



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(45) 공고일자 2018년02월14일  
(11) 등록번호 10-1829564  
(24) 등록일자 2018년02월08일

(51) 국제특허분류(Int. Cl.)  
G06F 19/18 (2011.01) C12Q 1/68 (2018.01)  
G06F 19/22 (2017.01)

(52) CPC특허분류  
G06F 19/18 (2013.01)  
C12Q 1/6827 (2013.01)

(21) 출원번호 10-2016-7021212(분할)

(22) 출원일자(국제) 2008년07월23일  
심사청구일자 2016년09월01일

(85) 번역문제출일자 2016년08월02일

(65) 공개번호 10-2016-0113146

(43) 공개일자 2016년09월28일

(62) 원출원 특허 10-2010-7003906  
원출원일자(국제) 2008년07월23일  
심사청구일자 2013년07월23일

(86) 국제출원번호 PCT/GB2008/002524

(87) 국제공개번호 WO 2009/013492  
국제공개일자 2009년01월29일

(30) 우선권주장  
60/951,438 2007년07월23일 미국(US)

(56) 선행기술조사문헌  
JP2007515947 A\*  
KR1020020064298 A\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
더 차이나이즈 유니버시티 오브 홍콩  
중국 홍콩 엔. 티. 새턴 더 차이나이즈 유니버시티 오브 홍콩

(72) 발명자  
로, 육-밍 데니스  
홍콩, 구룡, 호만탄, 7 킹 탁 스트리트, 4 플로어  
치우, 로사 와이 쿤  
홍콩, 뉴 테리토리스, 타이 포, 1 홍 램 드라이브, 콘스텔레이션 코브, 블록 1, 플랫 1에 이  
(뒷면에 계속)

(74) 대리인  
손민

전체 청구항 수 : 총 20 항

심사관 : 박진아

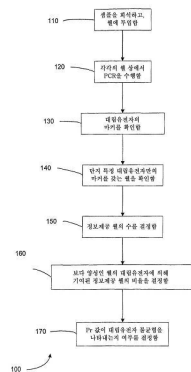
(54) 발명의 명칭 **핵산 서열 불균형의 결정**

**(57) 요약**

본 발명에서는, 핵산 서열 불균형이 생물학적 샘플 내에 존재하는지 여부를 결정하는 방법, 시스템 및 장치가 제공된다. 예를 들어, 2개의 서열 (또는 서열 세트)의 비의 불균형을 측정하기 위한 하나 이상의 분별값이 선택된다. 분별값은 적어도 부분적으로 샘플, 예컨대 모체 핵산 서열의 배경을 함유하는 모체 혈장 중 태아 DNA의 백

(뒷면에 계속)

**대표도** - 도1



분율을 기준으로 결정될 수 있다. 분별값은 또한 반응 당 서열의 평균 농도를 기준으로 결정될 수 있다. 한 측면에서, 분별값은 특정 핵산 서열을 함유하는 것으로 평가되는 정보제공 웰의 비율로부터 결정되며, 그 비율은 상기 언급된 백분율 및/또는 평균 농도를 기준으로 결정된다. 분별값은 다수의 상이한 유형의 방법, 예컨대 순차 확률비 검증 (SPRT)을 사용하여 결정될 수 있다.

(52) CPC특허분류

*C12Q 1/6883* (2013.01)

*C12Q 1/6888* (2013.01)

*G06F 19/22* (2013.01)

*C12Q 2600/112* (2013.01)

*C12Q 2600/154* (2013.01)

*C12Q 2600/156* (2013.01)

*G01N 2800/387* (2013.01)

(72) 발명자

**찬, 판 치**

홍콩, 구룡, 메이 푸 선 추엔, 브로드웨이 스트리트, 블록 34, 13/에프, 플랫 에이

**지, 베니 충 잉**

홍콩, 뉴 테리토리스, 마 온 산, 라 코스타, 타워 2, 플랫 18이

**총, 카 춘**

홍콩, 뉴 테리토리스, 타이 포, 푸 신 에스테이트, 신 킹 하우스, 29/에프, 플랫 06

## 명세서

### 청구범위

#### 청구항 1

핵산 서열 불균형이 태아를 임신한 여성 대상체의 생물학적 샘플 내에 존재하는지 여부를 결정하기 위한 방법으로서, 상기 생물학적 샘플은 여성 대상체 및 태아로부터의 무-세포 핵산 분자를 포함하고, 상기 생물학적 샘플은 모체 혈액이고, 상기 방법은 하기 단계를 포함하는 것인, 방법:

생물학적 샘플로부터의 무-세포 핵산 분자를 포함하는 다수의 반응으로부터 정량 데이터를 획득하는 단계로, 상기 데이터는

다수의 임상적으로 관련된 핵산 서열에 대해

(1) 다수의 반응에서 임상적으로 관련된 핵산 서열의 제1 총량을 나타내는 제1 정량 데이터; 및

다수의 배경 핵산 서열에 대해

(2) 다수의 반응에서 배경 핵산 서열의 제2 총량을 나타내는 제2 정량 데이터를 포함하고,

상기 배경 핵산 서열은 임상적으로 관련된 핵산 서열의 임의의 하나와 상이한 것인, 단계;

제1 정량 데이터로부터 제1 총량을 결정하는 단계;

제2 정량 데이터로부터 제2 총량을 결정하는 단계;

제1 총량과 제2 총량으로부터 파라미터를 결정하는 단계로, 상기 파라미터가 제1 총량과 제2 총량 사이의 상대적 양을 제공하는 것인, 단계;

파라미터를 하나 이상의 분별값(cutoff values)과 비교하는 단계; 및

비교를 기준으로, 핵산 서열 불균형이 임상적으로 관련된 핵산 서열에 대한 태아 계층에 존재하는지 여부를 결정하는 단계.

#### 청구항 2

제1항에 있어서, 제1 정량 데이터로부터 제1 총량을 결정하는 단계가

임상적으로 관련된 핵산 서열 각각에 대해

제1 정량 데이터를 분석하여 다수의 반응에서 각각의 임상적으로 관련된 핵산 서열의 각각의 제1 양을 결정하는 단계를 포함하고,

제2 정량 데이터로부터 제2 총량을 결정하는 단계가

배경 핵산 서열 각각에 대해

제2 정량 데이터를 분석하여 다수의 반응에서 각각의 배경 핵산 서열의 각각의 제2 양을 결정하는 단계를 포함하고,

각각의 제1 양을 합산하여 제1 총량을 계산하고,

각각의 제2 양을 합산하여 제2 총량을 계산하는 것인 방법.

#### 청구항 3

제2항에 있어서, 제1 합 및 제2 합을 계산하기 전에 각각의 제1 양 및 제2 양을 교정하는 단계를 더 포함하는 방법.

#### 청구항 4

제1항에 있어서, 반응이 서열분석 반응인 방법.

**청구항 5**

제4항에 있어서, 서열분석이 라이게이션을 포함하는 것인 방법.

**청구항 6**

제4항에 있어서, 서열분석이 합성을 포함하는 것인 방법.

**청구항 7**

제4항에 있어서, 제1 총량을 결정하는 단계가

서열 태그를 인간 게놈에 대해 정렬하는 단계; 및

다수의 임상적으로 관련된 핵산 서열에 대해 정렬된 서열 태그의 수를 계수하는 단계를 포함하는 것인 방법.

**청구항 8**

제4항에 있어서, 다수의 임상적으로 관련된 핵산 서열 및 다수의 배경 핵산 서열에 대해 생물학적 샘플을 농축시키는 단계; 및

이어서, 농축된 생물학적 샘플 중의 핵산 분자를 서열분석하는 단계를 더 포함하는 방법.

**청구항 9**

제8항에 있어서, 다수의 임상적으로 관련된 핵산 서열 및 다수의 배경 핵산 서열에 대해 생물학적 샘플을 농축시키는 단계가 혼성화를 기초로 하는 기술을 이용하는 단계를 포함하는 것인 방법.

**청구항 10**

제9항에 있어서, 혼성화를 기초로 하는 기술을 이용하는 단계가 혼성화로 다수의 임상적으로 관련된 핵산 분자 및 다수의 배경 핵산 서열을 선별하기 위해 올리고뉴클레오티드 어레이를 사용하는 단계를 포함하는 것인 방법.

**청구항 11**

제8항에 있어서, 다수의 임상적으로 관련된 핵산 서열 및 다수의 배경 핵산 서열에 대해 생물학적 샘플을 농축시키는 단계가 다수의 임상적으로 관련된 핵산 서열 및 다수의 배경 핵산 서열을 증폭시키는 단계를 포함하는 것인 방법.

**청구항 12**

제11항에 있어서, 동일한 프라이머 쌍이 임상적으로 관련된 핵산 서열 중 하나 및 배경 핵산 서열 중 하나를 증폭시키는데 사용되는 것인 방법.

**청구항 13**

제1항에 있어서, 다수의 반응을 수행하기 전에 다수의 임상적으로 관련된 핵산 서열 및 다수의 배경 핵산 서열에 대해 생물학적 샘플을 농축시키는 단계를 더 포함하는 방법.

**청구항 14**

제1항에 있어서, 생물학적 샘플 중의 태아 DNA의 분획 농도를 결정하는 단계;

분획 농도를 사용하여 하나 이상의 분별값을 결정하는 단계를 더 포함하는 방법.

**청구항 15**

제14항에 있어서, 생물학적 샘플 중의 태아 DNA의 분획 농도를 결정하는 단계가 여성 대상체와 태아 사이의 다형성 차이의 정량화를 포함하는 것인 방법.

**청구항 16**

제15항에 있어서, 정량화가

여성 대상체가 동형접합성이고 태아가 이형접합성인 표적 다형성 위치를 확인하는 단계; 및

표적 다형성 위치에서 태아-특이적 대립유전자의 양과 표적 다형성 위치에서 태아-특이적이지 않은 공통 대립유전자의 양을 비교하여 태아 DNA의 분획 농도를 결정하는 단계를 포함하는 것인 방법.

**청구항 17**

제14항에 있어서, 태아가 남성이며, 생물학적 샘플 중의 태아 DNA의 분획 농도를 결정하는 단계가 Y 염색체 DNA 농도를 결정하는 단계를 포함하는 것인 방법.

**청구항 18**

제14항에 있어서, 생물학적 샘플 중의 태아 DNA의 분획 농도를 결정하는 단계가 제1 유전자좌에서의 태아-특이적 메틸화 패턴을 나타내는 DNA 분자의 양을 제1 유전자좌에서의 DNA 분자의 총량과 비교하는 단계를 포함하는 것인 방법.

**청구항 19**

제18항에 있어서, 태아-유도된 DNA 분자가 과-메틸화되고, 모체-유도된 DNA 분자가 저-메틸화된 것인 방법.

**청구항 20**

핵산 서열 불균형이 태아를 임신한 여성 대상체의 생물학적 샘플 내에 존재하는지 여부를 결정하기 위한 작업을 수행하기 위한 컴퓨터 시스템 제어용의 다수의 지시로 코딩된 컴퓨터 판독가능 기록 매체로서, 생물학적 샘플은 여성 대상체 및 태아로부터의 무-세포 핵산 분자를 포함하고, 지시는 제1항 내지 제19항 중 어느 한 항에 따른 방법을 포함하는 것인, 컴퓨터 판독가능 기록 매체.

**발명의 설명**

**기술 분야**

- [0001] 우선권 주장
- [0002] 본 출원은 2007년 7월 23일 출원된, 발명의 명칭이 "핵산 서열 불균형의 결정(DETERMINING A NUCLEIC ACID SEQUENCE IMBALANCE)"인 미국 가출원 번호 제60/951438호 (사건관리 번호 제016285-005200US호)를 우선권으로 주장하며 이의 정규출원이고, 상기 가출원의 전체 내용은 모든 목적을 위해 참조로 본원에 포함된다.
- [0003] 관련 출원의 상호 참조
- [0004] 본 출원은 또한 동시에 출원된, 발명의 명칭이 "게놈 서열분석을 사용한 태아 염색체 이수성의 진단(DIAGNOSING FETAL CHROMOSOMAL ANEUPLOIDY USING GENOMIC SEQUENCING)"인 정규출원 (사건관리 번호 제016285-005220US호)에 관한 것이고, 이 가출원의 전체 내용은 모든 목적을 위해 참조로 본원에 포함된다.
- [0005] 발명의 분야
- [0006] 본 발명은 일반적으로 2개의 상이한 핵산 서열간 불균형을 결정하는 유전자형 및 질환의 진단 시험, 보다 특히 모체 혈액 샘플의 시험을 통한 태아에서의 다운 증후군, 다른 염색체 이수성(aneuploidy), 돌연변이 및 유전자형의 확인에 관한 것이다. 본 발명은 또한 암의 검출, 이식의 모니터링 및 감염성 질환의 모니터링에 관한 것이다.

**배경 기술**

- [0008] 유전 질환, 암 및 다른 상태는 종종 2개의 상응하는 염색체 또는 대립유전자 또는 다른 핵산 서열에서의 불균형으로부터 초래되거나 이러한 불균형을 생성한다. 즉, 다른 서열에 비해 한 서열의 양이 정상적인 경우보다 많거나 적다. 통상적으로, 정상 비율은 대등한 50/50 비율이다. 다운 증후군 (삼염색체 21)은 추가 염색체 21의 불균형을 갖는 질환이다.
- [0009] 삼염색체 21의 통상적인 출생 전 진단 방법은 한정된 태아 소실 위험을 낳는 침습적 절차, 예컨대 양수검사 또는 융모막 용모 생검에 의해 태아 물질을 샘플링하는 것을 포함한다. 비-침습적 절차, 예컨대 초음파검사 및 생화학 마커에 의한 스크리닝은 한정적인 침습적 진단 절차 전에 임신 여성을 위험-등급화하는데 사용되었다.

그러나, 이러한 스크리닝 방법은 전형적으로 핵심 염색체 이상 대신에 삼염색체 21과 관련된 부대징후를 측정하므로, 최적에 못미치는 진단 정확성 및 다른 단점들 (예컨대 재태 기간에 의해 많은 영향을 받음)을 갖는다.

[0010] 1997년에 모체 혈장 중 순환 무-세포 태아 DNA의 발견은 비-침습적 출생 전 진단의 새로운 가능성을 제공하였다 (문헌 [Lo, YMD and Chiu, RWK 2007 Nat Rev Genet 8, 71-77]). 이 방법이 성별-연계된 유전자 장애 (문헌 [Costa, JM et al. 2002 N Engl J Med 346, 1502]) 및 특정 단일 유전자 장애 (문헌 [Lo, YMD et al. 1998 N Engl J Med 339, 1734-1738])의 출생 전 진단에 용이하게 적용되었지만, 태아 염색체 이수성의 출생 전 검출에 적용하기에는 상당한 난점이 있었다 (상기 문헌 [Lo, YMD and Chiu, RWK 2007]). 첫째, 태아 핵산은 모체 혈장 중에서, 종종 분석에 방해가 될 수 있는 다량의 모체 기원 배경 핵산과 공존한다 (문헌 [Lo, YMD et al. 1998 Am J Hum Genet 62, 768-775]). 둘째, 태아 핵산은 모체 혈장 중에서 우세하게는 무-세포 형태로 순환하여, 태아 게놈 내 유전자 또는 염색체의 투여 정보 유도를 어렵게 만든다.

[0011] 이들 난점을 극복하는 유의한 개발이 최근에 이루어졌다 (문헌 [Benachi, A & Costa, JM 2007 Lancet 369, 440-442]). 한 접근법은 모체 혈장 중 태아-특이적 핵산을 검출하여, 모체 배경 방해 문제를 극복하는 것이다 (상기 문헌 [Lo, YMD and Chiu, RWK 2007]). 염색체 21의 투여는 태반-유래 DNA/RNA 분자에서 다형성 대립유전자의 비율로부터 추론되었다. 그러나, 이러한 방법은, 샘플이 보다 적은 양의 표적 유전자를 함유하는 경우 덜 정확하며, 하나의 다형성이 사용되는 경우 단지 그 집단의 아군인, 표적 다형성에 대해 이형접합성인 태아에게만 적용될 수 있다.

[0012] 달란 등(Dhallan et al)의 문헌 [Dhallan, R, et al. 2007, supra Dhallan, R, et al. 2007 Lancet 369, 474-481]에는, 모체 혈장에 포름알데히드를 첨가함으로써 순환 태아 DNA의 비율을 증가시키는 별법의 전략이 기재되어 있다. 모체 혈장 중 태아에 의해 기여된 염색체 21 서열의 비율은, 염색체 21 상의 단일 뉴클레오티드 다형성 (SNP)에 대한 부계-유전 태아-특이적 대립유전자 대 비-태아-특이적 대립유전자의 비율을 평가함으로써 결정되었다. SNP 비율은 참조 염색체에 대해 유사하게 추정되었다. 이어서, 태아 염색체 21의 불균형은, 염색체 21에 대한 SNP 비율과 참조 염색체에 대한 SNP 비율 사이의 통계적으로 유의한 차이를 검출함으로써 추론되었으며, 상기 유의성은 0.05 이하의 고정 p-값을 사용하여 정의된다. 고 집단 적용범위를 보장하기 위해, 염색체 당 500개 초과 SNP가 표적화되었다. 그러나, 고비율로 증가시키기 위한 포름알데히드의 효과와 관련하여 논쟁이 있었으므로 (문헌 [Chung, GTY, et al. 2005 Clin Chem 51, 655-658]), 방법의 재현성은 더 평가될 필요가 있다. 또한, 각각의 태아 및 모체는 각각의 염색체에 대해 상이한 수의 SNP의 정보를 제공하므로, SNP 비율 비교를 위한 통계적 시험의 영향력은 건마다 변할 것이다 (문헌 [Lo, YMD & Chiu, RWK. 2007 Lancet 369, 1997]). 또한, 이들 접근법은 유전자 다형성의 검출에 좌우되므로, 이들 다형성에 이형접합성인 태아에게 국한된다.

[0013] 삼염색체 21 및 정배수체 태아로부터 수득한 양수세포 배양물에서 염색체 21 유전자좌 및 참조 유전자좌의 중합효소 연쇄 반응 (PCR) 및 DNA 정량화를 사용하여, 짐메르만 등(Zimmermann et al) (문헌 [2002 Clin Chem 48, 362-363])은 삼염색체 21 태아에서 염색체 21 DNA 서열의 1.5배 증가를 기준으로 상기 2군의 태아를 구별할 수 있었다. DNA 주형 농도의 2배 차이는 단지 한 임계 사이클(threshold cycle (Ct))의 차이를 구성하기 때문에, 1.5배 차이의 구별은 통상적인 실시간 PCR의 한계였다. 보다 미세한 정도의 정량 구별을 달성하기 위해, 별법의 전략이 필요하다. 따라서, 본 발명의 일부 실시양태는 이러한 목적을 위해 디지털 PCR (문헌 [Vogelstein, B et al. 1999 Proc Natl Acad Sci USA 96, 9236-9241])을 사용한다.

[0014] 디지털 PCR은 핵산 샘플에서 대립유전자 비율이 벗어난 것을 검출하기 위해 개발되었다 (문헌 [Chang, HW et al. 2002 J Natl Cancer Inst 94, 1697-1703]). 임상적으로, 디지털 PCR은 종양 DNA 샘플에서 이형접합성의 소실 (LOH)을 검출하는데 유용한 것으로 나타났다 (문헌 [Zhou, W. et al. 2002 Lancet 359, 219-225]). 디지털 PCR 결과의 분석을 위해, 순차 확률비 검증 (SPRT)이 이전 연구에 의해 채택되어, 샘플 중 LOH의 존재를 제시하는 것과 제시하지 않는 것으로 실험 결과를 분류하였다 (문헌 [El Karoui et al. 2006 Stat Med 25, 3124-3133]). 이전 연구에서 사용된 방법에서, LOH를 결정하기 위한 분별값 (cutoff value)은 DNA 중 2개 대립유전자의 고정 참조 비율 2/3을 사용하였다. 모체 혈장 중 태아 핵산의 양, 비율 및 농도는 변할 수 있기 때문에, 이들 방법은 모체 혈장 중 모체 핵산의 배경에서 태아 핵산을 사용하여 삼염색체 21을 검출하는데 적합하지 않다.

[0015] 순환 태아 핵산 분석을 기준으로 하는 태아 삼염색체 21 (및 다른 불균형) 검출을 위한 비-침습적 시험, 특히 유전자 다형성 및/또는 태아-특이적 마커의 사용에 독립적인 비-침습적 시험을 수행하는 것이 바람직하다. 또한, 정확한 분별값 결정 및 서열 계수를 수행하여, 정확성에 필요한 모체 혈장 핵산 분자의 양 및/또는 데이터

의 웰 수를 감소시킬 수 있음으로써, 증가된 효율 및 비용-효과를 제공하는 것이 바람직하다. 또한, 비-침습적 시험이 고 감수성 및 특이성을 가져 오류 진단을 최소화하는 것이 바람직하다.

[0016] 모체 혈장 중 태아 DNA 검출에 대한 또다른 적용은 베타-지중해빈혈과 같은 단일 유전자 장애의 출생 전 진단에 대한 것이다. 그러나, 태아 DNA는 단지 모체 혈장 중 소수 분획의 DNA를 구성하므로, 이러한 접근법은 단지 태아에서 부계로부터 유전되었지만 모계로부터는 존재하지 않는 돌연변이를 검출할 수 있을 뿐이라고 생각된다. 이러한 예로는, 베타-지중해빈혈을 일으키는 베타-글로빈 유전자의 코돈 41/42에서의 4 bp 결실 (문헌 [Chiu RWK et al. 2002 Lancet, 360, 998-1000]) 및 낭포성 섬유증을 일으키는 낭포성 섬유증 막횡단 전도 조절자 유전자의 Q890X 돌연변이 (문헌 [Gonzalez-Gonzalez et al 2002 Prenat Diagn, 22, 946-8])가 포함된다. 그러나, 베타-지중해빈혈 및 낭포성 섬유증 둘 다는, 이 질환이 그 자체로 표출되기 전에 태아에게 각각의 부모로부터 돌연변이가 유전될 필요가 있는 상염색체 열성 상태이므로, 부계-유전 돌연변이만의 검출은 단지 상기 질환에 걸린 태아를 가질 위험을 25%에서 50%로 증가시킬 것이다. 진단적으로 이는 이상적이 아니다. 따라서, 현존하는 접근법의 주된 진단 적용은, 부계-유전된 태아 돌연변이가 모체 혈장에서 검출될 수 없는 경우, 이후 태아가 동형접합성 질환 상태에 걸리는 것으로부터 배제될 수 있는 경우의 시나리오를 위해서일 것이다. 그러나, 진단적으로, 이러한 접근법은 부계 돌연변이의 음성 검출을 기준으로 결론을 내린다는 단점을 갖는다. 따라서, 상기 제한 없이, 완전한 태아 유전자형 (동형접합성 정상, 동형접합성 돌연변이 또는 이형접합성)이 모체 혈장으로부터 결정되도록 하는 접근법이 매우 바람직할 것이다.

**발명의 내용**

**해결하려는 과제**

- [0018] 개요
- [0019] 본 발명의 실시양태는, 핵산 서열 불균형 (예를 들어, 대립유전자 불균형, 돌연변이 불균형 또는 염색체 불균형)이 생물학적 샘플 내에 존재하는지 여부를 결정하기 위한 방법, 시스템 및 장치를 제공한다. 예를 들어, 2개의 서열 (또는 서열 세트) 양의 비율의 불균형을 결정하기 위한 하나 이상의 분별값이 선택된다.
- [0020] 한 실시양태에서, 분별값은 적어도 부분적으로 생물학적 샘플, 예컨대 모체 핵산 서열의 배경을 함유하는 모체 혈장 또는 혈청 또는 노 중 태아 (임상적으로 관련된 핵산) 서열의 백분율을 기준으로 결정된다. 또다른 실시양태에서, 분별값은 다수의 반응에서 서열의 평균 농도를 기준으로 결정된다. 한 측면에서, 분별값은 특정 핵산 서열을 함유하는 것으로 평가되는 정보제공 웰의 비율로부터 결정되며, 그 비율은 상기 언급된 백분율 및/또는 평균 농도를 기준으로 결정된다.
- [0021] 분별값은 다수의 상이한 유형의 방법, 예컨대 SPRT, 오류 발견, 신뢰 구간, 수신자 판단 특성 (receiver operating characteristic, ROC)을 사용하여 결정될 수 있다. 이러한 전략은 신뢰성 있는 분류가 만들어질 수 있기 전에 필요한 시험의 양을 더 최소화시켰다. 이는 주형 양이 종종 제한적인 혈장 핵산 분석과 특히 관련되어 있다.
- [0022] 한 예시적 실시양태에 따라, (1) 임상적으로 관련된 핵산 서열의 제1 양을 나타내는 제1 세트의 정량 데이터; 및 (2) 임상적으로 관련된 핵산 서열과 상이한 배경 핵산 서열의 제2 양을 나타내는 제2 세트의 정량 데이터를 포함하는 데이터를 다수의 반응으로부터 획득하는 단계; 두 데이터 세트로부터 파라미터를 결정하는 단계; 다수의 반응 각각에서, 임상적으로 관련된 핵산 서열 또는 배경 핵산 서열인 참조 핵산 서열의 평균 농도로부터 제1 분별값을 유도하는 단계; 파라미터를 제1 분별값과 비교하는 단계; 및 비교를 기준으로, 핵산 서열 불균형이 존재하는지 여부를 분류를 결정하는 단계를 포함하는, 핵산 서열 불균형이 생물학적 샘플 내에 존재하는지 여부를 결정하기 위한 방법이 제공된다.
- [0023] 또다른 예시적 실시양태에 따라, (1) 임상적으로 관련된 핵산 서열의 제1 양을 나타내는 제1 세트의 정량 데이터; 및 (2) 임상적으로 관련된 핵산 서열과 상이한 배경 핵산 서열의 제2 양을 나타내는 제2 세트의 정량 데이터를 포함하며, 임상적으로 관련된 핵산 서열 및 배경 핵산 서열은 제1 유형 세포 및 하나 이상의 제2 유형 세포로부터 유래된 것인 데이터를 다수의 반응으로부터 획득하는 단계; 두 데이터 세트로부터 파라미터를 결정하는 단계; 생물학적 샘플에서, 제1 유형 세포로부터의 핵산 서열의 양을 측정하여 생성된 제1 백분율로부터 제1 분별값을 유도하는 단계; 파라미터를 분별값과 비교하는 단계; 및 비교를 기준으로, 핵산 서열 불균형이 존재하는지 여부를 분류를 결정하는 단계를 포함하는, 핵산 서열 불균형이 생물학적 샘플 내에 존재하는지 여부를 결정하기 위한 방법이 제공된다.

- [0024] 본 발명의 다른 실시양태는 본원에 기재된 방법과 관련된 시스템 및 컴퓨터 판독가능 매체에 관한 것이다.
- [0025] 하기 상세한 설명 및 수반되는 도면을 참조로 하여 본 발명의 성질 및 이점을 보다 잘 이해할 수 있다.

**과제의 해결 수단**

- [0027] 정의
- [0028] 본원에 사용된 용어 "생물학적 샘플"은 대상체 (예를 들어, 임신 여성과 같은 인간)로부터 취하며 하나 이상의 관심 핵산 분자(들)를 함유하는 임의의 샘플을 지칭한다.
- [0029] 용어 "핵산" 또는 "폴리뉴클레오티드" 단일가닥 또는 이중가닥 형태의 데옥시리보핵산 (DNA) 또는 리보핵산 (RNA) 및 이들의 중합체를 지칭한다. 구체적으로 한정하지 않는 한, 상기 용어는 참조 핵산과 유사한 결합 성질을 가지며 천연 발생 뉴클레오티드와 유사한 방식으로 대사되는 천연 뉴클레오티드의 공지된 유사체를 함유하는 핵산을 포함한다. 달리 나타내지 않는 한, 특정 핵산 서열은 또한 암묵적으로, 그의 보존적으로 변형된 변이체 (예를 들어, 퇴행 코돈 치환체), 대립유전자, 오소로그(ortholog), SNP 및 상보적 서열, 및 명백하게 나타내어진 서열을 포함한다. 구체적으로, 퇴행 코돈 치환체는 하나 이상의 선택된 (또는 모든) 코돈의 제3 위치가 혼합 염기 및/또는 데옥시이노신 잔기로 치환되는 서열을 생성함으로써 달성될 수 있다 (문헌 [Batzer et al., Nucleic Acid Res. 19:5081 (1991)], [Ohtsuka et al., J. Biol. Chem. 260:2605-2608 (1985)] 및 [Rossolini et al., Mol. Cell. Probes 8:91-98 (1994)]). 용어 핵산은 유전자, cDNA, mRNA, 소형 비코딩 RNA, 마이크로 RNA (miRNA), 피위(Piwi)-상호작용 RNA, 및 유전자 또는 유전자좌에 의해 코딩되는 단형 헤어핀 RNA (shRNA)와 상호교환적으로 사용된다.
- [0030] 용어 "유전자"는 폴리펩티드 쇄를 생성하는데 관여하는 DNA 절편을 의미한다. 코딩 영역 전 및 후 영역 (리더 및 트레일러) 뿐만 아니라 개별 코딩 절편 (엑손) 사이의 삽입 서열 (인트론)을 포함할 수 있다.
- [0031] 본원에 사용된 용어 "반응"은 관심있는 특정 폴리뉴클레오티드 서열의 존재 또는 부재를 나타내는 화학적, 효소적 또는 물리적 작용에 관여하는 임의의 과정을 지칭한다. "반응"의 한 예는 중합효소 연쇄 반응 (PCR)과 같은 증폭 반응이다. "반응"의 또다른 예는 합성 또는 라이게이션에 의한 서열분석 반응이다. "정보제공 반응"은 관심있는 하나 이상의 특정 폴리뉴클레오티드 서열의 존재를 나타내며 단지 하나의 관심 서열이 존재하는 한가지 경우에서의 반응이다. 본원에 사용된 용어 "웰"은 제한된 구조, 예를 들어 웰-형상 바이알, 셀, 또는 PCR 어레이 챔버 내의 예정된 위치에서의 반응을 지칭한다.
- [0032] 본원에 사용된 용어 "임상적으로 관련된 핵산 서열"은 잠재적 불균형에 대해 시험될 보다 큰 게놈 서열의 절편에 상응하는 폴리뉴클레오티드 서열 또는 보다 큰 게놈 서열 그 자체를 지칭할 수 있다. 한 예는 염색체 21의 서열이다. 다른 예는 염색체 18, 13, X 및 Y를 포함한다. 또다른 예는 돌연변이 유전자 서열 또는 유전자 다형성 또는 부모의 한쪽 또는 양쪽으로부터 태아에게 유전될 수 있는 카피 수 변형을 포함한다. 또다른 예는 악성 종양에서 돌연변이, 결실 또는 증폭된 서열, 예를 들어 이형접합성 소실 또는 유전자 복제가 일어나는 서열을 포함한다. 일부 실시양태에서, 다수의 임상적으로 관련된 핵산 서열 또는 동등하게는 임상적으로 관련된 핵산 서열의 다수의 마커는 불균형 검출을 위한 데이터를 제공하는데 사용될 수 있다. 예를 들어, 염색체 21 상의 5개의 비-연속 서열로부터의 데이터는 가능한 염색체 21 불균형의 결정을 위해 추가적인 방식으로 사용되어, 필요한 샘플 부피를 1/5로 유효하게 감소시킬 수 있다.
- [0033] 본원에 사용된 용어 "배경 핵산 서열"은 임상적으로 관련된 핵산 서열에 대한 정상 비율이 공지되어 있는 (예를 들어 1 대 1 비율인) 핵산 서열을 지칭한다. 한 예로서, 배경 핵산 서열 및 임상적으로 관련된 핵산 서열은 이형접합성으로 인해 구별되는, 동일한 염색체로부터의 2개의 대립유전자이다. 또다른 예로, 배경 핵산 서열은 임상적으로 관련된 핵산 서열인 다른 대립유전자에 대해 이형접합성인 한 대립유전자이다. 또한, 각각의 배경 핵산 서열 및 임상적으로 관련된 핵산 서열 중 일부는 상이한 개체로부터 유래될 수 있다.
- [0034] 본원에 사용된 용어 "참조 핵산 서열"은 반응 당 평균 농도가 공지되어 있거나 동등하게는 측정되는 핵산 서열을 지칭한다.
- [0035] 본원에 사용된 용어 "과대표출된(overrepresented) 핵산 서열"은 2개의 관심 서열 (예를 들어, 임상적으로 관련된 서열 및 배경 서열) 중 생물학적 샘플에서 다른 서열보다 풍부한 핵산 서열을 지칭한다.
- [0036] 본원에 사용된 용어 "~기준으로"는 "적어도 부분적으로 ~기준으로"를 의미하며, 방법의 투입과 방법의 산출의 관계에서 일어나는, 한 값 (또는 결과)이 또다른 값의 결정에 사용되는 것을 지칭한다. 본원에 사용된 용어 "



유도"는 그 유도가 식의 계산인 경우에 일어나는, 방법의 투입과 방법의 산출의 관계를 지칭한다.

- [0037] 본원에 사용된 용어 "정량 데이터"는 하나 이상의 수치 값을 제공하는 하나 이상의 반응으로부터 얻은 데이터를 의미한다. 예를 들어, 특정 서열에 대한 형광 마커를 나타내는 웰의 수는 정량 데이터일 것이다.
- [0038] 본원에 사용된 용어 "파라미터"는 정량 데이터 세트, 및/또는 정량 데이터 세트들 사이의 수치 관계를 특성화하는 수치 값을 의미한다. 예를 들어, 제1 핵산 서열의 제1 양과 제2 핵산 서열의 제2 양 사이의 비율 (또는 비율의 함수)이 파라미터이다.
- [0039] 본원에 사용된 용어 "분별값"은 생물학적 샘플에 대한 분류의 2가지 이상의 상태 (예를 들어 질환 상태 및 비질환 상태)를 중재하는데 사용되는 수치 값을 의미한다. 예를 들어, 파라미터가 분별값보다 큰 경우 정량 데이터의 제1 분류가 만들어지거나 (예를 들어 질환 상태), 또는 파라미터가 분별값보다 작은 경우 정량 데이터의 상이한 분류가 만들어진다 (예를 들어 비질환 상태).
- [0040] 본원에 사용된 용어 "불균형"은 하나 이상의 분별값으로 정의되는, 참조 정량으로부터 임상적으로 관련된 핵산 서열의 정량의 임의의 유의한 편차를 의미한다. 예를 들어, 참조 정량이 3/5의 비율일 수 있고, 따라서 측정 비율이 1:1인 경우에 불균형이 일어날 것이다.
- [0042] 발명의 상세한 설명
- [0043] 본 발명은, 다른 비-임상적으로 관련된 서열 (예를 들어, 염색체 또는 대립유전자 불균형)에 관하여 임상적으로 관련된 핵산 서열의 참조 (예를 들어 비질환 상태) 양에 비교되는 증가 또는 감소가 생물학적 샘플 내에 존재하는지 여부를 결정하기 위한 방법, 시스템 및 장치를 제공한다. 예를 들어 2개의 서열 (또는 서열 세트) 양의 비율과 관련하여 참조 양에 비교되는 변화 (즉 불균형)이 존재하는지 여부를 결정하기 위해 하나 이상의 분별값이 선택된다. 참조 양에서 검출된 변화는 다른 비-임상적 관련 서열에 대한 임상적 관련 핵산 서열에 관하여 임의의 편차 (상향 또는 하향)일 수 있다. 따라서, 참조 상태는 임의의 비율 또는 다른 양 (예를 들어 1-1 대응 이외의 것)일 수 있고, 변화를 의미하는 측정 상태는 하나 이상의 분별값에 의해 결정된 바와 같이 참조 양과 상이한 다른 양 또는 임의의 비율일 수 있다.
- [0044] 임상적으로 관련된 핵산 서열 및 배경 핵산 서열은 제1 유형 세포 및 하나 이상의 제2 유형 세포로부터 유래될 수 있다. 예를 들어, 태아/태반 세포로부터 유래된 태아 핵산 서열은 생물학적 샘플, 예컨대 모체 세포로부터 유래된 모체 핵산 서열의 배경을 함유하는 모체 혈장 중에 존재한다. 따라서, 한 실시양태에서, 분별값은 적어도 부분적으로, 생물학적 샘플 중 제1 유형 세포의 백분율을 기준으로 결정된다. 샘플 중 태아 서열의 백분율은 임의의 태아-유래 유전자좌에 의해 결정될 수 있으며, 임상적으로 관련된 핵산 서열을 측정하는 것에 국한되지 않는다는 것을 주목한다. 또다른 실시양태에서, 분별값은 적어도 부분적으로, 생물학적 샘플, 예컨대 체내 비-악성 세포로부터 유래된 핵산 서열의 배경을 함유하는 혈장, 혈청, 타액 또는 뇨 중 종양 서열의 백분율을 기준으로 결정된다.
- [0045] 또다른 실시양태에서, 분별값은 다수의 반응에서 서열의 평균 농도를 기준으로 결정된다. 한 측면에서, 분별값은 특정 핵산 서열을 함유하는 것으로 평가되는 정보제공 웰의 비율로부터 결정되며, 그 비율은 상기 언급된 백분율 및/또는 평균 농도를 기준으로 결정된다. 분별값은 다수의 상이한 유형의 방법, 예컨대 SPRT, 오류 발견, 신뢰 구간, 수신자 판단 특성 (ROC)을 사용하여 결정될 수 있다. 이러한 전략은 신뢰성 있는 분류가 만들어질 수 있기 전에 필요한 시험의 양을 더 최소화시킨다. 이는 주형 양이 종종 제한적인 혈장 핵산 분석과 특히 관련되어 있다. 디지털 PCR이 제시되더라도, 다른 방법이 사용될 수 있다.
- [0046] 디지털 PCR은 대부분의 양성 증폭이 단일 주형 분자로부터의 신호를 반영하도록 매우 희석된 핵산에 대한 멀티플렉스 PCR 분석을 포함한다. 이로써 디지털 PCR은 개별 주형 분자의 계수를 가능하게 한다. 분석된 PCR의 총 수 중 양성 증폭의 비율로 원래 또는 비-희석 샘플 중의 주형 농도를 평가할 수 있다. 이러한 기술은 다양한 유전 현상의 검출을 가능하게 하도록 제시되었으며 (상기 문헌 [Vogelstein, B et al. 1999]), 종양 샘플 중 이형접합성 소실의 검출 (상기 문헌 [Zhou, W. et al. 2002]) 및 암 환자의 혈장 중 이형접합성 소실의 검출 (상기 문헌 [Chang, HW et al. 2002])을 위해 이전부터 사용되었다. 디지털 PCR에 의한 주형 분자 정량화는 리포터 염료와 핵산 농도 사이의 투입량-반응 관계에 의존하지 않기 때문에, 그의 분석 정확도는 실시간 PCR의 분석 정확도보다 이론적으로 월등하여야 한다. 따라서, 디지털 PCR은 표적 유전자좌와 참조 유전자좌 사이의 보다 미세한 정도의 정량 차이를 잠재적으로 구별할 수 있게 한다.
- [0047] 이를 시험하기 위해, 본 발명자들은 먼저 디지털 PCR이 모체 혈장 중 염색체 21로부터의 태반 전사체인 PLAC4 mRNA (문헌 [Lo, YMD, et al. 2007 Nat Med 13, 218-223]) 대립유전자 비율을 결정하고, 이로써 삼염색체 21

및 정배수체 태아를 구별할 수 있는지 여부를 평가하였다. 이러한 접근법을 디지털 RNA-SNP 방법이라 칭한다. 이어서, 본 발명자들은 디지털 PCR의 증가된 정확도로 인해 유전자 다형성에 좌우되지 않으면서 태아 염색체 이수성을 검출할 수 있는지 여부를 평가하였다. 본 발명자들은 이를 디지털 상대적 염색체 투여 (RCD) 분석이라 칭한다. 전자의 접근법은 다형성-의존적이지만 정량 구별에 있어서 보다 낮은 정확도를 필요로 하는 반면, 후자의 접근법은 다형성-독립적이지만 정량 구별에 있어서 보다 높은 정확도를 필요로 한다.

[0049] I. 디지털 RNA-SNP

[0050] A. 개관

[0051] 디지털 PCR은 DNA 샘플 중 두 대립유전자의 벗어난(skewing) 대립유전자 비율의 존재를 검출할 수 있다. 예를 들어, 디지털 PCR은 종양 DNA 샘플 중 이형접합성 소실 (LOH)을 검출하는데 사용되어 왔다. DNA 샘플 중 두 대립유전자, 즉 A 및 G가 존재하고, A 대립유전자는 LOH가 있는 세포 중에 소실된다고 가정한다. LOH가 종양 샘플 중 50%의 세포에 존재하는 경우, DNA 샘플 중 G:A의 대립유전자 비율은 2:1일 것이다. 그러나, LOH가 종양 샘플에 존재하지 않는 경우, G:A의 대립유전자 비율은 1:1일 것이다.

[0052] 도 1은 디지털 PCR 실험을 설명하는 흐름도 100이다. 단계 110에서, DNA 샘플을 회석시킨 후, 독립된 웰로 분배한다. 본 발명자들은 일부 혈장 핵산 종이 원래의 샘플에서 이미 상당히 회석되도록 결정했다는 것을 주목한다. 따라서, 필요한 농도로 이미 존재하는 경우, 일부 주형에 대한 회석은 필요하지 않다. 이전의 연구 (예를 들어, 상기 문헌 [Zhou et al 2002])에서는, 특정 "주형 DNA"의 평균 농도가 웰 당 두 주형 중 하나의 대략 0.5 분자가 될 정도로 DNA 샘플을 회석시켰다. 용어 "주형 DNA"는 A 또는 G 대립유전자를 가리키는 것으로 보이고, 이러한 특정 농도에 대한 이론적 근거는 제공되지 않았다는 것을 주목한다.

[0053] 단계 120에서, 각각의 웰에서, PCR 과정은 A 및/또는 G 대립유전자를 동시에 검출하기 위해서 실시된다. 단계 130에서, 각각의 웰 중 마커가 식별된다 (예를 들어, 형광으로) (예를 들어 A, G, A 및 G, 또는 어느 쪽도 아님). LOH의 부재 하에, DNA 샘플 중 A 및 G 대립유전자의 과다는 동일하다 (셀 당 각각 한 카피). 따라서, 웰이 A 대립유전자 및 G 대립유전자에 대해 양성일 확률은 동일할 것이다. 이것은 유사한 수의 웰이 A 또는 G 대립유전자에 대해 양성인 것에 의해 반영될 것이다. 그러나, LOH가 종양 샘플 중 50% 또는 그 보다 많은 세포에 존재하는 경우, G 및 A 대립유전자의 대립유전자 비율은 적어도 2:1일 것이다. 이전의 방법들은 단순히 샘플이 적어도 50% 양성인 것으로 추정하였다. 따라서, 웰이 G 대립유전자에 대해 양성일 확률은 A 대립유전자에 대해 양성일 확률보다 높을 것이다. 그 결과, G 대립유전자에 대해 양성인 웰의 수는 A 대립유전자에 대해 양성인 웰의 수보다 더 많을 것이다.

[0054] 단계 140에서, 디지털 PCR 결과를 분류하기 위해, 각각의 대립유전자에 대해 양성이나 다른 대립유전자에 대해서는 양성인 웰의 수를 계수할 것이다. 상기 예에서, A 대립유전자에 대해 양성이나 G 대립유전자에 대해 음성인 웰의 수, 및 G 대립유전자에 대해 양성이나 A 대립유전자에 대해 음성인 웰의 수를 계수한다. 한 실시양태에서, 덜 양성인 웰을 나타내는 대립유전자는 참조 대립유전자로 간주된다.

[0055] 단계 150에서, 정보제공 웰의 총 수는 두 대립유전자 중 어느 하나에 대해 양성인 웰의 수의 합으로 결정된다. 단계 160에서, 더 많은 양성 웰의 대립유전자에 의해 기여되는 정보제공 웰의 비율 ( $P_r$ ) (파라미터의 한 예)이 계산된다.

[0056]  $P_r$  = 더 많은 양성 웰의 대립유전자에 대해서만 양성인 웰의 수/단지 하나의 대립유전자 (A 또는 G)에 대해 양성인 웰의 총 수.

[0057] 다른 실시양태는 하나의 대립유전자를 갖는 모든 웰을 하나 이상의 대립유전자를 갖는 모든 웰로 나눈 것을 사용할 수 있다.

[0058] 단계 170에서,  $P_r$  값이 대립유전자 불균형을 나타내는지 여부가 결정된다. 정확하고 효율적인 것이 바람직하므로, 이 작업은 간단하지 않다. 불균형을 결정하기 위한 한 방법은 베이지안(Bayesian)-유형 우도 방법, 순차 확률비 검증 (SPRT)을 사용한다. SPRT는 데이터가 축적되면서 두 확률 가설을 비교하도록 하는 방법이다. 즉, SPRT는 디지털 PCR의 결과를 분류하여 벗어난 대립유전자의 존재 또는 부재를 시사하는 통계적 방법이다. SPRT는 분석될 웰의 수를 최소화하여 소정의 통계적 검증력 및 정확도를 획득하도록 하는 이점을 갖는다.

[0059] 예시적 SPRT 분석에서, 실험 결과는 귀무가설 (null hypothesis) 및 대립가설 (alternative hypothesis)에 대해 시험될 것이다. 대립가설은 샘플 중 벗어난 대립유전자 비율이 있는 경우 수용된다. 귀무가설은 샘플 중

벗어난 대립유전자 비율이 없는 경우 수용된다.  $P_r$  값은 귀무가설 또는 대립가설을 수용하기 위해서 두 분별값과 비교될 것이다. 두 가설 모두 수용되지 않는 경우, 샘플은 관찰된 디지털 PCR 결과가 바람직한 통계적 신뢰도로 샘플을 분류하기에 충분하지 않다는 것을 의미하는 미분류로 표시될 것이다.

[0060] 귀무가설 또는 대립가설을 수용하기 위한 분별값은 전형적으로, 가설에서 언급된 가정 하에 고정된  $P_r$  값을 기준으로 계산되어 왔다. 귀무가설에서, 샘플은 벗어난 대립유전자 비율을 나타내지 않는 것으로 가정된다. 따라서, 각 웰이 A 및 G 대립유전자에 대해 양성일 확률은 동일할 것이므로, 예상  $P_r$  값은 1/2일 것이다. 대립가설에서, 예상  $P_r$  값은 2/3 또는 0.5와 2/3의 대략 중간 정도, 예를 들어 0.585로 여겨진다. 또한, 제한된 수의 실험으로 인해,  $(.585+3/N)$ 의 상한값 및  $(.585-3/N)$ 의 하한값이 선택될 수 있다.

[0062] **B. 다운 증후군의 검출**

[0063] 본 발명의 한 실시양태에서, 디지털 SNP는 임신 여성의 혈장으로부터 태아 다운 증후군을 검출하는데 사용된다. 태아/태반 세포에 특이적인 마커를 사용하여, 염색체 21 중 대립유전자의 비율을 측정할 수 있다. 예를 들어, PLAC4 대립유전자의 관찰된 과다표출 정도가 통계적으로 유의한지를 결정하기 위해서, SPRT가 사용된다.

[0064] 하나의 예시적 실시양태에 따라, 디지털 RNA-SNP는 염색체 21로부터 전사되고 태반에 의해 발현되는 PLAC4 mRNA에 위치한 A/G SNP의 다형성 대립유전자, rs8130833의 비율에서의 불균형을 결정한다. 이형접합 정배수체 태아에 대해, A 및 G 대립유전자는 태아 계놈에서 동등하게 표출되어야 하는 한편 (1:1 계놈 비율); 삼염색체 21에서, 삼염색체성 염색체 21은 태아 계놈 중 하나의 SNP 대립유전자의 추가 카피와 회합되어 2:1 비율을 나타낼 것이다. 디지털 PCR 분석의 목적은 분석된 샘플 중 두 PLAC4 대립유전자의 양이 동등한지 또는 다른지를 결정하는 것이다. 따라서, A 및 G PLAC4 대립유전자는 모두 표적 주형이다. 실시간 PCR 검정은 PLAC4 mRNA를 증폭시키도록 설계되었고, 두 SNP 대립유전자는 테크만(TaqMan) 형광 프로브에 의해 구별되었다. 분석 단계의 도식적 설명이 도 2a에 나타나 있다.

[0065] 도 2a는 본 발명의 실시양태에 따른 디지털 RNA-SNP 방법 200을 설명한다. 단계 210에서, 샘플을 얻는다. 단계 220에서, 핵산 서열, 예를 들어 PLAC4 mRNA를 추출된 RNA 샘플에서 정량화한다. 한 실시양태에서, 이것은 PLAC4 mRNA에 대해 실시간 PCR로 실시된다. 한 측면에서, 이 단계는 조작자에게 표적이 디지털 PCR 분석의 '영역'에 도달하기 전에 어느 만큼의 회석이 요구되는지에 대한 아이디어를 제공한다.

[0066] 단계 230에서, 샘플을 회석시킨다. 단계 240에서, 회석시킨 샘플의 농도를 측정한다. 회석시킨 샘플의 농도가 약 1 주형/웰 (즉, 참조 또는 비-참조 서열 또는 어느 하나의 대립유전자)인지 확인할 수 있다. 일부 실시양태는 이 측정을 위한 IV절에 기술된 기술을 사용한다. 예를 들어, 본 발명자들은 실시간 PCR 분석을 위해서 회석시킨 샘플을 96 웰로 배분하여 알맞게 회석되었는지를 확인하였다. 회석 농도는 또한 모르는 상태로 남겨두어 이 단계가 제거되도록 할 수 있고, 이는 추후 설명될 것이다.

[0067] 단계 250에서, 디지털 PCR을 어레이의 각 웰 상에서 실시하였다. 예를 들어, 동일하게 회석시킨 샘플을 실시간 PCR 분석을 위해서 384-웰에 배분하였다. PCR 결과로부터, 각 핵산 서열에 대한 마커의 양 및 정보제공 웰의 수를 확인한다. 정보제공 웰은 A 또는 G 대립유전자에 대해서만 양성이나, 둘 모두에 대해서 양성은 아닌 것으로 정의된다. 단계 260에서, 예상  $P_r$  값을 계산한다. 이러한 단계들은 추후 더욱 자세하게 설명될 것이다. 이 계산은 단계 250에서 측정된 값들로부터 파라미터를 결정하는 것을 포함한다. 예를 들어, 웰 당 실제 평균 주형 농도가 계산될 수 있다.

[0068] 단계 270에서, SPRT 또는 다른 우도-비율 시험은 불균형이 존재하는지 여부를 결정하기 위해서 실시될 수 있다. 정배수체 사례에 대해, 동일한 수의 A-양성 및 G-양성 웰이 예상된다. 그러나, 삼염색체 21 태아로부터의 주형 분자가 분석되는 경우, 하나의 대립유전자만을 함유하는 웰의 수는 다른 대립유전자만을 함유하는 웰의 수보다 많아야 한다. 요컨대, 삼염색체 21에 대한 대립유전자 불균형이 예상된다.

[0069] 상기 언급된 것과 같이, SPRT는 데이터가 축적되면서 두 확률 가설을 비교하도록 하는 베이지안-유형 우도 방법이다. 삼염색체 21 검출을 위한 디지털 PCR 분석에서, 대립가설은 대립유전자 불균형이 존재하는 경우 (즉, 삼염색체 21이 검출되는 경우) 수용되고; 귀무가설은 대립유전자 불균형이 존재하지 않는 경우 (즉, 삼염색체 21이 검출되지 않는 경우) 수용된다. 더 많은 수의 계수의 대립유전자는 잠재적으로 과다표출된 대립유전자로 지칭되고, 모든 정보제공 웰 중 그의 비율 ( $P_r$ )이 계산될 것이다. SPRT는,  $P_r$ 이 삼염색체 21 샘플에 대해 예상되는 충분한 정도의 대립유전자 불균형을 나타내는지를 결정하기 위해서 적용된다.

[0070] 조작 면에 있어서, SPRT는 어느 하나의 가설을 수용하거나 또는 거절하기 위한 확률 경계를 정의하기 위해서 구축된 한 쌍의 SPRT 곡선을 갖는 그래프를 사용하여 적용되고 해석될 수 있다. 도 3은 본 발명의 한 실시양태에 따라 다운 증후군을 결정하기 위한 SPRT 곡선을 갖는 그래프를 설명한다. SPRT 곡선은 신뢰성 있게 분류될 수 있는 경우 주어진 정보제공 웰의 총 수 ( $x$ -축)에 대해 잠재적으로 과다표출된 대립유전자에 대해 양성인 정보제공 웰의 요구되는 비율  $P_r$  ( $y$ -축)를 플롯팅한다. 도 3에 설명된 것과 같이, 상단 곡선은 대립가설을 수용하기 위한 확률 경계를 설정하고; 하단 곡선은 귀무가설을 수용하기 위한 확률 경계를 설정한다.

[0071] 실험적으로 유도된  $P_r$  값은 어느 하나의 가설을 수용하거나 또는 거절하기 위한 예상  $P_r$  값과 비교될 것이다. 귀무가설이 수용되는 경우, 샘플은 정배수체 태아를 가진 임신 여성으로부터 수득된 것으로 분류되었다. 대립가설이 수용되는 경우, 샘플은 삼염색체 21 태아를 가진 임신 여성으로부터 수득된 것으로 분류되었다. 별법으로, 주어진 정보제공 계수의 수에 대한  $P_r$ 이 질환 분류에 대해 요구되는 통계적 신뢰성 수준에 아직 도달하지 못한 경우, 어느 하나의 가설도 받아들여질 수 없다. 이러한 경우들은 더 많은 데이터가 이용가능하기 전까지는 분류불가능으로 간주된다. 질환 분류가 가능하지 않은 경우, 집계된 데이터가 SPRT에 의해 분류가능해지기 전까지 추가의 384-웰 플레이트를 실시할 수 있다.

[0072] 따라서, SPRT는 주어진 수준의 신뢰성을 위해서 다른 통계적 방법보다 더 적은 양의 시험이 요구된다는 이점을 제공한다. 실용적인 측면에서, SPRT는 요구되는 양의 데이터가 축적되어 불필요한 추가적 분석을 최소화하자마자 어느 하나의 가설의 수용 또는 거절을 가능하게 한다. 이 특징은 일반적으로 이용가능한 주형 분자의 수가 제한적인 경우 낮은 농도로 존재하는 혈장 핵산의 분석과 특히 관련이 있다. 엄격한 분류에 추가하여, 분류는 또한 백분율 정확성을 포함할 수 있다. 예를 들어, 분별값과의 비교로 인한 분류는 샘플이 특정 백분율의 핵산 서열 불균형의 우도를 나타내거나, 또는 동등하게는 측정된 불균형이 특정 백분율 또는 다른 값에 정확하다는 것을 제공할 수 있다.

[0073] 유사한 접근법을 적용하여, 모체 혈장 또는 혈청 중 태아 핵산을 사용하여 돌연변이 또는 유전자 다형성과 관련한 태아의 유전자형을 결정할 수 있다. 태아는 그의 모체로부터 그의 계놈의 절반을 물려받을 것이라는 것을 상기해야 한다. 설명을 위해서, 두 대립유전자 A 및 B가 있는 특정 유전자좌를 고려한다. 모체가 AB 유전자형의 이형접합체인 경우, 태아는 이론적으로는 AA, BB 또는 AB의 유전자형을 가질 수 있다. 태아가 AB 유전자형, 즉 모체와 같은 유전자형을 갖는 경우, 모체 혈장에는 (모체 및 태아 모두로부터의) AB 유전자형의 핵산만 존재할 것이다. 따라서, 핵산 또는 대립유전자 균형이 모체 혈장에서 보여진다. 반면, 태아가 AA 또는 BB의 유전자형을 갖는 경우, 모체 혈장에서 A 또는 B 대립유전자 각각이 과다표출된 대립유전자 불균형이 존재할 것이다. 이러한 생각은 또한 질환-유발성 돌연변이 (예를 들어, 낭포성 섬유증 또는 베타-지중해빈혈 또는 척수 근위축증을 유발하는 돌연변이)에도 적용가능한데, 각 경우 A는 야생형 대립유전자로, B는 돌연변이 대립유전자로 고려될 수 있다.

[0075] II. 디지털 RCD

[0076] 디지털 RNA-SNP의 단점은 디지털 RNA-SNP가 분석된 SNP에 대한 이형접합체의 경우에만 적용가능하다는 것이다. 순환성 태아 핵산 분석을 기준으로 한 삼염색체 21 태아 또는 다른 태아 염색체 이수성 (예를 들어, 삼염색체 18, 13 및 성 염색체 이수성)을 검출하기 위한 비침습성 시험이 유전자 다형성의 사용과 무관한 경우 디지털 RNA-SNP가 이상적이라는 것은 하나의 개선된 점이다. 따라서, 한 실시양태에서, 염색체 양은 참조 염색체, 즉 이 연구에서 염색체 1에 위치한 유전자좌에 상대적인 비-다형성 염색체 21 유전자좌의 디지털 PCR 분석에 의해 결정된다. 정배수체 태아의 계놈에서 염색체 21 대 염색체 1의 비율의 2:2로부터의 변화는 삼염색체 21의 경우와 구분된다. 삼염색체 21 검출을 위한 디지털 PCR 분석에서, 비교할 두 가설은 염색체 불균형이 존재하지 않는 (즉, 삼염색체 21이 검출되지 않는) 귀무가설 및 염색체 불균형이 존재하는 (즉, 삼염색체 21이 검출되는) 대립가설일 것이다.

[0077] 이 접근법은 다른 염색체 이수성, 예를 들어 삼염색체 18에서의 염색체 18, 삼염색체 13에서의 염색체 13, 터너 증후군에서의 염색체 X와 연관된 다른 염색체에 일반화될 수 있다. 또한, 염색체 1과는 별개로, 관련 이수성과 연관되지 않은 다른 염색체도 또한 참조 염색체로 사용될 수 있다. 유사한 접근법이 또한 부분적으로, 암에서 통상적으로 결실되는 염색체 대 참조 염색체의 비율 변화를 분석함으로써 암 검출에 적용될 수 있다. 전자의 예로는 결장직장암에서의 염색체 5q, 폐암에서의 염색체 3p 및 비인두 암종에서의 염색체 9p가 포함된다. 도 2b는 서열 불균형으로 나타나는 일부 통상적인 암-관련 염색체 이상을 열거한다.

[0078] 도 2a는 또한 본 발명의 한 실시양태에 따른 디지털 RCD 방법 205를 설명한다. 단계 220-230에 대한 한 실시양

태에서, 추출된 DNA를, 예를 들어, 나노드롭(Nanodrop) 기술로 정량화하고, 웰 당 염색체 21 또는 표준화 염색체 (예컨대, 염색체 1)로부터의 약 하나의 표적 주형의 농도로 희석시킨다. 단계 240의 한 실시양태에서, 384-웰 플레이트 중 테크만 프로브를 모두 사용하는 디지털 RCD 분석을 진행하기 전에 약 37% 수준의 웰이 음성인지 여부를 확인하기 위해서, 단지 96-웰 포맷에서 염색체 1 프로브를 사용하는 검정에 의해 희석시킨 DNA 샘플을 분석함으로써 확인할 수 있다. 37%의 유의성은 IV절에서 추후 기술될 것이다.

[0079] 단계 240의 시험 및 단계 250의 결과는 한 쌍의 테크만 프로브에 의해 구별되는 파라로그(paralogous) 서열 변이에 의해 식별되는 두 염색체 모두에 존재하는 파라로그 서열을 증폭시키도록 고안된 실시간 PCR 검정으로 수행될 수 있다 (문헌 [Deutsch, S. et al. 2004 J Med Genet 41, 908-915]). 이러한 맥락에서, 정보제공 웰은 염색체 21 또는 염색체 1 유전자좌에 대해 양성인, 그러나 둘 모두에 대해서 양성은 아닌 웰로서 정의된다. 정배수체 태아에 대해, 어느 하나의 유전자좌에 대해 양성인 정보제공 웰의 수는 대략 동등해야 한다. 삼염색체 21 태아에 대해, 염색체 1보다 염색체 21에 대해 양성인 웰의 과다표출이 있어야 한다. 과다표출의 정확한 비율은 하기 절에 기술되어 있다.

[0081] III. 태아 서열의 혼입률

[0082] 상기 기술된 방법 200 및 205의 실시양태의 단점은 태아 특이적 마커가 요구된다는 것이다. 따라서, 본 발명의 한 실시양태에서, 비-태아 특이적 마커가 사용된다. 비-태아 특이적 마커를 사용하기 위해서, 본 발명의 실시양태는 모체 혈장 (즉, 생물학적 샘플) 중 태아 DNA 분획 농도를 측정한다. 이러한 정보와 함께, 보다 유용한  $P_r$  값이 하기와 같이 계산될 수 있다.

[0083] 모체 혈장 중 태아 DNA의 낮은 분획 백분율로도, 삼염색체 21 태아는 모체 혈장으로 방출된 태아 DNA의 게놈-당 가물 (GE) 당 염색체 21 서열의 추가적 양에 기여할 것이다. 예를 들어, 태아에 의해 기여된 5 GE/mL DNA와 50 GE/mL의 총 DNA (즉, 10% 태아 DNA 분획 농도)를 함유하는 정배수체 임신으로부터의 모체 혈장 샘플은 모체 혈장 (밀리리터) 당 염색체 21 서열의 총 100개의 카피 (90개의 모체 카피 + 10개의 태아 카피)를 함유해야 한다. 삼염색체 21 임신에 대해, 태아 GE는 각각 3개 카피의 염색체 21에 기여하여, 모체 혈장 중 염색체 21 서열의 총 105개의 카피/mL (90개의 모체 카피 + 15개의 태아 카피)가 되도록 한다. 따라서, 10% 태아 DNA 농도에서, 삼염색체 임신의 모체 혈장 중 염색체 21 유래 서열의 양은 정배수체 경우의 1.05배일 것이다. 따라서, 분석적 접근법이 이러한 적은 정도의 정량 차이를 결정하도록 개발될 수 있다면, 삼염색체 21 태아의 비침습성 출생 진단에 대한 다형성-독립적 시험이 가능할 것이다.

[0084] 따라서, 과다표출의 정도는 분석된 DNA 샘플 중 태아 DNA 분획 농도에 좌우될 것이다. 예를 들어, 태반 DNA를 분석하는 경우, 태아 게놈 중 이론적 RCD 비율은 3:2, 즉 1.5배 차이이어야 한다. 그러나, 상기 기술된 것과 같이, 이론적 RCD 비율은 10% 태아 DNA를 함유하는 모체 혈장 샘플을 분석하는 경우 1.05로 감소할 것이다. 실험적으로 유도된  $P_r$ 은 단지 염색체 21 유전자좌에 양성인 웰의 수를 총 정보제공 웰의 수로 나누어 계산된다. 실험적으로 유도된  $P_r$ 은 계산된  $P_r$  및 이론적 RCD 비율과 함께 SPRT 분석에 적용된다.

[0085] 도 4는 본 발명의 실시양태에 따라 태아 핵산의 백분율을 사용하여 질환 상태를 결정하는 방법 400을 나타낸다. 단계 410에서, 태아 물질의 분획 백분율이 측정된다. 한 실시양태에서, 분획 백분율은 비-태아-특이적 마커 (즉, 모체 및 태아 모두에 존재하는 유전자 서열)와 관련하여 태아-특이적 마커 (예를 들어, Y-염색체, 유전자 다형성 마커 (예컨대, SNP), 태반 후성 기호 (placental epigenetic signatures))의 양을 측정하여 결정된다. 실시간 PCR, 디지털 PCR, 서열분석 반응 (대규모 병렬 게놈 서열분석을 포함) 또는 임의의 다른 정량 방법에 의해 실질적 측정을 수행할 수 있다. 한 측면에서, 이 측정에 대해 잠재적으로 대립유전자 불균형이 있을 수 있는 유전자 표적을 사용하지 않는 것이 바람직하다.

[0086] 단계 420에서, 샘플을 희석시키고, 희석시킨 샘플을 웰에 넣고, 각 웰에서 반응을 측정하는 것을 비롯한 디지털 PCR 또는 다른 측정 방법이 실시된다. 단계 430에서, PCR 결과는 상이한 참조 핵산 서열 (예컨대, 염색체 또는 대립유전자)의 마커를 식별하기 위해서 사용된다. 단계 440에서, 과다표출된 서열의 실제 비율 ( $P_r$ )이 계산된다. 단계 450에서, 질환 상태를 결정하기 위한 분별값은 샘플 중 태아 물질의 백분율을 사용하여 계산된다. 단계 460에서, 실제  $P_r$  및 분별값으로부터, 불균형이 존재하는지 여부가 결정된다.

[0087] 한 실시양태에서, 참조 핵산 서열의 분획 백분율은 디지털 RNA-SNP 방법에 혼입된다. 따라서, 암 세포로 인한 LOH를 연구하는 경우, 50% 미만 암 세포를 갖는 종양 샘플로 실시할 수 있다. 이것은 또한 50% 초과 암 세포를 갖는 샘플에 대해서도 사용되어, 더욱 정확한  $P_r$ 을 획득함으로써 부정확한 진단으로 이끌 수 있는 잘못된 양

성의 수를 감소시킬 수 있다. 또다른 실시양태에서, 태아 핵산 백분율은 디지털 PCR 방법에 혼입되어, 태아가 모체 혈장 핵산 분석으로부터 부모의 유전자 돌연변이 (예를 들어, 낭포성 섬유증 또는 베타-지중해빈혈 또는 척수 근위축증을 유발하는 돌연변이) 또는 다형성을 물려받는지 여부를 결정한다.

[0089] IV. 웰 당 평균 농도의 혼입

[0090] 이전 방법 (예를 들어, 상기 문헌 [Zhou, W. et al. 2002])의 또다른 단점은 웰 당 주형의 평균 농도 (m)가 웰 당 1로 요구된다는 것이다. 정확한 농도를 얻는 것이 어렵다는 것을 감안할 때, 이것은 부정확해질 수 있다. 또한, 웰 당 1 주형의 정확한 농도로도, 이전 방법은 웰 중 주형의 통계적 분포를 간과하였다. 이전 방법에서, 즉, 기존 알고리즘에서, 대립가설을 수용하기 위한 예상  $P_r$  값은 대립유전자 비율로 가정되어, 웰 당 주형 DNA 의 평균 농도와 무관하였다.

[0091] 그러나, 희석시킨 샘플에서 주형의 자연적인 통계적 편차로 인해, 정확하게 웰 당 1 주형은 아닐 것이다. 본 발명의 실시양태는 하나 이상의 서열의 평균 농도를 측정하여, 분별값, 즉 예상  $P_r$  을 계산하는데 사용된다. 한 측면에서, 이러한 계산은 상이한 핵산 서열을 함유하는 웰의 확률을 측정하여 예상  $P_r$  을 결정하기 위해서 사용되는 통계적 분포를 포함한다.

[0092] 한 실시양태에서, 평균 농도는 일례로 DNA 샘플 중 더 낮은 농도의 핵산 서열인 한 참조 핵산 서열로 고려된다. 불균형이 없는 샘플의 경우, 샘플 중 두 서열의 농도는 동일할 것이고, 어느 하나는 참조 대립유전자로 간주될 수 있다. 예를 들어 LOH가 있는 샘플의 경우, 암 세포에서 결실된 대립유전자는 참조 대립유전자로 간주될 것이다. 참조 대립유전자의 평균 농도는  $m_r$  로 표시될 것이다. 또다른 실시양태에서, 더 높은 농도의 서열이 참조 서열로 간주될 수 있다.

[0094] **A. 디지털-SNP, SPRT 및 디지털 PCR을 사용한 예**

[0095] 도 5는 본 발명의 한 실시양태에 따라 평균 주형 농도를 사용하여 질환 상태를 결정하는 방법 500을 보여준다. 단계 510에서, 상이한 서열의 양을 측정한다. 이것은, 예를 들어 상기 설명된 것과 같은 디지털 PCR 실험에서 마커를 계수하여 실시할 수 있다. 그러나, 이것은 증폭 단계를 포함하지 않거나 또는 형광 마커를 사용하지 않지만 다른 성질, 예컨대 질량, 특정 광학적 성질 또는 염기-쌍 성질과 같은 물성을 사용할 수 있는 다른 방법에 의해서도 실시할 수 있다.

[0096] 단계 520에서, 과대표출된 서열의 실제 비율이 결정된다. 이것은 상기 기술된 것과 같이 단지 그 서열만 나타내는 웰의 수를 취하고, 정보제공 웰의 수로 나누어 실시될 수 있다. 단계 530에서, 하나 이상의 서열의 평균 농도가 측정된다 (참조 서열). 한 실시양태에서, 참조 서열은 과대표출된 서열이다. 또다른 실시양태에서, 참조 서열은 과소대표출된(underrepresented) 서열이다. 측정은 디지털 PCR 실험에서 참조 서열에 대해 음성인 웰의 수를 계수하여 실시할 수 있다. 음성 웰의 비율과 평균 주형 농도 사이의 관계는 다음 소절에 기술된 것과 같은 푸아송 (Poisson) 분포에 의해 설명된다.

[0097] 단계 540에서, 상이한 서열에 대해 양성인 웰의 예상 양이, 예를 들어 푸아송 분포를 사용하여 계산된다. 예상 양은 웰 당 서열, 웰 당 평균 서열, 서열을 함유하는 웰의 수 또는 임의의 다른 적합한 양의 확률일 수 있다. 단계 550에서, 예상  $P_r$  은 예상 양으로부터 계산된다. 단계 560에서, 분별값은, 예를 들어 SPRT를 사용하여 예상  $P_r$  으로부터 계산된다. 단계 570에서, 핵산 서열 불균형의 분류가 결정된다. 방법 500의 특정 측면이 이제 기술된다.

[0099] 1. 서열 예상 양의 결정

[0100] 웰 (반응 또는 반응 혼합물) 당 평균 농도가 단계 530으로부터 알려지면, 서열을 나타내는 웰의 예상 수가 단계 540에서 계산될 수 있다. 이 양은 %, 분수 값, 또는 정수 값으로 표현될 수 있다. 설명을 위해 특정 예를 사용하면서, 웰 당 참조 주형의 평균 농도 ( $m_r$ )는 웰 당 0.5이고, PLAC4 SNP, rs8130833에서 삼염색체 21 태아의 유전자형은 AGG인 것으로 가정한다. 따라서, 참조 주형은 A 대립유전자일 것이고, 과대표출된 주형은 G 대립유전자일 것이다.

[0101] 한 실시양태에서, 푸아송 분포는 디지털 PCR과 같은 측정 과정의 웰의 반응 혼합물 중 A 대립유전자의 분포로 가정한다. 다른 실시양태에서는, 이항 분포와 같은 다른 분포 함수가 사용된다.

[0102] 푸아송 방정식은  $P(n) = \frac{m^n e^{-m}}{n!}$  이고, 여기서 n은 웰 당 주형 분자의 수이고; P(n)은 특정 웰 중 n 주형 분자의 확률이고; m은 특정 디지털 PCR 실험에서 한 웰 중 주형 분자의 평균 수이다.

[0103] 따라서, 0.5의 평균 A-대립유전자 농도에서 A 대립유전자의 임의의 분자를 함유하지 않는 임의의 웰의 확률은 하기와 같을 것이다:

[0104] 
$$P(0) = \frac{0.5^0 e^{-0.5}}{0!} = e^{-0.5} = 0.6065$$

[0105] 따라서, 하나 이상의 A 대립유전자의 분자를 함유하는 임의의 웰의 확률은:  $1 - 0.6065 = 0.3935$ 이다. 그러므로, 약 39%의 웰이 하나 이상의 A 대립유전자의 분자를 함유할 것으로 예상될 것이다.

[0106] 비-참조 핵산 서열의 경우, 삼염색체 21 태아의 각 세포에 대해서, A 대 G의 계놈 비율은 1:2일 것이다. 추출된 RNA 또는 DNA 샘플 중 A 대 G 비율이 변화하지 않는 것으로 가정하면, 웰 당 G 대립유전자의 평균 농도는 A 대립유전자의 평균 농도의 2배, 즉  $2 \times 0.5 = 1$ 일 것이다.

[0107] 따라서, 평균 G-대립유전자 농도인 1에서 G 대립유전자의 임의의 분자를 함유하지 않는 임의의 웰의 확률은 하기와 같을 것이다:

[0108] 
$$P(0) = \frac{1^0 e^{-1}}{0!} = e^{-1} = 0.3679$$

[0109] 따라서, 하나 이상의 G 대립유전자 분자를 함유하는 임의의 웰의 확률은:  $1 - 0.3679 = 0.6321$ 일 것이다. 그러므로, 약 63%의 웰이 하나 이상의 G 대립유전자 분자를 함유하는 것으로 예상될 것이다.

[0110] 2. 과다표출된 서열 비율의 결정

[0111] 예상 양이 계산된 후, 과다표출된 핵산 서열의 비율이 결정될 수 있다. A 대립유전자 및 G 대립유전자로 웰을 채우는 것이 독립적이라고 가정하면, 두 대립유전자를 모두 함유하는 웰의 확률은  $0.3935 \times 0.6321 = 0.2487$ 일 것이다. 그러므로, 약 25%의 웰이 두 대립유전자를 모두 함유하는 것으로 예상될 것이다.

[0112] A 대립유전자는 함유하되 G 대립유전자는 함유하지 않는 것으로 예상되는 웰의 비율은 하나 이상의 A 대립유전자를 함유하는 웰의 수에서 A 및 G 대립유전자를 모두 함유하는 웰의 수를 뺀 것이다:  $0.3935 - 0.2487 = 0.1448$ . 유사하게, G 대립유전자는 함유하되 A 대립유전자는 함유하지 않는 것으로 예상되는 웰의 비율은  $0.6321 - 0.2487 = 0.3834$ 일 것이다. 정보제공 웰은 A 대립유전자 또는 G 대립유전자에 대해 양성이나, 둘 모두에 대해서는 양성이 아닌 웰로서 정의된다.

[0113] 따라서, 디지털 RNA-SNP 분석에서 G 대립유전자에 상대적으로 A 대립유전자를 함유하는 웰의 예상 비율은  $0.1448/0.3834$ 이다. 즉, 단지 G 대립유전자에 대해서만 양성인 웰의 비율은 단지 A 대립유전자에 대해서만 양성인 웰의 2.65배이다. 이것은 과다표출된 대립유전자가 다른 대립유전자의 2배인 태아 계놈 비율과 대조적이다.

[0114] SPRT 분석에 대해, 과다표출된 대립유전자에 대해 양성인 정보제공 웰의 비율 ( $P_r$ )은 SPRT 곡선을 사용하여 계산되고 해석된다. 현 예에서, 정보제공 웰의 비율은  $0.1448 + 0.3834 = 0.5282$ 일 것이다. 따라서,  $m_r$  0.5에서 삼염색체 21 경우의 예상  $P_r$ 은  $0.3834/0.5282 = 0.73$ 이다.

[0115] 평균 주형 농도 (m)는 푸아송 방정식의 주요 파라미터이므로,  $P_r$ 은 m에 따라변할 것이다. 도 6은 본 발명의 실시양태에 따라 웰 당 평균 참조 주형 농도 ( $m_r$ )로 표현된 주형 농도의 범위에 대한 삼염색체 21 샘플의 예상 디지털 RNA-SNP 대립유전자 비율 및  $P_r$ 을 표로 만든 표 600을 나타낸다. 표 600은 일련의 웰 당 평균 참조 주형 농도 ( $m_r$ )에 대한 예상 대립유전자 비율 및 과다표출된 대립유전자에 대해 양성인 정보제공 웰의 비율 ( $P_r$ )을 나타낸다.

[0116] 예상  $P_r$  값은 비-선형 방식으로 웰 당 참조 대립유전자의 평균 농도 ( $m_r$ )에 따라 변한다. 표 600에 제시된 것과 같이, 대립가설을 수용하기 위한 예상  $P_r$  값은  $m_r$ 과 함께 증가할 것이다. 귀무가설을 수용하기 위한 예상  $P_r$  값이 0.5로 고정되어 있으므로, 대립유전자 불균형이 있는 샘플 및 대립유전자 불균형이 없는 샘플은  $m_r$ 이 증가할

경우,  $P_r$  값의 측면에서 보다 광범위하게 나뉘어질 것이다. 다른 실시양태에서는 귀무가설을 수용하기 위한 값이 0.5가 아닐 수 있다는 것을 주목한다. 이것은 정상 비율이 1:1과 다른 경우, 예를 들어, 5:3인 경우 일어날 수 있으며, 따라서 비율이 5:3으로부터 벗어나는 경우 불균형이 발생할 것이다. 두 상이한 핵산 서열 양의 차이는 개별 경우에 따라 결정될 것이다.

[0117] 그러나, 이전 방법 (예를 들어, 상기 문헌 [Zhou, W. et al. 2002])은 LOH 샘플에 대해 고정된 예상  $P_r$  값을 사용하였으므로, LOH가 있는 샘플에 대한  $P_r$  값을 과소평가하였다 (대립가설 수용). 과소평가의 정도는  $m_r$ 과 함께 증가할 것이다. 즉, DNA 샘플 중 참조 대립유전자의 평균 농도가 더 높을수록, 이전 방법은 더욱 부정확해질 것이다. 대립가설을 수용하기 위한  $P_r$ 의 과소평가는 귀무가설 및 대립가설 모두를 수용하기 위한 분별값의 계산을 부정확하게 할 것이다.

[0119] 3. 예상  $P_r$ 을 기준으로 한 분별값의 계산

[0120] SPRT를 사용하는 실시양태에 대해, 문헌 [El Karoui et al. (2006)]으로부터 SPRT 곡선의 상한 및 하한을 계산하기 위한 방정식을 사용할 수 있다. 또한, 귀무가설 또는 대립가설을 수용하기 위해서 바람직한 통계적 신뢰성의 수준은 방정식에서 임계값 우도 비율(threshold likelihood ratio)을 조정하면서 변할 수 있다. 이 연구에서, 임계값 우도 비율 8이 암 검출의 맥락에서 대립유전자 불균형이 있는 샘플과 대립유전자 불균형이 없는 샘플을 구별하는 만족할만한 성능을 제공하는 것으로 나타났기 때문에, 상기 값 8이 사용된다. 따라서, 한 실시양태에서, SPRT 곡선의 상한 및 하한을 계산하기 위한 방정식은 다음과 같다:

[0121] 상한 =  $[(\ln 8)/N - \ln \delta]/\ln \gamma$

[0122] 하한 =  $[(\ln 1/8)/N - \ln \delta]/\ln \gamma$

[0123] 여기서,  $\delta = (1 - \theta_1)/(1 - \theta_0)$ ,

[0124]  $\gamma = -(\theta_1(1 - \theta_0)/\theta_0(1 - \theta_1))$ ,

[0125]  $\theta_0$  = 귀무가설이 사실인 경우 비-참조 대립유전자를 함유하는 정보제공 웰의 비율 = 0.5 (하기 참조),

[0126]  $\theta_1$  = 대립가설이 사실인 경우 비-참조 (즉, 과다표출된) 대립유전자를 함유하는 정보제공 웰의 비율,

[0127]  $N$  = 정보제공 웰의 수 = 어느 하나의 대립유전자에만 양성인 웰의 수.

[0128] ( $\ln$ 은 자연 로그, 즉  $\log_e$ 를 나타내는 수학 기호임)

[0129] 귀무가설을 수용하기 위한  $\theta_0$ 의 결정을 위해서, 샘플은 정배수체 태아를 갖는 임신 여성으로부터 얻어지는 것으로 가정된다. 이러한 가정 하에, 어느 하나의 주형에 양성인 웰의 예상 수는 1:1일 것이므로, 비-참조 대립유전자를 함유하는 정보제공 웰의 예상 비율은 0.5일 것이다.

[0130] 대립가설을 수용하기 위한  $\theta_1$ 의 결정을 위해서, 샘플은 삼염색체 21 태아를 갖는 임신 여성으로부터 얻어지는 것으로 가정된다. 디지털 RNA-SNP 분석을 위한 삼염색체 21 경우의 예상  $P_r$ 의 계산은 표 600에 상술되어 있다. 따라서, 디지털 RNA-SNP 분석을 위한  $\theta_1$ 은 표 600의 마지막 열에 제시된 데이터를 참조한다.

[0132] 4. 평균 농도의 측정

[0133]  $m_r$ 의 측정은 당업자에게 공지되거나 또는 공지될 다양한 메커니즘에 의해 실시될 것이다. 한 실시양태에서,  $m_r$  값은 디지털 PCR 분석의 실험 과정 중 결정된다.  $m_r$  값과 참조 대립유전자에 대해 양성인 웰의 총 수 사이의 관계는 분포 (예를 들어, 푸아송 분포)에 의해 좌우될 수 있으므로,  $m_r$ 은 다음 식을 사용하여 참조 대립유전자에 대해 양성인 웰의 수로부터 계산될 수 있다:

[0134]  $m_r = -\ln(1 - \text{참조 대립유전자에 대해 양성인 웰의 비율})$ .

[0135]  $\ln$ 은 자연 로그, 즉  $\log_e$ 라는 것에 주목한다. 이 접근법은 디지털 PCR 실험을 위해 사용되는 DNA 샘플 중  $m_r$ 의 직접적이고 정확한 평가를 제공한다.



[0136] 이 방법은 바람직한 농도를 얻기 위해서 사용될 수 있다. 예를 들어, 샘플의 추출된 핵산은 방법 200의 단계 240에서 실시한 것과 같이 반응 웰 당 하나의 주형 분자와 같은 특정 농도로 희석될 수 있다. 푸아송 분포를 사용하는 한 실시양태에서, 주형이 없는 웰의 예상 비율은  $e^{-m}$ 으로 계산될 수 있고, 여기서  $m$ 은 웰 당 주형 분자의 평균 농도이다. 예를 들어, 웰 당 하나의 주형 분자의 평균 농도에서, 주형 분자가 없는 웰의 예상 비율은  $e^{-1}$ , 즉 0.37 (37%)로 주어진다. 나머지 63%의 웰은 하나 이상의 주형 분자를 함유할 것이다. 전형적으로, 디지털 PCR 실행에서 양성 웰 및 정보제공 웰의 수가 계수될 것이다. 정보제공 웰의 정의 및 디지털 PCR 데이터가 해석되는 방식은 용도에 좌우된다.

[0137] 다른 실시양태에서, 웰 당 평균 농도  $m_r$ 은 또다른 정량화 방법, 예를 들어 질량 분광법을 사용하는 정량 실시간 PCR, 반-정량 경쟁적 PCR, 실-경쟁적 PCR 등에 의해 측정된다.

[0139] **B. 디지털 RCD**

[0140] 평균 농도를 사용하는 디지털 RCD를 상기 기재된 디지털 SNP 방법과 유사한 방식으로 수행할 수 있다. 참조 염색체 (비-염색체 21) 마커, 염색체 21 마커, 및 두 마커 모두에 대해 양성인 웰의 수를 디지털 PCR에 의해 측정할 수 있다. 디지털 SNP 분석에 대한 웰 당 참조 마커의 평균 농도 ( $m_r$ )의 계산에서와 같은 푸아송 확률 함수에 따라, 염색체 21 마커의 양성에 상관없이, 웰 당 참조 마커의 평균 농도 ( $m_r$ )를 참조 마커에 대해 음성인 웰의 총 수로부터 계산할 수 있다.

[0141] 이후, 정배수체 또는 삼염색체 21 태아를 갖는 임신 여성으로부터 얻어진 것과 같은 혈장 샘플을 분류하기 위해 SPRT 분석을 사용할 수 있다. 태아가 정배수체인 경우 귀무가설이 수용될 것이다. 이 시나리오에서, 참조 마커 및 염색체 21 마커에 대해 양성인 웰의 예상 비율은 1:1일 것이며, 이에 따라, 염색체 21 마커에 대해 양성 신호를 갖는 정보제공 웰의 예상 비율은 0.5일 것이다. 태아가 염색체 21에 대해 삼염색체인 경우 대립가설이 수용될 것이다. 이 시나리오에서, 샘플 DNA가 전적으로 태아로부터 유래된 경우, 각 웰에서의 염색체 21 마커의 평균 농도는 참조 마커의 평균 농도 ( $m_r$ )의 3/2배일 것이다.

[0142] 디지털 RCD를 사용하여 태아-특이적 마커, 예를 들어 태반 후성 기호의 검출을 통해 염색체 양을 측정할 수 있지만 (문헌 [Chim, SSC. et al. 2005 Proc Natl Acad Sci USA 102, 14753-14758]), 디지털 RCD 분석의 실시양태는 비-태아-특이적 마커를 사용한다. 따라서, 비-태아 특이적 마커를 사용하는 경우 태아 물질의 백분율을 측정하는 추가의 단계가 나타날 것이다. 따라서, 웰 당 염색체 21 마커의 평균 농도는 샘플 중 태아 DNA의 비율에 의존적일 것이며,  $m_r [(200\% + \text{태아 DNA 백분율})/200\%]$ 을 사용하여 계산할 수 있다.

[0143] 예시용 특정 실시예를 다시 사용하여, 웰 당 참조 주형인 염색체 1의 평균 농도 ( $m_r$ )는 0.5인 것으로 가정하며, 50%의 DNA는 태아로부터 유래된 것으로 가정하고, 샘플 중 50%의 DNA는 모체로부터 유래된다.

[0144] 따라서, 푸아송 분포를 사용하여, 평균 농도가 웰 당 0.5인 경우 염색체 1 유전자좌의 어떠한 분자도 함유하지 않는 임의의 웰의 확률은 하기와 같을 것이다:

$$P(0) = \frac{0.5^0 e^{-0.5}}{0!} = e^{-0.5} = 0.6065$$

[0145]

[0146] 따라서, 염색체 1 유전자좌 중 하나 이상의 분자를 함유하는 임의의 웰의 확률은  $1 - 0.6065 = 0.3935$ 일 것이다. 따라서, 대략 39%의 웰이 유전자좌 중 하나 이상의 분자를 함유하는 것으로 예상될 것이다.

[0147] 이러한 삼염색체 21 태아의 각각의 세포에 대하여, 염색체 21 대 염색체 1의 계승 비율은 3:2일 것이다. DNA 샘플에서 염색체 21과 염색체 1 사이의 비율은 태아 DNA 분획 농도 (태아 DNA %)에 의존적일 것이며,  $3 \times \text{태아 DNA \%} + 2(1 - \text{태아 DNA \%}) : 2 \times \text{태아 DNA \%} + 2 \times (1 - \text{태아 DNA \%})$ 일 것이다. 따라서, 이 경우에서, 태아 DNA 분획 농도가 50%인 경우, 비는  $(3 \times 50\% + 2 \times 50\%) / (2 \times 50\% + 2 \times 50\%) = 1.25$ 일 것이다. 디지털 SNP 방법이 태아 특이적 마커를 사용하지 않은 경우, 이러한 계산을 사용하여 비-참조 서열의 평균 농도를 계산할 수도 있다.

[0148] 따라서, 웰 당 염색체 1 유전자좌의 평균 농도가 0.5인 경우, 웰 당 염색체 21 유전자좌의 평균 농도는  $1.25 \times 0.5 = 0.625$ 이다. 따라서, 평균 농도가 웰 당 0.625인 경우 염색체 21 유전자좌의 어떠한 분자도 함유하지 않는 임의의 웰의 확률은 하기와 같을 것이다:

$$P(0) = \frac{0.625^0 e^{-0.625}}{0!} = e^{-0.625} = 0.5353$$

- [0149]
- [0150] 따라서, 염색체 21 유전자좌 중 하나 이상의 분자를 함유하는 임의의 웰의 확률은  $1 - 0.5353 = 0.4647$ 일 것이다. 따라서, 대략 46%의 웰이 유전자좌 중 하나 이상의 분자를 함유하는 것으로 예상될 것이다. 유전자좌 중 하나로의 웰의 충전이 독립적이라고 가정하면, 유전자좌 모두를 함유하는 웰의 확률은  $0.3935 \times 0.4647 = 0.1829$ 일 것이다. 따라서, 대략 18%의 웰이 두 유전자좌 모두를 함유하는 것으로 예상될 것이다.
- [0151] 염색체 1 유전자좌를 함유하되 염색체 21 유전자좌는 함유하지 않는 것으로 예상되는 웰의 비율은, 하나 이상의 염색체 1 유전자좌를 함유하는 웰의 수에서 유전자좌 모두를 함유하는 웰의 수를 차감한 것, 즉  $0.3935 - 0.1829 = 0.2106$ 일 것이다. 유사하게, 염색체 21 유전자좌를 함유하되 유전자좌 모두는 함유하지 않는 것으로 예상되는 웰의 비율은  $0.4647 - 0.1829 = 0.2818$ 일 것이다. 정보제공 웰은 염색체 1 유전자좌 또는 염색체 21 유전자좌 중 하나에 대해 양성인 둘 모두에 대해서는 아닌 웰로서 정의된다.
- [0152] 따라서, 디지털 RCD 분석에서의 염색체 21 대 염색체 1 예상 비율은  $0.2818/0.2106 = 1.34$ 이다. 바꾸어 말하면, 염색체 21 유전자좌에 대해서만 양성인 웰의 비율은 염색체 1 유전자좌에 대해서만 양성인 웰의 비율의 1.34배이다. 이는 DNA 샘플에서의 1.25의 비율과는 대조적이다.
- [0153] SPRT 분석에 대하여, 염색체 21 유전자좌에 대해 양성인 정보제공 웰의 비율 ( $P_r$ )을 SPRT 곡선을 사용하여 계산 및 해석하는 것이 필요할 것이다. 현재 예에서, 정보제공 웰의 비율은  $0.2106 + 0.2818 = 0.4924$ 일 것이다. 따라서, 0.5의  $m_r$ 에서 50% 태아 DNA를 갖는 삼염색체 21 경우의 예상  $P_r$ 은  $0.2818/0.4924 = 0.57$ 이다.
- [0154] 평균 주형 농도 ( $m$ )가 푸아송 방정식에서의 중요 파라미터이기 때문에,  $P_r$ 은  $m$ 에 의해 변할 것이다. 도 7은 본 발명의 실시양태에 따른 웰 당 평균 참조 주형 농도 ( $m_r$ )로서 표현된 주형 농도의 범위에서 삼염색체 21 샘플 중 10%, 25%, 50% 및 100%의 태아 DNA 분획 농도에 대한 예상  $P_r$ 을 표로 만든 표 700을 나타낸다. 디지털 RCD 분석에 대한 삼염색체 21 경우의 예상  $P_r$ 에 대한 계산을 표 700에서 상술한다. 따라서, 표 700에서 상응하는 예상  $P_r$  값을 나타내는 열로부터 여러 태아 DNA 분획 농도를 갖는 샘플의 디지털 RCD 분석에 대한  $\Theta_1$ 을 얻을 수 있다.
- [0156] **C. 결과**
- [0157] 1. 상이한  $m_r$ 의 비교
- [0158] 대립유전자 또는 염색체 불균형의 이론값 (태아 계놈에서와 같음)과 실험적으로 예상되는 정도 사이의 차이, 및  $m_r$  값 범위에 대한 후자의 측정을 위한 계산값에 대한 근거를 표 600 및 700에 나타낸다. 삼염색체 21 샘플의 디지털 RNA-SNP 분석에서,  $m_r = 0.5$ 인 경우, 참조 대립유전자만을 함유하는 웰에 대한 과다표출된 대립유전자만을 함유하는 웰, 즉 디지털 RNA-SNP 비율은 2.65이다 (표 600). 100% 태아 DNA로 구성된 표본의 디지털 RCD 분석에서,  $m_r = 0.5$ 인 경우, 염색체 1 유전자좌에 대해 단독으로 양성인 웰에 대한 염색체 21 유전자좌에 대해 단독으로 양성인 웰, 즉 디지털 RCD 비율은 1.7이다 (표 700) ( $P_r = 0.63$ , 따라서 디지털 RCD 비율은  $0.63/(1 - 0.63) = 1.7$ 임). 태아 DNA 분획 농도가 감소하기 때문에, 디지털 RCD 비율은 동일한  $m_r$ 에 대해 감소한다 (표 700).
- [0159] 표 600 및 700에서 나타낸 것과 같이, 대립유전자 또는 염색체 과다표출 정도는  $m_r$ 에 따라 증가한다. 그러나, 정보제공 웰의 백분율은  $m_r = 0.5$  근처에서 그의 최대에 도달하며,  $m_r$ 이 추가로 증가함에 따라 점차 감소한다. 실제로, 표본에 대한 주형 분자의 양이 제한되지 않는 경우 정보제공 웰의 비율의 감소는 분석된 웰의 총 수의 증가에 의해 상쇄될 수 있지만, 추가의 웰은 시약 비용의 증가를 요구할 것이다. 따라서, 최적 디지털 PCR 수행은 주형 농도와 샘플 당 시험된 웰의 총 수 간에 상호 의존적이다.
- [0161] 2. SPRT 곡선을 사용하는 예
- [0162] 상기 논의된 것과 같이, 디지털 PCR 실험에 대한 예상 대립유전자 또는 염색체 불균형 정도는 반응 혼합물 (예를 들어, 웰) 당 실제 주형 농도에 의존적이다. 본 발명자들은 참조 대립유전자를 기준으로 한 주형 농도, 즉

웰 당 평균 참조 주형 농도 ( $m_r$ )를 기재한다. 상기 방정식에서 나타낸 것과 같이, 예상  $P_r$ 을 사용하여 상부 및 하부 SPRT 곡선의 플롯팅을 결정할 수 있다. 예상  $P_r$ 이 또한  $m_r$  값에 의존적이기 때문에, SPRT 곡선의 플롯팅은 본질적으로  $m_r$  값에 의존적일 것이다. 따라서, 실제로, 디지털 PCR 데이터 세트의 실제  $m_r$ 에 관련된 한 세트의 SPRT 곡선이 특정 실행으로부터  $P_r$ 의 해석을 위해 사용될 필요가 있다.

[0163] 도 8은 본 발명의 실시양태에 따른 디지털 RNA-SNP 분석에 있어서 0.1, 0.5 및 1.0의  $m_r$  값에 대한 SPRT 곡선에 서의 차이의 정도를 예시하는 플롯 800을 나타낸다. 각 세트의 디지털 PCR 데이터는 특정 실행의 정확한  $m_r$  값에 관련된 특정 곡선으로 해석되어야 한다. 디지털 RNA-SNP 및 RCD 방법에 대한 예상 대립유전자 또는 염색체 불균형 정도가 상이하기 때문에 (전자: 2:1, 후자: 3:2), 2개 디지털 PCR 시스템에 대하여 상이한 세트의 SPRT 곡선이 필요하다는 것을 주목한다. 실험적으로 유도된  $P_r$ 을 디지털 PCR 실행의 상응하는  $m_r$ 에 의해 선택된 관련 SPRT 곡선을 이용하여 해석한다. 이는 고정된 세트의 곡선이 사용된 경우 디지털 PCR에 의한 LOH의 분자 검출을 위해 앞서 보고된 SPRT를 사용한 것과는 대조적이다.

[0164] SPRT를 사용하여 디지털 PCR 데이터를 해석하기 위한 실제적 방법을 가상 디지털 RNA-SNP 실행을 사용하여 하기 예시한다. 각 경우의 디지털 RNA-SNP 분석 이후, A 대립유전자에 대해서만 양성인 웰, G 대립유전자에 대해서만 양성인 웰, 또는 두 대립유전자 모두에 대해서 양성인 웰의 수를 계수한다. 참조 대립유전자는 보다 적은 수의 양성 웰을 갖는 대립유전자로서 정의된다. 푸아송 확률 밀도 함수에 따라, 다른 대립유전자가 양성인 지에 관계없이, 참조 대립유전자에 대해 음성인 웰의 총 수를 사용하여  $m_r$  값을 계산한다. 본 발명자의 가설 예의 데이터는 다음과 같다:

[0165] 96-웰 반응에서, 20개의 웰은 A 대립유전자에 대해서만 양성이고, 24개의 웰은 G 대립유전자에 대해서만 양성이며, 33개의 웰은 두 대립유전자 모두에 대해 양성이다. A 대립유전자가 참조 대립유전자로 간주되는데, A-양성 웰이 G-양성 웰에 비해 적기 때문이다. 참조 대립유전자에 대해 음성인 웰의 수는  $96 - 20 - 33 = 43$ 이다. 따라서, 푸아송 방정식을 사용하여 다음과 같이  $m_r$ 을 계산할 수 있다:  $-\ln(43/96) = 0.80$ . 이 경우의 실험적으로 결정된  $P_r$ 은  $24/(20 + 24) = 0.55$ 이다.

[0166] 표 600에 따르면,  $m_r = 0.8$ 에서의 삼염색체 21 샘플의 예상  $P_r$ 은 0.76이다. 따라서, 이 경우  $\Theta_1$ 은 0.76이다.  $\Theta_1 = 0.76$ 을 기준으로 한 SPRT 곡선을 사용하여 실험적으로 유도된  $P_r$  (이 경우 0.55)을 해석할 것이다.  $P_r = 0.55$ 를 관련된 SPRT 곡선 상에 피팅하는 경우, 데이터 지점은 하부 곡선 아래에 있다. 따라서, 본 경우는 정배수체로 분류되며, 도 3을 참조한다.

[0168] 3. 기존 방법과의 비교

[0169] 도 9a는 96-웰 디지털 RNA-SNP 분석에서의 정배수체 및 삼염색체 21 경우를 분류하기 위한 신규 및 기존 SPRT 알고리즘의 효과의 비교 표 900을 나타낸다. 도 9b는 384-웰 디지털 RNA-SNP 분석에서 정배수체 및 삼염색체 21 경우를 분류하기 위한 신규 및 기존 SPRT 알고리즘의 효과의 비교의 표 950이다. 신규 알고리즘은 디지털 PCR 데이터로부터 유도된  $m_r$ 에 특이적인 SPRT 곡선의 선택을 의미한다. 기존 알고리즘은 모든 디지털 PCR 실행에 대한 고정된 세트의 SPRT 곡선의 사용을 의미한다. 분류의 정확도에 대한 분별값의 부정확한 계산의 효과는 표 900에 나타난 시뮬레이션 분석에 의해 밝혀진다.

[0170] 기존 연구에서 고정된 세트의 SPRT 곡선의 사용과 비교하여, 분류불가능한 데이터의 비율은 표 900 및 950에서 나타난 것과 같이 본 발명의 방법에서 훨씬 적다. 예를 들어, 본 발명의 방법을 사용하여,  $m_r = 0.5$ 에서, 14% 및 0%의 삼염색체 21 샘플이 각각 96-웰 및 384-웰 디지털 RNA-SNP 분석에 대하여 분류불가능할 것이나, 고정된 곡선의 사용에 대해서는 각각 62% 및 10%가 분류불가능할 것이다 (표 900). 따라서, 본 발명의 방법은 보다 적은 수의 정보제공 웰로 질환 분류를 가능하게 한다.

[0171] 표 900에서 나타난 것과 같이, 신규 알고리즘은 0.1 내지 2.0의 모든  $m_r$  값에 대한 대립유전자 비율이 벗어나거나 벗어나지 않은 샘플의 분류에서 보다 정확하다. 예를 들어,  $m_r$ 이 1.0과 같고, 96-웰 디지털 RNA-SNP 실행을 수행하는 경우, 신규 알고리즘은 각각 88% 및 92%의, 대립유전자 비율이 벗어난 샘플 및 벗어나지 않은 샘플을 정확하게 분류하며, 반면 기존 알고리즘을 사용하여 대립유전자 비율이 벗어난 샘플 및 벗어나지 않은 샘플의 정확한 분류의 백분율은 단지 각각 19% 및 36%이다.

- [0172] 신규 알고리즘을 사용하여, 대립유전자 비율이 벗어난 샘플과 벗어나지 않은 샘플의 분리는  $m_r$ 에 따라 증가할 것이다. 결과적으로, 분류 정확도는  $m_r$ 에 따라 증가할 것이다.  $m_r$ 이 2.0 초과로 증가하는 경우 분류 정확도에 대한 2개 샘플 그룹의 분리에서의 증가의 효과는 감소될 것인데, 이는 정보제공 웰의 백분율에서의 감소 때문이다. 반면, 기존 알고리즘의 사용에서,  $m_r$ 이 증가하는 경우 분류 정확도는 유의하게 감소하는데, 이는 예상  $P_r$  값과 그의 참값의 편차 증가 때문이다.
- [0173] 본 발명자의 실험 및 시뮬레이션 데이터는 디지털 RNA-SNP가 삼염색체 21 검출에 대해 효과적이며 정확한 방법임을 나타낸다. 모체 혈장에서의 *PLAC4* mRNA가 태아로부터 순수하게 유래되기 때문에, 시험된 13개 모체 혈장 샘플 중 12개에 대하여, 오직 1회의 384-웰 디지털 PCR 실험만이 정확한 분류에 필요하였다 (도 13b의 표 1350). 이에 따라, 상기 균질한 실시간 디지털 PCR-기반 방법은 RNA-SNP 분석에 대한 질량 분광법-기반 방법의 또다른 형태를 제공한다 (상기 문헌 [Lo, YMD, et al. 2007 Nat Med]). 태반-특이적 mRNA 전사체와는 별개로, 본 발명자들은 또한 모체 혈장에서의 다른 유형의 태아-특이적 핵산 종을 태아 염색체 이수성의 디지털 PCR-기반 검출에 사용할 수 있다는 것을 구상한다. 한 예는 태아 후성 마커이며 (문헌 [Chim, SSC et al. (2005) Proc Natl Acad Sci USA 102, 14753-14758]; [Chan, KCA et al. (2006) Clin Chem 52, 2211-2218]), 이는 최근에 후성 대립유전자 비율 (EAR) 방법을 사용하여 삼염색체 18의 비-침습적 출생 전 검출에 사용되었다 (문헌 [Tong, YK et al. (2006) Clin Chem 52, 2194-2202]). 따라서, 본 발명자들은 디지털 EAR이 가능한 분석 기법일 것으로 예상된다.
- [0175] V. 증가 %, 다중 마커 및 PCR 변형법
- [0176] 상기 기재된 것과 같이, 태아 DNA만이 임신 11주 내지 17주 사이에 대략 3%의 평균 분획 농도로 모체 혈장 DNA의 소수 분획을 구성하는 경우, 모체 혈장으로부터 추출된 DNA에 본 발명의 실시양태를 적용하는 것이 복잡해질 수 있다. 그러나, 본원에서 나타낸 것과 같이, 디지털 RCD는 심지어 이수체 DNA가 소수 집단으로서 존재하는 경우에도 이수성 검출을 가능하게 한다. 임신 초기 동안 존재할 수 있는, 태아 DNA 분획 농도의 감소에 대하여, 보다 많은 수의 정보제공 계수가 디지털 RCD에 대해 요구된다. 도 12의 표 1200에서 요약한 것과 같이, 본 작업의 의미는 본 발명자들이 한 세트의 기준 파라미터, 예를 들어 요구되는 태아 DNA 분획 및 총 주형 분자를 제공하였다는 것이며, 이들에 대해 진단 검정이 수립될 수 있다. 본 발명자의 견해에서, 25%의 태아 DNA 분획 농도에 대한 7680개의 총 수의 반응은 특히 매력적인 세트의 기준 파라미터이다. 이러한 파라미터는 표 1200에서 나타낸 것과 같이 정배수체 및 삼염색체 21 샘플이 정확하게 97%로 분류가능하게 한다.
- [0177] 모체 혈장의 단위 부피 당 존재하는 혈장 DNA 분자의 수는 제한된다 (문헌 [Lo, YMD. et al. 1998 Am J Hum Genet 62, 768-7758]). 예를 들어, 임신 초기에서, 태아 및 모체 모두로부터의 기여에 대하여, 삼염색체 유전자좌,  $\beta$ -글로빈 유전자의 중앙 모체 혈장 농도는 986 카피/mL인 것으로 나타났다 (문헌 [Lo, YMD. et al. 1998 Am J Hum Genet 62, 768-7758]). 7,680개의 분자를 포획하기 위해, 대략 8 mL의 모체 혈장으로부터 추출된 DNA가 필요할 것이다. 약 15 mL의 모체 혈액으로부터 얻을 수 있는 상기 혈장 부피는 일반적인 실시의 한도이다. 그러나, 본 발명자들은 다중 세트의 chr21 및 참조 염색체 표적을 디지털 RCD 분석을 위해 합할 수 있다는 것을 구상한다. 5쌍의 chr21 및 참조 염색체 표적에 대하여, 단지 1.6 mL의 모체 혈장이 분석에 필요한 주형 분자의 수를 제공하기 위해 필요할 것이다. 멀티플렉스 단일 분자 PCR을 수행할 수 있다. 이러한 멀티플렉스 단일 분자 분석의 견고성(robustness)은 단일 분자 일배체형화(haplotyping)에 대해 이미 입증되었다 (문헌 [Ding, C. and Cantor, CR. 2003 Proc Natl Acad Sci USA 100, 7449-7453]).
- [0178] 별법으로, 25%의 태아 DNA 분획 농도를 달성하기 위해, 방법은 모체 혈장에서 태아 DNA의 선택적 풍부화 (문헌 [Li, Y. et al. 2004 Clin Chem 50, 1002-1011]) 또는 모체 DNA 배경의 억제 (문헌 [Dhallan, R et al. 2004 JAMA 291, 1114-1119]) 또는 둘 모두를 가능하게 할 수 있다. 태아 DNA 풍부화 및 모체 DNA 억제를 위한 물리적 방법과는 별개로, 특정 DNA 메틸화 패턴을 나타내는 태아 DNA 분자의 표적화와 같은 분자 풍부화 전략을 사용하는 것이 또한 가능할 것이다 (문헌 [Chim, SSC et al. 2005 Proc Natl Acad Sci USA 102,14753-14758], [Chan, KCA et al. 2006 Clin Chem 52, 2211-2218]; [Chiu, RWK et al. 2007 Am J Pathol 170, 941-950].)
- [0179] 추가적으로, 디지털 PCR 수행에 대한 현행 연구에 사용되는 것과 같은 디지털 실시간 PCR 분석의 수동 설정을 위한 다수의 또다른 방법이 현재 존재한다. 이러한 다른 방법에는 미세유체 디지털 PCR 칩 (문헌 [Warren, L et al. 2006 Proc Natl Acad Sci USA 103, 17807-17812]; [Ottesen, EA et al. 2006 Science 314, 1464-1467]), 에멀전 PCR (문헌 [Dressman, D et al. 2003 Proc Natl Acad Sci USA 100, 8817-8822]), 및 예를 들어 로슈(Roche) 454 플랫폼, 일루미나 솔렉사(Illumina Solexa) 플랫폼, 및 어플라이드 바이오시스템즈

(Applied Biosystems)의 솔리드(SOLiD)<sup>TM</sup> 시스템을 사용하는 대규모 병렬 게놈 서열분석 (문헌 [Margulies, M. et al. 2005 Nature 437, 376-380]) 등이 있다. 후자에 대하여, 본 발명자의 방법은 또한 단일 DNA 분자에 대한 대규모 병렬 서열분석 방법에 적용가능하며, 이는 증폭 단계, 예컨대 헬리코스 트루(Helicos True) 단일 분자 DNA 서열분석 기술 (문헌 [Harris TD et al. 2008 Science, 320, 106-109]), 퍼시픽 바이오사이언시즈 (Pacific Biosciences)의 단일 분자, 실시간 (SMRT<sup>TM</sup>) 기술, 및 나노공극(nanopore) 서열분석 (문헌 [Soni GV and Meller A. 2007 Clin Chem 53, 1996-2001])을 요구하지 않는다. 이들 방법을 사용하여, 디지털 RNA-SNP 및 디지털 RCD를 다수의 샘플에 대하여 빠르게 수행함으로써, 비-침습적 출생 전 진단에 대해 본원에서 제안된 방법의 임상적 실행가능성을 향상시킬 수 있다.

**도면의 간단한 설명**

[0181]

- 도 1은 디지털 PCR 실험을 설명하는 흐름도이다.
- 도 2a는 본 발명의 실시양태에 따른 디지털 RNA-SNP 및 RCD 방법을 예시한다.
- 도 2b는 암에서 자주 검출되는 염색체 이상의 예에 대한 표를 도시한다.
- 도 3은 본 발명의 실시양태에 따른 다운 증후군을 결정하는데 사용된 SPRT 곡선을 갖는 그래프를 예시한다.
- 도 4는 본 발명의 실시양태에 따른 태아 세포의 백분율을 사용하여 질환 상태를 결정하는 방법을 도시한다.
- 도 5는 본 발명의 실시양태에 따른 평균 농도를 사용하여 질환 상태를 결정하는 방법을 도시한다.
- 도 6은 본 발명의 실시양태에 따른 웰 당 평균 참조 주형 농도 ( $m_r$ )로서 표현된 주형 농도 범위에 대해 삼염색체 21 샘플의 예상되는 디지털 RNA-SNP 대립유전자 비율 및  $P_r$ 을 표로 만든 표를 도시한다.
- 도 7은 본 발명의 실시양태에 따른 웰 당 평균 참조 주형 농도 ( $m_r$ )로서 표현된 주형 농도 범위에서 삼염색체 21 샘플 중 분획 태아 DNA 농도 10%, 25%, 50% 및 100%에 대해 예상되는  $P_r$ 을 표로 만든 표를 도시한다.
- 도 8은 본 발명의 실시양태에 따른 디지털 RNA-SNP 분석에서  $m_r$  값 0.1, 0.5 및 1.0에 대해 SPRT 곡선의 차이 정도를 예시하는 플롯을 도시한다.
- 도 9a는 본 발명의 실시양태에 따른 96-웰 디지털 RNA-SNP 분석에서 정배수체 및 삼염색체 21 사례를 분류하기 위한 신규/기존 SPRT 알고리즘 효과의 비교에 대한 표를 도시한다.
- 도 9b는 본 발명의 실시양태에 따른 384-웰 디지털 RNA-SNP 분석에서 정배수체 및 삼염색체 21 사례를 분류하기 위한 신규/기존 SPRT 알고리즘 효과의 비교에 대한 표를 도시한다.
- 도 10은 본 발명의 실시양태에 따른 소정의 정보제공 카운트에 대해 정배수체 또는 이수체로서 정확하게 및 부정확하게 분류된 태아 및 분류가능하지 않은 태아의 백분율을 도시하는 표이다.
- 도 11은 본 발명의 실시양태에 따른 순수 (100%) 태아 DNA 샘플에 대한 디지털 RCD 분석용 컴퓨터 시뮬레이션 을 도시하는 표 1100이다.
- 도 12는 본 발명의 실시양태에 따른 상이한 분획 농도의 태아 DNA를 갖는 정배수체 또는 삼염색체 21 태아로부터의 샘플을 분류하기 위해  $m_r=0.5$ 에서 디지털 RCD 분석의 정확성을 컴퓨터 시뮬레이션한 결과를 도시하는 표 1200이다.
- 도 13a는 본 발명의 실시양태에 따른 정배수체 및 삼염색체 21 임신 태반 조직에서의 디지털 RNA-SNP 분석에 대한 표 1300을 도시한다.
- 도 13b는 본 발명의 실시양태에 따른 정배수체 및 삼염색체 21 임신으로부터의 모체 혈장의 디지털 RNA-SNP 분석에 대한 표 1350을 도시한다.
- 도 14a 내지 14c는 본 발명의 실시양태에 따른 RCD 분석으로부터 생성된 분별값 곡선을 예시하는 플롯을 도시한다.
- 도 15a는 본 발명의 실시양태에 따른 정배수체 및 삼염색체 21 임신 태반 조직에서의 디지털 RNA-SNP 분석에 대한 표를 도시한다.

- 도 15b는 본 발명의 실시양태에 따른 한 모체 혈장 샘플로부터의 12개 반응 패널의 디지털 RNA-SNP 데이터에 대한 표를 도시한다.
- 도 15c는 본 발명의 실시양태에 따른 정배수체 및 삼염색체 21 임신으로부터의 모체 혈장의 디지털 RNA-SNP 분석에 대한 표를 도시한다.
- 도 16a는 본 발명의 실시양태에 따른 정배수체 및 삼염색체 18 태반의 디지털 RNA-SNP 분석에 대한 표를 도시한다.
- 도 16b는 본 발명의 실시양태에 따른 정배수체 및 삼염색체 18 태반에 대한 디지털 RNA-SNP 데이터의 SPRT 해석을 도시한다.
- 도 17은 본 발명의 실시양태에 따른 정배수체 및 삼염색체 21 임신의 50% 태반/모체 혈액 세포 DNA 혼합물의 디지털 RCD 분석에 대한 표를 도시한다.
- 도 18은 본 발명의 실시양태에 따른 정확한 분류에 대한 관정 경계값을 예시하는 SPRT 곡선을 도시한다.
- 도 19는 본 발명의 실시양태에 따른 정배수체 및 삼염색체 21 임신으로부터의 양수 샘플의 디지털 RCD 분석에 대한 표를 도시한다.
- 도 20은 본 발명의 실시양태에 따른 정배수체 및 삼염색체 18 임신 (E = 정배수체; T18 = 삼염색체 18)으로부터의 태반 DNA 샘플의 디지털 RCD 분석에 대한 표를 도시한다.
- 도 21은 본 발명의 실시양태에 따른 정배수체 및 삼염색체 21 임신 (E = 정배수체; T21 = 삼염색체 21; U = 미분류)의 50% 태반/모체 혈액 세포 DNA 혼합물의 멀티플렉스 디지털 RCD 분석에 대한 표를 도시한다.
- 도 22a 및 22b는 본 발명의 실시양태에 따른 50% 정배수체 또는 삼염색체 21 태반 계층 DNA/50% 모체 연막 (buffy coat) DNA 혼합물의 멀티플렉스 디지털 RCD 분석에 대한 표를 도시한다. 미분류(Unclass)는 분류가능하지 않음을 나타내고, T21은 삼염색체 21을 나타낸다.
- 도 23은 남성 및 여성 파트너 둘 다가 동일한 돌연변이를 전달하는 시나리오를 도시한다.
- 도 24a는 본 발명의 실시양태에 따른 여성/남성 및 남성/남성 DNA 혼합물의 디지털 RMD 분석에 대한 표를 도시한다.
- 도 24b는 본 발명의 실시양태에 따른 25% 여성 및 75% 남성 DNA를 갖는 혼합물의 디지털 RMD 분석에 대한 표를 도시한다.
- 도 25는 본 발명의 실시양태에 따른 HbE 돌연변이에 대해 모체 혈장 샘플을 모방하는 15% 내지 50% DNA 혼합물의 디지털 RMD 분석에 대한 표를 도시한다.
- 도 26a는 본 발명의 실시양태에 따른 CD41/42 돌연변이에 대해 모체 혈장 샘플을 모방하는 5% 내지 50% DNA 혼합물의 디지털 RMD 분석에 대한 표를 도시한다.
- 도 26b는 본 발명의 실시양태에 따른 CD41/42 돌연변이에 대해 모체 혈장 샘플을 모방하는 20% DNA 혼합물의 디지털 RMD 분석에 대한 표를 도시한다.
- 도 27은 본 발명의 실시양태에 따른 시스템 및 방법으로 이용가능한 예시적 컴퓨터 장치의 블록 다이어그램을 도시한다.

**발명을 실시하기 위한 구체적인 내용**

- [0182] **실시예**
- [0183] 하기 실시예는 청구된 발명을 예시하되 제한하지 않으려고 한다.
- [0185] I. 컴퓨터 시뮬레이션
- [0186] SPRT 방법을 사용하여 삼염색체 21 진단의 정확도를 평가하기 위해 컴퓨터 시뮬레이션을 수행하였다. 마이크로소프트 엑셀 2003 소프트웨어 (마이크로소프트 코퍼레이션(Microsoft Corp.; 미국)) 및 윈도우용 SAS 9.1 소프트웨어 (에스에이에스 인스티튜트 인크.(SAS Institute Inc., 미국 노스 캐롤라이나주))를 사용하여 컴퓨터 시뮬레이션을 수행하였다. 디지털 PCR의 수행은 참조 주형 농도 ( $m_r$ ), 정보제공 계수의 수, 및 대립유전자 또는

염색체 불균형의 추정 정도 ( $P_r$ ) 사이의 상호작용이다. 이들 변수 각각의 범위에 대하여 별도의 시뮬레이션을 수행하였다. 디지털 RNA-SNP 및 디지털 RCD에 대한 SPRT 곡선의 결정 경계가 상이하기 때문에, 2개 시스템에 대한 시뮬레이션 분석을 별도로 수행하였다.

[0187] 시뮬레이션된 각각의 디지털 PCR 조건 (즉,  $m_r$ , 태아 DNA 분획 농도, 총 웰 수)에 대하여, 2회의 시뮬레이션을 수행하였다. 제1회는 시험된 샘플을 정배수체 태아를 갖는 임신 여성으로부터 얻은 시나리오를 시뮬레이션하였다. 제2회는 시험된 샘플을 삼염색체 21 태아를 갖는 임신 여성으로부터 얻은 경우의 시나리오를 시뮬레이션하였다. 각각의 시뮬레이션에서, 5000명의 태아의 시험을 시뮬레이션하였다.

[0189] **A. RNA-SNP**

[0190] 디지털 RNA-SNP에 대하여,  $m_r = 0.1$  내지  $m_r = 2.0$ 으로 384-웰 실험의 시뮬레이션을 수행하였다. 각각의  $m_r$  값에서, 본 발명자들은 5000명의 정배수체 및 5000명의 삼염색체 21 태아를 시험하는 시나리오를 시뮬레이션하였다. 주어진  $m_r$ 에 적절한 SPRT 곡선을 사용하여 10,000명의 태아를 분류하였다. 도 10은 본 발명의 실시양태에 따른, 정배수체 또는 이수체로서 정확하게 및 부정확하게 분류된 태아, 및 주어진 정보제공 계수에 대해 분류가 능하지 않은 태아의 백분율을 나타내는 표 1000이다. 정배수체 및 이수체 경우의 진단에 대한 정확도는 0.5 내지 2.0 사이의  $m_r$ 에 대해 모두 100%이다.  $m_r = 0.1$ 인 경우, 57% 및 88%의 정배수체 및 삼염색체 21 태아만이 384-웰의 분석 이후 정확하게 분류될 수 있었다.

[0191] 하기 단계에 기재된 것과 같이 시뮬레이션 데이터를 생성하였다.

[0192] 단계 1에서, 각 웰에 대하여, SAS 프로그램의 무작위 (푸아송) 함수 ([www.sas.com/technologies/analytics/statistics/index.html](http://www.sas.com/technologies/analytics/statistics/index.html))를 사용하여 2개 난수를 생성하여 각각 A 및 G 대립유전자를 나타내었다. 무작위 (푸아송) 함수는 0으로부터 출발하는 양의 정수 (즉, 0, 1, 2, 3, ...)를 생성할 것이며, 생성되는 각 정수의 확률은 웰 당 대립유전자의 평균 농도를 나타내는 주어진 평균값에 대한 푸아송 확률 밀도 함수에 따른 상기 수의 확률에 의존하였다. A 대립유전자를 나타내는 난수가 0보다 큰 경우, 즉 A 대립유전자의 1개 이상의 분자를 함유한 경우, 웰은 A 대립유전자에 대해 양성으로서 간주되었다. 유사하게, G 대립유전자를 나타내는 난수가 0보다 큰 경우, 웰은 G 대립유전자에 대해 양성으로서 간주되었다.

[0193] 정배수체 태아를 갖는 임신 여성의 시나리오를 시뮬레이션하기 위해, A 대립유전자 및 G 대립유전자에 대한 난수를 생성하는데 동일한 평균 값을 사용하였다. 예를 들어,  $m_r = 0.5$ 에서의 디지털 RNA-SNP 분석을 시뮬레이션 하는 분석에서, A 대립유전자 또는 G 대립유전자 중 하나에 대한 평균 값을 웰 당 0.5 분자의 대립유전자 중 하나에 대한 평균 농도를 의미하는 0.5와 동일하게 설정하였다. 푸아송 방정식을 사용하여, 0.5의 평균 농도에서, A 또는 G 대립유전자에 대해 양성인 웰의 비율은 동일할 것이며, 0.3935였다 (표 600 참조).

[0194]  $m_r = 0.5$ 에서 삼염색체 21 태아를 갖는 임신 여성의 디지털 RNA-SNP 분석을 시뮬레이션한 경우, 웰 당 과다표출된 대립유전자의 평균 농도는 참조 대립유전자의 평균 농도의 2배, 즉 1이 되는 것으로 예상될 것이다. 이 상황에서, 과다표출된 대립유전자에 대해 양성인 웰의 확률은 0.6321이었다 (표 600 참조).

[0195] 디지털 PCR 웰에 대한 난수의 생성 이후, 웰을 하기 상태 중 하나로서 분류할 수 있었다:

[0196] a. A 및 G 대립유전자 모두에 대해 음성

[0197] b. A 및 G 대립유전자 모두에 대해 양성

[0198] c. A 대립유전자에 대해 양성이되, G 대립유전자에 대해 음성

[0199] d. G 대립유전자에 대해 양성이되, A 대립유전자에 대해 음성.

[0200] 단계 2에서, 원하는 수의 웰 (현재 시뮬레이션에 대하여 384-웰)을 생성할 때까지 단계 1을 반복하였다. A 대립유전자에 대해서만 양성인 웰의 수 및 G 대립유전자에 대해서만 양성인 웰의 수를 계수하였다. 덜 양성인 웰을 갖는 대립유전자는 참조 대립유전자로서 간주되며, 보다 양성인 웰을 갖는 대립유전자는 잠재적으로 과다표출된 대립유전자로서 간주되었다. 정보제공 웰의 수는 대립유전자 중 하나에 대해 양성이되, 모두에 대해서는 아닌 웰의 총 수였다. 이후, 잠재적으로 과다표출된 대립유전자를 함유하는 정보제공 웰의 비율 ( $P_r$ )을 계산하였다. 귀무가설 또는 대립가설을 수용하기 위한 관련 SPRT 곡선에 대한 상한 및 하한을 본 발명의 실시양태에 따라 계산하였다.

[0201] 단계 3에서, 정배수체 또는 삼염색체 21 태아를 갖는 임신 여성의 각각의 2개 시나리오에 대해 5000회 시뮬레이션을 수행하였다. 각각의 시뮬레이션은 임신 여성으로부터 얻어진 독립적인 생물학적 샘플로서 간주될 수 있었다. 표 1000에서, 정배수체 경우의 정확한 분류는 귀무가설이 수용되는 정배수체 경우를 의미하며, 정배수체 경우의 부정확한 분류는 대립가설이 수용되는 정배수체 경우를 의미한다. 유사하게, 대립가설이 수용되는 삼염색체 21 경우는 정확하게 분류된 것으로 간주되며, 귀무가설이 수용되는 삼염색체 21 경우는 부정확하게 분류된 것으로 간주된다. 두 군 모두에 대하여, 미리-명시된 웰의 총 수를 시뮬레이션한 이후 귀무가설 또는 대립가설 모두가 수용되지 않는 경우에, 미분류된 것으로 간주되었다.

[0202] 단계 4에서, 단계 1 내지 3을 0.1 내지 2.0 범위의  $m_r$ 에 대해 0.1 증가량으로 수행하였다.

[0204] **B. RCD**

[0205] 도 11은 본 발명의 실시양태에 따라 0.1 내지 2.0 범위의  $m_r$ 에서 순수한 (100%) 태아 DNA 샘플에 대한 디지털 RCD 분석용 컴퓨터 시뮬레이션을 나타내는 표 1100이다. 태아 DNA 분획 농도가 낮아졌기 때문에, 염색체 21 과다표출 정도는 감소되며, 이에 따라 정확한 질환 분류를 위해 보다 많은 수의 정보제공 웰이 요구된다. 따라서,  $m_r = 0.5$ 에서 384개 내지 7680개 웰 범위의 총 웰 수에 대하여 50%, 25% 및 10%의 태아 DNA 농도에 대해 추가로 시뮬레이션을 수행하였다.

[0206] 도 12는 본 발명의 실시양태에 따른 상이한 분획 농도의 태아 DNA를 갖는 정배수체 또는 삼염색체 21 태아로부터의 샘플의 분류에 대하여  $m_r = 0.5$ 에서 디지털 RCD 분석의 정확도의 컴퓨터 시뮬레이션 결과를 나타내는 표 1200이다. 디지털 RCD의 수행은 보다 높은 태아 DNA 분획 농도를 갖는 경우에 대해 보다 양호하다. 25% 태아 DNA 농도 및 7680개 총 수의 PCR 분석에서, 97%의 정배수체 및 이수체 경우 모두가 부정확한 분류 없이 분류가 가능할 것이다. 나머지 3%의 경우는 분류가 달성될 수 있을 때까지 추가 분석을 요구한다.

[0207] 디지털 RCD 분석을 시뮬레이션하기 위한 절차는 디지털 RNA-SNP 분석에 대해 기재된 것과 유사하였다. 시뮬레이션을 위한 단계를 하기 기재한다.

[0208] 단계 1에서, 푸아송 확률 밀도 함수 하의 2개의 난수를 생성하여 참조 유전자좌, 염색체 1 및 염색체 21 유전자좌를 나타내었다. 정배수체 태아를 갖는 대상체에 대하여, 염색체 1 및 염색체 21 유전자좌 모두의 평균 농도는 동일하였다. 상기 시뮬레이션 분석에서, 웰 당 각 유전자좌에 대한 0.5의 평균 주형 농도를 사용하였다. 삼염색체 21 태아를 갖는 대상체에 대하여, 상기 시뮬레이션에서의  $m_r$ 은 0.5였으나, 웰 당 염색체 21 유전자좌의 평균 농도는 표 700에서 나타낸 것과 같이 시험된 샘플에서의 태아 DNA 분획 농도에 의존할 것이다. 웰에 대한 참조 및/또는 염색체 21 유전자좌의 분포를 웰 당 유전자좌의 적절한 평균 농도를 갖는 푸아송 확률 밀도 함수에 따라 생성된 각각의 유전자좌를 나타내는 난수에 의해 측정하였다.

[0209] 단계 2에서, 원하는 수의 웰, 예를 들어 384-웰 플레이트 실험에 대하여 384개 웰이 생성될 때까지 단계 1을 반복하였다. 염색체 1 단독 및 염색체 21 단독에 대해서만 양성인 웰의 수를 계수하였다. 정보제공 웰의 수는 염색체 중 하나에 대해 양성인 웰, 모두에 대해서는 아닌 웰의 총 수였다. 이후, 염색체 21에 대해 양성인 정보제공 웰의 비율 ( $P_r$ )을 계산하였다. 귀무가설 또는 대립가설을 수용하기 위한 관련 SPRT 곡선의 상한 및 하한을 SPRT 분석에 대해 앞선 섹션에서 기재된 것과 같이 계산하였다.

[0210] 단계 3에서, 정배수체 또는 삼염색체 21 태아를 갖는 임신 여성의 2개 시나리오 각각에 대해 5000회의 시뮬레이션을 수행하였다. 각각의 시뮬레이션은 임신 여성으로부터 얻어진 독립적인 생물학적 샘플로서 간주될 수 있었다. 표 1100에서, 정배수체 경우의 정확한 분류는 귀무가설이 수용되는 정배수체 경우를 의미하며, 정배수체 경우의 부정확한 분류는 대립가설이 수용되는 정배수체 경우를 의미한다. 유사하게, 대립가설이 수용되는 삼염색체 21 경우는 정확하게 분류된 것으로 간주되며, 귀무가설이 수용되는 삼염색체 21 경우는 부정확하게 분류된 것으로 간주되었다. 두 군 모두에 대하여, 미리-명시된 웰의 총 수를 시뮬레이션한 이후 귀무가설 또는 대립가설 모두가 수용되지 않는 경우에, 미분류된 것으로 간주되었다.

[0211] 단계 4에서, 단계 1 내지 3을 384개 내지 7680개 범위의 총 웰 수에서 10% 25%, 50% 및 100% 태아 DNA를 갖는 샘플에 대해 반복하였다.

[0213] **II. 삼염색체 21 검출의 검증**

[0214] **A. PLAC4에 대한 RNA-SNP**



- [0215] 염색체 21 상의 PLAC4 유전자에 대한 rs8130833 SNP를 사용하여 디지털 RNA-SNP의 실제적 실행가능성을 입증하였다 (문헌 [Lo, YMD et al. 2007 Nat Med 13, 218-223]). 2개 정배수체 및 2개 삼염색체 21 이형접합 태반으로부터의 태반 DNA 및 RNA 샘플을 분석하였다. 디지털 RNA-SNP 프로토콜을 사용하되, 역전사 단계를 생략하였으며, 이에 따라 필수적으로 절차를 디지털 DNA-SNP 분석으로 전환하여, 태반 DNA 샘플을 분석하였다. 정확한 경우 분류의 가능성과 정보제공 웰의 비율 사이에 균형을 맞추기 위해, 본 발명자들은 웰 당 임의의 유형의 하나의 대립유전자를 목적으로 하는 샘플을 희석하고, 96-웰 디지털 PCR 분석에 의해 확인하였다. 이후, 384-웰 디지털 RNA-SNP 실험을 하였다.  $P_r$  및  $m_r$ 을 계산하고, 상기  $m_r$  값에 대한 SPRT 곡선을 데이터 해석을 위해 사용하였다.
- [0216] 도 13a는 본 발명의 실시양태에 따른 정배수체 및 삼염색체 21 임신의 태반 조직에서의 디지털 RNA-SNP 분석의 표 1300을 나타낸다. 질량 분광 검정에 의해 유전자형을 측정하였다. 실험적으로 얻어진  $P_r$ 이 분류불가능한 영역 아래에 있는 경우 "정배수체"로 지정하였고, 실험적으로 얻어진  $P_r$ 이 분류불가능한 영역 위에 있는 경우 삼염색체 21을 나타내는 "T21"로 지정하였다. T21, 삼염색체 21. 1회 384-웰 실험으로 DNA 및 RNA 샘플 모두를 사용하여 각각의 상기 경우를 정확하게 분류하였다.
- [0217] 추가적으로, 본 발명자들은 정배수체 태아를 갖는 9명의 여성 및 삼염색체 21 태아를 갖는 4명의 여성으로부터의 혈장 RNA 샘플을 시험하였다. 도 13b는 본 발명의 실시양태에 따른 정배수체 및 삼염색체 21 임신 여성으로부터의 모체 혈장의 디지털 RNA-SNP 분석의 표 1350을 나타낸다. 모든 경우를 정확하게 분류하였다. 한 삼염색체 21 경우 (M2272P)에 대한 초기 결과는 1회 384-웰 실험 이후 SPRT 곡선 사이의 분류불가능한 영역 내에 있었다. 따라서, 추가의 384-웰 실험을 수행하였다. 총 768개 웰로부터의 응집된 데이터로부터 신규  $m_r$  및  $P_r$  값을 계산하였으며, 상기  $m_r$  값을 기준으로 선택된 신규 세트의 SPRT 곡선을 사용하여 분류를 수행하였다. 이후, 상기 경우를 이수체로서 정확하게 스코어링하였다.
- [0218] 본 발명자의 실험 및 시뮬레이션 데이터는 디지털 RNA-SNP가 삼염색체 21 검출에 대한 효과적이며 정확한 방법임을 나타낸다. 모체 혈장 중 PLAC4 mRNA가 태아로부터 순수하게 유래되기 때문에, 시험된 13개 모체 혈장 샘플 중 12개에 대하여, 오직 1회의 384-웰 디지털 PCR 실험만이 정확한 분류를 위해 요구되었다. 이에 따라, 상기 균질한 실시간 디지털 PCR-기반 방법은 RNA-SNP 분석을 위한 질량 분광법-기반 방법과는 또다른 방법을 제공한다. 태반-특이적 mRNA 전사체와는 별개로, 본 발명자들은 또한 모체 혈장에서의 다른 유형의 태아-특이적 핵산 종이 태아 염색체 이수성의 디지털 PCR-기반 검출에 대해 사용될 수 있다는 것을 구상한다. 한 예는 최근에 후성 대립유전자 비 (EAR) 방법을 사용하여 삼염색체 18의 비-침습적 출생 전 검출에 사용된 태아 후성 마커이다 (문헌 [Tong YK et al. 2006 Clin Chem, 52, 2194-2202]). 따라서, 본 발명자들은 디지털 EAR이 가능한 분석 기술일 것이라고 예상한다.
- [0220] **B. RCD**
- [0221] 또한, 염색체 21 및 1에서의 파라로그 서열을 표적화하는 PCR 분석을 사용하여 삼염색체 21 검출에 대한 디지털 RCD의 실제적 실행가능성을 연구하였다. 파라로그 유전자좌를 본원에서 예로서 사용하였다. 염색체 21 및 임의의 다른 참조 염색체에 대한 비-파라로그 서열 또한 RCD에 대해 사용할 수 있다. 2개 정배수체 및 2개 삼염색체 21 태반으로부터의 태반 DNA 샘플을 웰 당 염색체 중 하나로부터 대략 1개 표적 주형으로 희석시켰으며, 96-웰 디지털 PCR 분석에 의해 확인하였다. 각각의 확인된 샘플을 384-웰 디지털 RCD 실험에 의해 분석하였으며,  $P_r$  및  $m_r$  값을 계산하였다. 디지털 RCD에 대하여, 염색체 1 파라로그는 참조 주형이었다. 상기  $m_r$  값을 사용하여 데이터 해석을 위한 상응하는 세트의 SPRT 곡선을 선택하였다. 모든 태반 샘플을 도 14a에서 나타낸 것과 같이 정확히 분류하였다.
- [0222] 모체 혈장에서의 태아 DNA의 시나리오와 같은 디지털 RCD 방법을 사용하여 과량의 정배수체 DNA와 혼합된 삼염색체 21 DNA를 검출할 수 있는지를 입증하기 위해, 정배수체 모체 혈액 세포 DNA의 배경 중 50% 및 25%의 삼염색체 21 태반 DNA를 함유하는 혼합물을 분석하였다. 10개의 삼염색체 21 및 10개의 정배수체 경우로부터의 태반 DNA를 각각 동일한 양의 정배수체 모체 혈액 세포 DNA와 혼합하였으며, 이에 따라 20개의 50% DNA 혼합물을 제조하였다. 도 14b는 본 발명의 실시양태에 따른 50% 태아 DNA 혼합물의 RCD 분석에 대한 SPRT 해석을 예시하는 플롯 (1440)을 나타낸다. 유사하게는, 5개의 삼염색체 21 및 5개의 정배수체 경우로부터의 태반 DNA를 각각 3배 과량의 정배수체 모체 혈액 세포 DNA와 혼합하였으며, 이에 따라 10개의 25% DNA 혼합물을 제조하였다. 도 14c는 25% 태아 DNA 혼합물의 RCD 분석에 대한 SPRT 해석을 예시하는 플롯 1470을 나타낸다. 도 14b

및 14c에서 나타난 것과 같이, 모든 정배수체 및 이수체 DNA 혼합물을 정확하게 분류하였다.

[0223] 각각의 샘플은 도 14b 및 14c에 표시된 것과 같이 여러 384-웰 디지털 PCR 분석 이후 분류가능한 지점에 도달하였다. 50% DNA 혼합물에 대하여, 384-웰 플레이트의 수는 1개 내지 5개의 범위가 요구된다. 25% DNA 혼합물에 대하여, 384-웰 플레이트의 수는 1개 내지 7개의 범위가 요구된다. 매 384회의 디지털 PCR 분석의 추가와 함께 정확하게 분류되는 경우의 누적 비율은 표 1200에서 제시된 컴퓨터 시뮬레이션에 의해 예상된 것과 같았다.

[0225] III. 디지털 PCR을 이용한 방법

[0226] **A. 디지털 RNA-SNP**

[0227] 모든 RNA 샘플을 먼저 써모스크립트(ThermoScript) 역전사효소 (인비트로젠(Invitrogen))를 사용하여 유전자-특이적 역전사 프라이머에 의해 역전사시켰다. 역전사 프라이머의 서열은 **5'-AGTATATAGAACCATGTTTAGGCCAGA-3'** (서열 1) (미국 아이오와주 코럴빌 소재의 인테그레이티드 DNA 테크놀로지스(Integrated DNA Technologies))였다. 디지털 RNA-SNP를 위한 역전사된 RNA (즉, cDNA) 샘플 및 DNA 샘플 (예를 들어, 태반 DNA)의 후속 처리는 본질적으로 동일하였다. 디지털 PCR 분석 전에, 먼저 DNA 및 cDNA 샘플을 프라이머 **5'-CCGCTAGGGTGTCTTTTAAGC-3'** (서열 2), **5'-GTGTTGCAATACAAAATGAGTTTCT-3'** (서열 3), 및 형광 프로브 **5'-(FAM)ATTGGAGCAAATTC(MGBNFQ)-3'** (서열 4) (미국 캘리포니아주 포스터 시티 소재의 어플라이드 바이오시스템즈) (여기서, FAM은 6-카르복시플루오레세인이고, MGBNFQ는 마이너 그루브 결합 비-형광 쉐퍼)로 이루어진, PLAC4에 대한 실시간 PCR 검정을 사용하여 정량화하였다.

[0228] 애플리콘이 특정된 HPLC-정제된 단일 가닥 합성 DNA 올리고뉴클레오티드 (싱가포르 소재의 프로리고(Proligo))의 계열 희석에 의해 보정 곡선을 제작하였다. 서열은 **5'-CGCCGCTAGGGTGTCTTTTAAGCTATTGGAGCAAATTCAAATTTGGCTTAAAGAA AAAGAACTCATTGTATTGCAACACCAGGAGTATCCCAAGGGACTCG-3'** (서열 5)였다. 반응 부피 25 µL로 2X 태크만 유니버설 PCR 마스타 믹스(TaqMan Universal PCR Master Mix) (어플라이드 바이오시스템즈)를 사용하여 반응을 설정하였다. 각각의 프라이머 400 nM 및 프로브 80 nM을 각각의 반응에 사용하였다. ABI PRISM 7900HT 서열 검출 시스템 (어플라이드 바이오시스템즈)에서, 반응을 50°C에서 2분 동안 개시하고, 이어서 95°C에서 10분 동안 진행하고, 95°C에서 15초 및 60°C에서 1분의 45주기를 수행하였다. 그 다음, 웰 당 대략 1개의 주형 분자에서 후속 디지털 PCR 증폭이 수행될 수 있도록 DNA 또는 cDNA 샘플의 계열 희석에 착수하였다. 이러한 농도에서는, 반응 웰의 대략 37%가 음성 증폭을 나타낼 것으로 예측되며, 이는 96-웰 디지털 실시간 PCR 분석을 수행함으로써 최초로 확인되었다. 이어서, 한 세트의 비-인트론 관통(spanning) 프라이머: 전방향 프라이머 **5'-TTTGTATTGCAACACCATTGG-3'** (서열 6) 및 앞서 기재된 유전자-특이적 역전사 프라이머를 사용하여 384-웰 플레이트에서 디지털 RNA-SNP 분석을 수행하였다.

[0229] PLAC4 서열 상 rs8130833 SNP의 2개의 대립유전자를 각각 표적으로 하는 2개의 대립유전자-특이적 태크만 프로브를 설계하였다. 그들의 서열은 G 및 A 대립유전자에 대하여 각각 **5'-(FAM)TCGTCGTCCTAACTTG(MGBNFQ)-3'** (서열 7) 및 **5'-(VIC)ATTCGTCATCTAACTTG(MGBNFQ)-3'** (서열 8)였다. 반응 부피 5 µL로 2X 태크만 유니버설 PCR 마스타 믹스를 사용하여 반응을 설정하였다. 각각의 반응은 1X 태크만 유니버설 PCR 마스타 믹스, 각각의 프라이머 572 nM, 대립유전자-G-특이적 프로브 107 nM 및 대립유전자-A-특이적 프로브 357 nM을 함유하였다. 반응은 ABI PRISM 7900HT 서열 검출 시스템에서 수행하였다. 반응을 50°C에서 2분 동안 개시하고, 이어서 95°C에서 10분 동안 진행하고, 95°C에서 15초 및 57°C에서 1분의 45주기를 수행하였다. 반응 동안, SDS 2.2.2 소프트웨어 (어플라이드 바이오시스템즈)의 "절대 정량화" 응용 프로그램에 의해 형광 데이터를 수집하였다. 상기 소프트웨어는 기준선 및 임계값을 자동으로 계산하였다. A 또는 G 대립유전자에 대해 양성인 웰의 수를 기록하여 SPRT 분석에 제공하였다.

[0231] **B. 디지털 RCD 분석**

[0232] 본 연구에 사용되는 모든 태반 및 모체 연막 DNA 샘플을 먼저 나노드롭 분광광도계 (미국 델라웨어주 윌밍턴 소재의 나노드롭 테크놀로지(NanoDrop Technology))에 의해 정량화하였다. 6.6 pg/세포의 전환을 사용하여 DNA

농도를 카피/μL로 전환하였다. 웰 당 대략 1개의 주형에 상응하는 DNA의 양을, DNA 샘플을 계열 희석함으로써 결정하고 96-웰 포맷에서 실시간 PCR 검정에 의해 확인하였는데, 여기서 웰의 대략 37%가 음성 증폭을 나타낼 것으로 예측되었다. 확인 플레이트에 대한 PCR 설정은 참조 염색체에 대한 프로브를 추가한 것만 제외하고는 이하에 기재된 것과 동일하였다. 디지털 RCD 분석에서, 염색체 21 및 1 상의 파라로그 유전자좌 (문헌 [Deutsch, S. et al. 2004 J Med Genet 41, 908-915])를 먼저 전방향 프라이머 **5'-GTTGTTCTGCAAAAAACCTTCGA-3'** (서열 9) 및 역방향 프라이머 **5'-CTTGCCAGAAATACTTCATTACCATAT-3'** (서열 10)에 의해 동시 증폭시켰다. 염색체 21 및 1 파라로그를 표적으로 하도록 2개의 염색체-특이적 태크만 프로브를 설계하였는데, 그들의 서열은 각각 **5'-(FAM)TACCTCCAT AATGAGTAAA(MGBNFQ)-3'** (서열 11) 및 **5'-(VIC)CGTACCTCTGTAATGTGTAA(MGBNFQ)-3'** (서열 12)였다. 각각의 반응은 1X 태크만 유니버설 PCR 마스타 믹스 (어플라이드 바이오시스템즈), 각각의 프라이머 450 nM 및 각각의 프로브 125 nM을 함유하였다. 총 반응 부피는 5 μL/웰이었다. 반응을 50°C에서 2분 동안 개시하고, 이어서 95°C에서 10분 동안 진행하고, 95°C에서 15초 및 60°C에서 1분의 50주기를 수행하였다. 모든 실시간 PCR 실험은 ABI PRISM 7900HT 서열 검출 시스템 (어플라이드 바이오시스템즈) 상에서 수행하였고, SDS 2.2.2 소프트웨어 (어플라이드 바이오시스템즈)의 "절대 정량화" 응용 프로그램에 의해 형광 데이터를 수집하였다. 디폴트 기준선 및 매뉴얼 임계값을 사용하였다. 염색체 21 또는 염색체 1에 대해 양성인 웰의 수를 기록하여 SPRT 분석에 제공하였다. SPRT에 의해 질환 분류가 가능해질 때까지 1개 이상의 384-웰 플레이트가 분석될 것이다.

[0234] IV. 미세유체-기반 디지털 PCR의 사용

[0235] **A. 디지털 RNA-SNP**

[0236] 본 실시예는 미세유체-기반 디지털 PCR을 사용한 디지털 PCR 분석의 성능을 증명한다. 플루이다임 바이오마크 (Fluidigm BioMark)<sup>TM</sup> 시스템을 사용하는 이 접근법의 변형이 그에 제한되지 않는 예로서 본원에 예시되었다. 상기 시스템은 실행 1회당 9000회 초과인 디지털 PCR을 수행할 수 있다.

[0237] 정배수체 또는 삼염색체 21 태아를 가진 임신 여성으로부터 태반 조직 및 모체 말초 혈액 샘플을 채취하였다. 프라이머 연장에 이은 질량 분광법에 의해 태반 DNA 샘플에서 PLAC4 유전자 상 rs8130833 SNP의 유전자형 결정을 수행하였다. 태반 및 모체 혈장 샘플로부터 RNA를 추출하였다.

[0238] 모든 RNA 샘플을 씨모스크립트 역전사효소 (인비트로젠)를 사용하여 유전자-특이적 역전사 프라이머 (**5'-AGTATATAGAACCATGTTTAGGCCAGA-3'** (서열 13))에 의해 역전사시켰다. 태반 cDNA 샘플에 대해, 웰 당 대략 1개의 주형 분자에서 후속 디지털 PCR 증폭이 수행될 수 있도록 계열 희석을 수행하였다.

[0239] 12.765 디지털 어레이 (플루이다임)에 의해 바이오마크 시스템<sup>TM</sup> (플루이다임) 상에서 디지털 PCR을 수행하였다. 각각의 디지털 어레이는 12개의 샘플-검정 혼합물을 수용하기 위한 12개의 패널로 이루어져 있다. 각각의 패널을 7 nL 반응물/웰을 수행하기 위한 765개 웰로 추가 분할하였다. PLAC4 유전자 상의 rs8130833 SNP 영역을 전방향 프라이머 (**5'-TTTGTATTGCAACACCATTTGG-3'** (서열 14)) 및 앞서 기재된 유전자-특이적 역전사 프라이머에 의해 증폭시켰다. rs8130833 SNP의 2개의 대립유전자를 각각 표적으로 하는 2개의 대립유전자-특이적 태크만 프로브를 설계하였다. 그들의 서열은 G 및 A 대립유전자에 대하여 각각 **5'-(FAM)TCGTCGCTAACTTG(MGBNFQ)-3'** (서열 15) 및 **5'-(VIC)ATTCGTCATCTAACTTG(MGBNFQ)-3'** (서열 16)였다. 반응 부피 10 μL로 2X 태크만 유니버설 PCR 마스타 믹스를 사용하여 1개의 어레이 패널에 대한 반응을 설정하였다. 각각의 반응은 1X 태크만 유니버설 PCR 마스타 믹스, 각각의 프라이머 572 nM, 대립유전자-G-특이적 프로브 53.5 nM, 대립유전자-A-특이적 프로브 178.5 nM 및 cDNA 샘플 3.5 μL를 함유하였다. 각각의 태반 cDNA 샘플에 대해서는 1개의 반응 패널이 사용된 반면, 각각의 모체 혈장 샘플에 대해서는 12개의 패널이 사용되었다. 샘플-검정 혼합물을 나노플렉스(NanoFlex)<sup>TM</sup> IFC 제어기 (플루이다임)에 의해 디지털 어레이에 로딩하였다. 반응은 바이오마크<sup>TM</sup> 시스템에서 수행하였다. 반응을 50°C에서 2분 동안 개시하고, 이어서 95°C에서 10분 동안 진행하고, 95°C에서 15초 및 57°C에서 1분의 40주기를 수행하였다.

[0240] 1개의 정배수체 및 2개의 T21 이형접합체 태반으로부터의 태반 RNA 샘플을 765-웰 반응 패널에서 분석하였다.

각각의 샘플에 대해, A 또는 G 대립유전자 (둘 다는 아님)에 대해 양성인 것을 포함하는 정보제공 웰의 수를 계수하였다. 모든 정보제공 웰 중 과다표출된 대립유전자의 비율 ( $P_r$ )을 결정하였다. 이들 실험의 웰 당 정확한 평균 참조 주형 농도 ( $m_r$ )에 적당한 SPRT 곡선을 적용하여, 실험적으로 얻은  $P_r$ 이 정배수체를 지시하는지 T21 샘플을 지시하는지를 결정하였다. 도 15a에 나타난 바와 같이, 상기 접근법을 사용하여 모든 RNA 샘플이 정확하게 분류되었다.

[0241] 정배수체 태아를 가진 4명의 여성 및 삼염색체 21 태아를 가진 1명의 여성으로부터 채취한 혈장 RNA 샘플을 추가로 시험하였다. 각각의 샘플을 12개의 765-웰 반응 패널, 즉 혈장 RNA 샘플 당 9180회의 반응으로 분석하였다. 도 15b는 상기 혈장 RNA 샘플의 12개 패널 각각에 대한 정보제공 웰의 수를 나타낸다. 표에 나타난 바와 같이, 혈장 샘플 중 주형 농도는 임의의 1개의 반응 패널 중의 정보제공 웰의 수로는 SPRT 분류에 충분하지 않을 정도로 묶였다. 상기 샘플을 정배수체 샘플로 분류하기 전에, 3개의 반응 패널로부터의 정보제공 웰을 합해야 했다 (도 15c). 도 15c는 2 내지 12개의 패널로부터 종합한 데이터를 사용하여 모든 혈장 사례가 정확하게 분류될 수 있었음을 나타낸다.

[0242] 디지털 PCR을 수행하기 위한 매뉴얼 방법과 비교해서, 상기 미세유체-기반 방법은 훨씬 더 빠르고 덜 노동 집약적이었다. 전체 과정이 2시간 반 내에 완료될 수 있었다.

[0244] **B. 삼염색체 18의 출생 전 검출을 위한 디지털 RNA-SNP 분석**

[0245] 본 실시예에서는, 삼염색체 18 태아에 대한 다형성 대립유전자의 비율 불균형을 검출하기 위해, 염색체 18 상의 태반-발현 전사체인 세르핀 펩티다제 억제제 클레이드 B (오브알부민) 멤버 2 (SERPINB2) mRNA에 대한 디지털 PCR-기반 대립유전자 식별 검정을 사용하였다. QIAamp DNA 미니 키트 (독일 힐텐 소재의 퀴아젠(Qiagen)) 및 TRIzol 시약 (미국 캘리포니아주 칼즈바드 소재의 인비트로젠)을 제조업자의 프로토콜에 기재된 바와 같이 각각 사용하여, 태반 조직 샘플로부터의 DNA 및 RNA 추출을 수행하였다. 추출한 태반 RNA 샘플을 DNase I (인비트로젠) 처리하여 오염된 게놈 DNA를 제거하였다. 앞서 기재된 바와 같은 매스어레이 컴팩트(MassARRAY Compact) (미국 샌디에고 소재의 시퀀놈(Sequenom))를 사용하여 호모지너스 매스익스텐드(Homogenous MassEXTEND) (hME) 검정에 의해 태반 조직 DNA 샘플에서 SERPINB2 유전자 상 rs6098 SNP의 유전자형 결정을 수행하였다.

[0246] 씨모스크립트 역전사효소 (인비트로젠)를 사용하여 유전자-특이적 프라이머

**5'-CGCAGACTTCTCACAAACA-3'** (서열 17) (미국 아이오와주 코럴빌 소재의 인테그레이티드 DNA 테크놀로지스)에 의해 태반 조직 RNA 샘플 상에서 SERPINB2 전사체에 대한 역전사를 수행하였다. 모든 cDNA 샘플을, 반응 웰 당 1개의 주형 분자의 평균 농도에서 후속 디지털 PCR 증폭이 수행될 수 있도록 하는 농도로 희석하였다. 태크만 유니버설 PCR 마스타 믹스 (미국 캘리포니아주 포스터 시티 소재의 어플라이드 바이오시스템즈) 및 바이오마크™ PCR 시약 (미국 샌프란시스코 소재의 플루이다임)을 사용하여 디지털 PCR을

실정하였다. 전방향 프라이머 (서열 18) (인테그레이티드 DNA 테크놀로지스) 및 역방향 프라이머 (역전사에 사용된 유전자-특이적 프라이머와 동일한 것)는 600 nM의 농도로 사용하였다. SERPINB2 서열 상 rs6098 SNP의 A 또는 G 대립유전자를 표적으로 하는 2개의 태크만 프로브는

**5'-(FAM)CCACAGGGAATTATTT (MGBNFQ)-3'** (서열 19) 및 5'-

(VIC)CCACAGGGGATTATTT(MGBNFQ)-3' (서열 20) (어플라이드 바이오시스템즈)였다. FAM은 6-카복시플루오레세인이고, MGBNFQ는 마이너 그루브 결합 비-형광 켄치이며, 각각 300 nM 및 500 nM의 농도로 사용하였다. 각각의 샘플-시약 혼합물을 나노플렉스™ IFC 제어기 (플루이다임)를 사용하여 바이오마크™ 12.765 디지털 어레이 상의 765개 반응 웰에 분할하였다. 분할 후, 어레이를 열 증폭 및 형광 검출을 위해 바이오마크™ 실시간 PCR 시스템 (플루이다임)에 넣었다. 반응을 50°C에서 2분 동안 개시하고, 95°C에서 5분 동안 지속한 후, 95°C에서 15초 및 59°C에서 1분의 45주기를 수행하였다. 증폭 후, 정보제공 웰 (A 또는 G 대립유전자 단독에 대해 양성인 것)의 수 및 대립유전자 둘 다에 대해 양성인 웰의 수를 계수하여 순차 확률비 검증 (SPRT) 분석에 제공하였다.

[0247] 이형접합 정배수체 태아에 대해, A 및 G 대립유전자는 태아 게놈에서 동등하게 표출되어야 하는 반면 (1:1), 삼염색체 18에 대해서는, 한 대립유전자의 추가 카피가 존재함으로써 태아 게놈 내에 2:1의 비율이 주어질 것이다. 상이한 샘플의 해석에 대한 일련의 SPRT 곡선을 생성시켰다. 이들 곡선은 분류에 필요한 정보제공 웰의 주어진 총 수 ( $x$ 축)에 대한, 과다표출된 대립유전자에 대해 양성인 정보제공 웰의 예상 비율  $P_r$  ( $y$ 축)을 나타낸다. 각각의 샘플에 대해, 실험적으로 유도된  $P_r$ 을 예상  $P_r$  값과 비교하였다. 상부 곡선 위쪽의 샘플은 삼염색체 18로 분류되고, 하부 곡선 아래쪽의 샘플은 정배수체로 분류되었다. 두 곡선 사이의 공간은 분류불가능

한 영역이다.

[0248] 태아 삼염색체 18의 검출을 위한 디지털 RNA-SNP 분석의 실행가능성을 SERPINB2 유전자 상 rs6098 SNP를 사용하여 증명하였다. 먼저 이형접합체 사례를 확인하기 위해서, 질량 분광법에 의해 정배수체 및 삼염색체 18 태아를 갖는 대상체로부터의 태반 조직 DNA 샘플의 유전자형을 결정하였다. 9개의 정배수체 및 3개의 삼염색체 18 이형접합체 태반을 발견하여 디지털 RNA-SNP 분석에 제공하였다. 각각의 샘플에 대해  $P_r$  및  $m_r$ 을 계산하고, 상기  $m_r$ 값에 대한 SPRT 곡선을 질량 분류를 위해 사용하였다. 도 16a에 나타난 바와 같이, 모든 샘플이 정확하게 분류되었다. 삼염색체 18 태반에 대한  $P_r$ 값은 분류불가능한 영역의 위쪽인 반면, 정배수체 태반에 대한  $P_r$ 값은 상기 영역의 아래쪽에 속하였다.

[0249]  $m_r = 0.1, 0.2$  및  $0.3$ 을 기준으로 한 SPRT 곡선을 갖는 샘플을 도 16b에 나타내었다. 이들 데이터는 디지털 RNA-SNP 방법이 삼염색체 18 임신에 대한 매우 유용한 진단 도구임을 시사한다. 2개의 곡선은 분류불가능한 영역에 대한 경계선을 나타낸다. 상부 곡선 위쪽의 데이터 지점을 갖는 샘플은 이수체로 분류되고, 하부 곡선 아래쪽의 데이터 지점을 갖는 샘플은 정배수체로 분류되었다.

[0251] **C. 디지털 RCD 분석**

[0252] 본 실시예는 미세유체-기반 디지털 PCR을 사용한 디지털 RCD 분석의 성능을 증명한다. 플루이다임 바이오마크™ 시스템을 사용하는 이 접근법의 변형이 그에 제한되지 않는 예로서 본원에 예시되었다. 상기 시스템은 실험 1회당 9000회 초과 디지털 PCR을 수행할 수 있다.

[0253] 정배수체 또는 삼염색체 21 (T21) 태아를 가진 임신 여성으로부터 태반 조직, 모체 혈액 세포 및 양수 샘플을 얻었다. 10개의 T21 및 10개의 정배수체 사례로부터의 태반 DNA를 각각 동량의 정배수체 모체 혈액 세포 DNA와 혼합함으로써 20개의 50% DNA 혼합물을 제조하였다. 혼합물 샘플 중 정확한 태아 비율을 확실히 하기 위해, 추출한 DNA를 먼저 260 nm에서 광학 밀도 (OD) 측정에 의해 정량화하였다. 그 다음, 12.765 디지털 어레이 (플루이다임)를 사용하여 바이오마크™ 시스템 (플루이다임)에 의해 이들을 디지털 정량화하였다. 샘플의 정량화 검정은 참조 염색체에 대한 프로브를 사용한 것만 제외하고는 이하에 기재된 것과 동일하였다.

[0254] 50% DNA 혼합물 및 양수 샘플 내 염색체 양을, 염색체 1 상에 위치한 유전자좌에 대한 비-다형성 염색체 21 유전자좌의 디지털 PCR 분석에 의해 결정하였다. 먼저 염색체 21 및 1 상의 1쌍의 파라로그 유전자좌의 101-bp 앰플리콘을 전방향 프라이머 **5'-GTTGTTCTGCAAAAAACCTTCGA-3'** (서열 21) 및 역방향 프라이머 **5'-CTTGCCAGAAATACTTCATTACCATAT-3'** (서열 22)에 의해 동시 증폭시켰다. 염색체 21 및 1 파라로그 사이를 구별하도록 2개의 염색체-특이적 태크만 프로브를 설계하였는데, 그들의 서열은 각각 **5'-(FAM)TACCTCCATAATGAGTAAA(MGBNFQ)-3'** (서열 23) 및 **5'-(VIC)CGTACCTCTGTAATGTGTAA(MGBNFQ)-3'** (서열 24)였다. 본원에서 파라로그 유전자좌는 단지 예로서 사용되었다. 환언하면, 비-파라로그 유전자좌도 이러한 분석에 대해 사용될 수 있다.

[0255] 삼염색체 18 (T18)을 검출하기 위한 디지털 RCD 접근법의 용도를 증명하기 위해, 염색체 21 및 18 상의 파라로그 서열을 표적으로 하는 또다른 검정을 설계하였다. 먼저 염색체 21 및 18 상의 파라로그 유전자좌의 128-bp 앰플리콘을 전방향 프라이머 **5'-GTACAGAAACCACAAACTGATCGG-3'** (서열 25) 및 역방향 프라이머 **5'-GTCCAGGCTGTGGGCCT-3'** (서열 26)에 의해 동시 증폭시켰다. 염색체 21 및 18 파라로그 사이를 구별하도록 2개의 염색체-특이적 태크만 프로브를 설계하였는데, 그들의 서열은 각각 **5'-(FAM)AAGAGGCGAGGCAA(MGBNFQ)-3'** (서열 27) 및 **5'-(VIC)AAGAGGACAGGCAAC(MGBNFQ)-3'** (서열 28)였다. 본원에서 파라로그 유전자좌는 단지 예로서 사용되었다. 환언하면, 비-파라로그 유전자좌도 이러한 분석에 대해 사용될 수 있다.

[0256] 모든 실험은 12.765 디지털 어레이 (플루이다임)를 사용하여 바이오마크™ 시스템 (플루이다임) 상에서 수행하였다. 반응 부피 10  $\mu$ L로 2X 태크만 유니버설 PCR 마스타 믹스 (어플라이드 바이오시스템즈)를 사용하여 1개의 패널에 대한 반응을 설정하였다. 각각의 반응은 1X 태크만 유니버설 PCR 마스타 믹스, 각각의 프라이머 900 nM, 각각의 프로브 125 nM 및 50% 태반/모체 혈액 세포 DNA 샘플 3.5  $\mu$ L를 함유하였다. 샘플/검정 혼합물을

나노플렉스™ IFC 제어기 (플루이다임)에 의해 디지털 어레이에 로딩하였다. 검출을 위해 바이오마크™ 시스템 상에서 반응을 수행하였다. 반응을 50℃에서 2분 동안 개시하고, 이어서 95℃에서 10분 동안 진행하고, 95℃에서 15초 및 57℃에서 1분의 40주기를 수행하였다.

[0257] 정배수체 및 T21 50% 태반/모체 혈액 세포 DNA 샘플을 디지털 어레이 상에서 chr21/chr1 검정에 의해 분석하였다. 각각의 샘플에 대해, chr21 또는 chr1 마커 단독 (둘 다는 아님)에 대해 양성인 웰을 포함하는 정보제공 웰의 수를 계수하였다. 모든 정보제공 웰 중 과다표출된 마커의 비율 ( $P_r$ )을 결정하였다. 디지털 PCR 패널 중 어느 하나에 대한 웰 당 정확한 평균 참조 주형 농도 ( $m_r$ )에 적당한 SPRT 곡선을 적용하여, 실험적으로 얻은  $P_r$  이 정배수체 샘플을 지시하는지 T21 샘플을 지시하는지를 결정하였다. 결정이 이루어질 수 있을 때까지, 분류되지 않은 채로 남아있는 샘플에 대한 여분의 패널로부터 데이터를 종합하였다. 도 17에 나타낸 바와 같이, 모든 50% 태반/모체 혈액 세포 DNA 샘플이 상기 접근법을 사용하여 필요한 1 내지 4개 패널 범위의 데이터에 의해 정확하게 분류되었다. 도 18에 나타낸 바와 같이, 정확한 분류를 위한 결정 경계선을 나타내기 위해 SPRT 곡선을 또한 플롯팅하였다.

[0258] 정배수체 태아를 가진 23명의 여성 및 T21 태아를 가진 6명의 여성으로부터 얻은 양수 샘플에 대한 RCD 분석을 추가로 적용하였다. 각각의 샘플을 단일 765-웰 반응 패널에서 chr21/chr1 검정에 의해 분석하였다. 도 19는 SPRT 분류 개요를 나타낸다. 도 19에 나타낸 바와 같이, 모든 29개 샘플이 정확하게 분류되었다. 따라서, 디지털 RCD 방법은 출생 전 진단에 사용되는 다양한 종류의 샘플 유형, 예컨대 양수 및 용모막 용모 생검에서 초위성체 (문헌 [Levett LJ, et al. A large-scale evaluation of amnio-PCR for the rapid prenatal diagnosis of fetal trisomy. *Ultrasound Obstet Gynecol* 2001; 17: 115-8]) 또는 단일 뉴클레오티드 다형성 (SNP) (문헌 [Tsui NB, et al. Detection of trisomy 21 by quantitative mass spectrometric analysis of single-nucleotide polymorphisms. *Clin Chem* 2005; 51: 2358-62]) 마커 또는 실시간 비-디지털 PCR (문헌 [Zimmermann B, et al. Novel real-time quantitative PCR test for trisomy 21. *Clin Chem* 2002; 48: 362-3])을 사용하는 삼염색체의 검출에 대한 대안적인 접근법이다.

[0259] T18 사례를 검출하기 위한 시도로, 3개의 정배수체 및 5개의 T18 태반 DNA 샘플 상에 chr21/chr18 검정을 적용하였다. 모든 정보제공 웰 중 과다표출된 마커의 비율 ( $P_r$ )을 계산하였다. 이들 모두는 1개의 T18 사례가 정배수체로 잘못 분류된 것을 제외하고는 정확하게 분류되었다. 결과를 도 20에 요약하였다.

[0261] V. 질량 분광법 플랫폼에 대한 멀티플렉스 디지털 RCD 검정의 이용

[0262] 모체 혈액의 단위 부피 당 존재하는 혈장 DNA 분자의 수는 제한적이다 (문헌 [Lo YMD, et al. 1998 *Am J Hum Genet* 62, 768-7758]). 예를 들어, 임신 초기에 상염색체 유전자좌,  $\beta$ -글로빈 유전자의 중앙 모체 혈장 농도는 태아 및 모체 모두에 기인하여 986개 카피/mL인 것으로 밝혀진 바 있다 (문헌 [Lo YMD, et al. 1998 *Am J Hum Genet* 62, 768-7758]). 7,680개의 분자를 포획하기 위해서, 대략 8 mL의 모체 혈장에서부터 추출된 DNA가 필요할 것이다. 약 15 mL의 모체 혈액으로부터 수득가능한 상기 부피의 혈장은 통상적인 실험의 한계이다. 그러나, 본 발명자들은 chr21 및 참조 염색체 표적의 다중 세트를 디지털 RCD 분석을 위해 합할 수 있는 것으로 생각한다. 5쌍의 chr21 및 참조 염색체 표적에 대해, 단지 1.6 mL의 모체 혈장이 분석에 필요한 주형 분자의 수를 제공하기 위해 필요할 것이다. 멀티플렉스 단일 분자 PCR을 수행할 수 있었다. 이러한 멀티플렉스 단일 분자 분석의 견고성은 단일 분자 일배체형화에 대해 미리 입증된 바 있다 (문헌 [Ding, C. and Cantor, CR. 2003 *Proc Natl Acad Sci U S A* 100, 7449-7453]).

[0263] 일례에서, 태반 조직 및 모체 혈액 세포 샘플을 정배수체 또는 삼염색체 21 (T21) 태아를 가진 임신 여성으로부터 수득하였다. 5개의 정배수체 및 5개의 T21 태반 DNA 샘플을 각각 동일한 비율의 모체 혈액 세포 DNA와 혼합하여 50% 태아 DNA를 갖는 혈장 샘플을 모방한 10개의 DNA 혼합물을 생성하였다. 혼합물 샘플 중 정확한 태아 비율을 확실하게 하기 위해, 추출된 DNA를 먼저 260 nm에서의 광학 밀도 (OD) 측정으로 정량화하였다. 이어서, 이를 384-웰 포맷에서 실시간 PCR에 의해 디지털 방식으로 정량화하였다. 샘플을 정량화하기 위한 검정은 상기 디지털 RCD 분석의 예에 기재된 바와 동일하였다.

[0264] 50% 믹스 내 염색체 양을 염색체 1 상에 위치한 유전자좌에 대한 비-다형성 염색체 21 유전자좌의 디지털 PCR 분석으로 결정하였다. 상기 방법은 디지털 상대적 염색체 양 (RCD) 분석으로 칭해진다. 염색체 21 및 1 상의 한 쌍의 파라로그 유전자좌의 121-bp 앰플리콘 (각각의 프라이머 상에 10-mer 포함)을 전방향 프라이머 5'-ACGTTGGATGGTTGTTCTGCAAAAAACCTTCGA-3' (서열 29) 및 역방향 프라이머 5'-ACGTTGGATGCTGGCCAGAAATACTTCATTACCATAT-3' (서열 30)를 사용하여 동시 증폭시켰다. 염색체 21 및 염색체 1

사이의 염기 차이를 표적으로 한 연장 프라이머를 고안하였고, 그의 서열은 5'-CTCATCCTCACTTCGTACCTC-3' (서열 31)이다.

[0265] T21 사례를 검출하는 멀티플렉싱 디지털 PCR 검정의 유용성을 입증하기 위해서, 염색체 21 및 18 상의 파라로그 서열을 표적으로 한 또다른 디지털 RCD 검정을 고안하였다. 염색체 21 및 18 상의 파라로그 유전자좌의 148-bp 앰플리콘 (각각의 프라이머 상에 10-mer 포함)을 전방향 프라이머 5'-ACGTTGGATGGTACAGAAACCACAAACTGATCGG-3' (서열 32) 및 역방향 프라이머 5'-ACGTTGGATGGTCC AGGCTGTGGGCCT-3' (서열 33)를 사용하여 동시 증폭시켰다. 염색체 21 및 염색체 18 사이의 염기 차이를 표적으로 한 연장 프라이머를 고안하였고, 그의 서열은 5'-ACAAAAGGGGAAGAGG-3' (서열 34)이다.

[0266] 멀티플렉스 디지털 RCD 분석은 프라이머 연장 프로토콜을 이용하여 수행되었다. PCR 반응은 진앰프(GeneAmp) PCR 코어 시약 키트 (어플라이드 바이오시스템즈)를 사용하여 반응 부피 5  $\mu$ L로 설정하였다. 각각의 반응물은 1X 완충액 II, 2 mM MgCl<sub>2</sub>, 200  $\mu$ M dNTP 믹스, 0.2 U 앰플리태크 골드(AmpliTaq Gold), 각각 200 nM의 4종의 프라이머 및 50% DNA 믹스를 함유하였다. 상기 검정/샘플 혼합물을 384-웰 PCR 플레이트에 분배하고, 50°C에서 2분 동안 반응을 개시한 후, 95°C에서 10분 동안 진행하고, 95°C에서 15초 및 57°C에서 1분의 40회 주기를 수행하였다.

[0267] PCR 생성물을 쉬립프 알칼린 포스파타제 (SAP) 처리하여 혼입되지 않은 dNTP를 제거하였다. 상기 혼합물을 37°C에서 40분 동안, 이어서 85°C에서 5분 동안 인큐베이션하였다. 이어서, 프라이머 연장 반응을 수행하였다. 간단히 말해서, 연장 카테일 중 chr21/chr1 검정으로부터의 771 nM의 연장 프라이머, chr21/chr18 검정으로부터의 1.54  $\mu$ M의 연장 프라이머, 0.67 U 썬모시퀀나제(Thermosequenase) (시퀀놈), 및 각각 64  $\mu$ M의 ddCTP, ddGTP, dATP 및 dTTP를 SAP-처리된 PCR 생성물에 첨가하였다. 반응 조건은 94°C에서 2분, 이어서 94°C에서 5초, 50°C에서 5초 및 72°C에서 5초의 80회 주기였다. 16  $\mu$ L의 물 및 3 mg의 클린 레진(Clean Resin) (시퀀놈)을 최종 처리(clean up)를 위해 연장 생성물에 첨가하였다. 상기 혼합물을 회전기에서 20 내지 30분 동안 혼합한 후, 361 g에서 5분 동안 원심 분리단계를 수행하였다. 15 내지 25 nL의 최종 생성물을 매스어레이 나노분산기 S (시퀀놈)로 스펙트로칩(SpectroCHIP) (시퀀놈) 상에 분배하였다. 스펙트로칩으로부터의 데이터 획득을 매스어레이 분석기 콤팩트 질량 분광기 (시퀀놈)에서 수행하였다. 분석을 위해 질량 데이터를 매스어레이 타이퍼(MassARRAY Typer) (시퀀놈) 소프트웨어로 импорт하였다.

[0268] 5개의 정배수체 및 5개의 T21 50% 태반/모체 DNA 샘플을 이중 RCD 검정으로 분석하였다. 각 샘플에 대해, 오직 chr21 또는 chr1 또는 chr18 마커에만 양성인 웰을 포함하여 각각의 검정으로부터 정보제공 웰의 수를 계수하였다. 전체 정보제공 웰 중 chr21 마커의 비율 ( $P_r$ )을 각각의 RCD 검정에 대해 개별적으로 계산하였다. 이어서, 순차 확률비 검증 (SPRT)을 적용하여  $P_r$ 이 정배수체 샘플을 나타내는 지 또는 T21 샘플을 나타내는 지 결정하였다. 이렇게 하여, 각각의 플레이트를 2회 계수함으로써 필요한 웰의 수를 감소시켰다.

[0269] chr21/chr1 검정을 통상 먼저 적용하였다. 샘플이 미분류 상태로 있는 경우, chr21/chr18 검정으로부터의 값을 추가 계산에 포함시킬 것이다. 여분의 플레이트를 결정이 이루어질 수 있을 때까지 미분류 상태로 있는 샘플에 사용하였다. 도 21에 나타난 바와 같이, 모든 정배수체 50% 믹스 샘플을 단일 384-웰 플레이트를 사용하여 정확히 분류하였다. 여러 T21 사례는 정확한 분류를 위해 2개 이상의 플레이트를 필요로 하였다. 하나의 검정만을 이용한 경우, 분류 달성시 필요한 정보제공 웰의 수에 이르기 위해서 더 많은 플레이트가 필요할 것이다. 예를 들어, RCD 검정 중 어느 하나를 단독 이용한 경우에는 사례 N0230에 대한 데이터가 분류될 수 없었다. 그러나, 2개의 검정으로부터의 데이터를 합한 경우에는 정확한 분류가 이루어졌다. 이중 RCD 검정을 이용하지 않은 경우, 추가의 분석 플레이트가 필요할 것이다. 본 발명자들은 보다 높은 수준의 멀티플렉싱 검정으로 웰 수의 추가 감소를 예상할 것이다.

[0270] 또다른 예에서, 본 발명자들은 염색체 21, 및 다른 비-염색체-21 상염색체 상에 위치한 그의 상응하는 파라로그 파트너에 대한 4종의 상이한 앰플리콘을 표적으로 하는 4-플렉스 검정을 개발하였다. 이러한 4-플렉스 검정은 정배수체 및 삼염색체 21 임신 여성으로부터의 샘플의 디지털 RCD 분석, 이어서 SPRT 분류에 사용되었다. 태반 샘플로부터의 DNA 추출은 퀴아앰프 조직 키트 (독일 힐덴 소재의 퀴아젠)를 사용하여 수행되었다.

[0271] 상기 연구에 사용된 모든 태반 및 모체 연막 DNA 샘플을 먼저 나노드롭 분광광도계 (미국 델라웨어주 윌밍턴 소재의 나노드롭 테크놀로지)로 정량화하였다. DNA 농도를 6.6 pg/세포의 전환을 이용하여 게놈 당량 (GE)/ $\mu$ L로 전환하였다. 웰 당 대략 1개의 주형에 해당하는 DNA의 양을 DNA 샘플의 계열 희석으로 결정하였다. 이러한 조건하에, 본 발명자들은 대략 37%의 웰이 음성 증폭을 나타낼 것으로 예상할 것이다. 멀티플렉스 디지털 RCD

분석에서, 4 세트의 파라로그 서열 표적을 선택하였다: 염색체 21 및 1 상의 파라로그 유전자좌를 전방향 프라이머 5'-ACGTTGGATGTTGATGAAGTCTCATCTCTACTTCG-3' (서열 35) 및 역방향 프라이머 5'-ACGTTGGATGCAATAAGCTTGGCCAGAAATACT-3' (서열 36)를 사용하여 동시 증폭시켜 81 bp의 앰플리콘을 생성하였다. 염색체 21 및 7 상의 파라로그 유전자좌를 전방향 프라이머 5'-ACGTTGGATGGAATTAAGCTAAATCAGCCTGAACTG-3' (서열 37) 및 역방향 프라이머 5'-ACGTTGGATGGTTTCTCATAGTTTCATCGTAGGCTTAT-3' (서열 38)를 사용하여 동시 증폭시켜 82 bp의 앰플리콘을 생성하였다. 염색체 21 및 2 상의 파라로그 유전자좌를 전방향 프라이머 5'-ACGTTGGATGTCAGGCAGGGTTCTATGCAG-3' (서열 39) 및 역방향 프라이머 5'-ACGTTGGATGAGGCGGCTTCTGCTTCTG-3' (서열 40)를 사용하여 동시 증폭시켜 101 bp의 앰플리콘을 생성하였다. 염색체 21 및 6 상의 파라로그 유전자좌를 전방향 프라이머 5'-ACGTTGGATGGCTCGTCTCAGGCTCGTAGTT-3' (서열 41) 및 역방향 프라이머 5'-ACGTTGGATGTTTCTTCGAGCCCTTCTGG-3' (서열 42)를 사용하여 동시 증폭시켜 102 bp의 앰플리콘을 생성하였다. 각각의 반응물은 10X 완충액 II (어플라이드 바이오시스템즈), MgCl<sub>2</sub> 및 각각 100 nM의 프라이머를 함유하였다. 총 반응 부피는 5 µL/웰이었다. 95°C에서 5분 동안 반응을 개시한 후, 95°C에서 30초, 62°C에서 30초 및 72°C에서 30초의 45회 주기를 수행하고, 72°C에서 7분 동안 최종 연장을 수행하였다. 모든 통상적인 PCR 증폭은 진 앰프 PCR 시스템 9700 (어플라이드 바이오시스템즈) 상에서 수행되었다. 혼입되지 않은 뉴클레오티드를 쉬텀프 알칼린 포스파타제 (SAP) 처리로 비활성화시켰다. 각각의 반응물은 10X SAP 완충액 (시퀀놈) 및 SAP 효소 (시퀀놈)를 함유하였다. 2 µL의 SAP 믹스를 각각의 PCR에 첨가하였다. SAP 반응물을 37°C에서 40분 동안 및 85°C에서 5분 동안 인큐베이션하였다. SAP 처리 후, iPLEX 골드(Go1d) 키트 (시퀀놈)를 사용하여 PCR 생성물 상에서 프라이머 연장 반응을 수행하였다. 염색체 21 및 1 상의 파라로그 유전자좌에 대한 파라로그 서열 미스매치 (PSM)는 연장 프라이머 5'-GTCTCATCTCTACTTCGTACTC-3' (서열 43)에 의해 검사되었다. 염색체 21 및 7 상의 파라로그 유전자좌에 대한 PSM은 연장 프라이머 5'-TTTACGCTGTCCCATTT-3' (서열 44)에 의해 검사되었다. 염색체 21 및 2 상의 파라로그 유전자좌에 대한 PSM은 연장 프라이머 5'-GGTCTATGCAGGAGCCGAC-3' (서열 45)에 의해 검사되었다. 염색체 21 및 6 상의 파라로그 유전자좌에 대한 PSM은 연장 프라이머 5'-TGGGCGGGAGCGGACTTCGCTGG-3' (서열 46)에 의해 검사되었다. 각각의 반응물은 10X iPLEX 완충액 (시퀀놈), iPLEX 종결 믹스 (시퀀놈), iPLEX 효소 (시퀀놈), 및 1.03 µM로 사용된 염색체 21 및 6에 대한 PSM용 연장 프라이머를 제외하고는 각각 343 nM의 연장 프라이머를 함유하였다. 2 µL의 iPLEX 믹스를 5 µL의 PCR 생성물에 첨가하였다. 200-쇼트-사이클 프로그램에 따라 iPLEX 반응을 순환시켰다. 간단히 말해서, 샘플을 먼저 94°C에서 35초 동안 변성시킨 후, 52°C에서 5초 동안 어닐링하고, 80°C에서 5초 동안 연장하였다. 어닐링 및 연장 주기를 총 5회의 주기 동안 4회 이상 반복하고, 이어서 94°C, 5초의 변성 단계로 역 순환시킨 후, 다시 5회 주기의 어닐링 및 연장 루프를 수행하였다. 1회의 변성 단계와 함께 5회의 어닐링 및 연장 주기를 39회 반복하여 총 40회에 이르게 하였다. 72°C에서 3분 동안 최종 연장을 수행하였다. iPLEX 반응 생성물을 16 µL의 물로 희석하고, 각각의 PCR에 대해 6 mg의 수지로 탈염시켰다. 384-웰 플레이트를 1600 g에서 3분 동안 원심분리한 후, 스펙트로칩 (시퀀놈) 및 매트릭스-보조 레이저 탈착/이온화 비행시간형 (MALDI-TOF) 질량 분광법 MS 분석기 (시퀀놈) 상에 분배하였다.

[0272] 염색체 21에 대해서만 또는 4개의 검정 각각에 대한 참조 염색체에 대해서만 양성인 웰의 수를 독립적으로 기록하였다. 각각의 검정에 있어서, 염색체 21 및 참조 염색체에 대한 푸아송 교정된 분자 수를 계산하였다. 염색체 21에 대한 푸아송 교정된 분자 수의 합 뿐만 아니라 전체 4개의 검정으로부터의 푸아송 교정된 참조 염색체 수의 합을 계산하였고, 이는 4-플렉스 검정에 대한 정보제공 계수로 간주되었다. P<sub>r</sub> 값은 4-플렉스 검정에 대한 염색체 21 계수를 4-플렉스 검정에 대한 염색체 21 및 참조 염색체 계수의 합으로 나눈 값이었다. 실험적으로 유도된 P<sub>r</sub> 값을 SPRT 분석에 적용하였다. SPRT에 의해 질환 분류가 가능해질 때까지 1개 이상의 384-웰 플레이트를 분석할 것이다. 총 2개의 50% 정배수체 태반 계놈 DNA/50% 모체 연막 DNA 믹스 및 2개의 50% 삼염색체 21 태반 계놈 DNA/50% 모체 연막 DNA 믹스를 분석하였다.

[0273] 실험적으로 유도된 P<sub>r</sub> 값을 귀무가설 또는 대립가설 시험시의 예상 P<sub>r</sub> 값과 비교할 것이다. 다르게는, 주어진 수의 정보제공 계수에 대한 P<sub>r</sub>이 질환 분류에 필요한 통계적 신뢰 수준에 도달하지 못한 경우, 귀무가설 또는 대립가설은 모두 수용될 수 없을 것이다. 이러한 경우는 보다 많은 데이터가 이용가능해질 때까지 분류불가능한 것으로 간주되었다.

[0274] 각 샘플의 결과 및 SPRT 분류를 도 22a 및 22b에 표로 작성하였다. 2개의 정배수체 샘플은 SPRT 분류에 도달할 수 있기 전에 2개 및 5개의 384-웰 멀티플렉스 디지털 RCD 분석을 필요로 하였다. 4-플렉스 검정의 개별 구성으로부터의 데이터는 어떤 것도 SPRT에 의한 질환 분류가 가능하지 않았다. 삼염색체 21 샘플은 둘 다 단지 하



나의 384-웰 멀티플렉스 디지털 RCD 분석으로 각각 정확히 분류되었다. 유사하게, 4-플렉스 검정의 개별 구성으로부터의 데이터는 어떤 것도 SPRT에 의한 질환 분류가 가능하지 않았다. 그러나, 4-플렉스 검정으로부터의 혼합 계수는 정확한 SPRT 분류를 가능하게 하였다. 이들 데이터는, 멀티플렉스 디지털 RCD의 이용에 의해 정보 제공 계수의 유효 수가 단일-플렉스 디지털 RCD 검정의 이용과 비교하여 주어진 수의 수행된 디지털 PCR 분석에 대해 상당히 증가하였음을 예증하였다.

[0276] VI. 디지털 후성 상대적 염색체 양의 사용

[0277] 여기서, 본 발명자들은, 염색체 이수성에 관여하는 염색체 (즉, 삼염색체 21에서 염색체 21) 및 참조 염색체 상의 태아-특이적 DNA 메틸화 패턴 또는 다른 후성 변화를 나타내는 후성 마커를 디지털 PCR 분석에 적용하는, 소위 디지털 후성 상대적 염색체 양 (디지털 ERCD)이라 불리는 접근법을 약술한다. 정상 태아를 가진 임신 여성으로부터 추출된 혈장 DNA 내 참조 염색체 후성 마커에 대해 양성인 웰 수에 대한 염색체 21 후성 마커에 대해 양성인 웰 수의 비율은 참조 범위를 제공할 것이다. 상기 비율은 태아가 삼염색체 21을 갖는 경우 증가할 것으로 예상될 것이다. 하나 초과인 염색체 21 마커 및 하나 초과인 참조 염색체 마커를 본 분석에서 사용할 수 있다는 것이 당업자에게 명백하다.

[0278] 태아 (태반)-특이적 메틸화 패턴을 나타내는 염색체 21 상의 유전자의 일례는 완전카복실라제 합성효소 (Holocarboxylase synthetase, HLCS) 유전자이다. HLCS는 태반에서 과-메틸화되지만, 모체 혈액 세포에서는 저-메틸화되며, 이는 본원에 참조로 포함된 미국 특허 출원 제11/784499호에 포함된다. 태아 (태반)-특이적 메틸화 패턴을 나타내는 참조 염색체 상의 유전자의 일례는 염색체 3 [10] 상의 RASSF1A 유전자이다. RASSF1A는 태반에서 과-메틸화되지만, 모체 혈액 세포에서는 저-메틸화되며, 본원에 참조로 포함된 미국 특허 출원 제 11/784501호를 참조한다.

[0279] 모체 혈장을 이용한 태아 내 삼염색체 21의 디지털 PCR 검출에 대한 과-메틸화된 HLCS 및 과-메틸화된 RASSF1A의 적용에서, 모체 말초 혈액을 먼저 수집한다. 이어서, 상기 혈액을 원심분리하고, 혈장을 수집한다. 이어서, 혈장으로부터의 DNA를 당업계에 널리 공지된 기술, 예컨대 쿼아엠프 블러드(Blood) 키트 (쿼아젠)를 사용하여 추출한다. 이어서, 혈장 DNA를 1종 이상의 메틸화-감수성 제한 효소, 예컨대 HpaII 및 BstUI를 사용하여 소화시킨다. 이들 메틸화-감수성 제한 효소는 태아의 과-메틸화된 서열을 무손상 상태로 두는 반면, 이들 유전자의 모체, 비-메틸화된 버전은 절단할 것이다. 이어서, 소화된 혈장 DNA 샘플을, 제한 효소 처리되었지만 무손상인 HLCS 또는 RASSF1A 서열의 분자가 반응 웰 당 평균 대략 0.2 내지 1개 검출될 정도로 희석한다. 2개의 실시간 PCR 시스템을 사용하여 희석된 DNA를 증폭시킬 것이며; 하나는 2개의 프라이머, 및 HLCS 유전자에 특이적인 1개의 태크만 프로브 (서열이 비-메틸화된 경우 제한 효소(들)에 의해 절단될 영역 포함)를 사용하고; RASSF1A 유전자에 관한 다른 하나는 유사하게 2개의 프라이머 및 1개의 태크만 프로브를 사용한다. 후자의 RASSF1A 프라이머/프로브 세트와 관련하여, 일례가 문헌 [Chan et al 2006, Clin Chem 52, 2211-2218]에 기재되어 있다. HLCS 및 RASSF1A 표적에 관한 태크만 프로브는 상이한 형광 리포터, 예컨대 각각 FAM 및 VIC를 가질 것이다. 이어서, 384-웰 플레이트를 사용하여 디지털 PCR 실험을 수행한다. HLCS에 대해서만 양성으로 스코어링된 웰의 수 및 RASSF1A에 대해서만 양성으로 스코어링된 웰의 수를 계수할 것이며, 이들 계수의 비율을 구할 것이다. HLCS:RASSF1A 비율은 정상 정배수체 태아를 가진 임신 여성과 비교하여 삼염색체 21 태아를 가진 임신 여성으로부터 얻은 모체 혈장에 대해 보다 높을 것으로 예상될 것이다. 과다표출의 정도는 디지털 PCR 수행시 웰 당 평균 참조 주형 농도에 따라 달라질 것이다.

[0280] 이들 결과를 스코어링 하는 다른 방법, 예를 들어 RASSF1A에 대해 동시에 양성인 것에 관계 없이 HLCS에 대해 양성인 웰의 수를 계수하고, 반대로 HLCS에 대해 동시에 양성인 것에 관계 없이 RASSF1A에 대해 양성인 웰의 수를 계수하는 방법이 가능할 것이다. 또한, 비율을 계산하는 대신 HLCS 및 RASSF1A 계수의 총 수 또는 이들의 차이를 이용하여 태아의 삼염색체 21 상태를 나타낼 수 있었다.

[0281] 플레이트에서 디지털 PCR을 수행하는 것 외에도, 디지털 PCR의 다양한 변형, 예를 들어 미세유체 칩, 나노리터 PCR 마이크로플레이트 시스템, 에멀전 PCR, 폴로니 PCR, 롤링-썬클 증폭법, 프라이머 연장법 및 질량 분광법 등을 이용할 수 있다는 것이 당업자에게 명백할 것이다. 디지털 PCR의 이러한 변형은 단지 예로서 제시되며, 제한적이지 않다.

[0282] 실시간 PCR 외에도, 질량 분광법과 같은 방법을 이용하여 디지털 PCR 결과를 스코어링할 수 있다는 것이 당업자에게 명백할 것이다.

[0283] HLCS 및 RASSF1A의 태아 및 모체 버전을 식별하기 위해 메틸화-감수성 제한 효소를 사용하는 것 외에도, 메틸화

상태를 확인하기 위한 다른 방법, 예를 들어 바이숄라이트 변형법, 메틸화-특이적 PCR, 메틸화된 시토신에 대한 항체를 사용하는 면역침강법, 질량 분광법 등을 적용할 수도 있음이 당업자에게 명백할 것이다.

[0284] 또한, 본 특허 출원의 상기 예 및 기타 예에서 설명된 접근법을 모체 뇨, 양수, 자궁경관 세척액, 용모막 용모, 모체 타액 등을 비롯한 태아 DNA가 발견될 수 있는 기타 체액에서도 사용할 수 있음이 당업자에게 명백할 것이다.

[0286] VII. 에멀전 PCR 및 기타 전략을 이용한 대규모 병렬 게놈 서열분석

[0287] 여기서, 본 발명자들은, 핵산 분자의 디지털 관독이 모체 혈장 중 태아 염색체의 이수성, 예를 들어 삼염색체 21을 검출하기 위해 사용될 수 있기 때문에 핵산 분자의 디지털 관독의 또다른 예를 기재할 것이다. 태아 염색체 이수성은 염색체 또는 염색체 영역의 비정상적 용량(들)으로부터 기인한다. 비침습성 시험법이 감수성 및 특이성이 높기 때문에 오진을 최소화하기에 바람직하다. 그러나, 태아 DNA는 낮은 절대 농도로 존재하고, 모체 혈장 및 혈청 중의 모든 DNA 서열에 대해 소량의 부분으로 나타난다. 따라서, 디지털 PCR 샘플링 표적화에 특이적인 유전자좌의 개수는 동일한 생물학적 표본 내에서 무한히 증가될 수 없다. 따라서, 많은 세트의 특이적 표적 유전자좌를 분석함으로써, 디지털 PCR 샘플링 수행 횟수를 증가시키지 않으면서도 표본으로부터 얻을 수 있는 데이터의 양을 증가시킬 수 있다.

[0288] 따라서, 본 실시양태는 모체 배경 핵산을 함유한 생물학적 샘플 중에 소량의 집단으로서 존재하는 제한된 양의 태아 핵산으로부터 추론될 수 있는 유전 정보의 양을 최대화함으로써 태아 염색체의 이수성을 비-침습적으로 검출한다. 한 측면에서, 얻어진 유전 정보의 양은 정확한 진단을 위해서는 충분하지만, 요구되는 생물학적 샘플의 투입량 및 비용을 포괄할 정도로 지나치게 과량인 것은 아니다.

[0289] 454 플랫폼 (로슈) (문헌 [Margulies, M. et al. 2005 Nature 437, 376-380]), 일루미나 게놈 분석기 (또는 솔렉사 플랫폼) 또는 솔리드 시스템 (어플라이드 바이오시스템즈) 또는 헬리코스 트루 싱글 몰레클 DNA 서열분석 기술 (문헌 [Harris TD et al. 2008 Science, 320, 106-109]), 퍼시픽 바이오사이언시즈의 단일 분자, 실시간 (SMRT™) 기술, 및 나노기공 서열분석 (문헌 [Soni GV and Meller A. 2007 Clin Chem 53: 1996-2001]) 상에서 달성가능한 것과 같은 대규모 병렬 서열분석은 고도로 다중화된 표본으로부터 단리된 많은 핵산 분자를 병렬 방식으로 서열분석한다 (문헌 [Dear Brief Funct Genomic Proteomic 2003; 1: 397-416]). 이들 각각의 플랫폼은, 클로닝에 의해 확장된 또는 비-중복된 핵산 절편의 단일 분자를 서열분석한다.

[0290] 대다수의 서열분석 관독이, 수십만 내지 수백만 또는 더 가능하게는 수백만 또는 수백억 정도로, 각각의 진행에서 각각의 샘플로부터 생성되기 때문에, 얻어진 서열분석된 관독은 원래 표본 중 핵산 중 혼합물의 대표적인 프로파일을 형성한다. 예를 들어, 서열분석된 관독의 일배체형, 전자체 및 메틸화 프로파일은 원래 표본의 그것과 유사하다 (문헌 [Brenner et al Nat Biotech 2000; 18: 630-634]; [Taylor et al Cancer Res 2007; 67: 8511-8518]). 각각의 표본으로부터 서열을 대량 샘플링하기 때문에, 핵산 풀(pool)의 서열분석으로부터 몇배의 적용범위 또는 고 중복성으로 생성된 것과 같은 동일한 서열의 수는 또한 원래 샘플 중 특정 핵산 종의 수 또는 유전자위 수에 대한 좋은 정량적 표출이다.

[0291] 한 실시양태에서, 랜덤 서열분석은 임신 여성의 혈장 중에 존재하는 DNA 절편에 대해 수행되고, 본래 태아 또는 모체로부터 유래되는 게놈 서열을 얻는다. 랜덤 서열분석은 생물학적 샘플에 존재하는 핵산 분자의 랜덤 부분을 샘플링 (서열분석)하는 것을 포함한다. 서열분석이 랜덤하게 이루어지기 때문에, 핵산 분자 (및 따라서 게놈)의 상이한 하위세트 (분획)를 각각의 분석에서 서열분석할 수 있다. 본 실시양태는 이러한 하위세트가 샘플마다 및 심지어 동일한 샘플을 사용하여 이루어질 수 있는 분석마다 다양한 경우라 해도 작용할 것이다. 분획의 예는 게놈의 약 0.1%, 0.5% 또는 1%이다. 다른 실시양태에서, 분획은 이들 값 중 적어도 하나이다.

[0292] 이어서, 바이오인포메틱스 절차를 이용하여 이들 DNA 서열 각각을 인간 게놈에 위치시킬 수 있다. 상기 서열의 비율은 이들이 인간 게놈의 반복 영역 또는 개인간 변이, 예를 들어 카피수 변이(copy number variation)에 적용되는 영역에 존재하기 때문에 추후 분석으로부터 감소될 수 있다. 이에 따라, 관심 염색체 및 하나 이상의 다른 염색체의 양을 측정할 수 있다.

[0293] 한 실시양태에서, 염색체 이수성, 예를 들어 염색체 21 또는 염색체 18 또는 염색체 13에 잠재적으로 포함되는 염색체의 파라미터 (예를 들어, 분획 표출)는 예를 들어 바이오인포메틱스 절차의 결과로부터 계산될 수 있다. 분획 표출은 모든 서열 (예를 들어, 모든 염색체의 일부 측정치) 또는 특정 하위세트의 염색체 (예를 들어, 시험된 것과는 다른 염색체)의 양을 기준으로 얻어질 수 있다.

[0294] 이후, 상기 분획 표출은 정상 태아 (즉, 정배수체)를 포함하는 임신에서 확립된 참조 범위와 비교한다. 일부

변형된 절차에서, 참조 범위를 특정 모계 혈장 샘플 중 태아 DNA의 분획 농도 (f)에 따라 조정하는 것이 가능하다. f 값은 태아가 남성인 경우 Y 염색체에 대해 맵핑가능한 서열을 이용하여 서열분석 데이터세트로부터 측정될 수 있다. f 값은 또한 별도 분석에서, 예를 들어 태아 후성 마커를 사용하여 (문헌 [Chan KCA et al 2006 Clin Chem 52, 2211-8]) 또는 단일 뉴클레오티드 다형성 분석으로부터 측정될 수 있다.

[0295] 한 측면에서, 표본 중의 핵산 풀을 100% 미만의 게놈 적용범위에서 서열분석하고, 포획된 핵산 분자의 비율 중에서, 대부분의 각각의 핵산 종만을 일단 서열분석하는 경우에, 특정 유전자위 또는 염색체의 용량 불균형이 또한 정량적으로 측정될 수 있다. 다시 말하면, 유전자위 또는 염색체의 용량 불균형은 표본의 모든 맵핑가능한 서열분석된 태그 중 상기 유전자위의 표출 백분율로부터 추론된다.

[0296] 대규모 병렬 게놈 서열분석 접근법에 대한 한 측면에서, 모든 염색체로부터의 대표적인 데이터는 동시에 생성될 수 있다. 특정 절편의 기원은 해당 시간 전에 선택되지 않는다. 서열분석은 랜덤하게 이루어지고, 이어서 데이터베이스 조사를 수행하여 상기 특정 절편이 어디로부터 유래되었는지 알 수 있다. 이것은 염색체 21로부터의 특정 절편 및 염색체 1로부터의 또다른 절편을 증폭시키는 상황과 대비된다.

[0297] 한 예에서, 이러한 서열의 부분은 상기 설명 예에서 염색체 21과 같은 이수성에 포함되는 염색체로부터 유래할 것이다. 이러한 서열분석 실행으로부터 다른 서열이 다른 염색체로부터 유래할 것이다. 다른 염색체와 비교되는 염색체 21의 상대적 크기를 고려함으로써, 참조 범위 내에서, 상기 서열분석 실행으로부터의 염색체 21-특이적 서열의 정규화 도수를 얻을 수 있다. 태아가 삼염색체 21을 갖는 경우, 서열분석 실행으로부터의 염색체 21-유래된 서열의 정규화 도수는 증가할 것이고, 따라서 삼염색체 21이 검출된다. 정규화 도수의 변화 정도는 분석된 샘플 중 태아 핵산의 분획 농도에 따라 좌우될 것이다.

[0298] 한 실시양태에서, 본 발명자들은 인간 게놈 DNA 및 인간 혈장 DNA 샘플의 단일-말단 서열분석을 위해 일루미나 게놈 분석기를 사용하였다. 일루미나 게놈 분석기는 유동 셀로 명명된 고체 표면 상에서 포획된, 클로닝에 의해 확장된 단일 DNA 분자를 서열분석한다. 각각의 유동 셀은 8개의 개별 표본 또는 표본 풀의 서열분석을 위한 8개 레인을 포함한다. 각각의 레인은 오직 인간 게놈 서열의 30억개의 염기쌍의 분획인 약 200Mb의 서열을 생성할 수 있다. 각각의 게놈 DNA 또는 혈장 DNA 샘플은 유동 셀의 한 레인을 이용하여 서열분석하였다. 생성된 짧은 서열 태그를 인간 참조 게놈에 대해 정렬시키고, 염색체 기원을 기록하였다. 각각의 염색체에 대해 정렬된 서열분석된 개별 태그의 총 갯수를 인간 참조 게놈 또는 비질환 대표 표본으로부터 예상되는 바와 같은 각각의 염색체의 상대적인 크기와 비교하여 표로 만들었다. 이렇게 하여, 염색체 획득 또는 소실을 확인하였다.

[0299] 기재된 접근법은 현재 기재된 유전자/염색체 용량 전략의 하나의 예시일 뿐이다. 별법으로, 페어링 말단의 서열분석을 수행할 수 있다. 문헌 [Campbell et al, Nat Genet 2008; 40: 722-729]에 기재된 바와 같이 참조 게놈에서 예상되는 것으로부터 서열분석된 절편의 길이를 비교하는 대신, 정렬된 서열분석된 태그의 갯수를 카운팅하고, 염색체 위치에 따라 분류하였다. 염색체 영역 또는 전체 염색체의 획득 또는 소실은 태그 카운트를 참조 게놈에서의 예상되는 염색체 크기 또는 비질환 대표 표본의 예상되는 염색체 크기와 비교함으로써 측정하였다.

[0300] 또다른 실시양태에서, 실행 시 서열분석되는 핵산 풀의 분획은 서열분석 전에 추가로 하위 선택된다. 예를 들어, 올리고뉴클레오티드 어레이와 같이 혼성화를 기초로 하는 기술을 이용하여, 특정 염색체, 예를 들어 잠재적 이수체 염색체 및 시험된 이수성에 포함되지 않는 다른 염색체(들)로부터의 핵산 서열을 먼저 하위 선택할 수 있다. 또다른 예는 샘플 풀로부터의 핵산 서열의 특정 하위-집단이 서열분석 이전에 하위 선택되거나 또는 풍부화된다는 것이다. 예를 들어, 모체 혈장 중 태아 DNA 분자가 모체 배경 DNA 분자보다 더 짧은 절편으로 구성되어 있는 것으로 보고되어 있다 (문헌 [Chan et al Clin Chem 2004; 50: 88-92]). 따라서, 분자 크기에 따라 샘플 내 핵산 서열을 분획화하기 위해 당업자들에게 공지된 하나 이상의 방법, 예를 들어 겔 전기영동 또는 크기 배제 컬럼, 또는 미세유체-기반 접근법에 의한 방법을 사용할 수 있다. 그러나, 별법으로, 모체 혈장 중의 무-세포 태아 DNA 분석 예에서는, 태아 핵산 부분이 모체 배경을 억제하는 방법에 의해, 예컨대 포름알데히드의 첨가에 의해 풍부화될 수 있다 (문헌 [Dhallan et al JAMA 2004; 291 : 1114-9]).

[0301] 로슈 454 플랫폼, 어플라이드 바이오시스템즈 솔리드 플랫폼, 헬리코스 트루 싱글 몰레클 DNA 서열분석 기술, 퍼시픽 바이오사이언시즈의 단일 분자, 실시간 (SMRT™) 기술 및 나노기공 서열분석에 의해서와 같은 다른 단일 분자 서열분석 전략이 본 출원에서 유사하게 사용될 수 있다.

[0302] 결과 및 (예를 들어, 서열분석 및 파라미터 계산을 위한) 추가 논의의 예는 공동 출원된 출원 "게놈 서열분석을 사용한 태아 염색체 이수성 진단" (대리인 관리 번호 016285-005220US)에서 발견될 수 있고, 이것은 본원에 참

고로 포함된다. 분별값을 측정하기 위해 본원에 기재된 방법은, 예를 들어 본 섹션에 기재된 바와 같이, 반응물을 서열분석하는 경우 실행될 수 있음을 주목한다.

[0303] 모체 혈장 중 태아 DNA의 분획 농도의 측정은 또한 서열분석 진행과 별도로 수행될 수 있다. 예를 들어, Y 염색체 DNA 농도는 실시간 PCR, 미세유체 PCR 또는 질량 분석기를 사용하여 미리 측정할 수 있다. 사실, 태아 DNA 농도는 여성 태아에 적용할 수 있는 Y 염색체 이외의 유전자좌를 사용하여 측정될 수 있다. 예를 들어, 첸 등(Chan et al)은 태아-유래된 메틸화된 RASSF1A 서열이 모체 유래의 비메틸화된 RASSF1A 서열을 배경으로 하여 임신 여성의 혈장에서 검출됨을 밝혀냈다(문헌 [Chan et al, Clin Chem 2006;52:2211-8]). 따라서, 태아 DNA의 분획 농도는 메틸화된 RASSF1A 서열의 양을 전체 RASSF1A (메틸화 및 비메틸화 RASSF1A) 서열의 양으로 나눔으로써 측정될 수 있다.

[0304] 혈액 응고 동안 모체 혈액 세포로부터 DNA가 방출되기 때문에, 본 발명을 실행하기 위해서는 모체 혈장이 모체 혈청보다 바람직할 것으로 예상된다. 따라서, 혈청을 사용하는 경우, 태아 DNA의 분획 농도가 모체 혈청에서보다 모체 혈장에서 더 낮을 것으로 예상된다. 다르게 말하면, 모체 혈청을 사용하는 경우, 동일한 임신 여성으로부터 동시에 얻어지는 혈장 샘플과 비교했을 때, 진단하고자 하는 태아 염색체 이수성에 대해 더 많은 서열이 생성될 필요가 있을 것으로 예상된다.

[0305] 그러나, 태아 DNA의 분획 농도를 측정하기 위한 또다른 대안 방법은 임신 여성 및 태아 사이의 다형성 차이의 정량화를 통하는 것일 것이다(문헌 [Dhallan R, et al. 2007 Lancet, 369, 474-481]). 이러한 방법의 예는 임신 여성이 동형접합성이고 태아가 이형접합성인 다형성 부위를 표적으로 할 것이다. 태아-특이적 대립유전자의 양을 공통 대립유전자의 양과 비교하여 태아 DNA의 분획 농도를 측정할 수 있다.

[0306] 비교 게놈 혼성화, 마이크로어레이 비교 게놈 혼성화, 정량적 실시간 중합효소 연쇄 반응 (하나 이상의 특이적 서열(들)을 검출하고 정량화함)을 비롯한, 염색체 이상을 검출하기 위한 기존의 기술과 달리, 대규모 병렬 서열 분석은 미리 측정되거나 또는 미리 한정된 세트의 DNA 서열의 검출 또는 분석에 따라 좌우되지 않는다. 표본 풀로부터의 DNA 분자의 랜덤한 대표 분획을 서열분석한다. 다양한 염색체 영역에 대해 정렬된 상이한 서열 태그의 갯수는 종양성 DNA를 함유하거나 또는 함유하지 않은 표본 사이에서 비교한다. 염색체 이상은 표본에서 임의의 소정 염색체 영역에 대해 정렬된 서열의 갯수 (또는 백분율)의 차이로 나타낼 것이다.

[0307] 또다른 예에서, 혈장 무-세포 DNA에 대한 서열분석 기술을 이용하여 특정 암의 검출을 위해 혈장 DNA에서의 염색체 이상을 검출할 수 있다. 상이한 암은 한 세트의 전형적인 염색체 이상을 갖는다. 다중 염색체 영역의 변화 (증폭 및 결실)가 사용될 수 있다. 따라서, 증폭된 영역에 대해 정렬된 서열의 비율이 증가하고, 감소된 영역에 대해 정렬된 서열의 비율이 감소될 것이다. 염색체 당 표출 백분율은 전체 게놈과 관련하여 임의의 소정 염색체의 게놈 표출의 백분율로서 표현된 참조 게놈 중 각각의 상응하는 염색체에 대한 크기와 비교할 수 있다. 직접 비교 또는 참조 염색체에 대한 비교가 또한 사용될 수 있다.

[0309] VIII. 돌연변이 검출

[0310] 모체 혈장 중의 태아 DNA는 소량의 집단으로 존재하며, 모체 혈장 DNA 중 평균 3% 내지 6%가 태아에 의한다. 이러한 이유로 인해서, 본 분야에서 행해진 기존 연구의 대부분은 태아가 부계로부터 물려받고 모체 혈장 중의 대다수의 모체 DNA 배경과 구별되는 DNA 표적을 검출하는데 집중되었다. 이와 같이 기존에 검출된 표적의 예는 Y 염색체상의 SRY 유전자 [Lo YMD et al. 1998 Am J Hum Genet, 62, 768-775] 및 모계가 RhD-음성인 경우의 RHD 유전자 [Lo YMD et al. 1998 N Engl J Med, 339, 1734-1738]를 포함한다.

[0311] 태아 돌연변이 검출을 위해서 모체 혈장을 사용하는 기존의 전략은 부계가 보인자(carrier)인 상염색체 우성 상태로 제한되며, 부계 및 모계가 상이한 돌연변이를 갖는 경우에는 직접적인 돌연변이 검출을 수행하거나 또는 연관 분석 [Ding C. et al 2004 Proc Natl Acad Sci USA 101, 10762-10767]을 실시하여 상염색체 열성 질환은 제외한다. 이러한 기존 전략에는 유의한 한계가 있었다. 예를 들어, 남성 파트너와 여성 파트너가 동일한 돌연변이를 갖는 쌍에서는 모체 혈장 중에서의 직접적인 돌연변이 검출로는 의미있는 태아 진단을 실시할 수 없다.

[0312] 이러한 시나리오는 도 23에 예시되어 있다. 상기 시나리오에서는, 3종의 가능한 태아 유전자형, NN, NM 및 MM 이 존재하는데, 여기서 N은 정상적인 대립유전자를 나타내고, M은 돌연변이 대립유전자를 나타낸다. 돌연변이 대립유전자의 예는 낭포성 섬유증, 베타-지중해빈혈, 알파-지중해빈혈, 겸상적혈구 빈혈, 척수 근위축증, 선천성 부신 과형성 등을 일으키는 것을 포함한다. 이러한 장애의 다른 예는 OMIM (Online Mendelian Inheritance in Man) [ www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM&itool=toolbar]에서 찾을 수 있다. 모체 혈장에서,

대부분의 DNA는 모계로부터 유래되며 NM일 것이다. 상기한 3종의 태아 유전자형 중 어느 것에서도, 모체 혈장 중에서의 독특한 검출을 허용하는 임의의 독특한 태아 대립유전자는 존재하지 않는다. 따라서, 통상적인 전략은 여기에 적용될 수 없다.

[0313] 본원에 기재한 실시양태는 이러한 시나리오를 취급할 수 있다. 모계 및 태아가 둘다 NM인 시나리오에서는 N 대립유전자 및 M 대립유전자가 대립유전자 균형을 이루고 있을 것이다. 그러나, 모계가 NM이고 태아가 NN인 경우에는 모체 혈장 중에서 N 대립유전자가 과다표출된 대립유전자 불균형이 일어날 것이다. 한편, 모계가 NM이고 태아가 MM인 경우에는 모체 혈장 중에서 M 대립유전자가 과다표출된 대립유전자 불균형이 일어날 것이다. 따라서, 태아 돌연변이 검출을 위한 귀부가설에서는 태아가 NM 유전자형인 경우에 대립유전자 불균형이 존재하지 않는다고 가정한다. 대립가설에서는, 대립유전자 불균형이 존재하고 N 또는 M 대립유전자 중 어느 것이 과다표출되었는지의 여부에 따라 태아 유전자형은 NN 또는 MM일 수 있다고 가정한다.

[0314] 대립유전자 불균형의 존재 또는 부재는 본원에 기재한 실시양태를 사용한 디지털 PCR로 결정될 수 있다. 첫번째 시나리오에서, 특정 부피의 모체 혈장은 모계로부터의 50개 세포 및 태아로부터의 50개 세포의 100개 세포로부터 방출된 DNA를 함유한다. 따라서, 이러한 부피의 혈장 중에서 태아 DNA의 분획 농도는 50%이다. 모계가 유전자형 NM인 경우에는 50개의 N 대립유전자 및 50개의 M 대립유전자가 모계에 의한다. 태아가 유전자형 NM인 경우에는 50개의 N 대립유전자 및 50개의 M 대립유전자가 태아에 의한다. 따라서, N 대립유전자와 M 대립유전자 (각각 총 100개 카피) 사이에는 대립유전자 불균형이 존재하지 않는다. 한편, 태아가 NN 유전자형인 경우에는 이러한 부피의 혈장 중에 100개의 태아-유래의 N 대립유전자가 존재할 것이다. 따라서, 총 150개의 N 대립유전자와 50개의 M 대립유전자가 존재할 것이다. 즉, N과 M 사이에는 N이 M에 대하여 3:1의 비율로 과다표출된 대립유전자 불균형이 존재할 것이다.

[0315] 반대의 상황에서, 태아가 MM 유전자형인 경우에는 이러한 부피의 혈장 중에 100개의 태아-유래의 M 대립유전자가 존재할 것이다. 따라서, 150개의 M 대립유전자와 50개의 N 대립유전자가 존재할 것이다. 즉, N과 M 사이에는 M이 N에 대하여 3:1의 비율로 과다표출된 대립유전자 불균형이 존재할 것이다. 이러한 대립유전자 불균형은 디지털 PCR로 결정될 수 있다. 양성 웰의 수가 더 적은 대립유전자가 참조 주형으로 간주된다. 디지털 RNA-SNP 및 디지털 RCD 분석과 유사하게, 디지털 PCR 실험에서 대립유전자의 실제 분포는 피아송 확률 밀도 함수에 의해 좌우된다. 따라서, 본 시나리오에서는 대립유전자 불균형의 이론적인 정도가 3:1이지만, 대립유전자 불균형의 예상 정도는 디지털 PCR 분석 동안에 웰 당 평균 주형 농도에 따라 달라질 것이다. 따라서, SPRT 분석에서와 같은, 웰 당 평균 참조 주형 농도 ( $m_r$ )에 적절한 해석 분별값이 사례 분류에 사용될 필요가 있다.

[0316] 추가로, 결정할 필요가 있는 대립유전자 불균형의 정도는 태아 DNA의 분획 농도에 따라 달라진다. 상기한 예와는 반대로, 특정 부피의 모체 혈장이 모계로부터의 90개 세포 및 태아로부터의 10개 세포의 100개 세포로부터 방출된 DNA를 함유하는 경우를 가정한다. 따라서, 이러한 부피의 혈장 중에서 태아 DNA의 분획 농도는 10%이다. 모계가 유전자형 NM인 경우에는 90개의 N 대립유전자 및 90개의 M 대립유전자가 모계에 의한다. 태아가 유전자형 NM인 경우에는 10개의 N 대립유전자 및 10개의 M 대립유전자가 태아에 의한다. 따라서, N 대립유전자와 M 대립유전자 (각각 총 100개 카피) 사이에는 대립유전자 불균형이 존재하지 않는다. 한편, 태아가 NN 유전자형인 경우에는 이러한 부피의 혈장 중에 20개의 태아-유래의 N 대립유전자가 존재할 것이다. 따라서, 총 110개의 N 대립유전자와 90개의 M 대립유전자가 존재할 것이다.

[0317] 즉, N과 M 사이에는 N이 M에 대하여 과다표출된 대립유전자 불균형이 존재할 것이다. 반대의 상황에서, 태아가 MM 유전자형인 경우에는 이러한 부피의 혈장 중에 20개의 태아-유래의 M 대립유전자가 존재할 것이다. 따라서, 110개의 M 대립유전자와 90개의 N 대립유전자가 존재할 것이다. 즉, N과 M 사이에는 M이 과다표출된 대립유전자 불균형이 존재할 것이다. 태아 DNA의 분획 농도가 10%인 경우에 대립유전자 불균형의 이론적인 정도는 110:90일 것이고, 이는 상기 예에서 나타난 바와 같이 50% 태아 DNA가 존재하는 경우의 3:1 비율과 상이하다. 따라서, SPRT 분석에서와 같은, 태아 DNA의 분획 농도에 적절한 해석 분별값이 사례 분류에 사용될 필요가 있다.

[0318] 따라서, 혈장 DNA를 추출할 것이다. 혈장 샘플 중 모체 및 태아 DNA의 양은 예를 들어 기준에 수립된 실시간 PCR 검정 [Lo, et al. 1998 Am J Hum Genet 62, 768-775] 또는 당업자에게 공지된 다른 유형의 정량화 수단, 예를 들어 SNP 마커 [Dhallan R et al. 2007 Lancet, 369, 474-481] 및 태아 후성 마커 [Chan KCA et al. 2006 Clin Chem, 52, 2211-2218]를 이용하여 정량한다. 태아 DNA의 백분율(%)을 계산한다. 이후, 디지털 PCR 분석 동안에 각 반응 웰이 평균 1개의 주형 분자 (N 대립유전자일 수도 있고, 또는 M 대립유전자일 수도 있음)를 함유하도록 정량된 혈장 DNA 샘플 (예를 들어 희석하거나 농축된 것)을 제조한다. 한쌍의 프라이머, 및 2개의

테크만 프로브 (1개는 N 대립유전자에 특이적이고, 다른 하나는 M 대립유전자에 특이적임)를 사용하여 디지털 PCR 분석을 수행한다. 오직 M에만 양성인 웰의 수 및 오직 N에만 양성인 웰의 수를 계수한다. 이들 웰의 비율을 사용하여 대립유전자 불균형의 증거가 있는지 여부를 결정한다. 대립유전자 불균형의 통계적인 증거는 당업자에게 공지된 방법, 예컨대 SPRT를 사용하여 찾을 수 있다. 이러한 분석의 한 변형법에서는 M에만 양성이거나 M 및 N에 양성인 웰의 수를 계수하는 것도 가능하고, N에만 양성이거나 M 및 N에 양성인 웰의 수를 계수하는 것도 가능하며, 이들 계수치의 비율을 도출하는 것도 가능하다. 한번 더, 대립유전자 불균형의 통계적인 증거는 당업자에게 공지된 방법, 예컨대 SPRT를 사용하여 찾을 수 있다.

- [0319] 디지털 RMD (상대적 돌연변이 양(Relative Mutant Dosage))라고 불리는 태아 유전자 돌연변이의 양 결정을 여성/남성 (XX/XY) DNA 혼합물을 사용하여 검증하였다. 남성 및 여성으로부터의 혈액 세포 DNA를 남성 DNA와 각각 혼합하여, 도 24a에 나타낸 바와 같이, XY의 배경에 XX 또는 XY 유전자형을 각각 25% 및 50%의 분획 농도로 갖는 샘플을 수득한다.
- [0320] 추가로, 혈액 세포 샘플을 또한 12명의 남성 및 12명의 여성 대상체로부터 수득하였다. 여성 혈액 세포 DNA (유전자형 XX)를 3배 과량의 남성 혈액 세포 DNA (유전자형 XY)와 각각 혼합하여, XY 유전자형을 갖는 DNA 75%의 배경에 XX 유전자형을 갖는 DNA 25%를 함유하는 12개의 DNA 혼합물을 생성하였고, 이 결과는 도 24b에 나타냈다.
- [0321] SPRT 분석의 목적은 배경 DNA에 존재하는 소량의 유전자형을 결정하는 것이었다. 75% XY DNA의 배경에 25% XX DNA를 함유하는 DNA 혼합물에서, 소량의 대립유전자는 DNA 75%에서 유래된 Y이다. 샘플 중의 DNA 25%가 XX 유전자형이기 때문에, 샘플 중에 총 200개의 DNA 분자가 존재하는 경우에는 150개의 분자가 XY 개체에서 유래된 것이다. 따라서, Y 대립유전자의 수는 75개일 것으로 예상된다. 또한, 남성 DNA (유전자형 XY)의 비율에 의한 X 대립유전자의 수 역시 75개이다. 여성 (유전자형 XX)에 의한 X 대립유전자의 수는 50 (= 2×25)이다. 따라서, X/Y 비율은  $125/75 = (1 + 25\%)/(1 - 25\%) = 5/3$ 이다.
- [0322] 본 연구의 두번째 부분에서, 혈액 세포 샘플을 베타-글로빈 유전자, 즉, 헤모글로빈, 베타 (HBB) 유전자에 HbE (G→A) 및 CD41/42 (CTTT/-) 돌연변이를 갖는 남성 및 여성 대상체로부터 수득하였다. 모든 가능한 유전자형 (MM, MN 또는 NN)을 갖는 남성 태아를 보유하는 이형접합 모계 (MN (여기서, M = 돌연변이, N = 야생형))로부터 수득한 모체 혈장 샘플을 모방하기 위해서, 야생형 대립유전자에 대해 동형접합 (NN)이거나 2종의 돌연변이 중 하나에 대해 이형접합 (MN)인 남성으로부터의 혈액 세포 DNA를 동일 돌연변이에 대하여 이형접합 (MN)인 여성으로부터 수집한 혈액 세포 DNA 샘플과 각각 혼합하였다. 이로써, 다양한 분획 농도의 남성/돌연변이 DNA를 갖는 DNA 혼합물이 생성되었다. 또한, CD41/42 결실에 대해 동형접합 (MM)인 여성으로부터의 혈액 세포 DNA 샘플도 사용하여 DNA 혼합물을 제조하였다. SPRT 분류를 위한 남성 비율을 정확하게 하기 위해서, 각 DNA 혼합물 중 남성 DNA의 분획 농도를 ZFY/X 검정으로 결정하였다.
- [0323] 디지털 ZFY/X 검정을 이용하여 SPRT를 검증하고 또한 DNA 혼합물 중 남성 DNA의 분획 농도를 결정하였다. 염색체 X에서의 아연 핑거(Zinc Finger) 단백질 서열 (ZFX) 및 염색체 Y에서의 아연 핑거 단백질 서열 (ZFY)의 양을 디지털 PCR 분석으로 결정하였다. 전방향 프라이머 **5'-CAAGTGCTGGACTCAGATGTAACCTG-3'** (서열 47) 및 역방향 프라이머 **5'-TGAAGTAATGTCAGAAGCTAAAACATCA-3'** (서열 48)를 사용하여 ZFX 및 ZFY 유전자좌의 87-bp 앰플리콘을 우선 동시 증폭시켰다. 염색체 X 및 Y 파라로그를 구별하는 2종의 염색체-특이적 테크만 프로브를 설계하였고, 이의 서열은 각각 **5'-(VIC)TCTTTAGCACATTGCA(MGBNFQ)-3'** (서열 49) 및 **5'-(FAM)TCTTTACCACACTGCAC(MGBNFQ)-3'** (서열 50)였다.
- [0324] DNA 혼합물 중의 돌연변이 양을 돌연변이 대립유전자에 대한 정상 대립유전자의 디지털 PCR 분석으로 결정하였다. HbE 돌연변이의 경우, 정상 및 돌연변이 대립유전자의 87-bp 앰플리콘을 전방향 프라이머 **5'-GGGCAAGGTGAACGTGGAT-3'** (서열 51) 및 역방향 프라이머 **5'-CTATTGGTCTCCTTAAACCTGTCTTGTAA-3'** (서열 52)로 우선 동시 증폭시켰다. 정상 (G) 및 돌연변이 (A) 대립유전자 사이를 구별하는 2종의 대립유전자-특이적 테크만 프로브를 설계하였고, 이의 서열은 각각 **5'-(VIC) TTGGTGGTGAGGCC (MGBNFQ)-3'** (서열 53) 및 **5'-(FAM) TTGGTGGTAAGGCC (MGBNFQ)-3'** (서열 54)였다. HbE 돌연변이에 대한 결과를 도 25에 나타냈다.
- [0325] CD41/42 결실 돌연변이의 경우, 정상 및 돌연변이 대립유전자의 87-bp 및 83-bp 앰플리콘을 전방향 프라이머 **5'-TTTTCCCACCCTTAGGCTGC-3'** (서열 55) 및 역방향 프라이머 **5'-ACAGCATCAGGAGTGGACAGATC-3'** (서열 56)

각각을 사용하여 우선 동시 증폭시켰다. 정상 (결실 없음) 및 돌연변이 (결실 있음) 대립유전자 사이를 구별하는 2종의 대립유전자-특이적 태크만 프로브를 설계하였고, 이의 서열은 각각 5'-(VIC) CAGAGGTTCTTTGAGTCCT (MGBNFQ)-3' (서열 57) 및 5'-(FAM) AGAGGTTGAGTCCTT (MGBNFQ)-3' (서열 58)였다. HbE 돌연변이에 대한 결과를 도 26a 및 도 26b에 나타냈다.

[0326] 이러한 실험은 12.765 디지털 어레이즈 (플루이다임)를 사용하는 바이오마크™ 시스템 (플루이다임)에서 수행하였다. 1개의 패널에 대한 반응은 2× 태크만 유니버설 PCR 마스터 믹스 (어플라이드 바이오시스템즈)를 사용하여 반응 부피 10 μl로 설정하였다. CD41/42 및 ZFY/X 검정의 경우, 각 반응물은 1× 태크만 유니버설 PCR 마스터 믹스, 900 nM씩의 각 프라이머, 125 nM씩의 각 프로브 및 1 ng/μl의 DNA 혼합물 3.5 μl를 함유하였다. HbE 검정의 경우, 정상 (G) 및 돌연변이 (A) 대립유전자를 표적으로 하는, 각각 250 nM 및 125 nM의 프로브를 첨가하였다. 샘플/검정 혼합물을 나노플렉스™ IFC 제어기 (플루이다임)로 디지털 어레이에 로딩하였다. 반응은 신호 검출을 위해 바이오마크™ 시스템에서 수행하였다. 반응을 50℃에서 2분 동안 개시한 후에 95℃에서 10분 동안 진행하고, 95℃, 15초 및 57℃ (ZFY/X 및 CD41/42의 경우) 또는 56℃ (HbE의 경우), 1분의 50회 주기를 수행하였다. 각 경우마다 적어도 1개의 반응 패널을 사용하였고, 결정이 이루어질 수 있을 때까지 미분류 상태로 있는 샘플을 위한 여분의 패널로부터 데이터를 합하였다.

[0327] 또한, 디지털 PCR이 당업자에게 공지된 방법, 예를 들어 미세유체 칩, 나노리터 PCR 마이크로플레이트 시스템, 에멀전 PCR, 플로니 PCR, 롤링-써클 증폭법, 프라이머 연장법 및 질량 분광법을 이용하여 수행될 수 있다는 것이 당업자에게 명백할 것이다.

[0329] IX. 암을 이용한 예

[0330] 한 실시양태에서, 본 발명을 수행하여 양성 종양에서 발생할 수 있는 것과 같이 대립유전자 비율이 벗어나거나 벗어나지 않은 샘플을 분류할 수 있다. 한 측면에서, 각 경우에 오직 A 대립유전자에 대해서만 양성 신호를 갖는 웰의 수, 오직 G 대립유전자에 대해서만 양성 신호를 갖는 웰의 수, 및 2종의 대립유전자 둘다에 대해서 양성 신호를 갖는 웰의 수를 디지털 PCR로 결정하였다. 참조 대립유전자는 양성 웰의 수가 더 적은 대립유전자로 정의된다 (2종의 대립유전자가 둘다 동일한 수의 양성 웰을 갖는 시나리오가 있다고 한다면, 둘 중 어느 것이라도 참조 대립유전자로 사용될 수 있음). 참조 대립유전자에 대해 음성인 웰의 총 수를 사용하여, 다른 대립유전자가 양성인지 여부와는 관계 없이 푸아송 확률 밀도 함수에 따라 웰 당 참조 대립유전자의 평균 농도 추론치 ( $m_r$ )를 계산하였다. 본 연구진은 상기 계산을 예시하는 가정적인 예를 사용한다.

[0331] 96웰 반응에서, 20개 웰은 A 대립유전자에 대해 양성이고, 24개 웰은 G 대립유전자에 대해 양성이며, 28개 웰은 이들 2개의 대립유전자 모두에 대해 양성이다. A 대립유전자에 대하여 양성인 웰의 수가 더 적기 때문에, A 대립유전자를 참조 대립유전자로 간주한다. 참조 대립유전자에 대해 음성인 웰의 수는  $96 - 20 - 28 = 48$ 개이다. 따라서,  $m_r$ 은 푸아송 분포로 계산될 수 있고,  $-\ln(48/96) = 0.693$ 이다.

[0332] LOH 검출과 관련하여, 귀무가설에서는 1개의 대립유전자 결실이 존재하여 샘플에서 대립유전자 비율이 벗어나지 않았다고 가정한다. 이러한 가정하에서, 2종의 대립유전자에 대한 양성 웰의 수의 예상 비율은 1:1이고, 이에 따라 잠재적으로 과다표출된 대립유전자를 함유하는 정보제공 웰 (오직 1종의 대립유전자에 대해서만 양성인 웰)의 예상 비율은 0.5이다.

[0333] 299LOH 검출과 관련하여, 대립가설에서는 샘플의 세포 중 50%에서 1개의 대립유전자 결실이 존재하여 샘플에서 대립유전자 비율이 벗어나 있다고 가정한다. 과다표출된 대립유전자와 참조 대립유전자 사이의 대립유전자 비율이 2:1이기 때문에, 웰 당 과다표출된 대립유전자의 평균 농도는 참조 대립유전자의 평균 농도의 2배일 것이다. 그러나, 과다표출된 대립유전자에 양성인 웰의 수는 참조 대립유전자에 양성인 웰의 수의 단순히 2배가 아니라 푸아송 분포에 따를 것이다.

[0334] 정보제공 웰은 A 또는 G 대립유전자에 양성이지만 상기 2종의 대립유전자 둘다에는 양성인 웰로 정의된다. 대립유전자 비율이 벗어난 샘플의 경우에 과다표출된 대립유전자를 함유하는 웰의 수의 예상 비율의 계산치는 표 600에 나타낸 바와 같다. 상기 예에서, 종양 세포의 50%에서 LOH가 존재하는 경우에는 웰 당 G 대립유전자의 평균 농도가  $2 \times 0.693 = 1.386$ 일 것이다. LOH가 종양 세포의 50% 초과에 존재하는 경우에는 웰 당 G 대립유전자의 평균 농도가 하기 식에 따를 것이다:  $1/[1 - (\text{LOH를 갖는 비율})] \times m_r$ .

[0335] G 대립유전자에 양성인 웰의 예상 비율은  $1 - e^{-1.386} = 0.75$  (즉, 75% 또는 72개의 웰)일 것이다. A 또는 G 대립유전자에 대한 웰의 양성인 비율이 독립적이라고 가정할 때,  $0.5 \times 0.75 = 0.375$ 의 웰이 A 및 G 대립유전자 둘다

에 양성일 것이다. 따라서,  $0.5 - 0.375 = 0.125$ 의 웰이 오직 A 대립유전자에만 양성일 것이고,  $0.75 - 0.375 = 0.375$ 의 웰이 오직 G 대립유전자에만 양성일 것일 것이다. 따라서, 정보제공 웰의 비율은  $0.125 + 0.375 = 0.5$ 일 것이다. G 대립유전자를 갖는 정보제공 웰의 예상 비율은  $0.375/0.5 = 0.75$ 이다. 이후,  $P_r$ 에 대한 이러한 예상 값을 사용하여, 샘플에 대립유전자 비율이 벗어난 경우 (즉, 여기서는 LOH)가 존재하는지의 여부를 결정하기 위한 적절한 SPRT 곡선을 구축할 수 있다.

[0336] 이어서, 디지털 PCR 분석을 통해 실험적으로 결정된, 비-참조 대립유전자를 갖는 정보제공 웰의 실제 비율 ( $P_r$ )을 사용하여, 귀무가설 또는 대립가설 중 어느 것이 받아들여질 지, 또는 더 많은 웰을 사용한 추가의 분석이 필요할 지를 결정하였다. 귀무가설 또는 대립가설을 받아들이기 위한  $P_r$ 의 결정 한계값은 임계값 우도 비율 8을 기준으로 하여 계산되는데, 이는 상기 값이 대립유전자 불균형이 존재하는 샘플과 존재하지 않는 샘플을 암 검출과 관련하여 식별하는데 만족스러운 성능을 제공하는 것으로 나타났기 때문이다 ([Zhou, W, et al. (2001) Nat Biotechnol 19, 78-81], [Zhou et al 2002 (상기 문헌)]). 상기 예에서, 정보제공 웰의 수는  $20 + 24 = 44$ 이고, 실험적으로 수득된  $P_r$ 은  $24/44 = 0.5455$ 이다. 결정 한계값은 귀무가설을 받아들이기 위해서는  $\leq 0.5879$ 이고, 대립가설을 받아들이기 위해서는  $\geq 0.6739$ 이다. 따라서, 이러한 예에서의 샘플은 대립유전자 비율이 벗어난 NOT으로 분류된다.

[0337] 결론적으로, 본 연구진은 샘플 중의 서열 불균형을 검출하는 접근법을 알아냈다. 한 실시양태에서, 본 발명은 태아 염색체 이수성의 비-침습적 검출에 사용될 수 있고, 예를 들어 모체 혈장 중 태아 핵산의 분석을 통해 삼염색체 21을 검출할 수 있다. 이러한 접근법은 또한 양수, 용모막 용모 샘플, 모체 뇨, 자궁경관 샘플, 모체 타액 등을 포함하는, 태아 핵산을 함유하는 다른 생물학적 물질에도 적용될 수 있다. 첫째, 본 연구진은 삼염색체 21 태아를 보유하는 여성의 모체 혈장 중 염색체 21에서의 태반-발현된 전사체인 PLAC4 mRNA에서 SNP의 대립유전자 불균형을 결정하기 위한 본 발명의 용도를 입증하였다. 둘째, 본 연구진은 본 발명이 삼염색체 21의 비-침습적 태아 검출을 위해서 RCD (상대적 염색체 양(Relative Chromosome Dosage)) 분석을 통해 비-다형성 기체의 방법으로 사용될 수 있음을 입증하였다. 이러한 디지털 RCD-기반의 접근법은 태아 DNA를 함유하는 샘플 중 염색체 21의 전체 카피수가 참조 염색체에 비하여 과다표출되었는지 여부를 직접 평가하는 것을 수반한다. 디지털 RCD는 정교한 장치 없이도 25% 태아 DNA를 함유하는 샘플에서 삼염색체 21을 검출할 수 있게 한다. 본 연구진은 순차 확률비 검증 (SPRT)을 적용하여 디지털 PCR 데이터를 해석하였다. 컴퓨터 시뮬레이션 분석을 실시하여 질환 분류 알고리즘의 높은 정확성을 확인하였다.

[0338] 추가로, 본 연구진은 상기 접근법이 염색체 이수성 이외의 다른 형태의 핵산 서열 불균형의 결정, 예를 들어 모체 혈장 중 태아 돌연변이의 검출 또는 다형성 검출 및 혈장 중 종양-유래의 핵산 분석에 의한 악성 세포 계층의 국소적인 획득 및 소실의 검출에 적용될 수 있음을 알아냈다.

[0339] 본 명세서에 기재한 임의의 소프트웨어 성분 또는 기능은 임의의 적합한 컴퓨터 언어, 예를 들어 자바(Java), C++ 또는 펄(Perl) (예를 들어 통상의 기술 또는 객체 지향(object-oriented) 기술을 사용한 것)을 사용한 프로그래밍 언어에 의해 실행되는 소프트웨어 코드로서 시행될 수 있다. 소프트웨어 코드는 일련의 지시 또는 명령어로서 저장 및/또는 전송을 위한 컴퓨터 판독가능 매체에 저장될 수 있으며, 적합한 매체는 랜덤 액세스 메모리 (RAM, random access memory), 리드 온리 메모리 (ROM, read only memory), 자성 매체, 예컨대 하드 드라이브 또는 플로피 디스크, 또는 광학 매체, 예컨대 콤팩트 디스크 (CD) 또는 DVD (디지털 버서블 디스크(digital versatile disk)), 플래쉬 메모리 등을 포함한다. 컴퓨터 판독가능 매체는 이러한 저장 또는 전송 장치의 임의의 조합일 수 있다.

[0340] 이러한 프로그램은 또한 다양한 프로토콜, 예를 들어 인터넷에 따라 유선, 광학 및/또는 무선 네트워크를 통해 전송되도록 맞춰진 캐리어 신호를 사용하여 암호화되고 전송될 수도 있다. 이와 같이, 본 발명의 실시양태에 따른 컴퓨터 판독가능 매체가 이러한 프로그램으로 암호화된 데이터 신호를 사용하여 생성될 수 있다. 프로그램 코드로 암호화된 컴퓨터 판독가능 매체는 호환가능한 장치와 함께 패키징될 수도 있고, 또는 다른 장치와 별도로 제공 (예를 들어, 인터넷 다운로드에 의함)될 수도 있다. 임의의 이러한 컴퓨터 판독가능 매체는 1개의 컴퓨터 프로그램 제품 (예를 들어 하드 드라이브 또는 전체 컴퓨터 시스템)상에 있거나 그 내부에 있을 수 있고, 시스템 또는 네트워크 내부의 상이한 컴퓨터 프로그램 제품상에 있거나 그 내부에 있을 수 있다. 컴퓨터 시스템은 모니터, 프린터 또는 본원에서 언급한 임의의 결과를 사용자에게 제공하기 위한 다른 적합한 디스플레이를 포함할 수 있다.

[0341] 컴퓨터 시스템의 예는 도 27에 나타냈다. 도 27에 나타낸 하위시스템은 시스템 버스 (2775)를 통해 상호연결된



다. 디스플레이 어댑터 (2782)에 커플링된 추가의 하위시스템, 예컨대 프린터 (2774), 키보드 (2778), 고정 디스크 (2779), 모니터 (2776) 등을 나타냈다. I/O 제어기 (2771)에 커플링된 주변기기 및 입력/출력 (I/O) 장치는 직렬 포트 (2777)과 같이 당업계에 공지된 임의의 수의 수단을 통해 컴퓨터 시스템에 연결될 수 있다. 예를 들어, 직렬 포트 (2777) 또는 외부 인터페이스 (2781)을 사용하여 컴퓨터 장치를 광범위한 범위의 네트워크, 예를 들어 인터넷, 마우스 입력 장치 또는 스캐너에 연결할 수 있다. 시스템 버스를 통한 상호연결은 중앙 프로세서 (2773)가 각각의 하위시스템과 소통하고 시스템 메모리 (2772) 또는 고정 디스크 (2779)로부터의 지시의 수행 및 하위시스템들 사이의 정보 교환을 제어하도록 한다. 시스템 메모리 (2772) 및/또는 고정 디스크 (2779)는 컴퓨터 관독가능 매체를 포함할 수 있다.

[0342] 본 발명의 예시적인 실시양태에 관한 상기 기재는 예시와 설명을 위해서 제공된 것이다. 이는 본 발명을 상기한 형태로만 한정하거나 제한하려는 것이 아니며, 상기한 기술에 비추어 많은 변형과 변화가 가능하다. 본 발명의 이론 및 이것의 실제적인 적용을 최적으로 설명하고 그에 따라 당업자가 본 발명을 고려되는 특별한 용도에 맞게 다양하게 변형시켜 다양한 실시양태에서 최적으로 사용할 수 있도록 하는 실시양태를 선택하여 기재하였다.

[0343] 본원에서 언급한 모든 간행물, 특허 및 특허 출원서는 어떠한 목적으로도 그 전문이 본원에 참고로 포함된다.

SEQUENCE LISTING

```

<110> Lo, Yuk-Ming Dennis
      Chiu, Rossa Mai Kwun
      Chan, Kwan Chee
      Zee, Benny Chung Ying
      Ka, Chun Chong
      The Chinese University of Hong Kong

<120> Determining Nucleic Acid Sequence
      Imbalance

<130> 016285-005210PC

<140> WO PCT/GB08/02524
<141> 2008-07-23

<150> US 60/951,498
<151> 2007-07-23

<160> 58

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic real-time PCR and digital RNA-SNP analysis
      PLAC4 gene-specific reverse transcription primer

<400> 1
agtatataga accatgttta ggccaga                27

<210> 2
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic real-time PCR PLAC4 primer

<400> 2
ccgctagggt gtcttttaag c                        21

<210> 3
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic real-time PCR PLAC4 primer

<400> 3
qtqttqcaat acaaaatqag ttct                    25
    
```

[0344]

```

<210> 4
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic real-time PCR PLAC4 fluorescent probe

<220>
<221> modified_base
<222> (1)...(11)
<223> a modified by 6-carboxyfluorescein (FAM)

<220>
<221> modified_base
<222> (14)...(14)
<223> c modified by minor groove binding non-fluorescent
quencher (MGBNFQ)

<400> 4
attgagcaa atc 14

<210> 5
<211> 104
<212> DNA
<213> Artificial Sequence

<220>
<223> single-stranded synthetic DNA oligonucleotide
amplicon

<400> 5
egcgcgtgg gtgttcttta agctattgga gcaattcaa atttgctta aagcaaaaga 60
aacTCatttt gtattgcaac accaggagta tcccsagggg ctcg 104

<210> 6
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RNA-SNP analysis non-intron
spanning forward primer

<400> 6
tttgatttgc aacaccattt gg 22

<210> 7
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic allele-G-specific TaqMan probe targeting
rs8130833 SNP of PLAC4

<220>
<221> modified_base
<222> (1)...(11)
<223> t modified by 6-carboxyfluorescein (FAM)

```

```

<220>
<221> modified_base
<222> (15)...(15)
<223> g modified by minor groove binding non-fluorescent
quencher (MGBNFO)

<400> 7
tcgtgtctta acttg 15

<210> 6
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic allele-A-specific TaqMan probe targeting
rs8130833 SNP of PLAC4

<220>
<221> modified_base
<222> (1)...(1)
<223> a modified by fluorescent reporter VIC

<220>
<221> modified_base
<222> (17)...(17)
<223> g modified by minor groove binding non-fluorescent
quencher (MGBNFQ)

<400> 8
attcgtcctc taacttg 17

<210> 9
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RCD analysis co-amplification
forward primer for paralogous loci on chromosome
21 and 1

<400> 9
gttgttctgc aaaaaacctt cga 23

<210> 10
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RCD analysis co-amplification
reverse primer for paralogous loci on chromosome
21 and 1

<400> 10
cttgccaga actacttcac taccatc 28

```

```

<210> 11
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic chromosome-specific TaqMan probe
      designed to target chromosome 21 paralog

<220>
<221> modified_base
<222> (1)...(1)
<223> t modified by 6-carboxyfluorescein (FAM)

<220>
<221> modified_base
<222> (19)...(19)
<223> a modified by minor groove binding non-fluorescent
      quencher (MGBNFQ)

<400> 11
taactcaata atgagtaaa 19

<210> 12
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic chromosome-specific TaqMan probe
      designed to target chromosome 1 paralog

<220>
<221> modified_base
<222> (1)...(1)
<223> c modified by fluorescent reporter VIC

<220>
<221> modified_base
<222> (20)...(20)
<223> a modified by minor groove binding non-fluorescent
      quencher (MGBNFQ)

<400> 12
cgtacctctg taatgtgtaa 20

<210> 13
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR PLAC4 gene-specific reverse
      transcription primer

<400> 13
agtatataga accatgttta gccacaga 27

```

```

<210> 14
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR amplification forward primer
for rs8130833 SNP region of PLAC4 gene

<400> 14
tttgrattgc aacaccattt gg                22

<210> 15
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic allele-G-specific TaqMan probe targeting
PLAC4 rs8130833 SNP

<220>
<221> modified_base
<222> {1}...{11}
<223> t modified by 5-carboxyfluorescein (FAM)

<220>
<221> modified_base
<222> {15}...{15}
<223> g modified by minor groove binding non-fluorescent
quencher (MGBNFQ)

<400> 15
tcgtcgtccta atttg                15

<210> 16
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic allele-A-specific TaqMan probe targeting
PLAC4 rs8130833 SNP

<220>
<221> modified_base
<222> {11}...{11}
<223> a modified by fluorescent reporter VIC

<220>
<221> modified_base
<222> {17}...{17}
<223> g modified by minor groove binding non-fluorescent
quencher (MGBNFQ)

<400> 16
atttcgcata taacttg                17

```

```

<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR amplification serpin
peptidase inhibitor clade B (ovalbumin) member 2
(SERPINB2) gene-specific reverse transcription
primer

<400> 17
cgcagacttc tcaccaasca                20

<210> 18
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR amplification serpin
peptidase inhibitor clade B (ovalbumin) member 2
(SERPINB2) forward primer

<400> 18
ctcagctctg caatcaatgc                20

<210> 19
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic allele-A-specific TaqMan probe targeting
serpin peptidase inhibitor clade B (ovalbumin)
member 2 (SERPINB2) rs6098 SNP

<220>
<221> modified base
<222> (1)...(1)
<223> c modified by 6-carboxyfluorescein (FAM)

<220>
<221> modified base
<222> (16)...(16)
<223> T modified by minor groove binding non-fluorescent
quencher (MGBNFQ)

<400> 19
ccaccgggaa ttattt                    16

<210> 20
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic allele-G-specific TaqMan probe targeting
serpin peptidase inhibitor clade B (ovalbumin)
member 2 (SERPINB2) rs6098 SNP

```

```

<220>
<221> modified_base
<222> {1}...{1}
<223> c modified by 6-carboxyfluorecein (FAM)

<220>
<221> modified_base
<222> {16}...{16}
<223> t modified by minor groove binding non-fluorescent
quencher (MGBNFQ)

<400> 20
ccacagggga ttattt 16

<210> 21
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR analysis co-amplification
forward primer for 101-bp amplicon for paralogous
loci on chromosome 21 and 1

<400> 21
gttqttctgc acaaacctt cga 23

<210> 22
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR analysis co-amplification
reverse primer for 101-bp amplicon for paralogous
loci on chromosome 21 and 1

<400> 22
cttgccaga aatcttcat taccat 28

<210> 23
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic chromosome-specific TaqMan probe
designed to target chromosome 21 paralog

<220>
<221> modified_base
<222> {1}...{1}
<223> t modified by 6-carboxyfluorescein (FAM)

<220>
<221> modified_base
<222> {19}...{19}
<223> a modified by minor groove binding non-fluorescent
quencher (MGBNFQ)

```

```

<400> 23
tacctccata atgagtaa 19

<210> 24
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> synthetic chromosome-specific TaqMan probe
designed to target chromosome 1 paralog

<220>
<221> modified_base
<222> [1]...[1]
<223> c modified by fluorescent reporter VIC

<220>
<221> modified_base
<222> [20]...[20]
<223> a modified by minor groove binding non-fluorescent
quencher (MGBNFO)

<400> 24
cgtacctctg taatgtata 20

<210> 25
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<221> synthetic digital RCD analysis co-amplification
forward primer for 120-bp amplicon for paralogous
loci on chromosome 21 and 18

<400> 25
gtacagaaac cacaaactga tcgg 24

<210> 26
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<221> synthetic digital RCD analysis co-amplification
reverse primer for 120-bp amplicon for paralogous
loci on chromosome 21 and 18

<400> 26
gtccaggctg tggcct 17

<210> 27
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<221> synthetic chromosome-specific TaqMan probe
designed to target chromosome 21 paralog

```



```

<220>
<221> modified_base
<222> (1)...(1)
<223> a modified by 6-carboxyfluorescein (FAM)

<220>
<221> modified_base
<222> (14)...(14)
<223> a modified by minor groove binding non-fluorescent
quencher (MGBNFQ)

<400> 27
aaqaqqcqaq qcaa 14

<210> 28
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic chromosome-specific TaqMan probe
designed to target chromosome 10 paralog

<220>
<221> modified base
<222> (1)...(1)
<223> a modified by fluorescent reporter VIC

<220>
<221> modified_base
<222> (15)...(15)
<223> a modified by minor groove binding non-fluorescent
quencher (MGBNFQ)

<400> 28
aaqaqqacag qcaac 15

<210> 29
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR analysis co-amplification
forward primer for 121-bp amplicon for paralogous
loci on chromosome 21 and 1

<400> 29
acgttgatg gttgttctgc aaaaaacctt cga 33

<210> 30
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR analysis co-amplification
reverse primer for 121-bp amplicon for paralogous
loci on chromosome 21 and 1

```

[0352]

```

<400> 30
acggtggatg cttggacaga aatacttcac taccatat
38

<210> 31
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR analysis co-amplification
extension primer for 121-bp amplicon for
paralogous loci on chromosome 21 and 1

<400> 31
ctcctcctca attgctacct c
21

<210> 32
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RCD analysis co-amplification
forward primer for 148-bp amplicon for paralogous
loci on chromosome 21 and 18

<400> 32
acggtggatg gtacagacac cacaaactga tcgg
34

<210> 33
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RCD analysis co-amplification
reverse primer for 148-bp amplicon for paralogous
loci on chromosome 21 and 18

<400> 33
acgttgatg gtcagagtg tgggct
27

<210> 34
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RCD analysis co-amplification
extension primer for 148-bp amplicon for
paralogous loci on chromosome 21 and 18

<400> 34
acaaaaggag gaagagg
17

<210> 35
<211> 35
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> synthetic digital RCD analysis co-amplification
forward primer for 81 bp amplicon for paralogous
loci on chromosome 21 and 1

<400> 35
acggttgatg ttgatgaagt ataatatata ctteg
35

<210> 36
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RCD analysis co-amplification
reverse primer for 81 bp amplicon for paralogous
loci on chromosome 21 and 1

<400> 36
acgttgatg caataagctt gaccagaaat act
33

<210> 37
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RCD analysis co-amplification
forward primer for 82 bp amplicon for paralogous
loci on chromosome 21 and 7

<400> 37
acgttgatg gaatttaagc taatcaagcc tgaatg
37

<210> 38
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RCD analysis co-amplification
reverse primer for 82 bp amplicon for paralogous
loci on chromosome 21 and 7

<400> 38
acgttgatg gtttataata gttcatagta ggattat
37

<210> 39
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RCD analysis co-amplification
forward primer for 101 bp amplicon for paralogous
loci on chromosome 21 and 2

<400> 39
acgttgatg taaggcaggg ttatatgaag
30

```

[0354]

```

<210> 40
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RCD analysis co-amplification
reverse primer for 101 bp amplicon for paralogous
loci on chromosome 21 and 2

<400> 40
acgcttggatg aggcggcttc ctgctctctg          30

<210> 41
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RCD analysis co-amplification
forward primer for 102 bp amplicon for paralogous
loci on chromosome 21 and 6

<400> 41
acgcttggatg gctcgtctca ggctcgtagt t          31

<210> 42
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital RCD analysis co-amplification
reverse primer for 102 bp amplicon for paralogous
loci on chromosome 21 and 6

<400> 42
acgcttggatg ttctctcggg cctctctctg          30

<210> 43
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic extension primer for paralogous sequence
mismatches (PSM) on paralogous loci on chromosomes
21 and 1

<400> 43
gtctcatctc tacttcgtac ctg                    23

<210> 44
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic extension primer for paralogous sequence
mismatches (PSM) on paralogous loci on chromosomes
21 and 7

```

```

<400> 44
ttttacgctg tccccattt
19

<210> 45
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic extension primer for paralogous sequence
mismatches (PSM) on paralogous loci on chromosomes
21 and 2

<400> 45
ggctctatgca ggagccgca
19

<210> 46
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic extension primer for paralogous sequence
mismatches (PSM) on paralogous loci on chromosomes
21 and 6

<400> 46
tgggagcggg agcggacttc gctgg
25

<210> 47
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR analysis co-amplification
forward primer for 87-bp amplicon for zinc finger
protein sequences on chromosome X (ZFX) and Y
(ZFY)

<400> 47
caagtgcctgg actcagatgt aactg
25

<210> 48
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR analysis co-amplification
reverse primer for 87-bp amplicon for zinc finger
protein sequences on chromosome X (ZFX) and Y
(ZFY)

<400> 48
tgaagtaatg taagaagota aaacatca
28

<210> 49
<211> 16
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> synthetic chromosome-specific TaqMan probe
      designed to target chromosome X paralog

<220>
<221> modified_base
<222> (1)...(1)
<223> t modified by fluorescent reporter VIC

<220>
<221> modified_base
<222> (16)...(16)
<223> a modified by minor groove binding non-fluorescent
      quencher (MGBNFQ)

<400> 49
tcttttagcac attgca                               16

<210> 50
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic chromosome-specific TaqMan probe
      designed to target chromosome Y paralog

<220>
<221> modified_base
<222> (1)...(1)
<223> t modified by 6-carboxyfluorescein (FAM)

<220>
<221> modified_base
<222> (17)...(17)
<223> c modified by minor groove binding non-fluorescent
      quencher (MGBNFQ)

<400> 50
tctttaccac actgca                               17

<210> 51
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR co-amplification forward
      primer for 87-bp amplicon of beta-globin
      (hemoglobin beta (HBB)) HbE mutation forward
      primer

<400> 51
gggcaagggtg aacgtggat                           19

<210> 52
<211> 29
<212> DNA
<213> Artificial Sequence

```

[0357]

```

<220>
<223> synthetic digital PCR co-amplification forward
primer for 87-bp amplicon of beta-globin
(hemoglobin beta (HBB)) HbE mutation reverse
primer

<400> 52
ctatgggct ccttaaacct gctctgtaa
29

<210> 53
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic allele-G-specific TaqMan probe targeting
beta-globin (hemoglobin beta (HBB)) HbE mutation
normal allele

<220>
<221> modified base
<222> [1]...[1]
<223> t modified by fluorescent reporter VIC

<220>
<221> modified base
<222> [14]...[14]
<223> c modified by minor groove binding non-fluorescent
quencher (MGBNFQ)

<400> 53
ttggtggtga gggc
14

<210> 54
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic allele-A-specific TaqMan probe targeting
beta-globin (hemoglobin beta (HBB)) HbE mutation
mutant allele

<220>
<221> modified base
<222> [1]...[1]
<223> t modified by 6-carboxyfluorescein (FAM)

<220>
<221> modified base
<222> [14]...[14]
<223> c modified by minor groove binding non-fluorescent
quencher (MGBNFQ)

<400> 54
ttggtggtga gggc
14

<210> 55
<211> 20
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> synthetic digital PCR co-amplification forward
primer for 87-bp and 83-bp amplicon of CD41/42
deletion mutation normal and mutant allele

<400> 55
ttttccacc cttaggctgc                20

<210> 56
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic digital PCR co-amplification reverse
primer for 87-bp and 83-bp amplicon of CD41/42
deletion mutation normal and mutant allele

<400> 56
aacgactcag gagtggacag atc          23

<210> 57
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic allele-specific TaqMan probe targeting
CD41/42 deletion mutation normal allele (without
deletion)

<220>
<221> modified_base
<222> {1}...{1}
<223> c modified by fluorescent reporter VIC

<220>
<221> modified_base
<222> {19}...{19}
<223> t modified by minor groove binding non-fluorescent
quencher (MGBNFQ)

<400> 57
cagaggttct ttgagtcct                19

<210> 58
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic allele-specific TaqMan probe targeting
CD41/42 deletion mutation mutant allele (with
deletion)

<220>
<221> modified_base
<222> {1}...{1}
<223> a modified by 6-carboxyfluorescein (FAM)

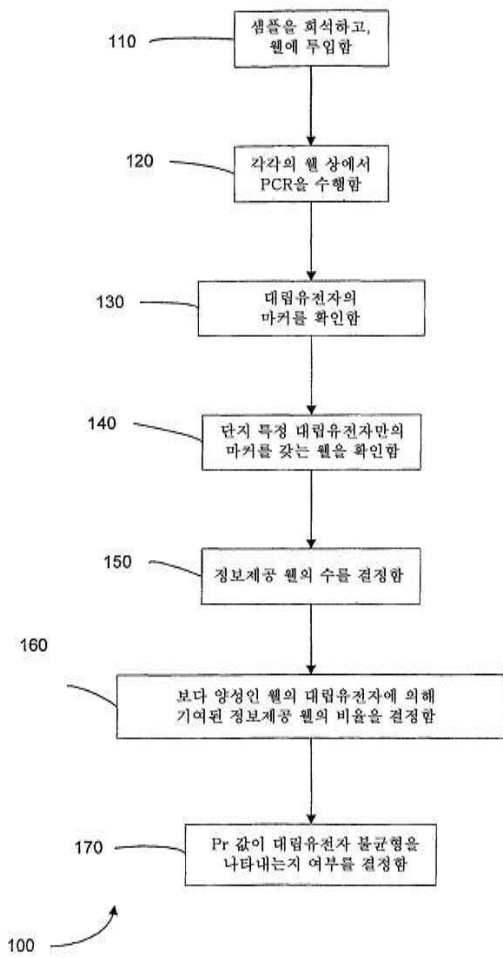
```

[0359]



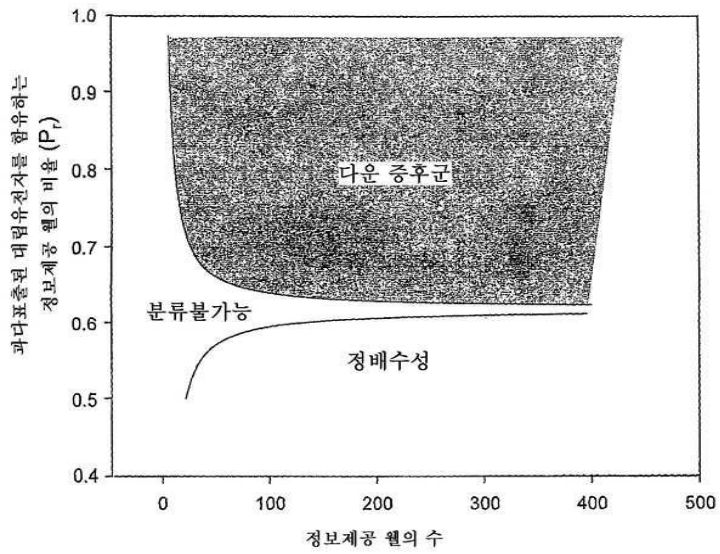
도면

도면1

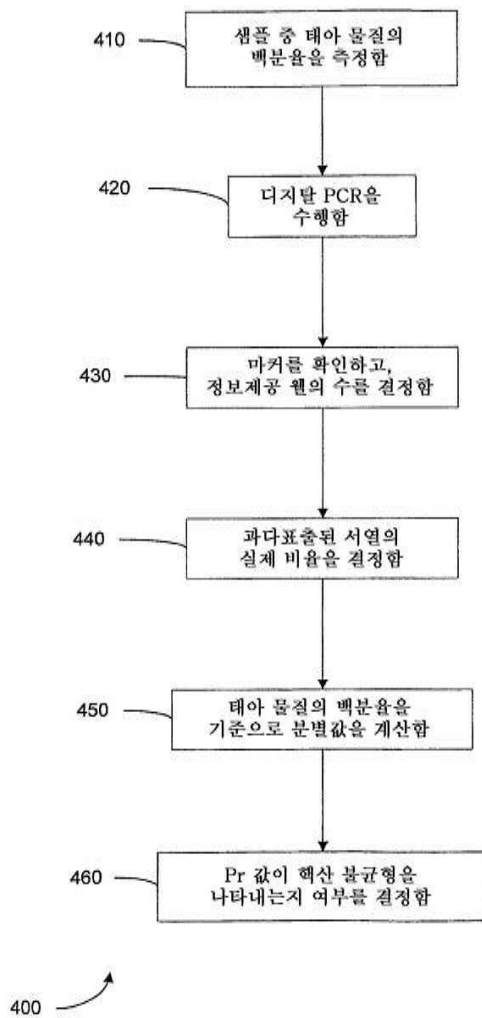




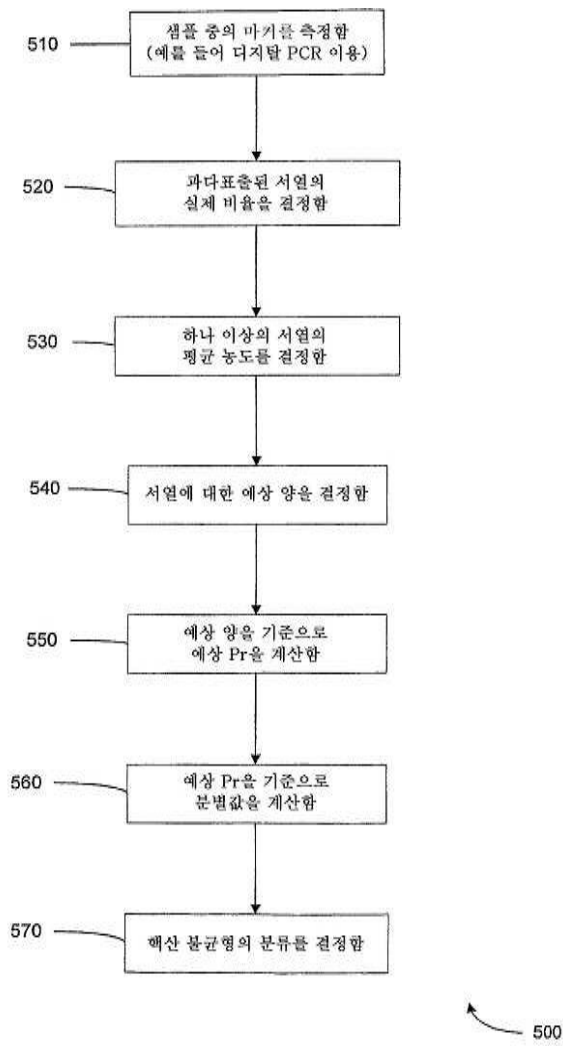
도면3



도면4



도면5



도면6

mv	참조 대립유전자에 대한 양성인 률의 비율	과다표출된 대립유전자에 대한 양성인 률의 비율	두 대립유전자 모두에 대한 양성인률의 비율	참조 대립유전자에 대해서만 양성인률의 비율	과다표출된 대립유전자에 대해서만 양성인률의 비율	디지털 RNA-SNP 대립유전자 비율	정보제공 률의 비율	P <sub>r</sub>
0.1	0.0952	0.1813	0.0173	0.0779	0.164	2.11	0.2419	0.68
0.2	0.1813	0.3297	0.0598	0.1215	0.2699	2.22	0.3914	0.69
0.3	0.2592	0.4512	0.1169	0.1422	0.3342	2.35	0.4763	0.7
0.4	0.3297	0.5507	0.1815	0.1481	0.3691	2.49	0.5173	0.71
0.5	0.3925	0.6321	0.2487	0.1448	0.3834	2.65	0.5282	0.73
0.6	0.4512	0.6988	0.3153	0.1359	0.3835	2.82	0.5194	0.74
0.7	0.5034	0.7534	0.3793	0.1241	0.3741	3.01	0.4983	0.75
0.8	0.5507	0.7981	0.4395	0.1112	0.3586	3.23	0.4698	0.76
0.9	0.5934	0.8347	0.4953	0.0981	0.3394	3.46	0.4375	0.78
1	0.6321	0.8647	0.5466	0.0855	0.3181	3.72	0.4036	0.79
1.1	0.6671	0.8892	0.5932	0.0739	0.296	4	0.3699	0.8
1.2	0.6988	0.9093	0.6354	0.0634	0.2739	4.32	0.3373	0.81
1.3	0.7275	0.9257	0.6734	0.054	0.2523	4.67	0.3063	0.82
1.4	0.7534	0.9392	0.7076	0.0458	0.2316	5.06	0.2774	0.83
1.5	0.7769	0.9502	0.7382	0.0387	0.212	5.48	0.2507	0.85
1.6	0.7981	0.9592	0.7656	0.0325	0.1937	5.95	0.2262	0.86
1.7	0.8173	0.9666	0.79	0.0273	0.1766	6.47	0.2039	0.87
1.8	0.8347	0.9727	0.8119	0.0228	0.1608	7.05	0.1836	0.88
1.9	0.8504	0.9776	0.8314	0.019	0.1462	7.69	0.1652	0.88
2	0.8647	0.9817	0.8488	0.0158	0.1329	8.39	0.1487	0.89

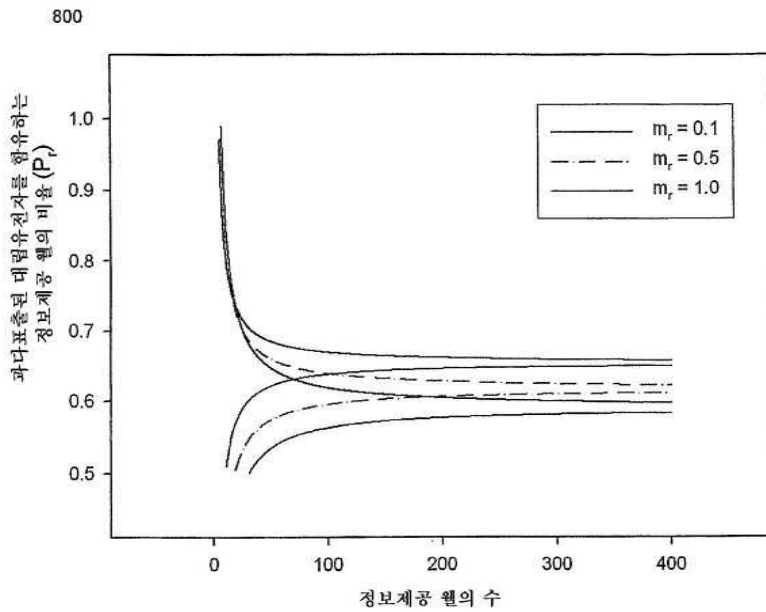
600

도면7

m <sub>r</sub>	샘플 중 태아 DNA의 백분율							
	10%		25%		50%		100%	
	정보제공 웰의 백분율	P <sub>r</sub>	정보제공 웰의 백분율	P <sub>r</sub>	정보제공 웰의 백분율	P <sub>r</sub>	정보제공 웰의 백분율	P <sub>r</sub>
0.1	18%	0.51	18%	0.53	19%	0.56	21%	0.61
0.2	30%	0.51	31%	0.53	32%	0.56	35%	0.61
0.3	39%	0.51	40%	0.53	41%	0.57	43%	0.62
0.4	45%	0.51	45%	0.54	46%	0.57	48%	0.63
0.5	48%	0.52	49%	0.54	49%	0.57	51%	0.63
0.6	50%	0.52	50%	0.54	50%	0.58	51%	0.64
0.7	50%	0.52	50%	0.54	50%	0.58	50%	0.65
0.8	49%	0.52	49%	0.54	49%	0.58	48%	0.65
0.9	48%	0.52	47%	0.55	47%	0.59	46%	0.66
1	46%	0.52	45%	0.55	44%	0.59	43%	0.67
1.1	44%	0.52	43%	0.55	42%	0.6	40%	0.68
1.2	41%	0.52	40%	0.55	39%	0.6	37%	0.69
1.3	39%	0.52	38%	0.55	36%	0.6	34%	0.69
1.4	36%	0.52	35%	0.56	33%	0.61	31%	0.7
1.5	34%	0.52	33%	0.56	31%	0.61	28%	0.71
1.6	31%	0.52	30%	0.56	28%	0.62	26%	0.72
1.7	29%	0.53	28%	0.56	26%	0.62	23%	0.73
1.8	27%	0.53	25%	0.57	24%	0.63	21%	0.73
1.9	24%	0.53	23%	0.57	21%	0.63	19%	0.74
2	22%	0.53	21%	0.57	20%	0.64	17%	0.75

700

도면8



도면9a

m <sub>r</sub>	분류	신규 알고리즘		기존 알고리즘		m <sub>r</sub>	분류	신규 알고리즘		기존 알고리즘	
		정태수제	상열식제 21	정태수제	상열식제 21			정태수제	상열식제 21	정태수제	상열식제 21
0.1	정확	0%	37%	0%	35%	1.1	정확	89%	92%	14%	35%
	부정확	3%	2%	3%	1%		부정확	1%	0%	4%	0%
	미분류	97%	61%	97%	63%		미분류	10%	7%	81%	65%
0.2	정확	6%	64%	0%	40%	1.2	정확	90%	92%	11%	35%
	부정확	4%	2%	4%	0%		부정확	1%	0%	4%	0%
	미분류	90%	34%	96%	60%		미분류	9%	8%	84%	65%
0.3	정확	34%	75%	6%	39%	1.3	정확	91%	92%	7%	34%
	부정확	4%	2%	4%	0%		부정확	1%	1%	4%	0%
	미분류	62%	24%	90%	61%		미분류	8%	7%	89%	66%
0.4	정확	54%	81%	14%	39%	1.4	정확	92%	91%	5%	34%
	부정확	3%	1%	4%	0%		부정확	1%	1%	4%	0%
	미분류	43%	19%	82%	61%		미분류	8%	9%	91%	66%
0.5	정확	64%	83%	20%	38%	1.5	정확	92%	90%	2%	32%
	부정확	3%	1%	4%	0%		부정확	1%	1%	4%	0%
	미분류	33%	14%	76%	62%		미분류	8%	9%	93%	68%
0.6	정확	74%	88%	23%	38%	1.6	정확	92%	89%	1%	31%
	부정확	2%	1%	4%	0%		부정확	1%	1%	4%	0%
	미분류	24%	11%	72%	62%		미분류	8%	10%	94%	69%
0.7	정확	80%	89%	25%	37%	1.7	정확	92%	89%	1%	30%
	부정확	2%	1%	4%	0%		부정확	1%	1%	4%	0%
	미분류	18%	10%	71%	63%		미분류	7%	10%	96%	70%
0.8	정확	83%	91%	23%	37%	1.8	정확	92%	88%	0%	28%
	부정확	2%	1%	4%	0%		부정확	1%	1%	4%	0%
	미분류	15%	9%	73%	63%		미분류	7%	11%	96%	72%
0.9	정확	86%	91%	21%	36%	1.9	정확	92%	87%	0%	26%
	부정확	1%	1%	4%	0%		부정확	1%	1%	3%	0%
	미분류	12%	8%	75%	64%		미분류	8%	12%	97%	74%
1	정확	88%	92%	19%	36%	2	정확	91%	86%	0%	24%
	부정확	1%	0%	4%	0%		부정확	1%	1%	4%	0%
	미분류	11%	8%	77%	64%		미분류	9%	13%	96%	76%

900

도면9b

m <sub>r</sub>	분류	신규 알고리즘		기존 알고리즘		m <sub>r</sub>	분류	신규 알고리즘		기존 알고리즘	
		정태수제	상열식제 21	정태수제	상열식제 21			정태수제	상열식제 21	정태수제	상열식제 21
0.1	정확	57%	88%	48%	89%	1.1	정확	100%	100%	92%	90%
	부정확	3%	1%	3%	1%		부정확	0%	0%	1%	0%
	미분류	40%	11%	49%	10%		미분류	0%	0%	7%	10%
0.2	정확	89%	98%	79%	90%	1.2	정확	100%	100%	90%	90%
	부정확	1%	0%	2%	0%		부정확	0%	0%	1%	0%
	미분류	10%	2%	19%	10%		미분류	0%	0%	9%	10%
0.3	정확	96%	99%	88%	90%	1.3	정확	100%	100%	89%	90%
	부정확	0%	0%	1%	0%		부정확	0%	0%	1%	0%
	미분류	4%	1%	11%	10%		미분류	0%	0%	10%	10%
0.4	정확	99%	100%	92%	90%	1.4	정확	100%	100%	87%	90%
	부정확	0%	0%	1%	0%		부정확	0%	0%	1%	0%
	미분류	1%	0%	7%	10%		미분류	0%	0%	12%	10%
0.5	정확	100%	100%	94%	90%	1.5	정확	100%	100%	85%	90%
	부정확	0%	0%	1%	0%		부정확	0%	0%	1%	0%
	미분류	0%	0%	6%	10%		미분류	0%	0%	13%	10%
0.6	정확	100%	100%	94%	90%	1.6	정확	100%	100%	82%	90%
	부정확	0%	0%	1%	0%		부정확	0%	0%	1%	0%
	미분류	0%	0%	5%	10%		미분류	0%	0%	17%	10%
0.7	정확	100%	100%	95%	90%	1.7	정확	100%	100%	79%	90%
	부정확	0%	0%	1%	0%		부정확	0%	0%	2%	0%
	미분류	0%	0%	5%	10%		미분류	0%	0%	19%	10%
0.8	정확	100%	100%	95%	90%	1.8	정확	100%	100%	74%	90%
	부정확	0%	0%	1%	0%		부정확	0%	0%	2%	0%
	미분류	0%	0%	5%	10%		미분류	0%	0%	24%	10%
0.9	정확	100%	100%	94%	90%	1.9	정확	100%	100%	70%	90%
	부정확	0%	0%	1%	0%		부정확	0%	0%	2%	0%
	미분류	0%	0%	6%	10%		미분류	0%	0%	28%	10%
1	정확	100%	100%	94%	90%	2	정확	100%	100%	67%	90%
	부정확	0%	0%	1%	0%		부정확	0%	0%	3%	0%
	미분류	0%	0%	6%	10%		미분류	0%	0%	30%	10%

950

도면10

$m_i$	분류 결과	정배수치	T21	$m_i$	분류 결과	정배수치	T21
0.1	정확	57%	88%	0.9	정확	100%	100%
	부정확	3%	1%		부정확	0%	0%
	미분류	40%	11%		미분류	0%	0%
0.2	정확	89%	98%	1	정확	100%	100%
	부정확	1%	0%		부정확	0%	0%
	미분류	10%	2%		미분류	0%	0%
0.3	정확	96%	99%	1.1	정확	100%	100%
	부정확	0%	0%		부정확	0%	0%
	미분류	4%	1%		미분류	0%	0%
0.4	정확	99%	100%	1.2	정확	100%	100%
	부정확	0%	0%		부정확	0%	0%
	미분류	1%	0%		미분류	0%	0%
0.5	정확	100%	100%	1.3	정확	100%	100%
	부정확	0%	0%		부정확	0%	0%
	미분류	0%	0%		미분류	0%	0%
0.6	정확	100%	100%	1.4	정확	100%	100%
	부정확	0%	0%		부정확	0%	0%
	미분류	0%	0%		미분류	0%	0%
0.7	정확	100%	100%	1.5	정확	100%	100%
	부정확	0%	0%		부정확	0%	0%
	미분류	0%	0%		미분류	0%	0%
0.8	정확	100%	100%	2	정확	100%	100%
	부정확	0%	0%		부정확	0%	0%
	미분류	0%	0%		미분류	0%	0%

1000

도면11

$m_i$	분류 결과	정배수치	참여석채 21	$m_i$	분류 결과	정배수치	참여석채 21
0.1	정확	38%	44%	0.7	정확	95%	94%
	부정확	2%	2%		부정확	0%	1%
	미분류	60%	54%		미분류	5%	5%
0.2	정확	64%	69%	0.8	정확	95%	95%
	부정확	2%	2%		부정확	1%	0%
	미분류	34%	29%		미분류	4%	5%
0.3	정확	78%	82%	0.9	정확	96%	96%
	부정확	2%	1%		부정확	0%	0%
	미분류	20%	17%		미분류	4%	4%
0.4	정확	85%	87%	1	정확	97%	96%
	부정확	1%	1%		부정확	0%	0%
	미분류	14%	12%		미분류	3%	4%
0.5	정확	89%	90%	1.5	정확	97%	96%
	부정확	1%	1%		부정확	0%	0%
	미분류	10%	9%		미분류	3%	4%
0.6	정확	92%	93%	2	정확	96%	95%
	부정확	1%	1%		부정확	0%	0%
	미분류	7%	7%		미분류	4%	5%

1100

도면12

필의 수	분류 결과	샘플 중 티아 DNA 백분율							
		10%		25%		50%		100%	
		정배수체	삼염색체 21	정배수체	삼염색체 21	정배수체	삼염색체 21	정배수체	삼염색체 21
384	정확	0%	0%	6%	7%	47%	49%	89%	90%
	부정확	0%	0%	0%	1%	2%	2%	1%	1%
	미분류	100%	100%	93%	92%	51%	49%	10%	9%
768	정확	0%	0%	23%	24%	73%	76%	98%	99%
	부정확	0%	0%	2%	2%	2%	1%	0%	0%
	미분류	100%	100%	75%	74%	25%	23%	2%	1%
1152	정확	1%	1%	39%	40%	86%	87%	100%	100%
	부정확	0%	0%	2%	2%	1%	1%	0%	0%
	미분류	99%	99%	59%	58%	13%	12%	0%	0%
1536	정확	2%	2%	50%	51%	93%	93%	100%	100%
	부정확	0%	0%	2%	2%	1%	1%	0%	0%
	미분류	98%	97%	48%	47%	7%	6%	0%	0%
1920	정확	4%	4%	60%	60%	96%	96%	100%	100%
	부정확	0%	1%	2%	2%	0%	0%	0%	0%
	미분류	96%	95%	38%	37%	4%	3%	0%	0%
3840	정확	18%	19%	84%	83%	100%	100%	100%	100%
	부정확	1%	1%	1%	1%	0%	0%	0%	0%
	미분류	81%	80%	15%	15%	18%	22%	0%	0%
7680	정확	43%	45%	97%	97%	100%	100%	100%	100%
	부정확	2%	2%	0%	0%	0%	0%	0%	0%
	미분류	55%	53%	3%	3%	0%	0%	0%	0%

1200

도면13a

샘플	유전자형*	필의 수	개별 대립유전자에 대해 양성인 필의 수				m <sup>b</sup>	P <sup>c</sup>	SPRT 결과	
			A 단독	G 단독	AG	모두 음성			분류불가능한 영역 <sup>d</sup>	분류*
<b>티아 DNA</b>										
N677	AG	384	85	83	126	90	0.79	0.51	0.63 - 0.65	정배수체
N710	AG	384	102	83	73	126	0.52	0.55	0.61 - 0.63	정배수체
N435	AGG	384	49	157	130	48	0.63	0.76	0.62 - 0.64	T21
N981	AAG	384	135	69	82	98	0.5	0.66	0.61 - 0.63	T21
<b>티아 RNA</b>										
V533	AG	384	103	93	71	117	0.56	0.53	0.61 - 0.63	정배수체
V943	AG	384	89	100	74	121	0.55	0.53	0.61 - 0.63	정배수체
N435	AGG	384	52	138	95	99	0.48	0.73	0.61 - 0.63	T21
T215	AAG	384	146	58	138	42	0.71	0.72	0.62 - 0.64	T21

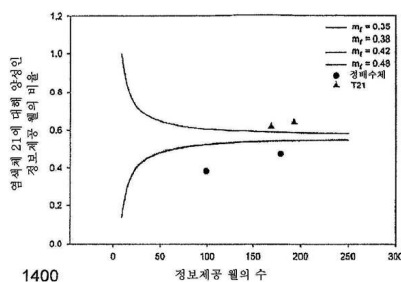
1300

도면13b

샘플	유전자형	필의 수	개별 대립유전자에 대해 양성인 필의 수				m	SPRT 결과		
			A 단독	G 단독	AG	모두 음성		비율	분류불가능한 영역	분류
M2390P	AG	384	90	100	97	97	0.67	0.526	0.62 - 0.64	정배수체
M2391P	AG	384	97	105	65	117	0.55	0.52	0.61 - 0.63	정배수체
M2473P	AG	384	66	92	34	192	0.3	0.582	0.59 - 0.62	정배수체
M2524P	AG	384	29	28	3	324	0.08	0.509	0.54 - 0.64	정배수체
M2528P	AG	384	112	85	44	143	0.41	0.569	0.60 - 0.62	정배수체
M2601P	AG	384	90	101	72	121	0.55	0.529	0.61 - 0.63	정배수체
M2607P	AG	384	73	91	57	163	0.41	0.555	0.60 - 0.63	정배수체
M2638P	AG	384	66	90	52	176	0.37	0.577	0.59 - 0.62	정배수체
M2639P	AG	384	71	56	17	240	0.21	0.559	0.58 - 0.62	정배수체
M2625P	AAG	384	110	53	21	200	0.21	0.675	0.58 - 0.61	T21
M2272P	AAG	768	246	127	112	283	0.37	0.66	0.60 - 0.61	T21
M2718P	AGG	384	66	114	66	138	0.42	0.633	0.60 - 0.62	T21
M1519P	AGG	384	56	130	54	142	0.34	0.691	0.59 - 0.62	T21

1350

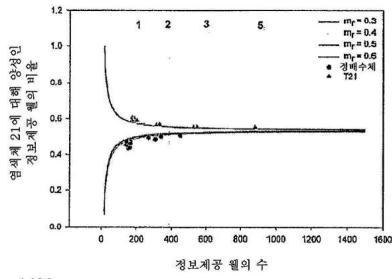
도면14a



1400

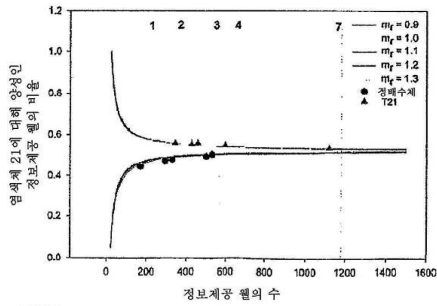


도면14b



1430

도면14c



1460

도면15a

샘플	유전자형	웰의 수	개별 대립유전자에 대해 양성인 웰의 수				SPRT 결과			
			A 단독	G 단독	GA	모두 음성	$m_r^a$	$P_r^b$	분류불가능한 영역	분류 <sup>c</sup>
V943	AG	765	168	209	117	271	0.47	0.554	0.609- 0.621	정체수체
N435	AGG	765	92	281	105	287	0.30	0.753	0.597- 0.610	T21
T21-5	AAG	765	267	126	141	231	0.43	0.679	0.607- 0.618	T21

<sup>a</sup> 웰 당 참조 분자의 수

<sup>b</sup>  $P_r$  값은 하기 식에 의해 계산하였음:

$$\frac{\text{과대표출된 대립유전자에 대해 양성인 웰의 수}}{\text{A에 대해서만 양성인 웰의 수} + \text{G에 대해서만 양성인 웰의 수}}$$

<sup>c</sup>  $P_r$ 이 분류불가능한 영역 아래에 있는 경우에 "정체수체"로 지정하였고;  $P_r$ 이 분류불가능한 영역 위에 있는 경우에 "T21"로 지정하였음. T21, 참조체 21

도면15b

패널 #	웰의 수	개별 대립유전자에 대해 양성인 웰의 수				합계
		A 단독	G 단독	GA	모두 음성	
S1	765	14	15	0	736	
S2	765	7	10	0	748	
S3	765	13	9	1	744	
S4	765	7	8	0	750	
S5	765	12	8	0	745	
S6	765	8	14	0	743	
S7	765	13	9	0	743	
S8	765	9	8	0	748	
S9	765	14	15	0	736	
S10	765	9	12	0	744	
S11	765	8	17	0	740	
S12	765	11	9	0	745	

도면15c

샘플	유전자형	패널 #	개별 대립유전자에 대해 양성인 웰의 수					SPRT 결과		
			웰의 수	A 단독	G 단독	GA	모두 음성	$P_T$	분류불가능한 영역	분류
M2879P	AG	5	3825	82	103	5	3835	0.557	0.570- 0.602	정배수체
M2875P	AG	3	2295	35	40	0	2220	0.533	0.546- 0.625	정배수체
M3078P	AG	3	2295	33	33	1	2228	0.500	0.541- 0.631	정배수체
M2876P	AG	12	9180	56	46	0	9078	0.549	0.556- 0.615	정배수체
M2831P	AAG	2	1530	42	13	0	1475	0.764	0.631- 0.640	T21

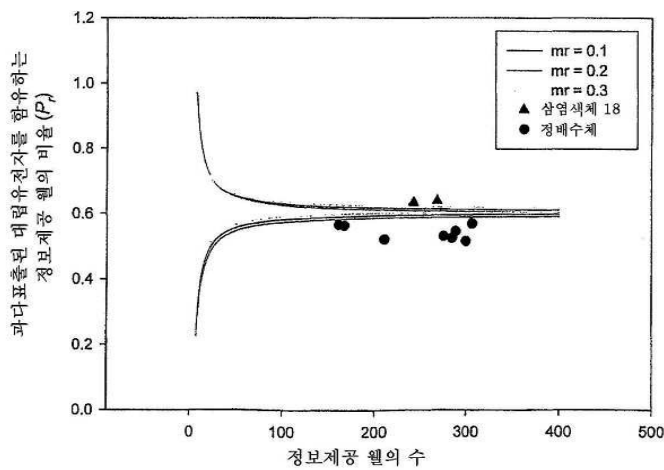
<sup>a</sup> 분류에 필요한 패널의 수

도면16a

샘플	유전자형 <sup>a</sup>	A 또는 G 대립유전자에 대해 양성인 웰의 수				SPRT 분석		
		A 단독	G 단독	둘 다	$m^b$	$P_T^c$	분류불가능한 영역 <sup>d</sup>	분류 <sup>e</sup>
N3505	AG	157	131	48	0.27	0.545	0.59-0.61	정배수체
N3530	AG	174	132	52	0.28	0.569	0.59-0.61	정배수체
N3531	AG	146	129	28	0.23	0.531	0.58-0.60	정배수체
N3533	AG	149	135	46	0.27	0.525	0.59-0.61	정배수체
N2772	AG	159	149	79	0.35	0.516	0.60-0.61	정배수체
N3506	AG	154	145	66	0.32	0.515	0.60-0.61	정배수체
N3555	AG	94	73	7	0.11	0.563	0.58-0.61	정배수체
N3560	AG	101	109	28	0.18	0.519	0.58-0.61	정배수체
N3561	AG	91	70	17	0.12	0.565	0.58-0.61	정배수체
N828	AAG	153	89	22	0.16	0.632	0.58-0.60	T18
CA2	AGG	12	28	0	0.016	0.700	0.51-0.66	T18
CA7	AGG	97	171	56	0.22	0.638	0.59-0.61	T18

- <sup>a</sup> 유전자형은 질량 분광법에 의해 결정하였음.
- <sup>b</sup>  $m^b$  값은 웰 당 참조 주형 분자의 평균 수를 나타냄.
- <sup>c</sup>  $P_T$  값은 과대표출된 대립유전자에 대해 양성인 정보 제공 웰의 비율을 나타냄.
- <sup>d</sup> 분류불가능한 영역은  $m^b$  값에 따라 변함.
- <sup>e</sup>  $P_T$  이 분류불가능한 영역 아래에 있는 경우에 "정배수체"로 지정하였고;  $P_T$  이 분류불가능한 영역 위에 있는 경우에 "T18"로 지정하였음.

도면16b

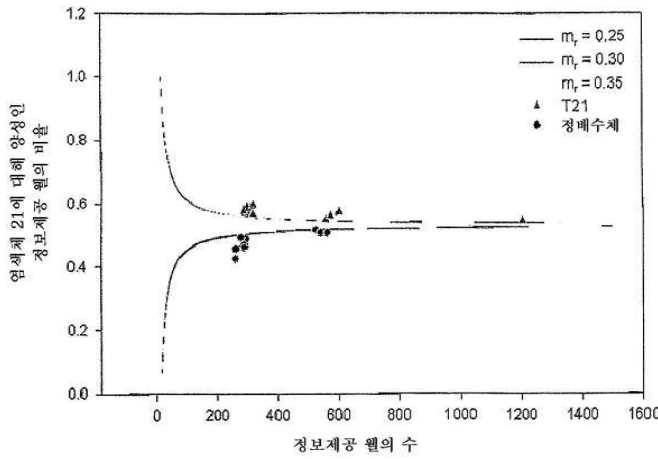


도면17

유형	사례	사용된 패널의 총 수	정보제공 패널의 총 수	계수한 패널의 수			SPRT 분석			
				문 다 양성	Chr21 단독	Chr1 단독	$m_i^a$	$P_i^b$	분류불가능한 영역	분류*
T21	N0224	1	303	68	172	131	0.30	0.588	0.507 - 0.569	T21
	N0274	2	568	119	307	251	0.28	0.550	0.518 - 0.547	T21
	N0352	1	293	73	189	125	0.30	0.573	0.508 - 0.569	T21
	N0349	2	574	107	322	252	0.27	0.581	0.519 - 0.546	T21
	N0495	1	299	67	176	123	0.28	0.589	0.508 - 0.559	T21
	N0459	1	319	88	189	129	0.30	0.564	0.508 - 0.558	T21
	N0667	2	602	134	345	257	0.30	0.573	0.519 - 0.546	T21
	N0794	1	298	62	167	122	0.28	0.578	0.505 - 0.580	T21
	N0890	4	1208	272	657	459	0.31	0.545	0.508 - 0.559	T21
	N0891	1	317	103	180	137	0.38	0.568	0.510 - 0.550	T21
정상	N0198	1	262	68	111	151	0.34	0.424	0.504 - 0.583	정배수체
	N0199	1	295	89	132	153	0.34	0.483	0.508 - 0.581	정배수체
	N0230	2	527	109	272	255	0.27	0.518	0.517 - 0.548	정배수체
	N0231	2	542	98	275	287	0.27	0.507	0.517 - 0.547	정배수체
	N0262	1	295	57	144	151	0.32	0.488	0.508 - 0.580	정배수체
	N0483	1	263	70	129	143	0.33	0.458	0.504 - 0.583	정배수체
	N0527	1	282	56	137	155	0.32	0.498	0.508 - 0.580	정배수체
	N0619	1	281	80	139	143	0.31	0.491	0.505 - 0.581	정배수체
	N0625	1	281	80	134	157	0.33	0.480	0.507 - 0.580	정배수체
	N0661	2	568	118	287	278	0.30	0.507	0.518 - 0.547	정배수체

<sup>a</sup> 될 당 참조 분자의 수  
<sup>b</sup>  $P_i$  값은 하기 식에 의해 계산하였음:  
 과다표출된 대립유전자 (chr21)에 대해 양성인 패널의 수  
 chr21에 대해서만 양성인 패널의 수 + chr1에 대해서만 양성인 패널의 수  
 \*  $P_i$ 이 분류불가능한 영역 아래에 있는 경우에 "정배수체"로 지정하였고;  $P_i$ 이 분류불가능한 영역 위에 있는 경우에 "T21"로 지정하였음.

도면18



도면19

유형	사례	사용된 패널의 수	정보제공 패널의 수	계수한 패널의 수			SPRT 분석			
				문 다 양성	Chr21 단독	Chr1 단독	$m_i^a$	$P_i^b$	분류불가능한 영역	분류*
T21	A0274	1	360	269	248	131	0.74	0.855	0.687 - 0.585	T21
	A2036	1	346	373	233	115	1.02	0.870	0.578 - 0.588	T21
	A2178	1	364	269	232	145	0.83	0.822	0.571 - 0.588	T21
	AS939A	1	384	109	244	120	0.38	0.670	0.551 - 0.573	T21
	AS979A	1	258	492	198	88	1.32	0.734	0.680 - 0.810	T21
	AS9457A	1	376	63	268	120	0.31	0.881	0.648 - 0.572	T21
	정상	A0236	1	387	275	195	122	0.94	0.604	0.578 - 0.582
A0237		1	392	228	184	208	0.84	0.489	0.672 - 0.588	정배수체
A0255		1	378	222	195	188	0.75	0.518	0.688 - 0.585	정배수체
A0283		1	254	49	127	127	0.25	0.500	0.541 - 0.578	정배수체
AS325A		1	379	135	192	179	0.53	0.519	0.558 - 0.579	정배수체
AS483A		1	354	84	188	188	0.38	0.525	0.550 - 0.574	정배수체
AS484A		1	372	271	184	178	0.83	0.522	0.573 - 0.580	정배수체
AS789A		1	419	188	200	213	0.74	0.494	0.588 - 0.584	정배수체
AS789A		1	318	67	159	157	0.35	0.503	0.548 - 0.575	정배수체
AS857A		1	189	8	67	111	0.17	0.438	0.532 - 0.578	정배수체
AS821A		1	152	8	59	53	0.13	0.454	0.623 - 0.588	정배수체
AS890A		1	394	137	191	203	0.58	0.485	0.581 - 0.580	정배수체
AS2282A		1	334	342	177	157	1.05	0.530	0.598 - 0.587	정배수체
AS2425A		1	128	8	64	64	0.10	0.530	0.518 - 0.581	정배수체
AS2494A		1	379	87	203	176	0.44	0.536	0.555 - 0.575	정배수체
A1119		1	354	123	185	188	0.48	0.523	0.555 - 0.577	정배수체
A1122		1	389	100	185	174	0.44	0.528	0.554 - 0.576	정배수체
A1152		1	258	49	128	128	0.28	0.538	0.542 - 0.578	정배수체
A1189		1	249	469	133	113	1.61	0.541	0.804 - 0.822	정배수체
A1275		1	373	65	201	172	0.43	0.538	0.554 - 0.575	정배수체
A1275		1	368	55	190	188	0.36	0.528	0.550 - 0.573	정배수체
A1388		1	357	67	185	182	0.47	0.482	0.585 - 0.577	정배수체
A1388		1	371	114	197	174	0.47	0.531	0.588 - 0.577	정배수체

<sup>a</sup> 될 당 참조 분자의 수  
<sup>b</sup>  $P_i$  값은 하기 식에 의해 계산하였음:  
 과다표출된 대립유전자 (chr21)에 대해 양성인 패널의 수  
 chr21에 대해서만 양성인 패널의 수 + chr1에 대해서만 양성인 패널의 수  
 \*  $P_i$ 이 분류불가능한 영역 아래에 있는 경우에 "정배수체"로 지정하였고;  $P_i$ 이 분류불가능한 영역 위에 있는 경우에 "T21"로 지정하였음.

도면20

유형	사례	사용된 패널의 총 수	정보제공 패널의 총 수	계수한 패널의 수			$m^a$	SPRT 분석		
				물 다 양성	Chr18 단독	Chr21 단독		$P_1^b$	분류불가능한 영역	분류 <sup>c</sup>
정배수체	N0230	1	360	243	181	189	0.86	0.476	0.573 - 0.588	E
	N0625	1	370	225	179	191	0.78	0.464	0.509 - 0.586	E
	N0651	1	358	206	192	165	0.67	0.536	0.564 - 0.563	E
T18	N0828	1	543	256	329	214	0.95	0.608	0.579 - 0.590	T18
	N0830	1	346	353	240	106	0.92	0.694	0.574 - 0.592	T18
	N0936	1	372	285	209	163	0.88	0.562	0.573 - 0.590	E
	N1678	1	343	115	221	122	0.37	0.644	0.551 - 0.574	T18
	N1701	1	303	94	191	112	0.31	0.630	0.546 - 0.575	T18

<sup>a</sup> 웰 당 참조 분자의 수

<sup>b</sup>  $P_1$  값은 하기 식에 의해 계산하였음:

$$\frac{\text{과대표출된 대립유전자 (chr18)에 대해 양성인 웰의 수}}{\text{chr21에 대해서만 양성인 웰의 수} + \text{chr18에 대해서만 양성인 웰의 수}}$$

<sup>c</sup>  $P_1$ 이 분류불가능한 영역 아래에 있는 경우에 "정배수체"로 지정하였고;  $P_1$ 이 분류불가능한 영역 위에 있는 경우에 "T18"로 지정하였음.

도면21

유형	사례	플래이트 수	검정	정보제공 패널의 총 수	계수한 패널의 수				$m^a$	SPRT 분석		
					물 다 양성	Chr21 단독	참조 단독	$m^a$		$P_1^b$	분류불가능한 영역	분류 <sup>c</sup>
경상	N0196	1	chr21chr1	148	78	71	77	0.52	0.480	0.489 - 0.884	E	
		1	chr21chr18	152	63	78	73	0.44	0.520	0.487 - 0.884	E	
			총합	300	141	150	150	0.48	0.600	0.512 - 0.560	E	
N0198	1	chr21chr1	162	73	70	92	0.58	0.432	0.490 - 0.591	E		
	1	chr21chr18	153	68	83	80	0.58	0.508	0.498 - 0.576	E		
		총합	345	141	163	182	0.55	0.472	0.517 - 0.557	E		
N0230	1	chr21chr1	162	58	81	81	0.45	0.500	0.490 - 0.581	U		
	1	chr21chr18	166	45	81	84	0.41	0.491	0.489 - 0.590	U		
		총합	327	103	162	165	0.43	0.495	0.512 - 0.558	E		
N0231	1	chr21chr1	173	53	81	92	0.47	0.468	0.494 - 0.578	E		
	1	chr21chr18	166	51	85	81	0.42	0.512	0.499 - 0.590	U		
		총합	339	104	166	173	0.45	0.490	0.514 - 0.557	E		
N0262	1	chr21chr1	186	97	101	85	0.84	0.643	0.503 - 0.575	U		
	1	chr21chr18	176	111	91	85	0.71	0.517	0.504 - 0.557	U		
		총합	362	208	192	170	0.88	0.535	0.532 - 0.557	U		
	2	chr21chr1	174	103	85	89	0.69	0.468	0.503 - 0.577	E		
		총합	536	311	277	259	0.88	0.617	0.523 - 0.592	E		
T21	N0234	1	chr21chr1	193	68	108	85	0.51	0.660	0.500 - 0.573	U	
		1	chr21chr18	178	66	92	86	0.50	0.617	0.498 - 0.576	U	
		총합	371	133	200	171	0.50	0.538	0.517 - 0.598	U		
	2	chr21chr1	176	73	92	86	0.53	0.517	0.498 - 0.576	U		
		총합	649	206	292	257	0.51	0.532	0.524 - 0.548	U		
	2	chr21chr18	178	82	95	83	0.47	0.634	0.495 - 0.576	U		
		총합	727	268	387	340	0.50	0.532	0.527 - 0.548	U		
	3	chr21chr1	184	108	103	81	0.68	0.660	0.504 - 0.578	U		
		총합	811	376	499	421	0.54	0.528	0.529 - 0.545	U		
	3	chr21chr18	198	102	120	78	0.63	0.608	0.505 - 0.572	T21		
	총합	1109	478	610	499	0.55	0.600	0.531 - 0.544	T21			
N0332	1	chr21chr1	173	109	103	70	0.83	0.696	0.500 - 0.577	T21		
	1	chr21chr18	191	82	110	81	0.55	0.576	0.501 - 0.574	T21		
		총합	364	191	213	151	0.59	0.585	0.519 - 0.567	T21		
N0348	1	chr21chr1	158	82	83	76	0.44	0.625	0.489 - 0.582	U		
	1	chr21chr18	166	46	87	79	0.39	0.524	0.489 - 0.580	U		
		총합	324	108	170	154	0.42	0.625	0.512 - 0.588	U		
	2	chr21chr1	140	66	89	83	0.41	0.577	0.485 - 0.585	U		
		총합	473	174	256	217	0.41	0.561	0.519 - 0.561	U		
2	chr21chr18	168	43	100	68	0.34	0.505	0.487 - 0.580	T21			
	총합	641	217	356	285	0.40	0.595	0.523 - 0.548	T21			
N0764	1	chr21chr1	182	82	106	78	0.53	0.662	0.486 - 0.575	T21		
	1	chr21chr18	178	83	87	81	0.66	0.545	0.498 - 0.578	U		
		총합	360	165	203	157	0.64	0.564	0.518 - 0.559	T21		
N0890	1	chr21chr1	186	103	108	72	0.61	0.620	0.501 - 0.576	T21		
	1	chr21chr18	188	68	104	84	0.50	0.523	0.499 - 0.574	U		
	총합	388	171	212	156	0.55	0.676	0.518 - 0.558	T21			

<sup>a</sup> 웰 당 참조 분자의 수

<sup>b</sup>  $P_1$  값은 하기 식에 의해 계산하였음:

$$\frac{\text{과대표출된 대립유전자 (chr21)에 대해 양성인 웰의 수}}{\text{chr21에 대해서만 양성인 웰의 수} + \text{참조에 대해서만 양성인 웰의 수}}$$

<sup>c</sup>  $P_1$ 이 분류불가능한 영역 아래에 있는 경우에 "정배수체"로 지정하였고;  $P_1$ 이 분류불가능한 영역 위에 있는 경우에 "T21"로 지정하였음.

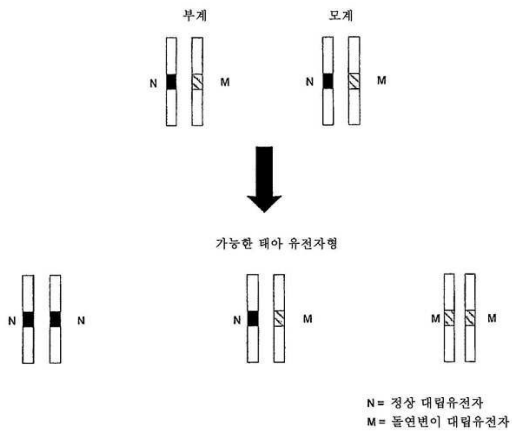
도면22a

		평가상 보정된 총 수		P <sub>i</sub>	분류
		업세서 21	참조 업세서		
정배수체 사례 1	검정 1	37	33	0.52345751	미분류 (Unclass)
	검정 2	31	42	0.42506276	미분류
	검정 3	29	37	0.44198785	미분류
	검정 4	39	26	0.6008691	미분류
	플레이트 1 소계	136	138	0.496821	미분류
	검정 1	32	25	0.56605316	미분류
	검정 2	27	24	0.53165424	미분류
	검정 3	10	17	0.3681535	미분류
	검정 4	24	18	0.5628764	미분류
	플레이트 2 소계	93	84	0.52474	미분류
총계	229	223	0.507182	정배수체	
사례 2	검정 1	45	35	0.56321585	미분류
	검정 2	31	45	0.41246525	미분류
	검정 3	32	22	0.59954763	미분류
	검정 4	38	35	0.5228175	미분류
	플레이트 1 소계	146	135	0.519153	미분류
	검정 1	28	23	0.55272135	미분류
	검정 2	23	22	0.51196923	미분류
	검정 3	23	15	0.59688187	미분류
	검정 4	29	26	0.5293349	미분류
	플레이트 2 소계	102	85	0.545204	미분류
플레이트 1+2 소계	248	221	0.529583	미분류	
검정 1	17	25	0.4122149	미분류	
검정 2	25	16	0.6027299	미분류	
검정 3	13	26	0.3382802	미분류	
검정 4	27	24	0.5316872	미분류	
플레이트 3 소계	82	91	0.475775	미분류	
플레이트 1-3 소계	331	311	0.515069	미분류	
검정 1	15	9	0.6269193	미분류	
검정 2	14	15	0.4824155	미분류	
검정 3	18	17	0.5146313	미분류	
검정 4	21	18	0.5270080	미분류	
플레이트 4 소계	69	60	0.532284	미분류	
플레이트 1-4 소계	399	372	0.517842	미분류	

도면22b

검정 1	10	13	0.4337957	미분류	
검정 2	11	14	0.4390082	미분류	
검정 3	7	16	0.3018022	미분류	
검정 4	19	10	0.6579085	미분류	
플레이트 5 소계	48	54	0.469951	미분류	
총계	447	426	0.5123444	정배수체	
T21					
사례 3	검정 1	50	33	0.60012253	미분류
	검정 2	56	47	0.54458055	미분류
	검정 3	46	28	0.6196626	미분류
	검정 4	56	36	0.6108731	미분류
	총계	208	144	0.59079	T21
사례 4	검정 1	56	36	0.61085513	미분류
	검정 2	38	32	0.53906912	미분류
	검정 3	40	24	0.62788068	미분류
	검정 4	47	27	0.634511	미분류
	총계	180	119	0.603476	T21

도면23



도면24a

사례	샘플	예상된 %	페닐의 수	ZFY	ZFX	ALL POS	NET ZFY	NET ZFX	참조 대립유전자	$m_r$	SPRT <sup>a</sup>
1	XY + XX (XX %)	50%	1	239	465	155	84	310	Y	0.37	XX
		25%	1	328	447	205	123	242	Y	0.56	XX
2	XY <sub>1</sub> + XY <sub>2</sub> (XY <sub>2</sub> %)	50%	1	414	405	220	194	185	X	0.85	XY
		25%	1	432	426	245	187	181	X	0.81	XY

도면24b

사례	페닐의 수	ZFY	ZFX	ALL POS	NET ZFY	NET ZFX	참조 대립유전자	$m_r$	SPRT
1	1	81	171	18	63	153	Y	0.11	XX
2	1	103	175	23	80	152	Y	0.14	XX
3	1	110	175	28	82	147	Y	0.16	XX
4	1	111	156	24	87	132	Y	0.16	XX
5	2	236	322	50	186	272	Y	0.17	XX
6	1	116	184	29	87	155	Y	0.16	XX
7	1	94	169	19	75	150	Y	0.13	XX
8	1	91	155	24	67	131	Y	0.13	XX
9	3	333	447	74	259	373	Y	0.16	XX
10	1	110	154	24	86	130	Y	0.16	XX
11	1	83	170	26	57	144	Y	0.11	XX
12	1	85	159	16	69	143	Y	0.12	XX

도면25

사례	패널의 수	모제 유전자형 <sup>a</sup>	태아 유전자형 <sup>a</sup>	예상된 태아 %	A	G	ALL POS	NET A	NET G	실제 태아 % <sup>b</sup>	참조 대립유전자 <sup>c</sup>	SPRT <sup>d</sup>
1	1	AG	GG	15%	264	318	119	148	202	9.8	A	0.42
2	1	AG	GG	15%	208	276	81	127	195	14.3	A	0.32
3	1	AG	AG	15%	248	251	87	161	164	15.4	A	0.39
4	1	AG	GG	15%	158	240	51	107	189	13.7	A	0.23
5	1	AG	AG	15%	226	245	70	156	175	16.8	A	0.35
6	1	AG	AG	15%	319	319	130	189	189	18.1	A	0.54
7	1	AG	AG	15%	254	279	94	160	185	18.8	A	0.40
8	1	AG	GG	15%	296	363	144	182	210	15.7	A	0.49
9	1	AG	AG	15%	279	285	100	179	185	13.0	A	0.45
10	1	AG	AA	15%	337	278	120	217	158	NA <sup>e</sup>	G	0.45
1	1	AG	GG	20%	248	285	98	152	199	13.0	A	0.38
2	1	AG	GG	20%	191	289	79	112	210	19.5	A	0.28
3	1	AG	AG	20%	248	218	75	173	143	18.2	G	0.34
4	1	AG	GG	20%	158	257	59	99	198	25.9	A	0.23
5	1	AG	AG	20%	231	215	71	160	144	20.3	G	0.33
6	1	AG	AG	20%	281	283	88	193	185	28.3	A	0.46
7	1	AG	AG	20%	258	231	88	168	143	22.7	G	0.38
8	1	AG	GG	20%	246	311	104	142	207	23.0	A	0.39
9	1	AG	AG	20%	287	249	83	184	166	21.9	G	0.39
10	1	AG	AA	20%	303	203	77	226	126	NA <sup>e</sup>	G	0.31
1	1	AG	GG	25%	244	304	99	145	205	15.7	A	0.38
2	1	AG	GG	25%	209	314	88	121	228	24.9	A	0.32
3	1	AG	AG	25%	204	212	55	149	157	28.0	A	0.31
4	1	AG	GG	25%	185	290	79	108	214	25.9	A	0.28
5	1	AG	AG	25%	233	243	77	156	168	24.5	A	0.38
6	1	AG	AG	25%	282	305	110	182	195	27.3	A	0.48
7	1	AG	AG	25%	291	275	95	166	160	22.5	A	0.42
8	1	AG	GG	25%	236	327	98	140	231	26.7	A	0.37
9	1	AG	AG	25%	271	280	104	167	178	25.2	A	0.44
10	1	AG	AA	25%	243	154	46	197	108	NA <sup>e</sup>	G	0.22
1	1	AG	GG	50%	136	293	57	79	236	33.6	A	0.20
2	1	AG	GG	50%	183	358	78	85	280	45.8	A	0.24
3	1	AG	AG	50%	200	208	53	147	155	35.7	A	0.30
4	1	AG	GG	50%	137	358	66	71	280	53.0	A	0.18
5	1	AG	AG	50%	247	248	91	159	155	57.7	G	0.39
6	1	AG	AG	50%	218	245	68	150	177	53.8	A	0.34
7	1	AG	AG	50%	255	282	109	162	179	51.3	A	0.41
8	1	AG	GG	50%	158	368	83	75	315	57.6	A	0.23
9	1	AG	AG	50%	235	225	74	161	151	55.7	G	0.35
10	1	AG	AA	50%	317	108	35	282	71	NA <sup>e</sup>	G	0.15

<sup>a</sup> A = 돌연변이 대립유전자, G = 야생형 대립유전자.  
<sup>b</sup> 실제 태아 DNA %는 디지털 ZFY/X 검정에 의해 결정함.  
<sup>c</sup> 태아는 여성이고, 디지털 ZFY/X 검정은 적용불가능함.  
<sup>d</sup> SPRT는 사례 10을 제외하고 디지털 ZFY/X 검정으로 결정된 태아 %로 수행하였음; U, 미분류.

도면26a

사례	모제 유전자형 <sup>a</sup>	태아 유전자형 <sup>a</sup>	패널의 수	I <sup>b</sup>		Z <sup>c</sup>		SPRT	디지털 ZFY/X의 의한 태아 %	참조 대립유전자 <sup>c</sup>	m <sub>z</sub>		
				M	N	ALL POS	NET M					NET N	예상된 태아 %
1	MN	MM	3	849	763	273	573	490	5%	MM	7.8%	M	0.40
	MN	MM	3	893	148	280	574	459	10%	MM	9.9%	M	0.39
	MN	MM	1	323	219	79	244	140	25%	MM	24.1%	M	0.34
	MN	MM	1	418	180	92	326	68	50%	MM	57.2%	M	0.23
2	MN	MN	9	2559	2476	887	1672	1589	5%	U	3.4%	U	0.45
	MN	MN	3	695	907	300	535	547	10%	MN	9.7%	MN	0.49
	MN	MN	1	277	158	97	180	168	25%	MN	24.2%	MN	0.42
	MN	MN	1	273	295	101	172	194	50%	MN	58.0%	MN	0.44

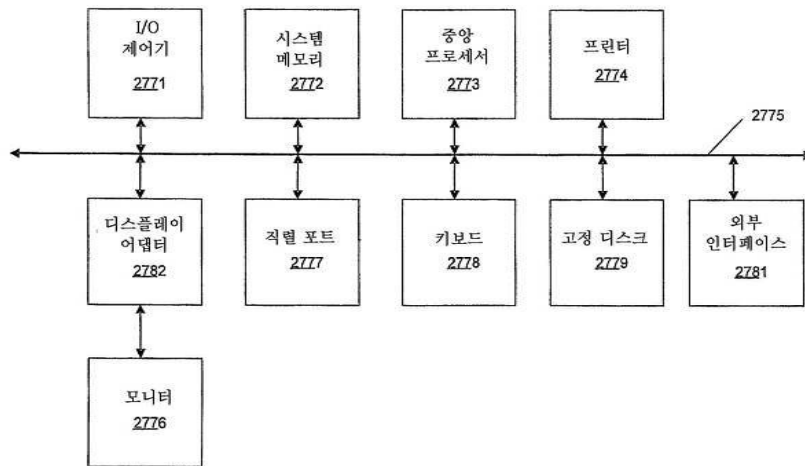
<sup>a</sup> M = CTTT 결실 존재, N = CTTT 결실 부재.  
<sup>b</sup> SPRT는 예상된 %로 수행하였음.  
<sup>c</sup> SPRT는 디지털 ZFY/X 검정으로 결정된 태아 %로 수행하였음; U, 미분류.

도면26b

사례	패널의 수	모제 유전자형 <sup>a</sup>	태아 유전자형 <sup>a</sup>	예상된 태아 %	M	N	ALL POS	NET M	NET N	실제 태아 % <sup>b</sup>	참조 대립유전자 <sup>c</sup>	m <sub>z</sub>	SPRT <sup>d</sup>
1	1	MN	MN	20%	203	200	58	145	142	29.0	N	0.30	MN
2	1	MN	NN	20%	149	204	39	110	165	29.7	M	0.22	NN
3	1	MN	NN	20%	128	212	31	97	181	30.7	M	0.18	NN
4	1	MN	MN	20%	173	196	45	128	151	27.0	M	0.28	MN
5	1	MN	NN	20%	147	230	47	100	183	24.3	M	0.21	NN
6	1	MN	MN	20%	199	206	45	154	161	19.2	M	0.30	MN
7	1	MN	NN	20%	165	237	58	107	179	18.7	M	0.24	NN
8	1	MN	MN	20%	203	180	43	160	137	23.4	N	0.27	MN
9	1	MN	NN	20%	173	241	61	112	180	16.9	M	0.26	NN
10	1	MN	MN	20%	174	190	52	122	138	16.8	M	0.26	U
11	1	MN	NN	20%	167	242	49	118	193	17.7	M	0.25	NN
12	1	MN	MN	20%	213	199	47	166	152	21.6	N	0.30	MN

<sup>a</sup> M = CTTT 결실 존재, N = CTTT 결실 부재.  
<sup>b</sup> 실제 태아 DNA %는 디지털 ZFY/X 검정으로 결정함.  
<sup>c</sup> SPRT는 디지털 ZFY/X 검정으로 결정된 태아 %로 수행하였음; U, 미분류.

도면27



서열 목록

SEQUENCE LISTING

- <110> Lo, Yuk-Ming Dennis  
 Chiu, Rossa Wai Kwun  
 Chan, Kwan Chee  
 Zee, Benny Chung Ying  
 Chong, Ka Chun  
 The Chinese University of Hong Kong
- <120> Determining a Nucleic Acid Sequence  
 Imbalance
- <130> IPA161075-CN
- <140> WO PCT/GB08/02524
- <141> 2008-07-23
- <150> US 60/951,438
- <151> 2007-07-23
- <160> 58
- <170> FastSEQ for Windows Version 4.0
- <210> 1
- <211> 27
- <212> DNA
- <213> Artificial Sequence
- <220>
- <223> synthetic real-time PCR and digital RNA-SNP analysis



PLAC4 gene-specific reverse transcription primer

<400> 1  
 agtatataga accatgttta ggccaga 27  
 <210> 2  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic real-time PCR PLAC4 primer  
 <400> 2  
 ccgctagggt gtctttaag c 21  
 <210> 3  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic real-time PCR PLAC4 primer  
 <400> 3  
 gigtgcaat acaaaatgag tttct 25  
 <210> 4  
 <211> 14  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic real-time PCR PLAC4 fluorescent probe  
 <220>  
 <221> modified\_base  
 <222> (1)...(1)  
 <223> a modified by 6-carboxyfluorescein (FAM)  
 <220>  
 <221> modified\_base  
 <222> (14)...(14)  
 <223> c modified by minor groove binding non-fluorescent  
 quencher (MGBNFQ)

<400> 4  
attggagcaa attc 14

<210> 5  
<211> 104  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> single-stranded synthetic DNA oligonucleotide  
amplicon

<400> 5  
cgccgctagg gtgtctttta agctattgga gcaaattcaa atttgctta aagaaaaaga 60  
aactcatttt gtattgcaac accaggagta tccaaggga ctcg 104

<210> 6  
<211> 22  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> synthetic digital RNA-SNP analysis non-intron  
spanning forward primer

<400> 6  
tttgattgc aacaccattt gg 22

<210> 7  
<211> 15  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> synthetic allele-G-specific TaqMan probe targeting  
rs8130833 SNP of PLAC4  
<220>  
<221> modified\_base  
<222> (1)...(1)  
<223> t modified by 6-carboxyfluorescein (FAM)  
<220>

<221> modified\_base  
 <222> (15)...(15)  
 <223> g modified by minor groove binding non-fluorescent  
 quencher (MGBNFQ)  
 <400> 7  
 tcgtcgtcta acttg 15  
 <210> 8  
 <211> 17  
 <212> DNA  
 <  
 213> Artificial Sequence  
 <220>  
 <223> synthetic allele-A-specific TaqMan probe targeting  
 rs8130833 SNP of PLAC4  
 <220>  
 <221> modified\_base  
 <222> (1)...(1)  
 <223> a modified by fluorescent reporter VIC  
 <220>  
 <221> modified\_base  
 <222> (17)...(17)  
 <223> g modified by minor groove binding non-fluorescent  
 quencher (MGBNFQ)  
 <400> 8  
 attcgtcatc taacttg 17  
 <210> 9  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital RCD analysis co-amplification  
 forward primer for paralogous loci on chromosome  
 21 and 1  
 <400> 9

gttgttctgc aaaaaacctt cga 23

<210> 10

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic digital RCD analysis co-amplification  
reverse primer for paralogous loci on chromosome  
21 and 1

<400> 10

cttggccaga aatacttcat taccatat 28

<210> 11

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic chromosome-specific TaqMan probe  
designed to target chromosome 21 paralog

<220>

<221> modified\_base

<222> (1)...(1)

<223> t modified by 6-carboxyfluorescein (FAM)

<220>

<221> modified\_base

<222> (19)...(19)

<223> a modified by minor groove binding non-fluorescent  
quencher (MGBNFQ)

<400> 11

tacctccata atgagtaaa 19

<210> 12

<211> 20

<212> DNA

<213> Artificial Sequence

<220>  
 <223> synthetic chromosome-specific TaqMan probe  
 designed to target chromosome 1 paralog  
 <220>  
 <221> modified\_base  
 <222> (1)...(1)  
 <223> c modified by fluorescent reporter VIC  
 <220>  
 <221> modified\_base  
 <222> (20)...(20)  
 <223> a modified by minor groove binding non-fluorescent  
 quencher (MGBNFQ)  
 <400> 12  
 cgtacctctg taatgtgtaa 20  
 <210> 13  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital PCR PLAC4 gene-specific reverse  
 transcription primer  
 <400> 13  
 agtatataga accatgttta ggccaga 27  
 <210> 14  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital PCR amplification forward primer  
 for rs8130833 SNP region of PLAC4 gene  
 <400> 14  
 tttgtattgc aacaccattt gg 22  
 <210> 15  
 <211> 15

<212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic allele-G-specific TaqMan probe targeting  
 PLAC4 rs8130833 SNP

<220>  
 <221> modified\_base  
 <222> (1)...(1)  
 <223> t modified by 6-carboxyfluorescein (FAM)  
 <220>  
 <221> modified\_base  
 <222> (15)...(15)  
 <223> g modified by minor groove binding non-fluorescent  
 quencher (MGBNFQ)

<400> 15

tcgtcgtcta acttg

15

<210> 16

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic allele-A-specific TaqMan probe targeting  
 PLAC4 rs8130833 SNP

<220>

<221> modified\_base

<222>

> (1)...(1)

<223> a modified by fluorescent reporter VIC

<220>

<221> modified\_base

<222> (17)...(17)

<223> g modified by minor groove binding non-fluorescent  
 quencher (MGBNFQ)

<400> 16

attcgtcatc taacttg 17

<210> 17

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic digital PCR amplification serpin  
 peptidase inhibitor clade B (ovalbumin) member 2  
 (SERPINB2) gene-specific reverse transcription  
 primer

<400> 17

cgcagacttc tcaccaaaca 20

<210> 18

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic digital PCR amplification serpin  
 peptidase inhibitor clade B (ovalbumin) member 2  
 (SERPINB2) forward primer

<400> 18

ctcagctctg caatcaatgc 20

<210> 19

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic allele-A-specific TaqMan probe targeting  
 serpin peptidase inhibitor clade B (ovalbumin)  
 member 2 (SERPINB2) rs6098 SNP

<220>

<221> modified\_base

<222> (1)...(1)

<223> c modified by 6-carboxyfluorescein (FAM)  
 <220>  
 <221> modified\_base  
 <222> (16)...(16)  
 <223> t modified by minor groove binding non-fluorescent  
 quencher (MGBNFQ)  
 <400> 19  
 ccacagggaa ttattt 16  
 <210> 20  
 <211> 16  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic allele-G-specific TaqMan probe targeting  
  
 serpin peptidase inhibitor clade B (ovalbumin)  
 member 2 (SERPINB2) rs6098 SNP  
 <220>  
 <221> modified\_base  
 <222> (1)...(1)  
 <223> c modified by 6-carboxyfluorescein (FAM)  
 <220>  
 <221> modified\_base  
 <222> (16)...(16)  
 <223> t modified by minor groove binding non-fluorescent  
 quencher (MGBNFQ)  
 <400> 20  
 ccacagggga ttattt 16  
 <210> 21  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital PCR analysis co-amplification



forward primer for 101-bp amplicon for paralogous  
 loci on chromosome 21 and 1

<400> 21  
 gttgttctgc aaaaaacctt cga 23

<210> 22

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic digital PCR analysis co-amplification  
 reverse primer for 101-bp amplicon for paralogous  
 loci on chromosome 21 and 1

<400> 22  
 cttggccaga aatacttcat taccatat 28

<210> 23

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic chromosome-specific TaqMan probe  
 designed to target chromosome 21 paralog

<220>

<221> modified\_base

<222> (1)...(1)

<223> t modified by 6-carboxyfluorescein (FAM)

<220>

<221> modified\_base

<222> (19)...(19)

<223> a modified by minor groove binding non-fluorescent  
 quencher (MGBNFQ)

<400> 23  
 tacctccata atgagtaaa 19

<210> 24

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic chromosome-specific TaqMan probe  
designed to target chromosome 1 paralog

<220>

<221> modified\_base

<222> (1)...(1)

<223> c modified by fluorescent reporter VIC

<220>

<221> modified\_base

<222> (20)...(20)

<223> a modified by minor groove binding non-fluorescent  
quencher (MGBNFQ)

<400> 24

cgtacctctg taatgtgtaa

20

<210> 25

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic digital RCD analysis co-amplification

forward primer for 128-bp amplicon for paralogous  
loci on chromosome 21 and 18

<400> 25

gtacagaaac cacaaactga tcgg

24

<210> 26

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic digital RCD analysis co-amplification  
reverse primer for 128-bp amplicon for paralogous

loci on chromosome 21 and 18

<400> 26

gtccaggctg tgggcct 17

<210> 27

<211> 14

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic chromosome-specific TaqMan probe  
designed to target chromosome 21 paralog

<220>

<221> modified\_base

<222> (1)...(1)

<223> a modified by 6-carboxyfluorescein (FAM)

<220>

<221> modified\_base

<222> (14)...(14)

<223> a modified by minor groove binding non-fluorescent  
quencher (MGBNFQ)

<400> 27

aagaggcgag gcaa 14

<210> 28

<211> 15

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic chromosome-specific TaqMan probe  
designed to target chromosome 18 paralog

<220>

<221> modified\_base

<222> (1)...(1)

<223> a modified by fluorescent reporter VIC

<220>

<221> modified\_base  
 <222> (15)...(15)  
 <223> c modified by minor groove binding non-fluorescent  
 quencher (MGBNFQ)  
 <400> 28  
 aagaggacag gcaac 15  
 <210> 29  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital PCR analysis co-amplification  
 forward primer for 121-bp amplicon for paralogous  
 loci on chromosome 21 and 1  
 <400> 29  
 acgttggatg gttgttctgc aaaaaacctt cga 33  
 <210> 30  
 <211> 38  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital PCR analysis co-amplification  
 reverse primer for 121-bp amplicon for paralogous  
 loci on chromosome 21 and 1  
 <400> 30  
 acgttggatg cttggccaga aataacttcat taccatat 38  
 <210> 31  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital PCR analysis co-amplification  
 extension primer for 121-bp amplicon for

paralogous loci on chromosome 21 and 1

<400> 31  
ctcatcctca cttcgtacct c 21

<210> 32  
<211> 34  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> synthetic digital RCD analysis co-amplification  
forward primer for 148-bp amplicon for paralogous  
loci on chromosome 21 and 18

<400> 32  
acgttggatg gtacagaaac cacaaactga tcgg 34

<210> 33  
<211> 27  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> synthetic digital RCD analysis co-amplification  
reverse primer for 148-bp amplicon for paralogous  
loci on chromosome 21 and 18

<400> 33  
acgttggatg gtccaggctg tgggcct 27

<210> 34  
<211> 17  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> synthetic digital RCD analysis co-amplification  
extension primer for 148-bp amplicon for  
paralogous loci on chromosome 21 and 18

<400> 34  
acaaaagggg gaagagg 17

<210> 35  
 <211> 35  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital RCD analysis co-amplification  
     forward primer for 81 bp amplicon for paralogous  
     loci on chromosome 21 and 1  
 <400> 35  
 acgttggatg ttgatgaagt ctcatctcta ctcg 35  
 <210> 36  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital RCD analysis co-amplification  
     reverse primer for 81 bp amplicon for paralogous  
     loci on chromosome 21 and 1  
 <400> 36  
 acgttggatg caataagctt ggccagaaat act 33  
 <210> 37  
 <211> 37  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital RCD analysis co-amplification  
     forward primer for 82 bp amplicon for paralogous  
     loci on chromosome 21 and 7  
 <400> 37  
 acgttggatg gaatttaagc taaatcagcc tgaactg 37  
 <210> 38  
 <211> 37  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> synthetic digital RCD analysis co-amplification  
 reverse primer for 82 bp amplicon for paralogous  
 loci on chromosome 21 and 7  
 <400> 38  
 acgttggatg gtttctcata gttcatcgta ggcttat 37  
 <210> 39  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital RCD analysis co-amplification  
 forward primer for 101 bp amplicon for paralogous  
 loci on chromosome 21 and 2  
 <400> 39  
 acgttggatg tcaggcaggg ttctatgcag 30  
 <210> 40  
 <211> 30  
 <212> DNA  
 <  
 213> Artificial Sequence  
 <220>  
 <223> synthetic digital RCD analysis co-amplification  
 reverse primer for 101 bp amplicon for paralogous  
 loci on chromosome 21 and 2  
 <400> 40  
 acgttggatg aggcggcttc ctggctcttg 30  
 <210> 41  
 <211> 31  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital RCD analysis co-amplification  
 forward primer for 102 bp amplicon for paralogous

loci on chromosome 21 and 6

<400> 41

acgttggatg gctcgtctca ggctcgtagt t 31

<210> 42

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic digital RCD analysis co-amplification  
reverse primer for 102 bp amplicon for paralogous  
loci on chromosome 21 and 6

<400> 42

acgttggatg tttcttcgag cccttcttgg 30

<210> 43

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic extension primer for paralogous sequence  
mismatches (PSM) on paralogous loci on chromosomes  
21 and 1

<400>

> 43

gtctcatctc tacttcgtac etc 23

<210> 44

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic extension primer for paralogous sequence  
mismatches (PSM) on paralogous loci on chromosomes  
21 and 7

<400> 44

ttttacgctg tccccattt 19



<210> 45  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic extension primer for paralogous sequence  
 mismatches (PSM) on paralogous loci on chromosomes  
 21 and 2  
 <400> 45  
 ggtctatgca ggagccgac 19  
 <210> 46  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic extension primer for paralogous sequence  
 mismatches (PSM) on paralogous loci on chromosomes  
 21 and 6  
 <400> 46  
 tgggcgcggg agcggacttc gctgg 25  
 <210> 47  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital PCR analysis co-amplification  
 forward primer for 87-bp amplicon for zinc finger  
 protein sequences on chromosome X (ZFX) and Y  
 (ZFY)  
 <400> 47  
 caagtgctgg actcagatgt aactg 25  
 <210> 48  
 <211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic digital PCR analysis co-amplification  
reverse primer for 87-bp amplicon for zinc finger  
protein sequences on chromosome X (ZFX) and Y  
(ZFY)

<400> 48

tgaagtaatg tcagaagcta aaacatca

28

<210> 49

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic chromosome-specific TaqMan probe  
designed to target chromosome X paralog

<220>

<221> modified\_base

<222> (1)...(1)

<223> t modified by fluorescent reporter VIC

<220>

<221> modified\_base

<222> (16)...(16)

<223> a modified by minor groove binding non-fluorescent  
quencher (MGBNFQ)

<400> 49

tctttagcac attgca

16

<210> 50

<211> 17

<212>

> DNA

<213> Artificial Sequence

<220>

<223> synthetic chromosome-specific TaqMan probe

designed to target chromosome Y paralog

<220>

<221> modified\_base

<222> (1)...(1)

<223> t modified by 6-carboxyfluorescein (FAM)

<220>

<221> modified\_base

<222> (17)...(17)

<223> c modified by minor groove binding non-fluorescent  
quencher (MGBNFQ)

<400> 50

tctttaccac actgcac 17

<210> 51

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic digital PCR co-amplification forward  
primer for 87-bp amplicon of beta-globin  
(hemoglobin beta (HBB)) HbE mutation forward  
primer

<400> 51

gggcaaggtg aacgtgat 19

<210> 52

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic digital PCR co-amplification forward  
primer for 87-bp amplicon of beta-globin  
(hemoglobin beta (HBB)) HbE mutation reverse  
primer

<400> 52

ctattggtct ccttaaacct gtcttgtaa 29

<210> 53

<211> 14

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic allele-G-specific TaqMan probe targeting  
 beta-globin (hemoglobin beta (HBB)) HbE mutation  
 normal allele

<220>

<221> modified\_base

<222> (1)...(1)

<223> t modified by fluorescent reporter VIC

<220>

<221> modified\_base

<222> (14)...(14)

<223> c modified by minor groove binding non-fluorescent  
 quencher (MGBNFQ)

<400> 53

ttggtggtga ggcc 14

<210> 54

<211> 14

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic allele-A-specific TaqMan probe targeting  
 beta-globin (hemoglobin beta (HBB)) HbE mutation  
 mutant allele

<220>

<221> modified\_base

<222> (1)...(1)

<223> t modified by 6-carboxyfluorescein (FAM)

<220>

<221> modified\_base

<222> (14)...(14)

<223> c modified by minor groove binding non-fluorescent quencher (MGBNFQ)

<400> 54  
 ttggtgtaa ggcc 14

<210> 55  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital PCR co-amplification forward primer for 87-bp and 83-bp amplicon of CD41/42 deletion mutation normal and mutant allele

<400> 55  
 ttttccacc ctaggctgc 20

<210> 56  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic digital PCR co-amplification reverse primer for 87-bp and 83-bp amplicon of CD41/42 deletion mutation normal and mutant allele

<400> 56  
 acagcatcag gaggacag atc 23

<210> 57  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> synthetic allele-specific TaqMan probe targeting CD41/42 deletion mutation normal allele (without deletion)

<220>

<221> modified\_base

<222> (1)...(1)

<223> c modified by fluorescent reporter VIC

<220>

<221> modified\_base

<222> (19)...(19)

<223> t modified by minor groove binding non-fluorescent  
quencher (MGBNFQ)

<400> 57

cagagtttct ttgagtcct

19

<210> 58

<211> 15

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic allele-specific TaqMan probe targeting  
CD41/42 deletion mutation mutant allele (with  
deletion)

<220>

<221> modified\_base

<222> (1)...(1)

<223> a modified by 6-carboxyfluorescein (FAM)

<220>

<221> modified\_base

<222> (15)...(15)

<223> t modified by minor groove binding non-fluorescent  
quencher (MGBNFQ)

<400> 58

agagttgag tcctt

15