Digital PCR for the molecular detection of fetal chromosomal aneuploidy

Y. M. Dennis Lo*†‡, Fiona M. F. Lun†, K. C. Allen Chan†, Nancy B. Y. Tsui†, Ka C. Chong§, Tze K. Lau¶, Tak Y. Leung¶, Benny C. Y. Zee§, Charles R. Cantor‡‖, and Rossa W. K. Chiu*†‡

*Li Ka Shing Institute of Health Sciences, Departments of †Chemical Pathology and ¶Obstetrics and Gynaecology and §School of Public Health, The Chinese University of Hong Kong, Sha Tin, New Territories, Hong Kong Special Administrative Region, People's Republic of China; and ‖Bioinformatics Program and Center for Advanced Biotechnology, Boston University, Boston, MA 02118

Contributed by Charles R. Cantor, June 21, 2007 (sent for review May 9, 2007)

Trisomy 21 is the most common reason that women opt for prenatal diagnosis. Conventional prenatal diagnostic methods involve the sampling of fetal materials by invasive procedures such as amniocentesis. Screening by ultrasonography and biochemical markers have been used to risk-stratify pregnant women before definitive invasive diagnostic procedures. However, these screening methods generally target epiphenomena, such as nuchal translucency, associated with trisomy 21. It would be ideal if noninvasive genetic methods were available for the direct detection of the core pathology of trisomy 21. Here we outline an approach using digital PCR for the noninvasive detection of fetal trisomy 21 by analysis of fetal nucleic acids in maternal plasma. First, we demonstrate the use of digital PCR to determine the allelic imbalance of a SNP on *PLAC4* mRNA, a placenta-expressed transcript on chromosome 21, in the maternal plasma of women bearing trisomy 21 fetuses. We named this the digital RNA SNP strategy. Second, we developed a nonpolymorphism-based method for the noninvasive prenatal detection of trisomy 21. We named this the digital relative chromosome dosage (RCD) method. Digital RCD involves the direct assessment of whether the total copy number of chromosome 21 in a sample containing fetal DNA is overrepresented with respect to a reference chromosome. Even without elaborate instrumentation, digital RCD allows the detection of trisomy 21 in samples containing 25% fetal DNA. We applied the sequential probability ratio test to interpret the digital PCR data. Computer simulation and empirical validation confirmed the high accuracy of the disease classification algorithm.

circulating fetal nucleic acids | noninvasive prenatal diagnosis | sequential probability ratio test | trisomy 21 | RNA SNP

The detection of fetal trisomy 21 (T21) is an important indication for prenatal diagnosis. The sampling of fetal materials by amniocentesis and chorionic villus sampling are invasive, with a finite risk of fetal loss (1). A variety of screening methods, such as ultrasound, have been investigated (2). However, these screening methods typically target T21-related epiphenomena instead of the core chromosomal abnormality and thus have suboptimal diagnostic accuracy and disadvantages, such as being highly influenced by gestational age.

The discovery of cell-free fetal DNA in maternal plasma in 1997 offered new possibilities for noninvasive prenatal diagnosis (3, 4). This method has been readily applied to sex-linked (5) and certain single-gene (6, 7) disorders, but its use for fetal chromosomal aneuploidies has been a challenge (4). First, fetal nucleic acids coexist in maternal plasma with a high background of maternal nucleic acids that can often interfere with analysis (8). Second, fetal nucleic acids circulate in maternal plasma in a cell-free form, making it difficult to derive chromosome dosage information. Significant developments have recently been made (9–11). One approach focuses on the detection of nucleic acid species that are fetal-specific, including DNA fragments with a placenta-specific DNA methylation pattern (10, 12) and RNA molecules expressed by the placenta (9). Because circulating fetal nucleic acids are mainly derived from the placenta, the problem of maternal background interference can be overcome by targeting such molecules in maternal plasma (4). Dosage of chromosome 21 (chr21) is then inferred from the ratios of polymorphic alleles in the placenta-derived DNA/RNA molecules. However, the dependence on genetic polymorphisms limits the use of these approaches to heterozygous fetuses.

It would be ideal if a noninvasive test for fetal T21 detection based on circulating fetal nucleic acid analysis were not dependent on the use of genetic polymorphisms. Theoretically, even with the small fractional concentration of fetal DNA (8), a T21 fetus would contribute an additional dose of chr21 sequences per genome equivalent (GE) of fetal DNA released into maternal plasma. For example, a maternal plasma sample from a euploid pregnancy containing 50 GEs per milliliter of total DNA with 5 GEs per milliliter of DNA contributed by the fetus (i.e., 10% fetal DNA) should contain a total of 100 copies (90 maternal copies plus 10 fetal copies) of chr21 sequences per milliliter of maternal plasma. For a T21 pregnancy, each fetal GE would contribute three copies of chr21, resulting in a total of 105 copies (90 maternal copies plus 15 fetal copies) of chr21 sequences per milliliter of maternal plasma. At 10% fetal DNA concentration, the amount of chr21-derived sequences in the maternal plasma of a T21 pregnancy would therefore be 1.05 times that of a euploid case. If an analytical approach could be developed to determine this small degree of quantitative difference, a polymorphism-independent test for noninvasive prenatal diagnosis of fetal T21 would be achieved.

Gene dosage assessment requiring 2-fold discrimination power can readily be attained with quantitative PCR (13). Through DNA quantification of a chr21 locus and a reference locus in amniocyte cultures, Zimmermann *et al.* (14) were able to detect the 1.5-fold increase in chr21 DNA sequences in T21 fetuses. Because a 2-fold difference in DNA template concentration constitutes a difference of only one threshold cycle (Ct), the discrimination of a 1.5-fold difference has been the limit of conventional real-time PCR. To achieve finer degrees of quantitative discrimination, alternative strategies are needed. Here, we explore the use of digital PCR (15) for this purpose.

Author contributions: Y.M.D.L., K.C.A.C., and R.W.K.C. designed research; F.M.F.L. and N.B.Y.T. performed research; T.K.L. and T.Y.L. collected clinical samples; K.C.C., B.C.Y.Z., and C.R.C. contributed new reagents/analytic tools; Y.M.D.L., F.M.F.L., K.C.A.C., N.B.Y.T., K.C.C., and R.W.K.C. analyzed data; and Y.M.D.L. and R.W.K.C. wrote the paper.

Conflict of interest statement: Y.M.D.L., F.M.F.L., K.C.A.C., N.B.Y.T., K.C.C., B.C.Y.Z., C.R.C., and R.W.K.C. have filed patent applications on aspects of noninvasive prenatal diagnostics. Y.M.D.L. has equity in Plasmagene Biosciences Limited. C.R.C. has equity in Sequenom, Inc., and is the Chief Scientific Officer of Sequenom, Inc.

Freely available online through the PNAS open access option.

Abbreviations: T21, trisomy 21; chr, chromosome; RCD, relative chromosome dosage; SPRT, sequential probability ratio test.

‡To whom correspondence may be addressed. E-mail: loym@cuhk.edu.hk, ccantor@sequenom.com, or rossachiu@cuhk.edu.hk.

This article contains supporting information online at www.pnas.org/cgi/content/full/0705765104/DC1.

Digital PCR involves multiple PCR analyses on extremely dilute nucleic acids such that most positive amplifications reflect the signal from a single template molecule (15), permitting the counting of individual template molecules. The proportion of positive amplifications among the total number of PCRs analyzed allows an estimation of the template concentration in the original nondiluted sample. This technique has been proposed to allow the detection of a variety of genetic phenomena (15), including the detection of loss of heterozygosity (LOH) in tumor samples (16) and plasma of cancer patients (17). Because template molecule quantification by digital PCR does not rely on dose–response relationships between reporter dyes and nucleic acid concentrations, its analytical precision should theoretically be superior to that of real-time PCR. To test whether this approach is precise enough to detect fetal chromosomal aneuploidies in maternal plasma, we first assessed whether digital PCR could measure the allelic ratio of *PLAC4* mRNA in maternal plasma (9), thereby distinguishing T21 from euploid fetuses. This is referred to as the digital RNA SNP method. We then evaluated whether the increased precision of digital PCR would allow the detection of fetal chromosomal aneuploidies without depending on genetic polymorphisms. We call this digital relative chromosome dosage (RCD) analysis.

Results

Principles of Digital PCR. The first step in digital PCR is the dilution of the extracted nucleic acids to a concentration such that, on average, one template molecule is present per reaction well. PCR is then set up so that a multitude of such single-molecule PCRs is analyzed per sample. We used 96-well and 384-well reaction plates and distributed each diluted nucleic acid sample to the reaction wells of one or more plates. Under these conditions, the actual number of template molecules distributed to each reaction well followed the Poisson distribution. Thus, an individual reaction well could contain zero, one, or more template molecules. The expected proportion of wells with no template is given by e^{-m}, where m is the average concentration of template molecules per well. For example, at an average concentration of one template molecule per well, the expected proportion of wells with no template molecule is given by e^{-1}, i.e., 0.37 (37%). The remaining 63% of wells will contain one or more template molecules. Typically, the number of positive and informative wells in a digital PCR run would then be counted. The definition of informative wells and the manner by which the digital PCR data are interpreted depend on the application (15) and are described below.

Principles of Digital RNA SNP. Digital RNA SNP is a digital version of our previously reported approach (9) for T21 detection by determining an imbalance in the ratio of polymorphic alleles of an A/G SNP, rs8130833, located on *PLAC4*. For a heterozygous euploid fetus, the A and G alleles should be equally represented in the fetal genome (1:1), whereas, in T21, an additional copy of one of the SNP alleles would give a 2:1 ratio. Digital RNA SNP analysis aims to determine whether the amounts of the two *PLAC4* alleles in the sample are equal or otherwise. Thus, both the A and G *PLAC4* alleles are the target templates. The analytical steps are schematically shown in Fig. 1.

After digital real-time PCR analysis of the *PLAC4* SNP alleles in 384-well plates, the number of informative wells was counted. An informative well is defined as one that was only positive for the A or G allele but not both (Fig. 1). For a euploid case, we expect an equal number of A-positive and G-positive wells (Fig. 1). However, when template molecules from a T21 fetus are analyzed, the number of wells containing just one allele should be higher than the number containing just the other allele (Fig. 1). In short, allelic imbalance is expected for T21. The same degree of imbalance would be expected if this approach were applied to the analysis of placental DNA, placental RNA, and maternal plasma RNA [*PLAC4* mRNA in maternal plasma being completely fetal in origin (9)].

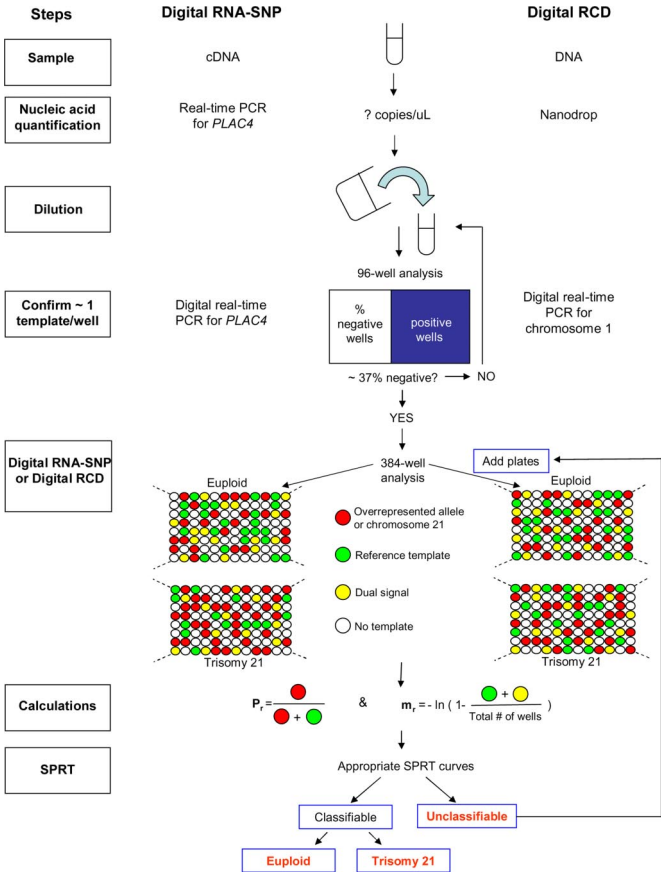

Fig. 1. Illustration of the analytical steps in digital RNA SNP and digital RCD analyses for T21 detection. Only a representative 96-well subset of the 384-well data is shown for one euploid and one T21 case for each of digital RNA SNP and digital RCD analyses, respectively. The T21 data depicted in the digital RNA SNP experiment represent a case where the G allele is overrepresented, i.e., a fetal genotype of AGG.

The allele with the higher number of counts is referred to as the overrepresented allele, and its proportion among all informative wells, P_r, was calculated (Fig. 1). The sequential probability ratio test (SPRT) (16, 18) (see below) was applied to determine whether the P_r indicated the degree of allelic imbalance that would be expected for a T21 sample. Alternatively, the SPRT analysis may indicate that the available data are not yet adequate for disease classification. When classification was not achieved, additional 384-well plates were analyzed until the aggregated data became classifiable by SPRT.

Principles of Digital RCD. We determined chromosome dosage by digital PCR analysis of a nonpolymorphic chr21 locus relative to one located on a reference chromosome, chr1. We aimed to differentiate a change in the ratio of chr21 to chr1 from 2:2 in the genome of a euploid fetus to 3:2 in a T21 fetus (Fig. 1). Here, an informative well is defined as one that is positive for either the chr21 or chr1 locus but not both. For a euploid fetus, the number of informative wells positive for either locus should be approximately equal (Fig. 1). For a T21 fetus, there should be an overrepresentation of wells positive for chr21 (Fig. 1). The degree of overrepresentation would depend on the fractional fetal DNA concentration in the sample. For example, when placental DNA is analyzed, the theoretical RCD ratio in the fetal genome should be 3:2, i.e., a 1.5-fold difference. However, as described earlier, the theoretical

RCD ratio would decrease to 1.05 when analyzing a maternal plasma sample containing 10% fetal DNA. The P_r was calculated by dividing the number of wells positive only for the chr21 locus by the total number of informative wells (Fig. 1). The P_r was subjected to SPRT analysis (16, 18) for disease classification. If the data were unclassifiable, one or more additional 384-well plates were analyzed.

Assessment of Allelic or Chromosomal Imbalance by Digital PCR. To determine whether the analyzed sample is from a T21 case, the observed RNA SNP or RCD ratio would be compared with that expected for a T21 case. The theoretical RNA SNP ratio is 2:1, and the RCD ratio is 3:2 for a pure T21 sample. However, due to the Poisson distribution, the exact ratios are not the same as those in the fetal genome. Furthermore, template concentration is a key variable in the Poisson equation. Thus, the exact ratios are dependent on the template concentration used in a particular experiment. Because the total number of template molecules for a given volume of sample from a T21 subject would be greater than that for a euploid case, we standardize our definition of the level of diluted template concentration as the average number of reference template molecules per reaction well, m_r. For digital RNA SNP analysis, the reference template would be the allele that was not overrepresented, whereas the reference template for digital RCD analysis would be the chr1 locus. Thus, the dilution of one target template molecule of any type per well for the digital PCR analysis of a euploid case equates to an m_r of 0.5.

The basis for the difference between the theoretical and expected degree of allelic or chromosomal imbalance and the calculations to determine the latter for a range of m_r values are shown in supporting information (SI) Tables 3 and 4. In digital RNA SNP analysis of a T21 sample, when the m_r value was 0.5, the digital RNA SNP ratio (namely, the ratio of wells containing just the overrepresented allele with respect to wells containing just the reference allele) was 2.65 (SI Table 3). In digital RCD analysis of a specimen composed of 100% fetal DNA, when the m_r value was 0.5, the digital RCD ratio (namely, the ratio of wells positive solely for the chr21 locus with respect to those positive solely for the chr1 locus) was 1.7 (SI Table 4). As the fractional fetal DNA concentration decreases, the digital RCD ratio decreases for the same m_r (SI Table 4). As shown in SI Tables 3 and 4, the extent of allelic or chromosomal overrepresentation increases with m_r. However, the percentage of informative wells approaches its maximum near an m_r value of 0.5 and decreases gradually with further increase in m_r. In practice, the decline in the proportion of informative wells could be compensated by increasing the total number of wells analyzed if the amount of specimen template molecules is not limiting, with an associated increase in reagent costs. Hence, optimal digital PCR performance is a tradeoff between the template concentration and total number of wells tested per sample.

SPRT Analysis. To determine whether an observed degree of overrepresentation of a *PLAC4* allele in digital RNA SNP, or the chr21 locus in digital RCD, is statistically significant, a SPRT-based approach was used (16, 18). SPRT is a method that allows testing of a hypothesis as data accumulate. SPRT has been used to interpret digital PCR data for loss of heterozygosity (LOH) in tumor samples (16, 18). In T21 detection, the null hypothesis is that there is no allelic or chromosomal imbalance (i.e., T21 is not detected). The alternative hypothesis is that allelic or chromosomal imbalance exists (i.e., T21 is detected). Operationally, SPRT can be performed with a pair of SPRT curves that are constructed to define the probabilistic boundaries for accepting or rejecting the null hypothesis (Fig. 2*A* and *SI Materials and Methods*). These curves show the required proportion of informative wells positive for the overrepresented allele or chr21, P_r (y axis, Fig. 2*A*), for a given total number of informative wells (x axis, Fig. 2*A*) needed for classification. Samples with data points that are above the top curve are classified as trisomic (Fig. 2*A*). Samples with data points that are below the bottom curve are classified as euploid. Samples with data points in between the two curves are unclassifiable and would require an increased total number of informative counts before classification. SPRT thus offers the advantage that a smaller amount of testing is required for a given level of confidence than other statistical methods. This feature is of particular relevance to the analysis of plasma nucleic acids in which the number of available template molecules is limited.

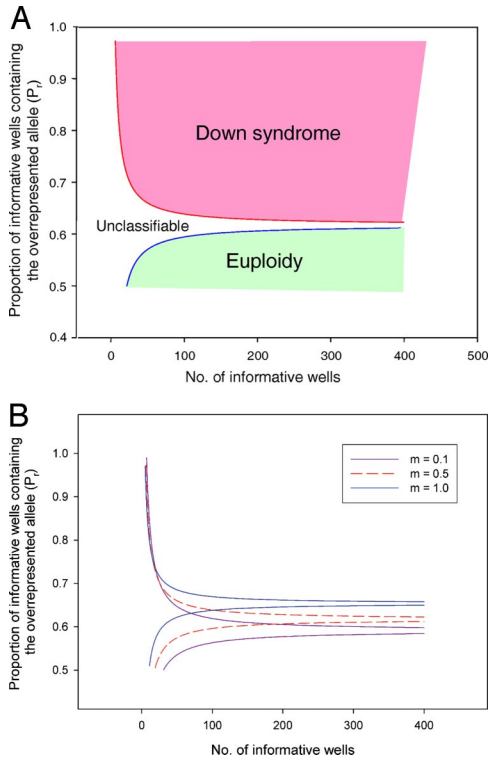

Fig. 2. SPRT analysis. (*A*) A pair of SPRT curves delimits the decision boundaries for accepting or rejecting the hypotheses that the sample belonged to a euploid or aneuploid fetus. (*B*) The decision boundaries of the SPRT curves would vary according to the template concentration. Curves applicable to digital RNA SNP analysis are shown.

As discussed above, the exact degree of allelic or chromosomal imbalance depends on the actual template concentration per experiment. We therefore constructed a series of SPRT curves for a range of m_r values (*SI Materials and Methods*). Each set of digital PCR data should be interpreted with the curves relevant to the m_r of that particular run. Thus, in practice, after digital RNA SNP or digital RCD analysis, m_r and P_r are calculated (Fig. 1). m_r is calculated by using the Poisson equation and the proportion of wells negative for the reference template (*SI Materials and Methods*). P_r is the proportion of informative wells positive just for the overrepresented template. The experimentally derived P_r is interpreted with the relevant SPRT curves selected by the corresponding m_r. This is in contrast to the previously reported use of SPRT for molecular detection of loss of heterozygosity (LOH) by digital PCR, where a fixed set of curves was used (16). Because the expected degrees of allelic or chromosomal imbalance for the digital RNA SNP and RCD approaches are different (2:1 for the former and 3:2 for the latter), different series of SPRT curves are needed. Fig. 2*B* illustrates the degree of differences in the SPRT curves for m_r values of 0.1, 0.5, and 1.0 for digital RNA SNP analysis. Compared with the use of a fixed set of SPRT curves in previous studies (*SI Materials and Methods*) (16, 18), the proportion of

Table 1. Digital RNA SNP analysis in placental tissues of euploid and T21 pregnancies

Sample	Genotype	No. of wells positive for individual alleles				m_r	SPRT result		
		A only	G only	AG	All negative		P_r	Unclassifiable region	Classification
Placental DNA									
N677	AG	85	83	126	90	0.79	0.51	0.63–0.65	Euploid
N710	AG	102	83	73	126	0.52	0.55	0.61–0.63	Euploid
N435	AGG	49	157	130	48	0.63	0.76	0.62–0.64	T21
N981	AAG	135	69	82	98	0.50	0.66	0.61–0.63	T21
Placental RNA									
V533	AG	103	93	71	117	0.56	0.53	0.61–0.63	Euploid
V943	AG	89	100	74	121	0.55	0.53	0.61–0.63	Euploid
N435	AGG	52	138	95	99	0.48	0.73	0.61–0.63	T21
T215	AAG	146	58	138	42	0.71	0.72	0.62–0.64	T21

The no. of wells for all samples was 384. Genotypes were determined by mass spectrometric assay. The m_r value indicates the average no. of reference molecules per reaction well. The P_r values were calculated by using the following equation: no. of wells positive for the overrepresented allele/(no. of wells positive for A only + no. of wells positive for G only). The unclassifiable region for the corresponding m_r is shown. "Euploid" was assigned when the P_r was below the unclassifiable region; "T21" was assigned when the P_r was above the unclassifiable region.

unclassifiable data is much lower with our approach (SI Tables 5 and 6). For example, when using our approach, at an m_r value of 0.5, 14% and 0% of T21 samples would be unclassifiable for 96-well and 384-well digital RNA SNP analyses, respectively, but 62% and 10%, respectively, would be unclassifiable when using fixed curves (SI Tables 5 and 6).

Computer Simulation of Classification Accuracies of Digital PCR Detection of T21. Computer simulation was performed to estimate the accuracy of diagnosing T21 by using the SPRT approach. Separate simulations were performed for different values of three parameters, namely, reference template concentration (m_r), number of informative counts, and projected degree of allelic or chromosomal imbalance (P_r). For digital RNA SNP, simulations of a 384-well experiment with m_r values of 0.1–2.0 were performed. At each m_r value, we simulated the scenario whereby 5,000 euploid and 5,000 T21 fetuses were tested (*SI Materials and Methods*). The SPRT curves appropriate for the given m_r were used to classify the 10,000 fetuses. The percentages of fetuses correctly and incorrectly classified as euploid or aneuploid and those unclassifiable for the given informative counts were determined (SI Table 7). The accuracies for diagnosing euploid and aneuploid cases are both 100%, for m_r values between 0.5 and 2.0. When the m_r value was 0.1, only 57% and 88% of euploid and T21 fetuses could be accurately classified by using 384 wells. Simulation results, using an illustrative repetition number of 300 times, are shown in SI Fig. 4.

Computer simulations for digital RCD analysis for a pure (100%) fetal DNA sample were similarly performed (SI Table 8 and SI Fig. 5). The extent of chr21 overrepresentation in digital RCD analysis depends on the fractional concentration of fetal DNA in the tested specimen. Because the fractional fetal DNA concentration becomes lower, the degree of chr21 overrepresentation diminishes, and thus a larger number of informative wells for accurate disease classification is required. Hence, simulations were further performed for fetal DNA concentrations of 50%, 25%, and 10% for a total well number ranging from 384 to 7,680 wells at an m_r value of 0.5 (SI Table 9). The performance of digital RCD is better for cases with a higher fetal DNA fractional concentration. At a fetal DNA concentration of 25% and with a total number of 7,680 PCR analyses, 97% of both euploid and aneuploid cases would be classifiable with no incorrect classification. The remaining 3% of cases require further analyses until classification can be achieved.

Validation of T21 Detection When Using Digital RNA SNP for *PLAC4*. The practical feasibility of digital RNA SNP was demonstrated by using the rs8130833 SNP on the *PLAC4* gene (*SI Materials and Methods*) (9). Placental DNA and RNA samples from two euploid and two T21 heterozygous placentas were analyzed. The placental DNA samples were analyzed with the omission of the reverse transcription step, thus essentially converting the procedure to digital DNA SNP analysis. We diluted the samples, aiming for approximately one allele of any type per well, and confirmed this

Table 2. Digital RNA SNP analysis of maternal plasma from euploid and T21 pregnancies

Sample	Genotype	No. of wells positive for individual alleles				m_r	SPRT result		
		A only	G only	AG	All negative		P_r	Unclassifiable region	Classification
M2390P	AG	90	100	97	97	0.67	0.526	0.62–0.64	Euploid
M2391P	AG	97	105	65	117	0.55	0.520	0.61–0.63	Euploid
M2473P	AG	66	92	34	192	0.30	0.582	0.59–0.62	Euploid
M2524P	AG	29	28	3	324	0.08	0.509	0.54–0.64	Euploid
M2528P	AG	112	85	44	143	0.41	0.569	0.60–0.62	Euploid
M2601P	AG	90	101	72	121	0.55	0.529	0.61–0.63	Euploid
M2607P	AG	73	91	57	163	0.41	0.555	0.60–0.63	Euploid
M2638P	AG	66	90	52	176	0.37	0.577	0.59–0.62	Euploid
M2639P	AG	71	56	17	240	0.21	0.559	0.58–0.62	Euploid
M2525P	AAG	110	53	21	200	0.21	0.675	0.58–0.61	T21
M2272P	AAG	246	127	112	283	0.37	0.660	0.60–0.61	T21
M2718P	AGG	66	114	66	138	0.42	0.633	0.60–0.62	T21
M1519P	AGG	58	130	54	142	0.34	0.691	0.59–0.62	T21

The number of wells for all samples except M2272P was 384. The number of wells for sample M2272P was 768. Genotypes were determined by mass spectrometric assay. The m_r value indicates the average no. of reference molecules per reaction well. The P_r values were calculated by using the following equation: no. of wells positive for the overrepresented allele/(no. of wells positive for A only + no. of wells positive for G only). The unclassifiable region for the corresponding m_r is shown. "Euploid" was assigned when the P_r was below the unclassifiable region; "T21" was assigned when the P_r was above the unclassifiable region.

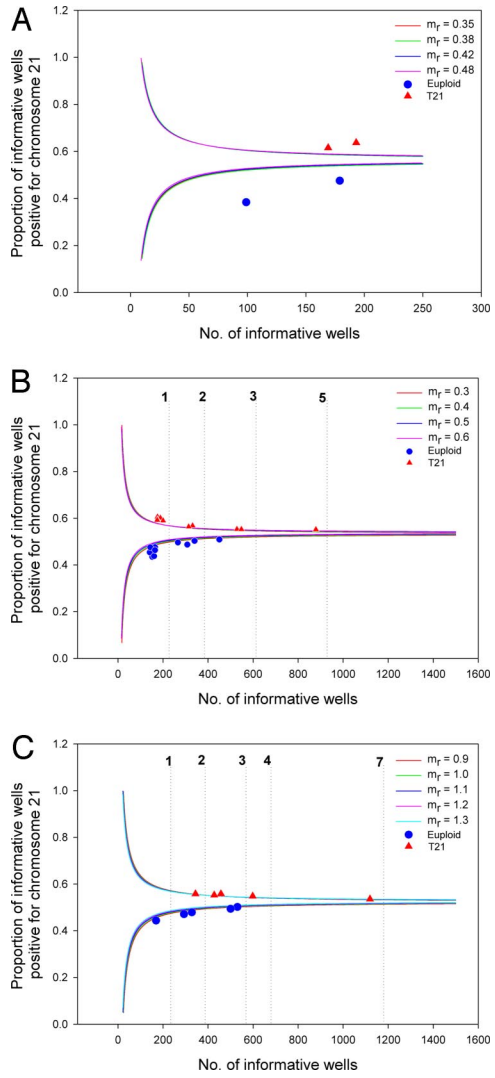

Fig. 3. SPRT interpretation of digital RCD analyses. (*A*) Placental DNA samples. (*B*) DNA mixtures of 50% placenta/maternal buffy coat. (*C*) DNA mixtures of 25% placenta/maternal buffy coat. Numbers at the top of *B* and *C* indicate the number of 384-well plates required before the data set was classifiable for the cases delimited by the dotted lines surrounding each number.

by a 96-well digital PCR analysis (Fig. 1). This was followed by a 384-well digital RNA SNP experiment. P_r and m_r were calculated, and the SPRT curve for this m_r value was used for data interpretation. The results are shown in Table 1. Each of these DNA and RNA samples was correctly classified with one 384-well experiment. We further tested plasma RNA samples from nine women carrying euploid fetuses and four women carrying T21 fetuses. All cases were correctly classified (Table 2). Initial results for one T21 case (M2272P) fell within the unclassifiable region between the SPRT curves after one 384-well experiment. Thus, we performed an additional 384-well run. New m_r and P_r values were calculated from the aggregated data of 768 wells, and the classification was performed by using a new set of SPRT curves selected based on this m_r value. The case was then scored correctly as aneuploid.

Validation of T21 Detection When Using Digital RCD. Placental DNA samples from two euploid and two T21 placentas were diluted to approximately one target template for either chromosome per well and confirmed by a 96-well digital PCR analysis (*SI Materials and Methods*). Each confirmed sample was analyzed by a 384-well digital RCD experiment, and the P_r and m_r values were calculated. For digital RCD, the chr1 paralog (19) was the reference template. This m_r value was used to select a corresponding set of SPRT curves for data interpretation. All of the placental samples were correctly classified (Fig. 3*A*). To demonstrate that digital RCD is applicable to nonpure fetal DNA samples (e.g., fetal DNA in maternal plasma), mixtures containing 50% and 25% of T21 placental DNA in a background of euploid maternal blood cell DNA were analyzed. Placental DNA from 10 T21 and 10 euploid cases was mixed with an equal amount of euploid maternal blood cell DNA, thus producing 20 DNA mixtures of 50%. Similarly, placental DNA from five T21 and five euploid cases was each mixed with a 3-fold excess of euploid maternal blood cell DNA, thus producing 10 DNA mixtures of 25%. All of the euploid and aneuploid DNA mixtures were correctly classified (Fig. 3 *B* and *C*). Each sample reached the point of being classifiable after a number of 384-well digital PCR analyses (Fig. 3 *B* and *C*). For the 50% DNA mixtures, the number of 384-well plates required ranged from one to five. For the 25% DNA mixtures, the number of 384-well plates required ranged from one to seven. The cumulative proportion of cases correctly classified increased progressively with the addition of each 384-well digital PCR analysis, as predicted in SI Table 9.

Discussion

In this study we have outlined and demonstrated the principle of digital PCR-based detection of chromosomal aneuploidy, using T21 as an example. As the statistical tool, we used SPRT, previously used for digital PCR-based detection of loss of heterozygosity (LOH) in samples with 50% tumor-derived DNA. In this setting, 50% of target DNA is contributed by normal cells with two copies of target chromosomes, and the other 50% is contributed by cancer cells where one target chromosome is lost (16, 18). We realized that a fetal trisomic cell is analogous to the combination of one noncancer cell and one cancer cell. The degree of allelic imbalance in a cancer sample containing 50% tumor-derived DNA is the same as that in a clinical sample containing pure fetal DNA (e.g., amniotic fluid) or RNA [e.g., *PLAC4* mRNA in maternal plasma (9)] from a pregnancy involving a T21 fetus. In both the cancer and the prenatal diagnosis scenarios, the ratio of the more abundant allele to the less abundant allele is 2:1. We further refined the SPRT analysis by constructing specific SPRT curves appropriate for the exact template concentration for any given digital PCR run and extended this strategy for the polymorphism-independent digital RCD approach. Alternative statistical methods, such as that based on the false discovery rate (20), could be further evaluated in future studies.

Our experimental and simulation data show that digital RNA SNP is an effective and accurate method for T21 detection. Because *PLAC4* mRNA in maternal plasma is derived purely from the fetus, for 12 of the 13 maternal plasma samples tested, only one 384-well digital PCR experiment was required for correct classification. This homogenous, real-time PCR-based approach thus offers an alternative to the previously described mass spectrometry-based approach for RNA SNP analysis (9). Apart from placental-specific mRNA transcripts, other types of fetal-specific nucleic acid species in maternal plasma could be used. One example is fetal epigenetic markers (12, 21) which have recently been used for the noninvasive prenatal detection of trisomy 18 via the epigenetic allelic ratio (EAR) approach (10). Thus, we predict that digital EAR would be a possible analytical technique.

Digital RCD was developed to overcome the requirement of heterozygosity for a polymorphism-based approach such as digital RNA SNP. Digital RCD could readily discriminate T21 and euploid placental DNA samples, thus supporting its applications to samples containing virtually pure fetal DNA, e.g., amniotic fluid and chorionic villus samples.

The application of digital RCD to DNA extracted from maternal plasma is complicated by the fact that fetal DNA constitutes only

a minor fraction of maternal plasma DNA, with a mean fractional concentration of some 3% between weeks 11 and 17 of gestation (8). Nevertheless, we have shown that digital RCD allows aneuploidy detection even when the fetal fraction is a minor population. With a decreasing fractional concentration of fetal DNA, e.g., during early gestation, a larger number of informative counts is needed for digital RCD. The significance of the present work, as summarized in SI Table 9, is that we have provided a set of benchmark parameters, e.g., fractional fetal DNA and total template molecules required, toward which future research can work. In our opinion, 7,680 reactions for a fractional fetal DNA concentration of 25% should be achievable (SI Table 9) and allows correct disease classification 97% of the time.

To achieve a fractional fetal DNA concentration of 25%, methods are needed to allow the selective enrichment of fetal DNA (22) or the suppression of the maternal DNA background (11, 23) in maternal plasma. For example, although the effect of formaldehyde has not been universally observed by all groups (24, 25), Dhallan *et al*. (11) reported that 85% (51 of 60) of their formaldehyde-treated plasma samples had fractional fetal DNA concentrations ≥25% and Benachi *et al*. (26) reported a mean fetal DNA concentration of 36.8% in their formaldehyde-treated plasma samples. Besides physical methods for fetal DNA enrichment and maternal DNA suppression, molecular enrichment strategies, such as targeting fetal DNA molecules that exhibit a particular DNA methylation pattern (12, 21, 27), may be possible. In this regard, placenta-specific DNA methylation markers from chr21 have recently been identified (S. S. C. Chim, S. Jin, T. Y. H. Lee, F.M.F.L., W. S. Lee, L. Y. S. Chan, Y. Jin, N. Yang, Y. K. Tong, T. Y. Leung, *et al.*, unpublished work).

The number of plasma DNA molecules that are present per unit volume of maternal plasma is limited (8). For example, in early pregnancy, the median maternal plasma concentration of an autosomal locus, the β-globin gene, has been shown to be 986 copies per milliliter, with contributions from both the fetus and mother (8). To capture 7,680 molecules, DNA extracted from some 8 ml of maternal plasma would be needed. This volume of plasma, obtainable from ≈15 ml of maternal blood, is at the limit of routine practice. However, we envision that multiple sets of chr21 and reference chromosome targets can be combined for digital RCD analysis. For five pairs of chr21 and reference chromosome targets, just 1.6 ml of maternal plasma would be needed to provide the number of template molecules needed for analysis. Multiplex single-molecule PCR would thus be needed. The robustness of such multiplex single-molecule analysis has been demonstrated previously for single-molecule haplotyping (28). Thus, the SPRT approach outlined here could be modified for the analysis of multiple target loci by methods like mass spectrometry (28).

The implementation of digital PCR, as illustrated in this proof-of-principle study, is rather labor-intensive, requiring one or more 384-well PCR plates to be set up per case. However, alternative approaches for conducting digital PCR, such as using microfluidic digital PCR chips (29, 30), emulsion PCR (31), and massively parallel genomic sequencing (32), are now available. These latter methods would greatly enhance the clinical applicability of the methods proposed here for noninvasive prenatal diagnosis and for other applications in which allelic or chromosome imbalance is seen.

Materials and Methods

Digital RNA SNP Analysis. A real-time PCR assay was designed to amplify *PLAC4* mRNA, with the two SNP alleles being discriminated by TaqMan probes. *PLAC4* mRNA concentrations were quantified in extracted RNA samples followed by dilutions to approximately one target template molecule of either type (i.e., either allele) per well. We distributed the diluted sample to 96 wells for real-time PCR analysis to confirm that a usable dilution has been achieved. When ≈37% (i.e., e^{-1}) of the wells were shown to be negative for any amplification, we proceeded to the digital RNA SNP analysis using the same diluted sample for 384-well analyses. Details are given in the *SI Materials and Methods*.

Digital RCD Analysis. Extracted DNA was quantified by spectrophotometry (NanoDrop Technologies, Wilmington, DE) and diluted to a concentration of approximately one target template from either chr21 or chr1 per well. A real-time PCR assay was designed to amplify a paralogous sequence (19) present on both chromosomes, distinguishable by a pair of TaqMan probes. The diluted DNA sample was first analyzed by the assay using the chr1 probe only in a 96-well format to confirm whether ≈37% of the wells were negative; then we proceeded to digital RCD analysis using both TaqMan probes in 384-well plates. Details are given in the *SI Materials and Methods*.

Computer Simulation of Classification Accuracy. The computer simulation was performed with Microsoft Office Excel 2003 software (Microsoft, Redmond, WA) and SAS 9.1 for Windows software (SAS Institute, Cary, NC). Details are given in the *SI Materials and Methods*.

This work was supported by Central Allocation Grant CUHK1/03C from the Research Grants Council of the Hong Kong Special Administrative Region (China) (to Y.M.D.L., R.W.K.C., and C.R.C.) and by the Li Ka Shing Foundation (to Y.M.D.L.).

1. Tabor A, Philip J, Madsen M, Bang J, Obel EB, Norgaard-Pedersen B (1986) *Lancet* 1:1287–1293.
2. Malone FD, Canick JA, Ball RH, Nyberg DA, Comstock CH, Bukowski R, Berkowitz RL, Gross SJ, Dugoff L, Craigo SD, *et al*. (2005) *N Engl J Med* 353:2001–2011.
3. Lo YM, Corbetta N, Chamberlain PF, Rai V, Sargent IL, Redman CW, Wainscoat JS (1997) *Lancet* 350:485–487.
4. Lo YMD, Chiu RWK (2007) *Nat Rev Genet* 8:71–77.
5. Costa JM, Benachi A, Gautier E (2002) *N Engl J Med* 346:1502 (lett).
6. Lo YMD, Hjelm NM, Fidler C, Sargent IL, Murphy MF, Chamberlain PF, Poon PM, Redman CW, Wainscoat JS (1998) *N Engl J Med* 339:1734–1738.
7. Chiu RWK, Lau TK, Leung TN, Chow KCK, Chui DHK, Lo YMD (2002) *Lancet* 360:998–1000.
8. Lo YMD, Tein MS, Lau TK, Haines CJ, Leung TN, Poon PM, Wainscoat JS, Johnson PJ, Chang AM, Hjelm NM (1998) *Am J Hum Genet* 62:768–775.
9. Lo YMD, Tsui NBY, Chiu RWK, Lau TK, Leung TN, Heung MM, Gerovassili A, Jin Y, Nicolaides KH, Cantor CR, *et al*. (2007) *Nat Med* 13:218–223.
10. Tong YK, Ding C, Chiu RWK, Gerovassili A, Chim SSC, Leung TY, Leung TN, Lau TK, Nicolaides KH, Lo YMD (2006) *Clin Chem* 52:2194–2202.
11. Dhallan R, Guo X, Emche S, Damewood M, Bayliss P, Cronin M, Barry J, Betz J, Franz K, Gold K, *et al*. (2007) *Lancet* 369:474–481.
12. Chim SSC, Tong YK, Chiu RWK, Lau TK, Leung TN, Chan LYS, Oudejans CB, Ding C, Lo YMD (2005) *Proc Natl Acad Sci USA* 102:14753–14758.
13. Chiu RWK, Murphy MF, Fidler C, Zee BC, Wainscoat JS, Lo YMD (2001) *Clin Chem* 47:667–672.
14. Zimmermann B, Holzgreve W, Wenzel F, Hahn S (2002) *Clin Chem* 48:362–363.
15. Vogelstein B, Kinzler KW (1999) *Proc Natl Acad Sci USA* 96:9236–9241.
16. Zhou W, Galizia G, Lieto E, Goodman SN, Romans KE, Kinzler KW, Vogelstein B, Choti MA, Montgomery EA (2001) *Nat Biotechnol* 19:78–81.
17. Chang HW, Lee SM, Goodman SN, Singer G, Cho SK, Sokoll LJ, Montz FJ, Roden R, Zhang Z, Chan DW, *et al*. (2002) *J Natl Cancer Inst* 94:1697–1703.
18. Zhou W, Goodman SN, Galizia G, Lieto E, Ferraraccio F, Pignatelli C, Purdie CA, Piris J, Morris R, Harrison DJ, *et al*. (2002) *Lancet* 359:219–225.
19. Deutsch S, Choudhury U, Merla G, Howald C, Sylvan A, Antonarakis SE (2004) *J Med Genet* 41:908–915.
20. El Karoui N, Zhou W, Whittemore AS (2006) *Stat Med* 25:3124–3133.
21. Chan KCA, Ding C, Gerovassili A, Yeung SW, Chiu RWK, Leung TN, Lau TK, Chim SSC, Chung GTY, Nicolaides KH, *et al*. (2006) *Clin Chem* 52:2211–2218.
22. Li Y, Zimmermann B, Rusterholz C, Kang A, Holzgreve W, Hahn S (2004) *Clin Chem* 50:1002–1011.
23. Dhallan R, Au WC, Mattagajasingh S, Emche S, Bayliss P, Damewood M, Cronin M, Chou V, Mohr M (2004) *J Am Med Assoc* 291:1114–1119.
24. Chung GTY, Chiu RWK, Chan KCA, Lau TK, Leung TN, Lo YMD (2005) *Clin Chem* 51:655–658.
25. Chinnapapagari SK, Holzgreve W, Lapaire O, Zimmermann B, Hahn S (2005) *Clin Chem* 51:652–655.
26. Benachi A, Yamgnane A, Olivi M, Dumez Y, Gautier E, Costa JM (2005) *Clin Chem* 51:242–244.
27. Chiu RWK, Chim SSC, Wong IHN, Wong CS, Lee WS, To KF, Tong JH, Yuen RKC, Shum AS, Chan JK, *et al*. (2007) *Am J Pathol* 170:941–950.
28. Ding C, Cantor CR (2003) *Proc Natl Acad Sci USA* 100:7449–7453.
29. Warren L, Bryder D, Weissman IL, Quake SR (2006) *Proc Natl Acad Sci USA* 103:17807–17812.
30. Ottesen EA, Hong JW, Quake SR, Leadbetter JR (2006) *Science* 314:1464–1467.
31. Dressman D, Yan H, Traverso G, Kinzler KW, Vogelstein B (2003) *Proc Natl Acad Sci USA* 100:8817–8822.
32. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, *et al*. (2005) *Nature* 437:376–380.