

명세서

발명의 명칭: 모체 혈액 유래 무세포 DNA 단편을 이용한 비침습적 산전 검사 방법

기술분야

- [1] 본 발명은 모체 혈액 유래 무세포 DNA 단편을 이용한 비침습적 산전 검사 방법에 관한 것으로, 구체적으로 모체와 태아의 DNA 단편 길이 차이를 기초로 모체 혈액 유래 무세포 DNA 단편 및 태아 유래 무세포 DNA 단편을 분리하고, 각각에 대한 z-스코어를 기반으로 태아의 염색체 이수성 또는 모체 모자이시즘을 판단하는 비침습적 산전 검사 방법에 관한 것이다.

[2]

배경기술

- [3] 인간 의학 연구에서 중요한 노력 중 하나는, 불리한 건강결과에 근간이되는 유전적 기형의 발견이다. 이와 같은 유전적 기형의 발견을 위한 산전 검사는 출생 전 태아의 질병 유무를 태아의 염색체 이수성을 통해 판단 및 진단한다.
- [4] 일반적으로, 만 35세 이상의 임산부, 본인 또는 직계가족이 유전적 질환이나 선천성 기형 병력이 있는 임산부, 다태아 임신 경험에 있는 임산부는 고위험 임산부로 분류된다. 최근 고위험 임산부의 수가 지속적으로 증가하는 경향을 보이는데, 이는 여성의 평균 출산 연령이 높아진 것이 가장 큰 원인이다. 고위험 임산부로 분류되면 모체와 태아의 안전을 위해 산전 관리에 많은 주의가 요구되고 산전 검사의 필요가 대두된다.
- [5] 산전 검사는 침습적 산전 검사(invasive prenatal testing) 및 비침습적 산전 검사(non-invasive prenatal testing; NIPT)로 나뉜다. 침습적 산전 검사는 양수천자(amniocentesis), 융모막 검사(chorionic villi sampling) 및 탯줄 천자(cordocentesis) 등이 있다. 그러나, 침습적 산전 검사는 검사 과정에서 태아에게 충격을 가하여 유산이나, 질병 또는 기형을 유발할 수 있어, 이와 같은 문제점들을 극복하기 위해 최근 비침습적 산전 검사 방법이 개발되고 있다.
- [6] 한편, 일반적인 비침습적 산전 검사 방법은 각 염색체에 배열된 염기서열 단편의 수를 전체 염기서열 단편의 수로 나누는 정규화 과정을 통해 z-스코어를 계산하여 정상 염색체와 이수성 염색체를 구별한다. 그러나, 태아 분율(fetal fraction, FF)이 낮거나 모체로부터 유래된 DNA가 증폭되거나 결실이 있으면 정확한 판정을 하기 어렵다.
- [7] 최근에는 차세대 염기서열 분석(next generation sequencing, NGS) 기술인 대규모 병렬형 염기서열 분석(massively parallel signature sequencing, MPSS) 방법이 도입되고, 모체 혈액 내에 존재하는 무세포 DNA(cell-free DNA, cfDNA)에서 무세포 태아 DNA(cell-free fetal DNA, cffDNA)를 발견함에 따라, 이를 활용한 비침습적 산전 검사 방법이 개발되었다. 그러나, 모체의 혈액 내에

존재하는 태아의 cffDNA의 양은 상대적으로 매우 적기 때문에 많은 수의 NGS 리드(read)를 생성하여 판별하는 방식이 일반적으로 사용되고 있어 경제적인 문제가 대두되고 있다. 또한, 위양성(false positive, FP) 또는 위음성(false negative, FN)과 같은 태아 기형의 진단 오류가 없는 더욱 정확하고 민감한 검사 방법의 개발이 필요하다.

- [8] 이와 관련하여, 대한민국 특허공개 제10-2017-0036648호는 모체의 혈액 시료로부터 얻어진 DNA 서열정보로부터 태아의 염색체 이수성 여부를 판별하기 위해, 이수성 여부를 판별하고자 하는 특정 염색체의 평균 리드 수와 상기 염색체를 제외한 다른 염색체의 평균 리드 수를 비교하여 실험간 편차를 제거하고, CV(coefficient of variation) 값으로 가중 평균된 염색체간 리드 수의 비율을 이용함으로써, 결과의 신뢰도 및 특이도를 향상시켜 위양성 확률을 줄일 수 있는 비침습적 태아 염색체 분석 방법을 제공한다.

- [9]
- [10] 이에, 본 발명자들은 정확한 비침습적 산전 검사 방법을 개발하기 위해 연구하던 중, 모체 혈액으로부터 분리된 전체 무세포 DNA 단편을 DNA 단편의 길이에 의존하여 모체 또는 태아 유래 무세포 DNA 단편으로 분리하고, 이들 각각으로부터 계산된 z-스코어를 이용하여 태아의 염색체 이수성 또는 모체 모자이시즘을 더욱 정확하게 판단할 수 있는 분석 방법을 개발함으로써, 본 발명을 완성하였다.

- [11]
- 발명의 상세한 설명**
기술적 과제

- [12] 본 발명의 목적은 모체 혈액 유래 무세포 DNA 단편을 이용한 비침습적 산전 검사 방법 및 장치를 제공하는 것이다.

- [13]
- 과제 해결 수단**

- [14] 상기 목적을 달성하기 위해, 본 발명은 모체의 혈액에서 추출된 무세포 DNA로부터 전체 염기서열 단편을 생산하는 단계; 상기 생산된 전체 염기서열 단편을 인간 참조 유전체 서열의 상동성 위치에 배열하는 단계; 상기 배열된 전체 염기서열 단편을 길이에 따라 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편으로 나누고, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각에 대하여 z-스코어(z-score)를 산출하는 단계; 및 상기 산출된, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각의 z-스코어를 비교하여 모체 모자이시즘 또는 태아의 염색체 이수성을 판단하는 단계를 포함하는 비침습적 산전 검사 방법을 제공한다.

- [15] 또한, 본 발명은 모체의 혈액에서 추출된 무세포 DNA로부터 전체 염기서열

단편을 생산하는 생산부; 상기 생산된 전체 염기서열 단편을 인간 참조 유전체 서열의 상동성 위치에 배열하는 배열부; 상기 배열된 전체 염기서열 단편을 길이에 따라 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편으로 나누고, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각에 대하여 z-스코어(z-score)를 산출하는 산출부; 및 상기 산출된, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각의 z-스코어를 비교하여 모체 모자이시즘 또는 태아의 염색체 이수성을 판단하는 판정부를 포함하는 비침습적 산전 검사 장치를 제공한다.

[16]

발명의 효과

[17] 본 발명에 따른 방법은 모체 혈액으로부터 분리된 전체 무세포 DNA 단편을 DNA 단편의 길이에 의존하여 모체 또는 태아 유래 무세포 DNA 단편으로 분리하고, 이들 각각으로부터 계산된 z-스코어를 이용하여 태아의 염색체 이수성 또는 모체 모자이시즘을 더욱 정확하게 판단할 수 있으므로, 비침습적 산전 검사에 유용하게 사용될 수 있다.

[18]

도면의 간단한 설명

[19] 도 1은 임신부 또는 비임신 여성으로부터 분리된 전체 무세포 DNA로부터 취득된 단편의 길이를 확인한 그래프이다(non-pregnancy.avar: 비임신 여성의 평균값, pregnancy.avar: 임신부의 평균값).

[20] 도 2는 한 임신부의 모체 혈액으로부터 분리된 전체 무세포 DNA, 모체 유래 무세포 DNA 또는 태아 유래 무세포 DNA로부터 취득된 단편의 수에 대한 z-스코어를 이용하여 모체의 21번 염색체에 존재하는 삼염색체를 확인한 결과를 히트맵(heatmap)으로 나타낸 도면이다(Original: 원본데이터, Fetus: 태아 유래 데이터, Mom: 모체 유래 데이터).

[21] 도 3은 한 임신부의 모체 혈액으로부터 분리된 전체 무세포 DNA, 모체 유래 무세포 DNA 또는 태아 유래 무세포 DNA로부터 취득된 단편의 수에 대한 z-스코어를 이용하여 모체의 21번 염색체에 존재하는 삼염색체를 확인한 결과를 히트맵(heatmap)으로 나타낸 도면이다(Original: 원본데이터, Fetus: 태아 유래 데이터, Mom: 모체 유래 데이터).

[22] 도 4는 모체 혈액으로부터 분리된 전체 무세포 DNA, 모체 유래 무세포 DNA 또는 태아 유래 무세포 DNA로부터 취득된 단편의 수에 대한 z-스코어의 평균값을 삼염색체 또는 모체 모자이시즘으로 판단된 경우로 나누어 나타낸 그래프이다(Original: 원본데이터, Fetus: 태아 유래 데이터, Mom: 모체 유래 데이터).

[23] 도 5는 한 임신부의 모체 혈액으로부터 분리된 전체 무세포 DNA, 모체 유래 무세포 DNA 또는 태아 유래 무세포 DNA로부터 취득된 단편의 수에 대한

z -스코어를 이용하여 태아의 X 염색체에 존재하는 일염색체를 확인한 결과를 히트맵으로 나타낸 도면이다(Original: 원본데이터, Fetus: 태아 유래 데이터, Mom: 모체 유래 데이터).

[24] 도 6은 한 임신부의 모체 혈액으로부터 분리된 전체 무세포 DNA, 모체 유래 무세포 DNA 또는 태아 유래 무세포 DNA로부터 획득된 단편의 수에 대한 z -스코어를 이용하여 모체의 X 염색체에 존재하는 일염색체를 확인한 결과를 히트맵으로 나타낸 도면이다(Original: 원본데이터, Fetus: 태아 유래 데이터, Mom: 모체 유래 데이터).

[25] 도 7은 모체 혈액으로부터 분리된 전체 무세포 DNA, 모체 유래 무세포 DNA 또는 태아 유래 무세포 DNA로부터 획득된 단편의 수에 대한 z -스코어의 평균값을 일염색체 또는 모체 모자이시즘으로 판단된 경우로 나누어 나타낸 그래프이다(Original: 원본데이터, Fetus: 태아 유래 데이터, Mom: 모체 유래 데이터).

[26]

발명의 실시를 위한 최선의 형태

[27] 이하, 본 발명을 상세히 설명한다.

[28] 본 발명은 모체의 혈액에서 추출된 무세포 DNA로부터 전체 염기서열 단편을 생산하는 단계; 상기 생산된 전체 염기서열 단편을 인간 참조 유전체 서열의 상동성 위치에 배열하는 단계; 상기 배열된 전체 염기서열 단편을 길이에 따라 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편으로 나누고, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각에 대하여 z -스코어(z -score)를 산출하는 단계; 및 상기 산출된, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각의 z -스코어를 비교하여 모체 모자이시즘 또는 태아의 염색체 이수성을 판단하는 단계를 포함하는 비침습적 산전 검사 방법을 제공한다.

[29] 일 측면에서, 본 발명에 따른 비침습적 산전 검사 방법은 모체의 혈액에서 추출된 무세포 DNA로부터 전체 염기서열 단편을 생산하는 단계를 제공한다.

[30] 본 명세서에서 사용된 용어, "비침습적 산전 검사 방법(non-invasive prenatal testing)"은 양수검사 또는 융모막 검사와 같은 산전 검사 방법과는 달리 침습적 수술절차 없이 태아의 염색체 이상을 확인할 수 있는 검사 방법을 의미한다. 상기 비침습적 산전 검사 방법은 모체의 혈액 속에 존재하는 무세포 태아 DNA의 염기서열 분석을 통해 수행될 수 있다.

[31] 본 발명에 따른 방법에서, 상기 혈액은 혈청 또는 혈장, 구체적으로는 혈장일 수 있다.

[32] 본 명세서에서 사용된 용어, "무세포 DNA(cell-free DNA, cfDNA)"는 세포의 핵안에 존재하지 않고 혈액 내에 떠돌아다니는 DNA 조각을 의미하며, 현재 암의 발병, 전이, 재발, 항암제 반응성 등의 예측에 사용되고 있다. 상기 cfDNA는

통상의 기술분야에 알려진 방법으로 혈액으로부터 분리될 수 있다. 분리된 cfDNA는 대규모 병렬형 염기서열 분석 방법을 사용함으로써 염기서열 단편으로 생산될 수 있다. 이때, 생산된 염기서열 단편은 100만 내지 1,000만개일 수 있다.

- [33] 상기 용어, "대규모 병렬형 염기서열 분석(massively parallel signature sequencing, MPSS)"은 차세대 염기서열 분석(next generation sequencing, NGS) 방법으로서, DNA, RNA 등의 염기서열을 단편으로 생산한 다음 이를 증폭시켜 여러 단편을 동시에 읽어내고 이렇게 획득된 서열 데이터를 생물정보학적 기법(bioinformatics)를 이용하여 조합함으로써 방대한 유전체 정보를 빠르게 해독하는 염기서열 분석 방법을 의미한다.
- [34] 상기 대규모 병렬형 염기서열 분석은 통상의 기술분야에 잘 알려진 방법에 따라 수행될 수 있으며, 상기 방법은 경우에 따라 통상의 기술자에 의해 적절히 변형될 수 있다. 구체적으로, 대규모 병렬형 염기서열 분석은 454 플랫폼(platform)(Margulies 등, Nature, 2005, 437:376-380), 일루미나 게놈 분석기(Illumina Genome Analyzer), 일루미나 HiSeq2000, HiSeq2500, MiSeq, NextSeq500, Life Tech Ion PGM, Ion Proton, Ion S5, Ion S5, Ion S5XL, SOLiD(Applied Biosystems), Helicos True Single Molecule DNA 서열분석 기술(Harris 등, Science, 2008, 320:106-109), Pacific Biosciences의 단일분자 또는 실시간(SMART™) 기술 등에 의해 수행될 수 있다.
- [35] 이때, 염기서열 분석은 페어드-엔드 서열분석(paired-end sequencing) 또는 싱글-리드 서열분석(single-read sequencing) 방법으로 수행될 수 있고, 구체적으로는 페어드-엔드 서열분석 방법으로 수행될 수 있다.
- [36] 다른 측면에서, 본 발명에 따른 비침습적 산전 검사 방법은 상기 생산된 전체 염기서열 단편을 인간 참조 유전체 서열의 상동성 위치에 배열하는 단계를 제공한다.
- [37] 상기 인간 참조 유전체 서열은 인간 유전체 서열로 공지되어 염기서열 데이터베이스에 저장된 서열로서, 본 발명에 따른 방법으로 생산된 염기서열 단편의 서열을 비교하기 위해 사용될 수 있다. 상기 인간 참조 유전체 서열은 빌드 37(build 37, GRCh37), hg18, hg19, hg38 등과 같이 공지된 서열일 수 있고, 본 발명의 일 실시예에 따르면 hg18일 수 있다.
- [38] 상기 단계에서, 전체 염기서열 단편을 인간 참조 유전체 서열의 상동성 위치에 배열하면서 중복된 염기서열 단편을 제거할 수 있다. 또한, 이때, 한 단편의 염기서열 증폭과정에서 순방향(forward) 및 역방향(reverse)으로 배열된 짧은 서열이 동일한 염기서열 단편에서 유래한 데이터만을 취하여 후속 단계에 사용할 수 있다.
- [39] 또 다른 측면에서, 본 발명에 따른 비침습적 산전 검사 방법은 상기 배열된 전체 염기서열 단편을 길이에 따라 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편으로 나누고, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각에 대하여 z-스코어(z-score)를 산출하는 단계를 제공한다.

- [40] 상기 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편은 배열된 염기서열 단편의 길이에 따라 나뉠 수 있다. 이때, 태아 유래 염기서열 단편은 150 bp 이하일 수 있고, 모체 유래 염기서열 단편은 185 bp 이상일 수 있다. 염기서열 단편의 길이에 따라 태아 또는 모체 유래의 염기서열 단편으로 나눔으로써, 본 발명에 따른 비침습적 산전 검사 방법으로 도출된 결과가 모체 모자이시즘인지 태아의 염색체 이수성인지 판단하는데 사용될 수 있다.
- [41] 상기 z-스코어는 통상의 기술분야에 잘 알려진 방법으로 산출될 수 있고, 경우에 따라 적절하게 변형된 방법으로 산출될 수 있다. 본 발명에 따른 검사 방법에서 z-스코어는 배열된 염기서열에서 염색체 이수성을 확인하고자 하는 염색체의 DNA 조각 수를 이용하여 산출될 수 있다. 상기 염색체는 상염색체 또는 성염색체일 수 있다. 상기 상염색체는 인간의 경우 1 내지 22번 염색체로 구성된 군으로부터 선택되는 어느 하나 이상일 수 있고, 성염색체는 인간의 경우 X 또는 Y 염색체일 수 있다.
- [42] 구체적으로, 상기 z-스코어는 하기 수학적 식 1로 계산될 수 있다:
- [43] [수학적 식 1]
- [44]
- $$Z\text{-score}_{(chromosome\ i)} = \frac{Normalized\ Value_{(target\ i)} - Mean_{(reference\ i)}}{Standard\ Deviation_{(reference\ i)}}$$
- [45] (여기서, i는 1 내지 22번 염색체, X 염색체 또는 Y 염색체이고,
- [46] reference는 정상샘플군이며,
- [47] target은 검사하고자 하는 모체 샘플을 의미한다).
- [48] 또 다른 측면에서, 본 발명에 따른 비침습적 산전 검사 방법은 상기 산출된, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편에 대한 z-스코어로부터 모체 모자이시즘 또는 태아의 염색체 이수성을 판단하는 단계를 제공한다.
- [49] 본 명세서에서 사용된 용어, "모체 모자이시즘(maternal mosaicism)"은 태아의 염색체는 정상이지만 모체가 비정상적인 핵형을 가져 비침습적 산전 검사 결과 염색체 이수성이 나타난 것으로 판단되는 경우를 의미한다. 즉, 본 발명에 따른 비침습적 산전 검사 방법은 모체 모자이시즘에 의한 위양성 오류를 감소시켜 결과의 정확도를 향상시킬 수 있다.
- [50] 본 명세서에서 사용된 용어, "염색체 이수성(aneuploidy)"은 염색체의 수가 정상 염색체의 수(2개)와 상이한 것, 즉, 염색체가 0개, 1개, 3개 또는 그 이상 존재하는 것을 의미한다. 이와 같은 염색체 이수성은 태아의 희귀 유전질환을 판단하는데 사용될 수 있다.
- [51] 본 발명에 따른 검사 방법에서, 모체 모자이시즘 또는 태아의 염색체 이수성은 하기에 의해 판단될 수 있다:
- [52] i) 태아 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 크면 태아가 삼염색체인 것으로 판단;

- [53] ii) 모체 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 크면 모체 모자이시즘인 것으로 판단;
- [54] iii) 태아 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 작으면 태아가 일염색체인 것으로 판단; 또는
- [55] iv) 모체 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 작으면 모체 모자이시즘인 것으로 판단.
- [56] 즉, 모체 모자이시즘 또는 태아의 염색체 이수성은 하기에 의해 판단될 수 있다:
- [57] i) 태아 유래 염기서열 단편에 대한 z-스코어에서 전체 염기서열 단편에 대한 z-스코어를 뺀 값이 0 초과이면 태아가 삼염색체인 것으로 판단;
- [58] ii) 모체 유래 염기서열 단편에 대한 z-스코어에서 전체 염기서열 단편에 대한 z-스코어를 뺀 값이 0 초과이면 모체 모자이시즘인 것으로 판단;
- [59] iii) 태아 유래 염기서열 단편에 대한 z-스코어에서 전체 염기서열 단편에 대한 z-스코어를 뺀 값이 0 미만이면 태아가 일염색체인 것으로 판단; 또는
- [60] iv) 모체 유래 염기서열 단편에 대한 z-스코어에서 전체 염기서열 단편에 대한 z-스코어를 뺀 값이 0 미만이면 모체 모자이시즘인 것으로 판단.
- [61] 구체적으로, 태아의 염색체 이수성은 하기에 의해 판단될 수 있다:
- [62] i) 태아 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 크고,
- [63] ii) 모체 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 작으면 태아가 삼염색체인 것으로 판단.
- [64] 또한, 태아의 염색체 이수성은 하기에 의해 판단될 수 있다:
- [65] i) 태아 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 작고,
- [66] ii) 모체 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 크면 태아가 일염색체인 것으로 판단.
- [67] 본 발명에 따른 검사방법은 상염색체 이상, 성염색체 이상 또는 상염색체 결실을 판단할 수 있다. 예를 들어, 상기 상염색체 이상은 염색체의 수가 3개 또는 그 이상, 구체적으로 3개인 경우일 수 있다. 상염색체 이상으로 나타날 수 있는 질병은 다운 증후군, 에드워드 증후군, 파타우 증후군, 9번 삼염색체증, 16번 삼염색체증 또는 22번 삼염색체증일 수 있다. 한편, 상기 성염색체 이상은 X 또는 Y 염색체가 추가 또는 결실된 것일 수 있다. 성염색체 이상으로 나타날 수 있는 질병은 구체적으로 터너 증후군, 트리플X 증후군, 클라인펠터 증후군 또는 제이콥스 증후군일 수 있다. 또한, 상기 상염색체 결실은 상염색체의 일부가 결실되어 나타날 수 있는 질병일 수 있다. 상염색체 결실로 나타날 수 있는 질병은 구체적으로 1번 단완 염색체 결실 증후군, 2q33.1 결실 증후군, 월프-허쉬호른 증후군, 고양이울음 증후군, 윌리엄스 증후군, 제이콥슨 증후군, 프라더-윌리/엔젤만 증후군 또는 디조지 증후군일 수 있다.

- [68]
- [69] 또한, 본 발명은 모체의 혈액에서 추출된 무세포 DNA로부터 전체 염기서열 단편을 생산하는 생산부; 상기 생산된 전체 염기서열 단편을 인간 참조 유전체 서열의 상동성 위치에 배열하는 배열부; 상기 배열된 전체 염기서열 단편을 길이에 따라 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편으로 나누고, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각에 대하여 z-스코어(z-score)를 산출하는 산출부; 및 상기 산출된, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각의 z-스코어를 비교하여 모체 모자이시즘 또는 태아의 염색체 이수성을 판단하는 판정부를 포함하는 비침습적 산전 검사 장치를 제공한다.
- [70] 상기 비침습적 산전 검사 장치는 상술한 비침습적 산전 검사 방법과 동일한 특징을 가질 수 있다. 일례로, 상기 염기서열 단편은 대규모 병렬형 염기서열 분석 방법으로 생산될 수 있고, 이때 생산된 염기서열 단편은 그 길이에 따라 태아 또는 모체 유래 염기서열 단편으로 구분할 수 있다. 또한, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각으로부터 산출된 z-스코어를 이용하여 모체 모자이시즘 또는 태아의 염색체 이수성은 하기에 의해 판단될 수 있다:
- [71] i) 태아 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 크면 태아가 삼염색체인 것으로 판단;
- [72] ii) 모체 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 크면 모체 모자이시즘인 것으로 판단;
- [73] iii) 태아 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 작으면 태아가 일염색체인 것으로 판단; 또는
- [74] iv) 모체 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 작으면 모체 모자이시즘인 것으로 판단.
- [75] 본 발명에 따른 비침습적 산전 검사 장치는 상염색체 이상, 성염색체 이상 또는 염색체 결실을 판단할 수 있다. 구체적으로, 상기 상염색체 이상은 다운 증후군, 에드워드 증후군, 파타우 증후군, 9번 삼염색체증, 16번 삼염색체증 또는 22번 삼염색체증을 유발할 수 있고, 상기 성염색체 이상은 터너 증후군, 트리플X 증후군, 클라인펠터 증후군 또는 제이콥스 증후군을 유발할 수 있으며, 상기 상염색체 결실은 1번 단완 염색체 결실 증후군, 2q33.1 결실 증후군, 월프-허쉬호른 증후군, 고양이울음 증후군, 윌리엄스 증후군, 제이콥슨 증후군, 프라더-윌리/엔젤만 증후군 또는 디조지 증후군을 유발할 수 있다.

[76]

발명의 실시를 위한 형태

- [77] 이하, 본 발명을 하기 실시예에 의해 상세히 설명한다. 단, 하기 실시예는 본 발명을 예시하는 것일 뿐, 본 발명이 이들에 의해 제한되는 것은 아니다. 본

발명의 청구범위에 기재된 기술적 사상과 실질적으로 동일한 구성을 갖고 동일한 작용 효과를 이루는 것은 어떠한 것이라도 본 발명의 기술적 범위에 포함된다.

[78]

[79] 실시예 1. 임신부 및 비임신 여성으로부터 수득된 무세포 DNA(cell free DNA, cfDNA) 확인

[80]

[81] 1-1. 혈액으로부터 cfDNA의 분리

[82] 임신부와 비임신 여성으로부터 수득된 cfDNA의 단편 길이를 비교하기 위해 먼저, 임신부 또는 비임신 여성의 혈액을 수집하였다.

[83] 구체적으로, 피검자로부터 10 ml의 혈액을 취하여 VanGenes사의 무세포 DNA 혈액 수집 튜브에 넣고, 이를 상온에서 1,900 g의 조건으로 15분 동안 원심분리하였다. 원심분리 후, 상청액을 1.5 ml 튜브에 옮겨서 이를 다시 상온에서 16,000 g의 조건으로 15분 동안 원심분리하여 상청액을 다시 수득하였다. QIAasympy DSP Virus/Pathogen Midi Kit(Qiagen)를 사용하여 제조사의 프로토콜에 따라 상기 수득된 2 ml의 상청액(혈장)으로부터 cfDNA를 분리하였다.

[84]

[85] 1-2. cfDNA 단편의 길이 확인

[86] 실시예 1-1.에서 임신부 또는 비임신 여성으로부터 분리한 100 ng의 cfDNA를 Illumina NextSeq500을 사용하여 75회 페어드-엔드 서열분석(75-cycle paired-end sequencing)을 통해 각각 단편 라이브러리 세트를 제조하였다. 이후, BWA(버전 0.7.10) 프로그램을 사용하여 제조된 단편을 인간 참조 유전체 서열(hg19)의 상동성 위치에 배열하였다. 이때, Picard(버전 1.81) 프로그램을 사용하여 중복된 염기서열 단편을 제거하였다. 배열된 단편 중에서 순방향(forward) 및 역방향(reverse)으로 배열된 짧은 서열이 동일한 염기서열 단편에서 유래한 경우의 데이터만 취합하였다(이를 properly paired read라 함).

[87] 이와 같이 구해진 bam 파일 및 SAMtools(버전 0.1.18) 프로그램을 사용하여 각 염색체에 배열된 염기서열 단편의 수를 산출하고, 각각의 단편 길이에 대한 분포를 도 1에 그래프로 나타내었다.

[88] 도 1에 나타난 바와 같이, 160 bp 이하의 길이를 갖는 염기서열 단편이 비임신 여성에 비해 임신부로부터 분리된 cfDNA에서 높은 비율로 존재했다. 반면, 160 bp 이하의 짧은 길이 단편을 많이 갖는 임신부는 비임신 여성에 비해 160 bp 이상의 길이를 갖는 단편은 상대적으로 낮았다(도 1). 즉, 상기로부터 임신부의 cfDNA에 포함된 cffDNA에 의해 임신부 cfDNA 단편의 길이가 짧은 것을 알 수 있었다.

[89]

[90] 실시예 2. 모체 혈액으로부터 cfDNA의 분리

[91] 비침습적 산전 검사를 통해 태아의 염색체 이수성을 진단하기 위해, 먼저 모체 혈액으로부터 cfDNA를 다음과 같은 방법으로 분리하였다. 이때, 630명의 임신부로부터 혈액을 수집하였고, 이중 195명은 21번 염색체에 삼염색체(trisomy) 이수성 염색체를 가졌다. cfDNA는 임신부로부터 혈액을 수득한 것을 제외하고는 실시예 1-1.에 기재된 바와 조건 및 방법으로 분리하였다.

[92]

[93] 실시예 3. cfDNA 단편의 배열 및 염기서열 단편의 수 산출

[94] 실시예 2에서 분리한 100 ng의 cfDNA를 이용하여 실시예 1-2.에 기재된 바와 동일한 조건 및 방법으로 cfDNA 단편을 배열하고, 배열된 염기서열 단편의 수를 산출하였다. 이때, 산출된 데이터는 원본 데이터로서 보관하였다. 한편, bam 파일 및 SAMtools(버전 0.1.18) 프로그램을 사용하여 염기서열의 길이가 150 bp 보다 짧은 서열만을 취하여 단편의 수를 산출하고, 태아 유래 데이터로서 보관하였다. 또한, 염기서열의 길이가 185 bp 보다 긴 서열만을 취하여 단편의 수를 산출하고, 모체 유래 데이터로서 보관하였다. 이때, 150 bp 내지 185 bp의 길이를 갖는 염기서열은 태아와 모체를 구분하기에 적합하지 않은 혼합 영역으로, 태아 또는 모체 유래 데이터 분석에는 사용하지 않았다.

[95]

[96] 실험예 1. 21번 염색체의 삼염색체(trisomy 21) 분석

[97] 실시예 3에서 산출된 원본데이터, 태아 유래 데이터 및 모체 유래 데이터를 이용하여 계산된 z-스코어를 이용하여 21번 염색체의 삼염색체를 분석하였다. 일반적인 z-스코어는, 태아의 염색체 이수성이 발견되지 않은 모체의 혈액 샘플을 정상샘플군으로서 이용하여 특정 염색체에 대해 하기 수학적 1을 사용하여 계산하였고, 이때 분석 알고리즘은 공지의 문헌을 참조하여 수행하였다(Hyuk Jung Kwon et al., Open Journal of Genetics, 2017, 7, p.1-8).

[98] [수학적 1]

[99]
$$Z\text{-score}_{(chromosome\ i)} = \frac{Normalized\ Value_{(target\ i)} - Mean_{(reference\ i)}}{Standard\ Deviation_{(reference\ i)}}$$

[100] (여기서, i는 1 내지 22번 염색체, X 염색체 또는 Y 염색체이고,

[101] reference는 정상샘플군이며,

[102] target은 검사하고자 하는 모체 샘플을 의미한다)

[103]

[104] 상기로부터 수득된 결과에서, 태아 유래 데이터로부터 계산된 z-스코어에서 원본데이터로부터 계산된 z-스코어를 뺀 값이 0 초과면 태아가 삼염색체를 갖는 것으로 판단하였고, 모체 유래 데이터로부터 계산된 z-스코어에서 원본데이터로부터 계산된 z-스코어를 뺀 값이 0 초과면 모체 모자이시즘(maternal mosaicism)으로 판단하였다. 그 결과는 히트맵(heatmap)

그래프로 작성하여 도 1에 나타내었다. 이때, 21번 염색체 구획의 한칸은 각각의 z-스코어 값을 나타내고 이는 임계값과 비교하여 이수성을 판정하는 근거가 된다.

- [105] 도 2에 나타난 바와 같이, 21번 염색체에 삼염색체를 갖는 한명의 임신부 cfDNA로부터 취득된 단편은 21번 염색체 구획(가로기준)에서 z-스코어 값이 높았다(도 2).
- [106] 한편, 도 3에 나타난 바와 같이, 원본데이터의 21번 염색체 구획(가로기준)에는 z-스코어가 태아의 염색체 이수성을 판단할만큼 명확하게 나타나지 않았다. 그러나, 이를 모체 또는 태아 유래 데이터로 나누어 확인한 경우에, 태아 유래 데이터의 21번 염색체 구획(가로기준)에는 z-스코어가 유의적으로 높은 반면, 모체 유래 데이터에서는 높지 않음을 확인하였다(도 3). 따라서, 이로부터 태아가 21번 염색체에 이수성을 가짐을 확인할 수 있었다.
- [107] 이에, 상기와 같이 각각의 임신부 혈액으로부터 취득된 원본데이터, 모체 또는 태아 유래 데이터의 z-스코어를 확인하고, 태아의 삼염색체 또는 모체 모자이시즘으로 판단된 경우로 나누어 z-스코어에 대한 그래프를 작성하여 도 4에 나타내었다.
- [108] 도 4에 나타난 바와 같이, 삼염색체 또는 모체 모자이시즘으로 판단된 모든 경우의 z-스코어가 유효한 값을 보였다. 모체 모자이시즘으로 판단된 경우, 원본데이터 및 태아 유래 데이터에서는 z-스코어 차이를 보이지 않았으나, 모체 유래 데이터의 z-스코어가 현저히 높았다(파란선). 반면, 태아의 삼염색체로 판단된 경우, 태아 유래 데이터의 z-스코어가 유의적으로 높았다(빨간선).
- [109]
- [110] **실험예 2. 성염색체의 일염색체(monosomy X) 분석**
- [111] 먼저, 실시예 3에서 산출된 원본데이터, 태아 유래 데이터 및 모체 유래 데이터를 이용하여 계산된 z-스코어를 이용하여 성염색체인 X 염색체의 일염색체를 분석하였다. 상기로부터 취득된 결과에서, 태아 유래 데이터로부터 계산된 z-스코어에서 원본데이터로부터 계산된 z-스코어를 뺀 값이 0 미만이면 태아의 X 염색체가 일염색체인 것으로 판단하였고, 모체 유래 데이터로부터 계산된 z-스코어에서 원본데이터로부터 계산된 z-스코어를 뺀 값이 0 미만이면 모체 모자이시즘으로 판단하였다. 그 결과는 히트맵 그래프로 작성하여 도 5 및 6에 나타내었다.
- [112] 도 5에 나타난 바와 같이, 태아의 X 염색체에 일염색체를 갖는 한명의 임신부 cfDNA로부터 취득된 단편은 X 염색체 구획(가로기준)에서 z-스코어 값이 낮았다. 즉, 상기 결과는 태아의 X 염색체가 일염색체임을 나타내었다. 또한, 원본데이터를 태아 또는 모체 유래 데이터로 나누어 비교분석한 결과, 모체 유래 데이터에 비해 태아 유래 데이터에서 더 낮은 z-스코어를 보임으로써, 상기 결과가 태아에게 나타나는 이수성 염색체임을 확인하였다(도 5).
- [113] 한편, 도 6에 나타난 바와 같이, 모체의 X 염색체에 일염색체를 갖는 한명의

임산부 cfDNA로부터 수득된 단편은 X 염색체 구획(가로기준)에서 z-스코어 값이, 태아 유래 데이터에 비해 모체 유래 데이터에서 더 낮음을 보임으로써, 태아는 정상의 염색체를 가짐을 확인하였다(도 6).

- [114] 또한, 상기와 같이 각각의 임산부 혈액으로부터 수득된 원본데이터, 모체 또는 태아 유래 데이터의 z-스코어를 확인하고, 태아의 일염색체 또는 모체 모자이시즘으로 판단된 경우로 나누어 z-스코어에 대한 그래프를 작성하여 도 7에 나타내었다.
- [115] 도 7에 나타난 바와 같이, 일염색체 또는 모체 모자이시즘으로 판단된 모든 경우의 z-스코어가 유효한 값을 보였다. 모체 모자이시즘으로 판단된 경우, 원본데이터 및 태아 유래 데이터에서는 z-스코어 차이를 보이지 않았으나, 모체 유래 데이터의 z-스코어가 현저히 낮았다(파란선). 반면, 태아의 일염색체로 판단된 경우, 태아 유래 데이터의 z-스코어가 유의적으로 낮았다(빨간선).
- [116]
- [117] 따라서, 결론적으로, 본 발명에 따른 방법은 21번 염색체뿐만 아니라, 9, 13, 18번 염색체 등의 상염색체 및 성염색체의 이상이나 결실을 판단하는데 적용할 수 있다.

청구범위

- [청구항 1] 1) 모체의 혈액에서 추출된 무세포 DNA로부터 전체 염기서열 단편을 생산하는 단계;
 2) 상기 단계 1)에서 생산된 전체 염기서열 단편을 인간 참조 유전체 서열의 상동성 위치에 배열하는 단계;
 3) 상기 단계 2)에서 배열된 전체 염기서열 단편을 길이에 따라 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편으로 나누고, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각에 대하여 z-스코어(z-score)를 산출하는 단계; 및
 4) 상기 단계 3)에서 산출된, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각의 z-스코어를 비교하여 모체 모자이시즘 또는 태아의 염색체 이수성을 판단하는 단계를 포함하는 비침습적 산전 검사 방법.
- [청구항 2] 제1항에 있어서, 염기서열 단편은 대규모 병렬형 염기서열 분석 방법으로 생산되는, 비침습적 산전 검사 방법.
- [청구항 3] 제1항에 있어서, 전체 염기서열 단편을 길이에 따라 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편으로 나누는 과정에서, 상기 태아 유래 염기서열 단편은 150 bp 이하로 분리하는, 비침습적 산전 검사 방법.
- [청구항 4] 제1항에 있어서, 전체 염기서열 단편을 길이에 따라 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편으로 나누는 과정에서, 상기 모체 유래 염기서열 단편은 185 bp 이상으로 분리하는, 비침습적 산전 검사 방법.
- [청구항 5] 제1항에 있어서, 모체 모자이시즘 또는 태아의 염색체 이수성은 하기에 의해 판단되는, 비침습적 산전 검사 방법:
 i) 태아 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 크면 태아가 삼염색체인 것으로 판단;
 ii) 모체 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 크면 모체 모자이시즘인 것으로 판단;
 iii) 태아 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 작으면 태아가 일염색체인 것으로 판단; 또는
 iv) 모체 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 작으면 모체 모자이시즘인 것으로 판단.
- [청구항 6] 제1항에 있어서, 태아의 염색체 이수성은 하기에 의해 판단되는, 비침습적 산전 검사 방법:
 i) 태아 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 크고,
 ii) 모체 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 작으면 태아가 삼염색체인 것으로 판단.

- [청구항 7] 제1항에 있어서, 태아의 염색체 이수성은 하기에 의해 판단되는, 비침습적 산전 검사 방법:
 i) 태아 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 작고,
 ii) 모체 유래 염기서열 단편에 대한 z-스코어가 전체 염기서열 단편에 대한 z-스코어보다 크면 태아가 일염색체인 것으로 판단.
- [청구항 8] 제1항에 있어서, z-스코어는 하기 수학적 식 1로 계산되는, 비침습적 산전 검사 방법:
 [수학적 식 1]

$$Z\text{-score}_{(\text{chromosome } i)} = \frac{\text{Normalized Value}_{(\text{target } i)} - \text{Mean}_{(\text{reference } i)}}{\text{Standard Deviation}_{(\text{reference } i)}}$$

(여기서, i는 1 내지 22번 염색체, X 염색체 또는 Y 염색체이고, reference는 정상샘플군이며, target은 검사하고자 하는 모체 샘플을 의미한다).

- [청구항 9] 제1항에 있어서, 상기 비침습적 산전 검사 방법은 상염색체 이상, 성염색체 이상 또는 상염색체 결실을 판단하는, 비침습적 산전 검사 방법.
- [청구항 10] 제9항에 있어서, 상기 상염색체 이상은 다운 증후군, 에드워드 증후군, 파타우 증후군, 9번 삼염색체증, 16번 삼염색체증 또는 22번 삼염색체증을 유발하는, 비침습적 산전 검사 방법.
- [청구항 11] 제9항에 있어서, 상기 성염색체 이상은 터너 증후군, 트리플X 증후군, 클라인펠터 증후군 또는 제이콥스 증후군을 유발하는, 비침습적 산전 검사 방법.
- [청구항 12] 제9항에 있어서, 상기 상염색체 결실은 1번 단완 염색체 결실 증후군, 2q33.1 결실 증후군, 윌프-허쉬호른 증후군, 고양이울음 증후군, 윌리암스 증후군, 제이콥슨 증후군, 프라더-윌리-엔젤만 증후군 또는 디조지 증후군을 유발하는, 비침습적 산전 검사 방법.
- [청구항 13] 1) 모체의 혈액에서 추출된 무세포 DNA로부터 전체 염기서열 단편을 생산하는 생산부;
 2) 상기 단계 1)에서 생산된 전체 염기서열 단편을 인간 참조 유전체 서열의 상동성 위치에 배열하는 배열부;
 3) 상기 단계 2)에서 배열된 전체 염기서열 단편을 길이에 따라 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편으로 나누고, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각에 대하여 z-스코어(z-score)를 산출하는 산출부; 및
 4) 상기 단계 3)에서 산출된, 전체 염기서열 단편, 태아 유래 염기서열 단편 및 모체 유래 염기서열 단편 각각의 z-스코어를 비교하여 모체 모자이시즘 또는 태아의 염색체 이수성을 판단하는 판정부를 포함하는

비침습적 산전 검사 장치.