respect to the wild-type allele. The expected allelic ratio is given by (1 + fetal DNA fraction)/(1 − fetal DNA fraction).

We apply a statistical approach termed sequential probability ratio test (SPRT) to objectively determine if allelic imbalance is present (5, 23). Imbalance suggests that the fetus is homozygous for the allele (mutant or wild-type). Balance suggests that the fetus is heterozygous. Alternatively, the SPRT analysis may indicate that there is inadequate statistical evidence for genotype classification (i.e., unclassifiable) and more digital PCR runs would be needed for the sample until statistical confidence could be reached for genotype classification.

Model System. We assessed the RMD concept in a model system involving mixtures of male and female DNA. We obtained blood cell samples from 12 male and 12 female subjects. We mixed each of the female blood cell DNA (genotype XX, to denote the sex chromosome status) with a 3-fold excess of male blood cell DNA (genotype XY), thus producing 12 DNA mixtures with 25% of DNA with the XX genotype in a background of 75% DNA with the XY genotype. We determined the amounts of sequences from chromosomes X and Y by a reported digital real-time PCR assay targeting the homologous *ZFX* and *ZFY* genes, respectively (7). Genotype classification findings are shown in supporting informa-

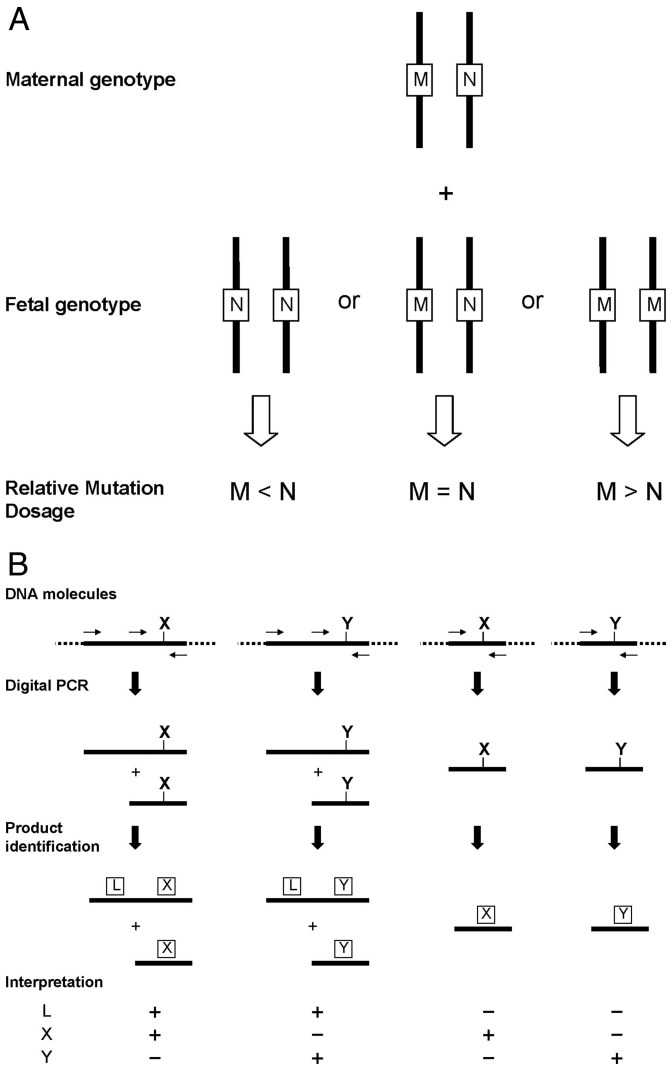

Fig. 1. Schematic illustration of the principles of (*A*) digital RMD and (*B*) digital NASS. (*A*) When a pregnant woman and her fetus are both heterozygous for a gene mutation, the amounts of the mutant allele (M) and wild-type allele (N) would be in allelic balance in maternal plasma. When the fetus is homozygous for the wild-type or mutant allele, there would be an underrepresentation or overrepresentation of the mutant allele, respectively. Digital RMD determines if the mutant and wild-type alleles in maternal plasma are in allelic balance or imbalance. (*B*) The scheme of the *ZFY/X* digital NASS assay is shown. The assay can discriminate between *ZFX*, denoted by X, and *ZFY*, denoted by Y, DNA molecules. In addition, the assay can distinguish if the *ZFX* and *ZFY* DNA molecules are long or short. Digital PCR is performed using two forward primers and one reverse primer (arrows), or vice versa, that are oriented to produce a short amplicon overlapping with the long amplicon. When a single DNA molecule at least as long as that specified by the long amplicon is captured in the reaction well, both the long and short PCR products would be generated. When a single DNA molecule shorter than the span of the long amplicon is captured, only the short amplicon would be generated. The presence of the long and/or short amplicons can be detected by strategically located hybridization probes or extension primers. An extension primer, L (boxed), is designed to detect the presence of the long amplicon. An extension primer is located within the short amplicon and the extension products are used to discriminate the *ZFX* and *ZFY* alleles (shown as boxed X and boxed Y, respectively). The identities of the DNA molecules could be interpreted by counting the products present within each well. +, present; −, absent.

Table 1. Digital RMD and fetal genotype classification of CD41/42 (–CTTT)

Cases	No. of panels*	Fetal genotype	M count	N count	Both positive†	Net M‡	Net N‡	Actual fetal %§	Reference allele¶	m_r‖	SPRT**
M62	24	NN	424	517	14	410	503	10.7	M	0.02	NN
M66	12	NN	201	264	7	194	257	10.7	M	0.02	NN
M74	12	MN	393	842	30	363	812	10.0	M	0.04	**NN**
M93	12	NN	269	344	8	261	336	15.3	M	0.03	NN
M99	24	MN	709	696	21	688	675	6.2	N	0.04	U

M, mutant allele; N, wild-type allele.
*Each panel consists of 765 wells.
†Wells positive for both alleles.
‡Wells only positive for the mutant allele and not wild-type allele or vice versa.
§Determined by the digital *ZFY/X* assay.
¶Reference allele is the one with the lesser count in each sample.
‖Template concentration of the reference allele.
**U, unclassified; bold font, misclassified.

tion (SI) Table S1. SPRT analysis assessed if *ZFX* was equally represented or overrepresented with respect to *ZFY*. As shown in Table S1, the XX genotype in the minor DNA population was correctly classified in all 12 DNA mixtures.

Digital RMD Analysis of Hemoglobinopathy-Causing Mutations in DNA Mixtures. We developed digital real-time allelic discrimination PCR assays to target two mutations on *HBB*, namely a β-thalassemia mutation, CD41/42 (–CTTT) (12), and the hemoglobin E (HbE) (G → A) mutation (24). We assessed the use of digital RMD in determining the genotype of the minority DNA population in genomic DNA mixtures comprised of ≈5% to ≈50% minor DNA in a heterozygous background. Heterozygous maternal blood cell DNA was mixed with blood cell DNA from heterozygous males, or males homozygous for the mutant or wild-type alleles. We used the *ZFY/X* digital assays to measure the actual male DNA fractional concentration in the DNA mixtures (7). We defined the reference allele as the allele with the lesser count in each digital PCR experiment. For digital PCR data interpretation, we defined the average template concentration per well of each experiment based on that of the reference allele (m_r). SPRT curves relevant for the m_r and male DNA concentration of each specific sample was used for genotype classification of that sample. Digital RMD results for the CD41/42 genotyping are shown in Table S2. All samples were correctly classified.

Blood cell DNA from nine males with different genotypes with respect to the HbE mutation were each mixed at proportions ranging from ≈10% to ≈50% in a background of female heterozygous DNA. We used the *ZFY/X* digital assays to measure the actual male DNA fractional concentration in the DNA mixtures. Digital RMD results for HbE genotyping are shown in Table S3. Case 2 at <10% male DNA was misclassified. All other cases were correctly classified over the full range of tested male DNA concentrations.

Fetal *HBB* Genotyping by Digital RMD in Maternal Plasma. We applied digital RMD to determine the fetal genotype in maternal plasma samples collected from second-trimester (median gestational age [GA]: 19 weeks; range: 18–20 weeks) pregnant women who were carriers for the CD41/42 or HbE mutations. All cases involved male fetuses. Digital PCR analysis was performed until genotype classification by SPRT was possible or until the maternal plasma was exhausted. Findings are shown in Table 1 and Table S4. The fetal genotypes were correctly classified in 5 of 10 cases. There was an incorrect classification for case M74. Cases M99, M53, M56, and M86, all with fetal DNA concentrations <10%, remained unclassified after 11–24 panels of 765-well digital PCR analysis.

Computer Simulation of Digital RMD. Computer simulation was performed to estimate the accuracy of fetal genotyping by digital RMD and SPRT. Separate simulations were performed for three reference template concentrations, namely at $m_r = 0.5$ (Table S5), $m_r = 0.1$ (Table S6), and $m_r = 0.05$ (Table S7). Simulations were done for fetal DNA concentrations ranging from 5% to 50% and digital PCR runs ranging from 765 wells to 15,300 wells. We performed simulations for the testing of 5,000 heterozygous and 5,000 homozygous fetuses at each combination of m_r, fetal DNA concentration, and digital PCR well number. The percentages of fetuses correctly and incorrectly classified as homozygous or heterozygous and those not classifiable were determined (Tables S4–S7).

According to Table S5, at $m_r = 0.5$ and a fetal DNA concentration of 10% with 1,530 wells of digital PCR analysis, 0.8% of homozygous fetuses would be incorrectly classified, and this may explain the misclassified result of case 2 at 10% concentration in the HbE genomic DNA mixtures (Table S3). According to Table S7, at $m_r = 0.05$ and a fetal DNA concentration of 10% with 7,650 wells of digital PCR analysis, 3.5% of heterozygous fetuses would be incorrectly classified, and this may explain the misclassified result of the maternal plasma case M74 (Table 1). Based on computer simulation, 22.9% and 50.8% of homozygous and heterozygous fetuses, respectively, would remain unclassified even after 7,650 wells of digital PCR analysis at a fetal DNA concentration of 10% and $m_r = 0.05$ (Table S7). Thus, the unclassified cases seen in the digital RMD analyses for maternal plasma (Tables 1 and S4) were comparable to the simulation results.

Digital NASS. The digital RMD findings suggest that when the volume of maternal plasma is limited, which prohibits the continued addition of digital PCR data, a proportion of cases (4 of 10 in Table 1 and Table S4) at the typical fetal DNA concentrations in early pregnancy, i.e., median concentration of ≈10% (7), would remain unclassified (Table 1 and Tables S4–S7). To increase the proportion of classifiable results, one could obtain more maternal plasma for additional digital PCR analysis. Alternatively, one could increase the fractional fetal DNA concentration by physical enrichment of fetal DNA (17, 18). However, here we develop and use a locus-specific method for fetal DNA enrichment, termed digital NASS.

Digital NASS exploits the fact that fetal DNA molecules are shorter than the maternal DNA fragments in maternal plasma (16). Digital NASS involves the performance of duplex digital PCR for the analysis of single DNA molecules. Its principles are schematically illustrated in Fig. 1*B*. Digital PCR is performed using two forward primers and one reverse primer, or vice versa, that are oriented to produce a short amplicon overlapping with a long amplicon. When a single DNA molecule at least as long as that specified by the long amplicon is captured in the reaction well, both the long and short PCR products are generated. When a single DNA molecule shorter than the span of the long amplicon is

Table 2. Combined digital RMD and NASS for determining the fetal *PLAC4* genotype in first trimester maternal plasma

Genotype												All DNA					Short DNA				
Fetal	Maternal	Sample	Wells*	No. negative wells	L	A	G	LA	LG	AG	LAG	A + LA	G + LG	m_r	Fetal %†	SPRT	A	G	m_r	Fetal %‡	SPRT
AA	AG	M4176P	383	299	0	34	12	16	19	0	3	50	31	0.08	18	AA	34	12	0.03	23	AA
GG	AG	M3848P	384	291	0	20	38	10	21	2	2	30	59	0.09	18	GG	20	38	0.06	24	GG
AG	AG	M3960P	2,653	2,243	0	134	131	64	65	8	8	198	196	0.08	11	AG	134	131	0.05	13	U
AA	AG	M3146P	381	325	0	27	15	9	1	3	1	36	16	0.05	15	AA	27	15	0.05	18	U
AG	AG	M2098P	2,296	1,896	0	117	120	78	64	12	9	195	184	0.09	11	U	117	120	0.06	14	AG

L, extension product generated only by the long amplicon; A, extension product of the A allele; G, extension product of the G-allele; U, unclassified.
*The digital PCR and primer extension reactions were performed in 384-well plates. Extension products were identified by mass spectrometry. Occasionally, the software, SpectroTYPER (Sequenom), that acquired the mass spectrometry data would identify and indicate certain wells as producing a "bad spectrum." These wells were excluded from further analysis, and therefore the number of digital PCR was not always multiples of 384 wells.
†Fetal DNA percentage based on the *ZFY/X* counts from all DNA (long and short) molecules.
‡Fetal DNA percentage based on the *ZFY/X* counts from the short DNA molecules.

captured, only the short amplicon is generated. The presence of the long and/or short amplicons can be detected by strategically located hybridization probes or extension primers. We used the latter in our digital NASS assays. An extension primer, L (Fig. 1*B*), is designed to detect the presence of the long amplicon only. Another extension primer is located within the short amplicon and would detect the presence of both the short and long amplicons. This primer extension reaction further discriminates between the alleles tested. In summary, a short molecule is indicated when the L extension product is not detected. As fetal DNA is shorter than its maternal background, it would be more highly represented among the short DNA pool. By focusing on the counting of the short DNA molecules only, a higher effective fetal DNA fraction would be obtained for digital RMD analysis.

Digital NASS Analysis of Sonicated DNA. We obtained a male buffy coat sample and sonicated half of the DNA. The absence of high molecular-weight DNA after sonication was confirmed by running on a Bioanalyzer (Agilent Technologies). The male DNA aliquots with and without sonication were each mixed with a female buffy coat DNA (not sonicated) to produce mixtures containing 15%, 25%, and 50% male DNA. Portions of the sonicated and nonsonicated male DNA were also analyzed individually and unmixed. The samples were analyzed using a digital NASS assay targeting *ZFX* and *ZFY*. Findings are shown in Fig. S1 and Table S8. The digital NASS assay indicated that less than 5% of all DNA in the samples without sonication were short, whereas the short DNA counts increased with the increasing concentration of sonicated male DNA.

Optimal Size Window Assessment. To achieve the optimal discrimination of short fetal DNA molecules among the long maternal DNA molecules in maternal plasma and to achieve the greatest degree of locus-specific fetal DNA enrichment, we investigated various combinations of PCR amplicon sizes for constructing the digital NASS assays. Six digital NASS assays using a combination of amplicon sizes ranging from 51 bp to 213 bp (Table S9) were designed for *ZFX* and *ZFY* and aimed at the testing of maternal plasma from pregnancies with male fetuses.

Three third-trimester maternal plasma samples were analyzed. The male fetal DNA fractional concentrations were calculated twice, with and without consideration of the amplicon sizes. We first calculated the concentrations using all wells containing only *ZFX* or only *ZFY* signals regardless of whether the signal was obtained from a long or short molecule. We then recalculated the concentrations using the wells showing signals of the short *ZFX* and *ZFY* amplicons only. As shown in Fig. S2*A*, the calculated fractional fetal DNA concentrations were higher by using only the short DNA molecules, compared with those calculated using both the long and short molecules. The increments in the fractional fetal DNA concentrations achieved were further calculated as percentage enrichment. According to the averaged results of the tested plasma samples (Fig. S2*A*), assays 179_64 and 213_82 showed the greatest enrichments, and assay 213_51 showed the highest fetal DNA percentage. These three assays were selected for further study in four first-trimester maternal plasma samples.

The fractional fetal DNA concentrations and the percentage enrichment in first-trimester maternal plasma samples are shown in Fig. S2*B*. The combination of 179 bp and 64 bp amplicons showed the greatest power to discriminate between maternal and fetal molecules in maternal plasma and thus resulted in the highest degree of fetal DNA enrichment.

Digital NASS Measurement of Fetal DNA Concentrations in Maternal Plasma. The *ZFY/X* 179_64 digital NASS assay was applied to 10 first-trimester maternal plasma samples (median GA: 13 weeks; range: 12–13 weeks) of women pregnant with male fetuses (Table S10). The median fetal DNA concentrations of counting all molecules and that of the short molecules were 18.1% and 22.9%, respectively. The median enrichment in fetal DNA attained was 38%.

We also designed a digital NASS assay to target a polymorphic SNP (rs8130833) on *PLAC4*. Such an assay would allow the measurement of circulating fetal DNA for both male and female fetuses when the mother is homozygous and the fetus is heterozygous. Fetal DNA concentration could be determined by counting signals from the paternal allele. We designed a duplex PCR assay with amplicon sizes of 179 bp and 63 bp. For 10 first-trimester pregnancies (median GA: 12 weeks; range: 12–13 weeks), the median fetal DNA concentrations counting all molecules and only short molecules were 9.6% and 12.5%, respectively (Table S11). The median enrichment in fetal DNA attained was 32.6%.

Combination of Digital RMD and NASS. Any increment in fetal DNA concentration should improve the percentage of classifiable cases for a given number of digital PCR analysis (Tables S5–S7). Using digital NASS assays for digital RMD would allow two chances of genotype classification from the same dataset. For example, genotype classification could be based on the mutant and wild-type allele counts of all DNA molecules. In addition, genotype classification could also be based on the allele counts from just the short DNA molecules. Because of the effective enrichment in fetal DNA achievable by interpreting the short DNA molecules only, the combination of NASS with RMD should allow more cases to be classifiable for a given total number of DNA molecules analyzed.

We analyzed maternal plasma samples from five pregnancies (median GA: 13; range: 12–17 weeks) where the mother was heterozygous for the *PLAC4* SNP and the fetus was a male. The

Table 3. *ZFY/X* digital NASS was performed on the maternal plasma samples to determine the fetal DNA concentration to select the relevant SPRT allelic ratio cutoff points for genotype classification

Sample	Wells	No. negative wells	L	X	Y	LX	LY	XY	LXY*	All DNA† X copies	All DNA† Y copies	All DNA† Fetal %	Short DNA‡ X copies	Short DNA‡ Y copies	Short DNA‡ Fetal %	Enrichment, % (Short–all)/all
M4176P	376	285	0	57	7	25	1	1	0	94	9	18	63	8	23	28
M3848P	384	295	0	54	6	26	1	2	0	92	9	18	61	8	24	31
M3960P	2,681	2,196	0	361	21	96	2	3	2	507	28	11	391	26	13	19
M3146P	383	310	0	57	5	10	0	0	1	75	6	15	62	6	18	19
M2098P	2,271	1,768	0	339	22	134	2	4	2	538	30	11	372	28	14	33

L, extension product generated only by the long amplicon; X, extension product of *ZFX*; Y, extension product of *ZFY*.

*As these are maternal plasma samples, wells positive for LXY are assumed to contain one copy of LX (i.e., long X DNA molecule) and one copy of Y (i.e., short Y DNA molecule).

†All DNA Y copies = $\{-\ln(N - Y - LY - XY - LXY)/N]\}N$

All DNA X copies = $\{-\ln(N - X - LX - XY - LXY)/N]\}N$

‡Short DNA Y copies = $\{-\ln(N - Y - XY - LXY)/N]\}N$

Short DNA X copies = $\{-\ln(N - X - XY)/N]\}N$

where N = the total number of digital PCR wells analyzed and ln is the natural logarithm.

plasma samples were analyzed with the *PLAC4* (Table 2) and *ZFY/X* (Table 3) digital NASS assays. The latter was used to determine the fractional fetal DNA concentration in the sample. *PLAC4* SNP genotype classification by SPRT analysis using informative counts from the long and short DNA molecules was based on the fetal DNA concentration derived from all *ZFY* and *ZFX* counts. Similarly, we also determined the *PLAC4* SNP genotype in each case using the informative short DNA counts only. The SPRT analysis was based on the fetal DNA concentration derived from the *ZFY* and *ZFX* short DNA counts only. The correct fetal genotype was determined for four of the five cases when the SPRT analysis was based on all DNA molecules (Table 2). For case M2098P, the fetal genotype was unclassified when all informative DNA molecules were considered, but was correctly classified when the short DNA counts were used for SPRT classification. Hence, by using a digital NASS assay, RMD classification was made possible for one case, out of the currently tested five cases, that would otherwise be unclassified.

Computer Simulation for Combined Digital RMD and NASS. We performed a computer simulation to determine the genotyping efficiency of combining digital NASS with digital RMD. We modeled the simulation against the data of the five maternal plasma samples shown in Tables 2 and 3. The parameters used included the median fetal DNA concentration without specifying for the short DNA molecules (15%), the median average reference template concentration per well of 0.08 copy/ml, the long fetal DNA fraction of 7%, and the long maternal DNA fraction of 30%. These figures translated to percentage enrichment in fetal DNA of 27% (*SI Text*). Using these parameters, we simulated the testing of 5,000 each of heterozygous and homozygous fetuses. As shown in Table S12, at any given number of digital PCR analyses, a proportion of cases that would otherwise be unclassified using information from all DNA molecules would become correctly classifiable when focusing on the short DNA molecules. For example, when classifying a heterozygous fetus with 3,456 wells of digital PCR, NASS adds 10% of correctly genotyped cases to the 74% correctly classified by RMD alone. When classifying homozygous fetuses with 2,304 digital PCRs, NASS adds 10% of cases correctly genotyped to the 75% of cases correctly classified by RMD alone. In short, NASS allows more cases to be classifiable by RMD when a limited amount of maternal plasma prevents the continual addition and accumulation of digital PCR data.

Discussion

Digital RMD may bring noninvasive prenatal diagnosis of monogenic diseases closer to clinical applicability. Previously, this area of research has been limited by the inability to assess the fetal inheritance of maternally transmitted disease-causing mutations (13). Researchers have exploited epigenetic differences, such as in DNA methylation, between maternal and fetal tissues to develop markers that allow the discrimination of fetal from maternal sequences despite their sharing the same primary genetic sequence (25–27). However, such differentially methylated regions may not coincide with loci harboring disease-causing mutations. In conjunction with previous approaches of the detection of paternally inherited mutations in maternal plasma, the development of digital RMD may allow noninvasive prenatal diagnosis of monogenic diseases achievable for all pregnancies.

Digital RMD is based on the quantitative discrimination of small imbalances in concentrations between the mutant and wild-type alleles in maternal plasma. It is therefore in principle similar to the digital RCD approach for the noninvasive detection of fetal trisomy 21 (5). The main distinction between RCD and RMD lies in the nature of the allelic imbalance that needs to be detected. For Down syndrome detection, RCD aims to assess the change from two copies of chromosome 21 to three copies. RMD aims to assess the change from one copy of the mutant/wild-type allele in the heterozygote to two copies of the allele in a homozygote. Theoretically, RMD interrogates a larger degree of allelic imbalance than for RCD and should be more easily achievable.

To improve the precision and reliability for detecting the imbalance between mutant and wild-type alleles when present, allelic ratio cutoff points for segregating and classifying fetal genotypes are specifically selected for each case based on the experimentally observed digital PCR average template concentration per well and the fractional fetal DNA concentration. As the absolute amount of DNA molecules present in maternal plasma tends to be limiting, we continued to use SPRT statistics for data interpretation, as we did for digital RCD previously (5). SPRT is a Bayesian type of statistics that allows the testing of hypotheses as data accumulate and could thus save unnecessary testing when evidence is sufficient for genotype classification (23). We have previously shown that greater diagnostic efficiency and efficacy is achieved by using SPRT curves and cutoff values specific for the data of each case, i.e., case-specific SPRT (5), as opposed to fixed cutoff values (23). For this approach to be applicable to female fetuses also, we have shown that a SNP assay on *PLAC4* (Table S11) and certainly other SNPs, could be used for fetal DNA quantification in informative fetal-maternal pairs (mother homozygous and fetus heterozygous). Future studies could explore the diagnostic performance of non-SPRT statistics or effects of increasing the stringency of the SPRT statistics.

Though digital RMD widens the spectrum of pregnancies at risk for monogenic diseases assessable by noninvasive means, genotype

classification is more difficult at low fractional fetal DNA concentrations (7). We developed the digital NASS approach by exploiting the shorter size of fetal DNA compared with maternal DNA in maternal plasma. NASS differs from nested PCR because the amplicons are generated at the same time in NASS but sequentially in nested PCR. The NASS system works only in a digital PCR format because when more than one DNA molecule is captured in a reaction well the presence of both the long and short amplicons may indicate the presence of both long and short DNA fragments, instead of just one long DNA molecule. Thus, NASS represents a duplex/multiplex PCR assay conducted on a single DNA molecule. The advantage of NASS is that it allows locus-specific enrichment of fetal DNA without the need for extra sampling of maternal plasma or additional experimental time. Furthermore, it is complementary to other means that physically enrich fetal DNA. For example, digital NASS assays could be used to analyze size fractionated maternal plasma samples to provide further fetal DNA enrichment.

In this study, we used a microfluidics platform (7) for the performance of digital RMD. Much operational ease is offered by such systems when compared with the manual performance of digital PCR (5). Other platforms are also available (28). However, for digital NASS, because of the need to discriminate multiple amplicons in each digital PCR and the need for precise allelic discrimination, the NASS assays are based on performing conventional digital PCR followed by multiplexed primer extension reactions and product identification by mass spectrometry. The specificity of the NASS assay configuration and analytical system is supported by the lack of occurrence of the long amplicon, L, alone. Besides digital PCR, other methods of single molecule analysis, such as sequencing (29), visualization, or physical detection, could be adapted for digital RMD and digital size selection.

Though digital RMD and digital NASS each solve different but complementary bottlenecks in plasma nucleic acids research, it is of significance that they could be combined to work synergistically (Tables 2 and 3), bringing noninvasive prenatal diagnosis of monogenic diseases closer to practical feasibility. With the gradual dismantling of the obstacles to noninvasive prenatal diagnosis, clinical trials could be set up to investigate the clinical performance of these new tests.

Materials and Methods

Details are in *SI Text* and Tables S13–S15.

Digital Relative Mutation Dosage (RMD). All experiments were carried out on the BioMark™ System (Fluidigm) using the 12.765 Digital Arrays (Fluidigm), which consists of 12 panels, each of 765 wells. A total of six panels were used for fetal DNA quantification using the *ZFY/X* assay (7). For digital RMD, at least one reaction panel was used for each case, and data were aggregated from extra panels for samples that remained unclassified until a genotype classified or when we exhausted the sample (Tables 1 and 2).

Digital Nucleic Acid Size Selection (NASS). Except specifically stated in the experiment conducted to select the optimal combination of amplicon sizes for discriminating fetal from maternal DNA, the *ZFY/X* digital NASS assay used in all other NASS experiments was the one that generated amplicons of 179 bp and 64 bp. The digital NASS assay targeting the polymorphic SNP (rs8130833) on *PLAC4* generates amplicons of 179 bp and 63 bp. All digital NASS assays were performed using MassARRAY hME reagents (Sequenom) with essentially the same protocol. A total of 384 wells were analyzed in each digital PCR run. One of a number of outcomes could be scored for each digital PCR after digital NASS analysis, and details of the scoring are described in Fig. 1*B* and *SI Text*.

SPRT Analysis. We constructed SPRT curves on a case-specific basis by calculating the expected degree of allelic imbalance based on the experimentally derived average template concentration per well and fetal DNA proportion for genotype classification of each assessed case.

Computer Simulation of Classification Accuracy by SPRT. The computer simulations were performed with the Microsoft Excel 2003 software (Microsoft Corp.) and SAS 9.1 for Windows software (SAS Institute, Inc.).

ACKNOWLEDGMENTS. This study was supported by the University Grants Committee of the Government of the Hong Kong Special Administration Region, China, under the Areas of Excellence Scheme (AoE/M-04/06). Y.M.D.L. was supported by the Chair Professorship scheme of the Li Ka Shing Foundation.

1. Lo YMD, *et al.* (1997) Presence of fetal DNA in maternal plasma and serum. *Lancet* 350:485–487.
2. Lo YMD, Chiu RWK (2007) Prenatal diagnosis: Progress through plasma nucleic acids. *Nat Rev Genet* 8:71–77.
3. Finning KM, Martin PG, Soothill PW, Avent ND (2002) Prediction of fetal D status from maternal plasma: Introduction of a new noninvasive fetal RHD genotyping service. *Transfusion* 42:1079–1085.
4. Lo YMD, *et al.* (2007) Plasma placental RNA allelic ratio permits noninvasive prenatal chromosomal aneuploidy detection. *Nat Med* 13:218–223.
5. Lo YMD, *et al.* (2007) Digital PCR for the molecular detection of fetal chromosomal aneuploidy. *Proc Natl Acad Sci USA* 104:13116–13121.
6. Weatherall DJ (2004) Thalassaemia: The long road from bedside to genome. *Nat Rev Genet* 5:625–631.
7. Lun FMF, *et al.* (2008) Microfluidics digital PCR reveals a higher than expected fraction of fetal DNA in maternal plasma. *Clin Chem* 54:1664–1672.
8. Chiu RWK, Lo YMD (2004) The biology and diagnostic applications of fetal DNA and RNA in maternal plasma. *Curr Top* Dev Biol 61:81–111.
9. Amicucci P, Gennarelli M, Novelli G, Dallapiccola B (2000) Prenatal diagnosis of myotonic dystrophy using fetal DNA obtained from maternal plasma. *Clin Chem* 46:301–302.
10. Saito H, *et al.* (2000) Prenatal DNA diagnosis of a single-gene disorder from maternal plasma. *Lancet* 356:1170.
11. Gonzalez-Gonzalez MC, *et al.* (2003) Huntington disease-unaffected fetus diagnosed from maternal plasma using QF-PCR. *Prenat Diagn* 23:232–234.
12. Chiu RWK, *et al.* (2002) Prenatal exclusion of beta thalassaemia major by examination of maternal plasma. *Lancet* 360:998–1000.
13. Ding C, *et al.* (2004) MS analysis of single-nucleotide differences in circulating nucleic acids: Application to noninvasive prenatal diagnosis. *Proc Natl Acad Sci USA* 101:10762–10767.
14. Lo YMD, *et al.* (1998) Quantitative analysis of fetal DNA in maternal plasma and serum: Implications for noninvasive prenatal diagnosis. *Am J Hum Genet* 62:768–775.
15. Chan KCA, *et al.* (2006) Hypermethylated RASSF1A in maternal plasma: A universal fetal DNA marker that improves the reliability of noninvasive prenatal diagnosis. *Clin Chem* 52:2211–2218.
16. Chan KCA, *et al.* (2004) Size distributions of maternal and fetal DNA in maternal plasma. *Clin Chem* 50:88–92.
17. Li Y, *et al.* (2005) Detection of paternally inherited fetal point mutations for beta-thalassemia using size-fractionated cell-free DNA in maternal plasma. *JAMA* 293:843–849.
18. Dhallan R, *et al.* (2004) Methods to increase the percentage of free fetal DNA recovered from the maternal circulation. *JAMA* 291:1114–1119.
19. Chung GTY, *et al.* (2005) Lack of dramatic enrichment of fetal DNA in maternal plasma by formaldehyde treatment. *Clin Chem* 51:655–658.
20. Chinnapapagari SK, *et al.* (2005) Treatment of maternal blood samples with formaldehyde does not alter the proportion of circulatory fetal nucleic acids (DNA and mRNA) in maternal plasma. *Clin Chem* 51:652–655.
21. Benachi A, *et al.* (2005) Impact of formaldehyde on the in vitro proportion of fetal DNA in maternal plasma and serum. *Clin Chem* 51:242–244.
22. Vogelstein B, Kinzler KW (1999) Digital PCR. *Proc Natl Acad Sci USA* 96:9236–9241.
23. Zhou W, *et al.* (2002) Counting alleles to predict recurrence of early-stage colorectal cancers. *Lancet* 359:219–225.
24. Tsang JCH, *et al.* (2007) Mass spectrometry-based detection of hemoglobin E mutation by allele-specific base extension reaction. *Clin Chem* 53:2205–2209.
25. Chim SSC, *et al.* (2005) Detection of the placental epigenetic signature of the maspin gene in maternal plasma. *Proc Natl Acad Sci USA* 102:14753–14758.
26. Chiu RWK, *et al.* (2007) Hypermethylation of RASSF1A in human and rhesus placentas. *Am J Pathol* 170:941–950.
27. Chim SSC, *et al.* (2008) Systematic search for placental DNA-methylation markers on chromosome 21: Toward a maternal plasma-based epigenetic test for fetal trisomy 21. *Clin Chem* 54:500–511.
28. Morrison T, *et al.* (2006) Nanoliter high throughput quantitative PCR. *Nucleic Acids Res* 34:e123.
29. Harris TD, *et al.* (2008) Single-molecule DNA sequencing of a viral genome. *Science* 320:106–109.