



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0021402
(43) 공개일자 2024년02월19일

(51) 국제특허분류(Int. Cl.)
G16B 20/20 (2019.01) G16B 45/00 (2019.01)
G16B 50/00 (2019.01)
(52) CPC특허분류
G16B 20/20 (2019.02)
G16B 45/00 (2019.02)
(21) 출원번호 10-2022-0099671
(22) 출원일자 2022년08월10일
심사청구일자 2022년08월10일

(71) 출원인
한림대학교 산학협력단
강원도 춘천시 한림대학길 1, 한림대학교(옥천동)
(72) 발명자
신재문
전라북도 남원시 대산면 남문로 125-9
전준범
강원도 춘천시 봉의산길22번길 29, 304호 (옥천동, 봉향원룸)
(74) 대리인
특허법인본

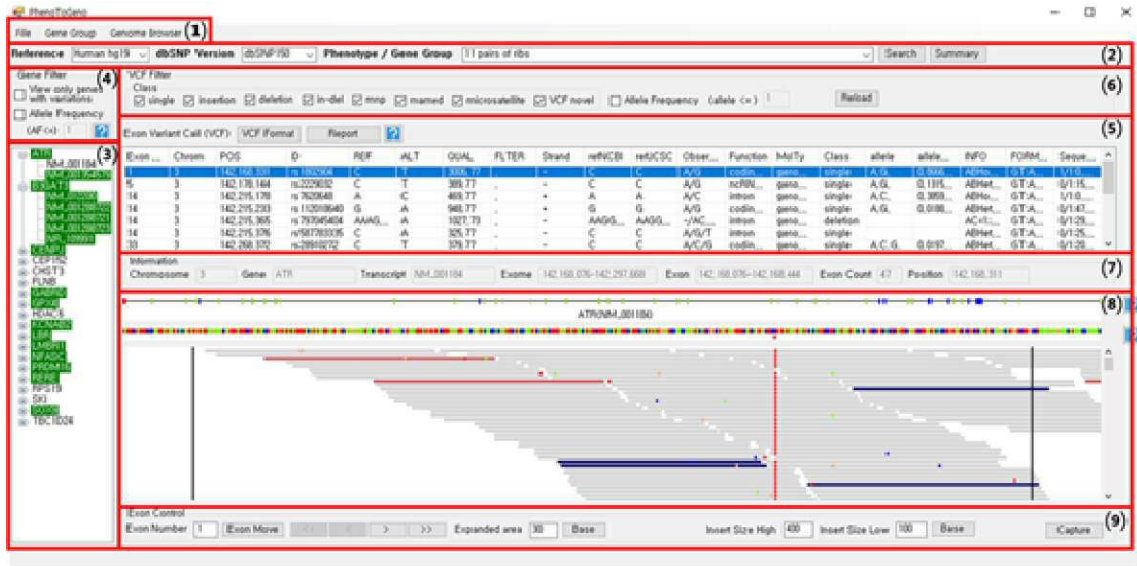
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 사용자 친화적인 유전자 변이 검출을 수행하고 정밀 의학을 위한 시각화 도구를 제공하는 전자 장치의 제어 방법

(57) 요약

전자 장치의 제어 방법이 개시된다. 본 제어 방법은, HPO(Human Phenotype Ontology) 데이터베이스를 바탕으로, 사용자 입력에 따라 선택된 표현형(phenotype)과 관련된 유전자 셋을 추출하는 단계, 추출된 유전자 셋을 reference data와 비교하여, 적어도 하나의 유전자에 대한 변이 정보를 획득하는 단계, 유전자 셋을 시각적으로 출력하는 단계를 포함한다. 유전자 셋을 시각적으로 출력하는 단계는, 유전자 셋 중 변이가 발생한 유전자와 변이가 발생하지 않은 유전자가 구분되도록 표시한다.

대표도 - 도1



(52) CPC특허분류

G16B 50/00 (2019.02)

(72) 발명자

정다운

서울특별시 동작구 만양로12길 49, 303호

김기용

서울특별시 서초구 사임당로20길 48-8

김윤중

경기도 용인시 기흥구 언동로 217번길 31, 205동
903 (신동백서해그랑블 2차)

정동훈

강원도 춘천시 승지골길16번길 47, 1010동 1304호
(뜨란채아파트)

윤지희

서울특별시 송파구 송파대로 345, 406동 302호 (헬
리오시티아파트)

명세서

청구범위

청구항 1

전자 장치의 제어 방법에 있어서,

HP0(Human Phenotype Ontology) 데이터베이스를 바탕으로, 사용자 입력에 따라 선택된 표현형(phenotype)과 관련된 유전자 셋을 추출하는 단계;

상기 추출된 유전자 셋을 reference data와 비교하여, 적어도 하나의 유전자에 대한 변이 정보를 획득하는 단계;

상기 유전자 셋을 시각적으로 출력하는 단계;를 포함하고,

상기 유전자 셋을 시각적으로 출력하는 단계는,

상기 유전자 셋 중 변이가 발생한 유전자와 변이가 발생하지 않은 유전자가 구분되도록 표시하는, 전자 장치의 제어 방법.

청구항 2

제1항에 있어서,

상기 유전자 셋을 시각적으로 출력하는 단계는,

변이 종류 및 변이 수치 중 적어도 하나에 대한 필터링 조건을 설정하는 사용자 입력이 수신되면, 상기 유전자 셋 중 상기 필터링 조건에 매칭되는 변이가 발생한 유전자만 시각적으로 출력하는, 전자 장치의 제어 방법.

청구항 3

제1항에 있어서,

상기 전자 장치의 제어 방법은,

상기 유전자 셋에 포함되는 적어도 하나의 유전자를 선택하는 사용자 입력이 수신되면, 상기 선택된 유전자에 매칭되는 전사체(transcript)에 대한 정보를 시각적으로 출력하는 단계; 및

상기 시각적으로 출력된 전사체를 선택하는 사용자 입력이 수신되면, 상기 전사체에 대한 엑손(exon) 구조, 인트론(intron) 구조, 및 변이 정보를 시각적으로 출력하는 단계;를 포함하는, 전자 장치의 제어 방법.

청구항 4

제3항에 있어서,

상기 전자 장치의 제어 방법은,

상기 변이 정보에 포함되는 적어도 하나의 변이를 선택하는 사용자 입력이 수신되면, 상기 변이를 포함하는 전사체의 엑손-인트론 구조를 시각적으로 출력하는 단계;

상기 엑손-인트론 구조에서 적어도 하나의 엑손을 선택하는 사용자 입력이 수신되면, 상기 엑손을 포함하는 엑손 영역 내 변이의 위치를 reference sequence와 함께 시각적으로 출력하는 단계; 및

상기 엑손-인트론 구조에서 적어도 하나의 엑손을 선택하는 사용자 입력이 수신되면, 상기 엑손을 포함하는 엑손 영역을 나타내는 read sequence와 상기 reference sequence를 정렬하여 변이가 존재하는 부분이 구분되도록 시각적으로 출력하는 단계;를 포함하는, 전자 장치의 제어 방법.

청구항 5

제3항에 있어서,

상기 전자 장치의 제어 방법은,

상기 변이 정보에 포함되는 적어도 하나의 변이를 선택하는 사용자 입력이 수신되면, 상기 변이가 발생한 위치를 포함하는 엑손 영역 및 상기 변이가 발생한 위치를 중심으로 하는 게놈 브라우저(Genome Browser)를 제공하는 단계;를 더 포함하는, 전자 장치의 제어 방법.

청구항 6

제5항에 있어서,

상기 게놈 브라우저를 제공하는 단계는,

분석을 위한 데이터 및 환경을 설정하기 위한 control 패널, 브라우저의 대상인 유전체의 cytoband를 포함하는 cytoband 패널, 각 유전체 위치에 대한 read depth가 동적으로 계산된 결과에 해당하는 coverage graph를 포함하는 coverage graph 패널, 브라우저의 대상인 유전체의 유전자 구조를 포함하는 genetic structure 패널, 및 리드 정렬과 변이 추출 결과를 포함하는 main genome view 패널을 시각적으로 출력하는, 전자 장치의 제어 방법.

청구항 7

제6항에 있어서,

상기 게놈 브라우저를 제공하는 단계는,

사용자 입력에 따라 0에서 22까지의 값을 가지는 해상도 레벨(n)을 설정하고,

이하 수학적식에 따라 상기 게놈 브라우저를 구성하는 복수의 패널 각각의 해상도를 변경하고,

$$S = \frac{\text{ChrLength}}{2^n}$$

상기 S는 해상도 레벨 n의 값에 따라 화면에 출력되는 뉴클레오타이드 bps 단위의 유전체 영역의 최대 길이이고, 상기 ChrLength는 염색체의 최대 길이인, 전자 장치의 제어 방법.

청구항 8

제7항에 있어서,

상기 게놈 브라우저를 제공하는 단계는,

상기 해상도 레벨이 저해상도인 경우, 상기 genetic structure 패널을 통해 엑손-인트론 구조를 시각적으로 출력하고,

상기 해상도 레벨이 고해상도인 경우, 아미노산 시퀀스를 시각적으로 출력하고,

상기 아미노산 시퀀스는,

서로 다른 컬러의 사각형이 반복되는 엑손 영역과 얇은 선으로 나타나는 인트론 영역을 포함하는, 전자 장치의 제어 방법.

청구항 9

제6항에 있어서,

상기 게놈 브라우저를 제공하는 단계는,

상기 main genome view 패널을 통해, reference sequence 상의 변이 위치를 염기 단위로 제공하고,

상기 main genome view 패널을 통해, reference sequence에 대한 read sequence의 정렬 결과 및 상기 정렬 결과에 따른 변이 정보를 염기 단위로 제공하는, 전자 장치의 제어 방법.

청구항 10

전자 장치의 프로세서에 의해 실행되어, 상기 전자 장치로 하여금 제1항의 제어 방법을 수행하도록 하는 컴퓨터

프로그램이 저장된, 컴퓨터 판독 가능 매체.

발명의 설명

기술 분야

[0001] 본 개시는 전자 장치의 제어 방법에 관한 것으로, 보다 상세하게는 대규모 유전체 데이터에 대한 질병과 유전자 변이의 관계를 시각화하는 전자 장치의 제어 방법에 관한 것이다.

배경 기술

[0002] 개인 맞춤 의학(Personalized Medicine)이라고도 하는 정밀의학(Precision Medicine)은 환자 개인의 임상적 표현형, 건강기록, 다양한 오믹스(omics) 데이터를 결합하여 인간의 질병에 대한 맞춤형 진단, 치료, 예방을 위한 새로운 의학 분야이다. 인간 게놈 프로젝트의 완료와 차세대 DNA 시퀀싱(NGS) 기술의 발전으로, 이제는 단일 뉴클레오타이드 분해능에서 환자의 유전적 변이를 식별하고 인간 질병에 대한 표현형 결과를 해석할 수 있게 되었다.

[0003] 여러 NGS 접근법이 복잡한 인간 질병과 관련된 유전적 변이의 식별에 사용되었다.

[0004] 여기에는 전체 게놈 시퀀싱, 전체 엑솜 시퀀싱 및 표적 유전자 패널 시퀀싱이 포함된다. 최근 알츠하이머병, 암, 관상동맥 질환 및 제2형 당뇨병과 같은 인간 질병과 관련된 유전적 변이체를 식별하기 위한 GWAS(genome-wide association Studies)의 메타 분석이 대규모로 시도되고 있다. 예를 들어 APOE, CD33, CLU, CR1, EPHA1, PICALM 및 TREM2와 같은 여러 유전자가 알츠하이머병의 주요 유전적 위험 인자로 확인되었다.

[0005] BRCA1 및 BRCA2 돌연변이는 전체 유방암 사례의 약 5%를 차지한다. 그러나 BRCA1과 BRCA2의 병원성 돌연변이는 유방암의 위험을 각각 65%와 45% 증가시킨다. 이는, 이들 유전자의 다형성으로 설명되는 낮은 총 유전율에도 불구하고 BRCA1 및 BRCA2의 유전적 변이를 모니터링하는 것이 고위험 환자의 유방암 진단에 유리하다는 것을 의미한다.

[0006] 따라서, 유전적 위험 예측을 용이하게 하기 위해서는 특정 인간 질병 및 표현형과 관련된 유전자 목록을 얻는 것이 필수적이다. 이미 MeSH(Medical Subject Headings), NCI(National Cancer Institute)의 Thesaurus, OMIM(Online Mendelian 상속 in man), SNOMEDCT, United Medical Language System(UMLS) 및 Human Phenotype Ontology(HPO)와 같은 많은 임상 데이터베이스가 있다. 이 중에서 HPO는 질병 유전자에 대한 의미론적 링크와 인간 표현형의 컴퓨터 분석을 위한 온톨로지를 포함하는 보다 포괄적인 리소스를 제공한다.

[0007] 개별 환자의 NGS 데이터에는 단일 뉴클레오타이드 변이체(SNV), 삽입 및 결실(indels)과 같은 다양한 유형의 유전적 변이 외에도 역전 및 전위와 같은 유전적 이상이 포함된다.

[0008] DeepVariant, DELLY, FermiKit, GATK HaplotypeCaller(GATK HC), Pindel, Platypus, Strelka2 및 VarScan과 같은 NGS 데이터에서 변이 호출을 위한 많은 계산 도구가 있다. 이러한 변이 검출 도구를 사용하여 분자 진단을 위한 질병 관련 유전자의 유전적 변이를 검출할 수 있다.

[0009] NGS 데이터에서 식별된 유전자 변이체는 VCF(변이체 호출 형식), MAF(돌연변이 주석 파일), MUT(돌연변이 파일 형식) 또는 기타 형식의 파일로 저장된다. 이 중 VCF는 돌연변이 데이터를 저장하는 데 가장 널리 사용되는 커뮤니티 표준이다. 그러나, VCF 파일의 분석에는 집중적인 생물정보학 전문 지식이 필요하다.

[0010] 따라서, VCF 데이터를 관리하고 시각화하기 위한 도구가 개발될 필요가 있다. 현재 사용 가능한 VCF 처리 프로그램에는 SNVerGUI, database.bio, DaMold, mirVAFC, GAVIN 및 gNOME이 포함된다. 또한 NGS 데이터에서 질병 관련 유전 변이를 식별하기 위해 IGV(통합 게놈 뷰어), VCF-Miner, VCF.Filter, myVCF, BrowseVCF 및 VCF-Server와 같은 다른 도구도 개발되었다. 이 중 IGV는 게놈 변이 데이터를 시각화하는 데 가장 널리 사용되는 도구이다. 그러나, 생물정보학자가 아닌 연구원의 진단 및 임상 의사 결정 지원을 위해 환자의 게놈 데이터를 분석할 수 있는 쉽고 포괄적인 도구가 여전히 필요하다.

발명의 내용

해결하려는 과제

[0011] 본 개시는 직관적인 사용자 인터페이스로서 Phenotype to Genotype Variation(PhenGenVar) 브라우저 애플리케이션

이션을 활용하는 전자 장치의 제어 방법을 제공한다. 본 전자 장치는, 의사나 연구자가 특정 질병이나 표현형과 관련된 선택된 유전자 셋에서 환자의 유전적 변이를 모니터링하도록 설계되었다.

[0012] 본 개시는, PhenGenVar 응용 프로그램 기반의 Exom Browser 및 Genome Browser를 바탕으로 단일 뉴클레오타이드 수준 뿐만 아니라 유전자 또는 엑손 수준에서 유전적 변이를 감지할 수 있는 전자 장치의 제어 방법을 제공한다.

[0013] 본 개시의 목적들은 이상에서 언급한 목적으로 제한되지 않으며, 언급되지 않은 본 개시의 다른 목적 및 장점들은 하기의 설명에 의해서 이해될 수 있고, 본 개시의 실시 예에 의해 보다 분명하게 이해될 것이다. 또한, 본 개시의 목적 및 장점들은 특허 청구 범위에 나타낸 수단 및 그 조합에 의해 실현될 수 있음을 쉽게 알 수 있을 것이다.

[0014] 본 개시는 "교비 학술연구과제"로서 과제번호 "H20200350"에 해당하는 "차세대 시퀀싱 데이터를 위한 변이 분석 브라우저"와 관련된 것이다.

과제의 해결 수단

[0015] 본 개시의 일 실시 예에 따른 전자 장치의 제어 방법은, HPO(Human Phenotype Ontology) 데이터베이스를 바탕으로, 사용자 입력에 따라 선택된 표현형(phenotype)과 관련된 유전자 셋을 추출하는 단계, 상기 추출된 유전자 셋을 reference data와 비교하여, 적어도 하나의 유전자에 대한 변이 정보를 획득하는 단계, 상기 유전자 셋을 시각적으로 출력하는 단계를 포함한다. 상기 유전자 셋을 시각적으로 출력하는 단계는, 상기 유전자 셋 중 변이가 발생한 유전자와 변이가 발생하지 않은 유전자가 구분되도록 표시한다.

[0016] 상기 유전자 셋을 시각적으로 출력하는 단계는, 변이 종류 및 변이 수치 중 적어도 하나에 대한 필터링 조건을 설정하는 사용자 입력이 수신되면, 상기 유전자 셋 중 상기 필터링 조건에 매칭되는 변이가 발생한 유전자만 시각적으로 출력할 수 있다.

[0017] 상기 전자 장치의 제어 방법은, 상기 유전자 셋에 포함되는 적어도 하나의 유전자를 선택하는 사용자 입력이 수신되면, 상기 선택된 유전자에 매칭되는 전사체(transcript)에 대한 정보를 시각적으로 출력하는 단계, 상기 시각적으로 출력된 전사체를 선택하는 사용자 입력이 수신되면, 상기 전사체에 대한 엑손(exon) 구조, 인트론(intron) 구조, 및 변이 정보를 시각적으로 출력하는 단계를 포함할 수 있다.

[0018] 이때, 상기 전자 장치의 제어 방법은, 상기 변이 정보에 포함되는 적어도 하나의 변이를 선택하는 사용자 입력이 수신되면, 상기 변이를 포함하는 전사체의 엑손-인트론 구조를 시각적으로 출력하는 단계, 상기 엑손-인트론 구조에서 적어도 하나의 엑손을 선택하는 사용자 입력이 수신되면, 상기 엑손을 포함하는 엑손 영역 내 변이의 위치를 reference sequence와 함께 시각적으로 출력하는 단계, 상기 엑손-인트론 구조에서 적어도 하나의 엑손을 선택하는 사용자 입력이 수신되면, 상기 엑손을 포함하는 엑손 영역을 나타내는 read sequence와 상기 reference sequence를 정렬하여 변이가 존재하는 부분이 구분되도록 시각적으로 출력하는 단계를 포함할 수 있다.

[0019] 또한, 상기 전자 장치의 제어 방법은, 상기 변이 정보에 포함되는 적어도 하나의 변이를 선택하는 사용자 입력이 수신되면, 상기 변이가 발생한 위치를 포함하는 엑손 영역 및 상기 변이가 발생한 위치를 중심으로 하는 게놈 브라우저(Genome Browser)를 제공하는 단계를 더 포함할 수 있다.

[0020] 상기 게놈 브라우저를 제공하는 단계는, 분석을 위한 데이터 및 환경을 설정하기 위한 control 패널, 브라우저의 대상인 유전체의 cytoband를 포함하는 cytoband 패널, 각 유전체 위치에 대한 read depth가 동적으로 계산된 결과에 해당하는 coverage graph를 포함하는 coverage graph 패널, 브라우저의 대상인 유전체의 유전자 구조를 포함하는 genetic structure 패널, 및 리드 정렬과 변이 추출 결과를 포함하는 main genome view 패널을 시각적으로 출력할 수 있다.

[0021] 또한, 상기 게놈 브라우저를 제공하는 단계는, 사용자 입력에 따라 0에서 22까지의 값을 가지는 해상도 레벨(n)을 설정하고, 이하 수확식에 따라 상기 게놈 브라우저를 구성하는 복수의 패널 각각의 해상도를 변경할 수

있다.
$$S = \frac{ChrLength}{2^n}$$
 상기 S는 해상도 레벨 n의 값에 따라 화면에 출력되는 뉴클레오타이드 bps 단위의 유전체 영역의 최대 길이이고, 상기 ChrLength는 염색체의 최대 길이이다.

[0022] 또한, 상기 게놈 브라우저를 제공하는 단계는, 상기 해상도 레벨이 저해상도인 경우, 상기 genetic structure

패널을 통해 엑손-인트론 구조를 시각적으로 출력하고, 상기 해상도 레벨이 고해상도인 경우, 아미노산 시퀀스를 시각적으로 출력할 수 있다. 상기 아미노산 시퀀스는, 서로 다른 컬러의 사각형이 반복되는 엑손 영역과 얇은 선으로 나타나는 인트론 영역을 포함할 수 있다.

[0023] 또한, 상기 게놈 브라우저를 제공하는 단계는, 상기 main genome view 패널을 통해, reference sequence 상의 변이 위치를 염기 단위로 제공하고, 상기 main genome view 패널을 통해, reference sequence에 대한 read sequence의 정렬 결과 및 상기 정렬 결과에 따른 변이 정보를 염기 단위로 제공할 수 있다.

발명의 효과

[0024] 본 개시에 따른 전자 장치의 제어 방법은, 지식 기반 유전자 셋에서 유전자 변이를 신속하게 식별하고, 특정 질병과 관련된 유전자 셋을 식별하여 다양한 형태로 시각화할 수 있다.

[0025] 종래의 프로그램들과 비교할 때 PhenGenVar 프로그램은 질병에 민감한 유전자의 유해한 돌연변이에 중요한 각 엑손 및 엑손-인트론 접합 영역의 유전적 변이를 신속하게 모니터링할 수 있으며, 저비용 고효율의 시각화 서비스가 가능하다.

도면의 간단한 설명

[0026] 도 1은 본 개시의 일 실시 예에 따른 전자 장치가 제공하는 Exome Browser의 메인 화면을 도시한 도면,
 도 2는 본 개시의 일 실시 예에 따른 Exome Browser의 menu bar를 도시한 도면,
 도 3은 본 개시의 일 실시 예에 따른 file 메뉴의 구성을 도시한 도면,
 도 4 내지 도 6은 각각 gene group의 관리, 추가, 수정을 위한 대화상자를 도시한 도면,
 도 7은 본 개시의 일 실시 예에 따른 control panel의 구성을 도시한 도면,
 도 8은 본 개시의 일 실시 예에 따른 전자 장치가 제공하는 통계 결과 출력 화면을 도시한 도면,
 도 9는 본 개시의 일 실시 예에 따른 전자 장치가 제공하는 gene filter panel 및 gene/transcript view panel을 도시한 도면,
 도 10은 본 개시의 일 실시 예에 따른 전자 장치가 제공하는 exon variant call panel을 도시한 도면,
 도 11은 본 개시의 일 실시 예에 따른 전자 장치가 제공하는 VCF format 정보 확인 창을 도시한 도면,
 도 12는 본 개시의 일 실시 예에 따른 전자 장치가 제공하는 information panel을 도시한 도면,
 도 13은 본 개시의 일 실시 예에 따른 전자 장치가 제공하는 main exon view panel을 도시한 도면,
 도 14는 본 개시의 일 실시 예에 따른 전자 장치가 transcript 구조를 표현하는 동작을 설명하기 위한 도면,
 도 15는 본 개시의 일 실시 예에 따른 전자 장치가 제공하는 exon 영역을 도시한 도면,
 도 16은 본 개시의 일 실시 예에 따른 전자 장치가 reference sequence 상의 변이 위치를 표현하는 동작을 설명하기 위한 도면,
 도 17은 본 개시의 일 실시 예에 따른 전자 장치가 제공하는 read alignment viewer를 도시한 도면,
 도 18은 본 개시의 일 실시 예에 따른 전자 장치가 SNP 변이 영역을 표현하는 동작을 설명하기 위한 도면,
 도 19는 본 개시의 일 실시 예에 따른 전자 장치가 deletions 영역을 표현하는 동작을 설명하기 위한 도면,
 도 20은 본 개시의 일 실시 예에 따른 전자 장치가 insertion 영역을 표현하는 동작을 설명하기 위한 도면,
 도 21은 본 개시의 일 실시 예에 따른 전자 장치가 Genome Browser 상에서 insertion 영역을 표현하는 동작을 설명하기 위한 도면,
 도 22는 본 개시의 일 실시 예에 따른 전자 장치가 view control panel을 제공하는 동작을 설명하기 위한 도면,
 도 23은 본 개시의 일 실시 예에 따른 전자 장치가 capture 버튼에 의하여 생성할 수 있는 캡처 이미지를 도시한 도면,

도 24는 본 개시의 일 실시 예에 따른 전자 장치가 Genome Browser를 호출하는 메뉴 바를 설명하기 위한 도면,

도 25는 본 개시의 일 실시 예에 따른 전자 장치가 exon variant call panel에서 Genome Browser를 호출하는 동작을 설명하기 위한 도면,

도 26은 본 개시의 일 실시 예에 따른 전자 장치가 exon variant call panel에서 호출된 Genome Browser를 실행하는 화면을 도시한 도면,

도 27은 본 개시의 일 실시 예에 따른 전자 장치가 Exome Browser의 main view panel에서 Genome Browser를 호출하는 동작을 설명하기 위한 도면,

도 28은 본 개시의 일 실시 예에 따른 전자 장치가 Genome Browser의 Main Window를 제공하는 동작을 설명하기 위한 도면,

도 29는 본 개시의 일 실시 예에 따른 전자 장치가 Genome Browser의 control panel을 제공하는 동작을 설명하기 위한 도면,

도 30은 본 개시의 일 실시 예에 따른 전자 장치가 cytoband panel을 제공하는 동작을 설명하기 위한 도면,

도 31 내지 도 32는 본 개시의 일 실시 예에 따른 전자 장치가 coverage graph 및 상세 정보를 표시하는 동작을 설명하기 위한 도면,

도 33 내지 도 34는 본 개시의 일 실시 예에 따른 전자 장치가 저해상도 및 고해상도 각각에서 유전자 구조를 표시하는 동작을 설명하기 위한 도면,

도 35는 본 개시의 일 실시 예에 따른 전자 장치가 main genome view panel을 표시하는 동작을 설명하기 위한 도면,

도 36은 본 개시의 일 실시 예에 따른 전자 장치가 최소 축소 및 최대 확대 각각의 경우에 있어 제공하는 variants along the reference sequence 영역을 도시한 도면,

도 37은 본 개시의 일 실시 예에 따른 전자 장치가 확대 표현된 리드 정렬 결과를 출력하는 동작을 설명하기 위한 도면,

도 38은 본 개시의 일 실시 예에 따른 전자 장치가 축소 표현된 리드 정렬 결과를 출력하는 동작을 설명하기 위한 도면,

도 39는 본 개시의 일 실시 예에 따른 전자 장치가 read sequence의 방향을 표시하는 동작을 설명하기 위한 도면,

도 40은 본 개시의 일 실시 예에 따른 전자 장치가 read sequence의 insert size를 표시하는 동작을 설명하기 위한 도면,

도 41은 본 개시의 일 실시 예에 따른 전자 장치가 SNP 변이 영역을 표현하는 동작을 설명하기 위한 도면,

도 42는 본 개시의 일 실시 예에 따른 전자 장치가 VCF 파일 정보가 있는 경우의 SNP 변이 영역을 표현하는 동작을 설명하기 위한 도면,

도 43은 본 개시의 일 실시 예에 따른 전자 장치가 read sequence의 deletion을 표현하는 동작을 설명하기 위한 도면, 그리고

도 44는 본 개시의 일 실시 예에 따른 전자 장치가 insertions 영역을 표현하는 동작을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0027] 본 개시에 대하여 구체적으로 설명하기에 앞서, 본 명세서 및 도면의 기재 방법에 대하여 설명한다.

[0028] 먼저, 본 명세서 및 청구범위에서 사용되는 용어는 본 개시의 다양한 실시 예들에서의 기능을 고려하여 일반적인 용어들을 선택하였다. 하지만, 이러한 용어들은 당해 기술 분야에 종사하는 기술자의 의도나 법률적 또는 기술적 해석 및 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 일부 용어는 출원인이 임의로 선정한 용어도 있다. 이러한 용어에 대해서는 본 명세서에서 정의된 의미로 해석될 수 있으며, 구체적인 용어 정의가 없으

면 본 명세서의 전반적인 내용 및 당해 기술 분야의 통상적인 기술 상식을 토대로 해석될 수도 있다.

- [0029] 또한, 본 명세서에 첨부된 각 도면에 기재된 동일한 참조번호 또는 부호는 실질적으로 동일한 기능을 수행하는 부품 또는 구성요소를 나타낸다. 설명 및 이해의 편의를 위해서 서로 다른 실시 예들에서도 동일한 참조번호 또는 부호를 사용하여 설명한다. 즉, 복수의 도면에서 동일한 참조 번호를 가지는 구성요소가 모두 도시되어 있다고 하더라도, 복수의 도면들이 하나의 실시 예를 의미하는 것은 아니다.
- [0030] 또한, 본 명세서 및 청구범위에서는 구성요소들 간의 구별을 위하여 "제1", "제2" 등과 같이 서수를 포함하는 용어가 사용될 수 있다. 이러한 서수는 동일 또는 유사한 구성요소들을 서로 구별하기 위하여 사용하는 것이며 이러한 서수 사용으로 인하여 용어의 의미가 한정 해석되어서는 안 된다. 일 예로, 이러한 서수와 결합된 구성 요소는 그 숫자에 의해 사용 순서나 배치 순서 등이 제한되어서는 안 된다. 필요에 따라서는, 각 서수들은 서로 교체되어 사용될 수도 있다.
- [0031] 본 명세서에서 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "구성되다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0032] 본 개시의 실시 예에서 "모듈", "유닛", "부(Part)" 등과 같은 용어는 적어도 하나의 기능이나 동작을 수행하는 구성요소를 지칭하기 위한 용어이며, 이러한 구성요소는 하드웨어 또는 소프트웨어로 구현되거나 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다. 또한, 복수의 "모듈", "유닛", "부(Part)" 등은 각각이 개별적인 특정한 하드웨어로 구현될 필요가 있는 경우를 제외하고는, 적어도 하나의 모듈이나 칩으로 일체화되어 적어도 하나의 프로세서로 구현될 수 있다.
- [0033] 또한, 본 개시의 실시 예에서, 어떤 부분이 다른 부분과 연결되어 있다고 할 때, 이는 직접적인 연결뿐 아니라, 다른 매체를 통한 간접적인 연결의 경우도 포함한다. 또한, 어떤 부분이 어떤 구성요소를 포함한다는 의미는, 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0034] 본 개시에 따른 전자 장치의 제어 방법은, PhenGenVar라는 소프트웨어를 실행하여 이하 다양한 동작을 수행할 수 있다. PhenGenVar는 Exome Browser, Genome Browser 등을 제공할 수 있으며, 이하 설명들을 통해 순차적으로 설명한다.
- [0035] PhenGenVar 브라우저 애플리케이션은 Microsoft.Net 프레임워크를 사용하여 C# 프로그래밍 언어로 작성되었다. 응용 프로그램은 Windows 운영 체제에서 수행할 수 있는 그래픽 인터페이스로 개발되었다. 위치, 주식 및 아미노산 조성을 포함하는 인간 참조 게놈 데이터(human reference genome data) 셋은 UCSC 게놈 브라우저 데이터베이스에서 획득될 수 있다. 중복된 읽기를 제거하고 시퀀스 정렬의 오류를 최소화하기 위해 플래그와 간결한 CIGAR(idiosyncratic gapped alignment report) 문자열 필드가 BAM 파일에서 읽기 데이터의 오류 수정 및 재정렬에 사용되었다.
- [0036] PhenGenVar에서, 대규모 게놈 데이터의 효율적인 시각화를 위해 데이터 분할 및 인덱싱 방법이 채택되었다. 게놈 데이터는 효율적인 인덱싱 체계를 사용하여 별도로 관리 및 액세스할 수 있는 파티션으로 나뉜다. 분할로 인한 정보 손실을 줄이기 위해 분할된 각 영역에는 인접 분할 영역 사이에 겹치는 영역이 포함되었다. 디스플레이 처리에서 인덱싱된 파티션이 메모리에 로드되고 해당 게놈 영역이 해상도에 따라 화면에 표시된다. 브라우저에서 시퀀스 정렬의 해상도는 100bp에서 100kb 창으로 조정되었다. 다중 해상도의 많은 출력 이미지를 렌더링하는 속도 저하를 최소화하기 위해 Direct2D 기반 렌더링이 사용되었다.
- [0037] 인간 참조 게놈 데이터는 hg19 및 hg38 버전으로 PhenGenVar 애플리케이션에 다운로드되어 사전 설치될 수 있다. SNP 데이터베이스(dbSNP)는 National Center for Biotechnology Information에서 다운로드 될 수 있다. hg19 인간 게놈에 사용할 수 있는 dbSNP 버전은 SNP138, SNP141, SNP142, SNP144, SNP146, SNP147, SNP150 및 SNP151에 해당한다. hg38에 사용할 수 있는 dbSNP 버전은 SNP141, SNP142, SNP144, SNP146, SNP147, SNP150 및 SNP151이다. 유전자 셋은 HPO에서 다운로드된 것일 수 있으며, 인간 질병 및 표현형과 관련된 유전자 목록을 제공한다.
- [0038] 환자 진단을 위한 PhenGenVar 애플리케이션을 테스트하기 위해 공개적으로 사용 가능한 개인의 전체 게놈 시퀀스 데이터가 샘플 수탁 번호 NA11995인 1000 게놈 프로젝트에서 검색되었다.

- [0040] 본 개시에 따른 PhenGenVar를 사용하는 전자 장치는, 특정 표현형과 관련된 게놈 영역 및 변이체를 탐색할 수 있다.
- [0041] PhenGenVar는 Exom Browser와 Genome Browser의 두 가지 개별 브라우저로 구성된다. Exom Browser는 사용자가 정의한 표현형과 관련 유전자 셋을 선택하고 특정 유전자의 유전적 변이를 모니터링하여 유전자 수준 분석을 위해 설계된 메인 프로그램이다. Exome Browser 출력은 보고된 변이체 및 자세한 정보와 함께 해당 게놈 영역을 표시한다.
- [0042] Genome Browser는 이후에 Exome Browser에서 호출될 수 있으며 읽기 정렬을 기반으로 한 변이체 호출 결과를 면밀히 탐색하는 데 사용될 수 있다. 따라서 PhenGenVar 응용 프로그램은 다양한 수준의 게놈 분해능에서 편리하고 직관적인 사용자 인터페이스를 제공할 수 있다.
- [0044] 유전자 수준 변이 분석을 위한 PhenGenVar의 Exome Browser는 엑솜 시퀀싱 또는 전체 게놈 시퀀싱과 같은 환자 게놈 데이터의 유전적 변이에 대한 유전자 수준 분석을 위한 개요 페이지이다. 본 브라우저 페이지는 주로 데이터 업로드 영역, 유전자 및 변이체 필터링 패널, 엑손 뷰어 패널로 구성된다(도 1).
- [0045] 데이터 분석을 시작하려면, 환자의 엑솜 또는 전체 게놈 시퀀싱 데이터에서 생성된 BAM(Binary Sequence Alignment Map) 파일과 VCF 파일을 업로드해야 한다. 환자의 질병 표현형과 관련된 유전자의 유전적 변이를 탐색하기 위해 유전자 목록을 만들거나 HUGO Gene Nomenclature Committee(HNGC) 공식 유전자 기호 이름으로 개별적으로 입력할 수 있다. 또한, Exome Browser는 인간 질병에서 발견되는 표현형 이상과 관련된 13,000개 이상의 용어에 대한 포괄적인 리소스인 인간 표현형 온톨로지(HPO)에 등록된 유전자 그룹을 제공한다. VCF/BAM 파일 데이터를 사용하여 각 유전자의 유전적 변이를 분석하기 위해 사용자는 특정 인간 참조 게놈 서열 버전과 dbSNP 버전을 유연하게 선택할 수 있다.
- [0047] Exome Browser는 대규모 유전체 데이터에 대한 질병-유전자-변이 관계를 효율적으로 시각화한다. 특히, HPO(Human Phenotype Ontology) 데이터베이스를 기반으로 특정 phenotype과 관련된 유전자 영역을 자동추출한 후, 각 유전자 영역에 대한 변이 분석을 수행할 수 있는 탐-다운 방식의 탐색 방법을 제공한다. Exome Browser의 메인 화면은 도 1과 같다. 메인 화면은 크게 9개의 영역으로 구성된다. 9개 영역의 각 기능을 <표 1>에 보인다.

표 1

번호	영역	기능
(1)	Menu Bar	메뉴 표시
(2)	Control Panel	BAM/VCF 파일 분석을 위한 제어
(3)	Gene/Transcript View Panel	해당하는 Gene/Transcript List 출력
(4)	Gene Filter Panel	Gene/Transcript View Panel에 출력되는 Gene/Transcript List에 대한 필터링 조건 설정
(5)	Exon Variant Call Panel	선택한 Transcript에 속한 엑손 영역의 변이 정보 출력
(6)	VCF Filter Panel	선택한 Transcript에 속한 변이 정보에 대한 필터링 조건 설정
(7)	Information Panel	현재 선택된 변이에 관련된 정보 출력
(8)	Main Exon View Panel	선택한 Transcript에 대하여 엑손 단위로 리드 정렬 결과 및 출현 변이를 시각화하여 보임
(9)	Exon View Control Panel	Main Exon View Panel의 제어

- [0050] (1) Menu Bar
- [0051] Menu Bar는 File, Gene Group, Genome Browser 메뉴로 구성된다(도 2 참조).
- [0052] ① File menu

- [0053] Exome Browser에서 사용되는 파일을 열기 위한 메뉴이다. File 메뉴는 BAM file open과 VCF file open을 하위 메뉴로 갖는다 (도 3 참조). 사용자가 제공하는 리드 정렬 결과는 BAM 표준 포맷의 파일로 입력된다. 또한 유전적 구조 변이 정보는 VCF 표준 포맷의 파일로 제공되어야 한다.
- [0054] ▶ BAM file open: 리드 정렬 결과 분석을 위하여 BAM 파일을 입력하여야 한다. BAM 파일을 불러오기 위해서 File 메뉴의 BAM file open 메뉴를 선택한다. 메뉴 선택 시 파일 선택 창이 생성되며, 하나의 BAM 파일을 선택한 후 열기 버튼을 누르면 BAM 파일이 입력된다. Browser는 정렬 결과 분석을 위하여 BAM 파일의 인덱싱 파일인 BAI 파일을 참고하게 되며, 따라서 BAM 파일과 BAM.BAI 파일이 동일한 폴더에 있어야 한다. 만약 BAI 파일이 없으면 Browser는 BAM 파일을 열 수 없으며, 프로그램을 실행할 수 없다.
- [0055] ▶ VCF file open: 유전적 구조 변이 결과를 분석하기 위하여 VCF 파일을 입력하여야 한다. VCF 파일을 불러오기 위해서 File 메뉴의 VCF file open 메뉴를 선택한다. 메뉴 선택 시 Browser는 파일 선택 창을 생성한다. 하나의 VCF 파일을 선택한 후 열기 버튼을 누르면 VCF 파일을 Browser에 불러온다. Browser는 VCF 파일의 인덱싱 파일인 IDX 파일을 참조한다. 따라서 VCF 파일과 VCF.IDX 파일은 동일한 폴더에 있어야 한다. 만약 IDX 파일이 없으면 Browser는 VCF 파일을 열 수 없다.
- [0056] ② Gene Group menu
- [0057] 사용자 정의 Gene Group을 생성/수정/삭제하는 기능을 제공한다. Gene Group 메뉴를 실행하면 Gene Group 관리 대화상자가 생성된다 (도 4 참조). Gene Group 관리 대화상자는 두 개의 ListView와 세 개의 버튼으로 구성되어 있다. 두 개의 ListView 중 Group Name ListView는 Gene Group 관리 대화상자의 왼쪽에 위치하며 사용자가 이전에 생성한 Gene Group 이름들이 출력된다. Gene Group 관리 대화상자의 오른쪽에 위치한 Gene Name ListView에는 선택한 Group 이름으로 등록된 Gene Symbol들이 출력된다. Gene Group 관리 대화상자의 하단에 있는 New, Edit, Delete 버튼의 기능은 다음과 같다.
- [0058] ▶ New Button: New Button을 클릭하면 새로운 Gene Group을 추가하기 위한 대화상자가 생성된다 (도 5 참조). 대화상자의 Group Name 입력창에 생성하고자 하는 Gene Group의 이름을 입력한다. 다음 하단의 ListView에 Gene Symbol들을 입력한 후 OK 버튼을 누르면 새로운 이름의 Gene Group이 생성된다. 단, 입력되는 모든 Gene Symbol은 HGNC Official Symbol을 사용하여야 한다.
- [0059] ▶ Edit Button: Gene Group 관리 대화상자의 Group Name ListView에서 임의의 Group을 선택한 후 Edit 버튼을 클릭하면 선택한 Group에 등록된 Gene Symbol들을 수정할 수 있다 (도 6 참조). 대화상자의 하단에 위치한 Gene Symbol을 제거하거나 새로운 Gene Symbol을 등록한 후 OK 버튼을 클릭하면 해당하는 Gene Group에 수정된 Gene Symbol들이 등록되게 된다.
- [0060] ▶ Delete Button: Gene Group 관리 대화상자의 Group Name ListView에서 임의의 Gene Group을 선택한 후 Delete 버튼을 클릭하면 삭제 확인 대화상자가 생성된다. 생성된 대화상자에서 OK 버튼을 클릭하면 해당 Group이 삭제된다. 이때 삭제된 Group은 복구할 수 없다.
- [0061] ③ Genome Browser 메뉴
- [0062] Genome Browser를 호출하기 위한 메뉴이다. Genome Browser를 호출하기 전에 사용자가 제공하는 BAM 파일이 로드되어 있어야 하며, File 메뉴를 사용하여 BAM 파일을 입력할 수 있다. 유전적 구조 변이 정보를 포함하는 VCF 파일은 선택적으로 입력할 수 있으나, VCF 파일이 입력되면 보다 세밀한 변이 정보 분석 결과를 얻을 수 있다 (Genome Browser 매뉴얼 참조).
- [0064] (2) Control Panel
- [0065] Control Panel은 BAM/VCF 파일 분석을 위한 데이터 선택 패널로써 5개의 영역으로 구성된다(도 7 참조). 5개 영역의 각 기능은 <표 2>와 같다.

표 2

번호	영역	기능
①	Reference build Combo Box	표준 참조 서열 선택

②	dbSNP Version Combo Box	dbSNP 버전 선택
③	Phenotype/Gene Group Combo Box	분석하고자 하는 Phenotype 또는 Gene Group 선택
④	Search Button	선택한 Phenotype 또는 Gene_Group에 해당하는 Gene Set 추출
⑤	Summary Button	선택한 Phenotype 또는 Gene_Group에 해당하는 Gene Set에서 검출된 변이의 요약 정보 출력

[0068] ① Reference build Combo BoxReference build Combo Box는 분석에 필요한 Reference sequence 버전을 선택하기 위하여 사용한다. 본 Browser에서는 Reference sequence 버전으로서 Human hg_19와 Human hg_38을 사용할 수 있다.

[0069] ② dbSNP Version Combo Box

[0070] dbSNP Version Combo Box는 분석에 필요한 dbSNP 버전을 선택하기 위하여 사용한다. dbSNP 버전은 선택한 Reference sequence에 따라 제공하는 버전이 다르다. Reference sequence에 따라 선택 가능한 dbSNP 버전은 <표 3>과 같다.

표 3

Reference sequence	dbSNP 버전
Human hg_19	SNP_138, SNP_141, SNP_142, SNP_144, SNP_146, SNP_147, SNP_150, SNP_151
Human hg_38	SNP_141, SNP_142, SNP_144, SNP_146, SNP_147, SNP_150, SNP_151

[0072] ③ Phenotype/Gene_GroupPhenotype/Gene_Group 영역은 사용자가 분석하고자 하는 Phenotype 또는 Gene_Group을 입력/선택하기 위하여 사용된다. Phenotype의 경우, 검색 대상의 키워드를 (부분)입력하면 자동완성(word completion) 기능에 의하여 HPO 데이터베이스를 기반으로 관련 phenotype list가 출력되며, 이 중 해당하는 phenotype을 선택할 수 있다. 또한 Gene_Group의 경우, 검색 대상의 Gene_Group명을 (부분)입력하면 자동완성(word completion) 기능에 의하여 등록된 Gene_Group을 선택하여 사용할 수 있다.

[0073] ④ Search Button

[0074] Search 버튼은 Phenotype/Gene_Group 영역에서 입력/선택한 Phenotype과 관련된 Gene set (HPO에 관련 Gene set으로 등록된 것) 또는 Gene_Group에 포함되어 있는 Gene set을 불러오기 위하여 사용된다. 이 들 Gene Set은 Gene/Transcript View 영역에 출력된다. 또한 사용자가 Gene Filter를 사용하여 필터링 조건을 새로이 입력하여 변경된 Gene Set을 Gene/Transcript View 영역에 불러오기 위하여 하는 경우, Search Button을 다시 클릭하여야 한다.

[0075] ⑤ Summary Button

[0076] Summary 버튼은 사용자가 입력한 변이 정보 (VCF 파일)의 통계 정보를 요약하여 출력하기 위하여 사용한다. 통계 정보는 사용자가 선택한 Phenotype 또는 Gene_Group의 Gene Set을 대상으로 하며, Gene 단위로 요약 정보가 생성된다. 세부 내용은 <표 4>와 같다. 통계 정보를 출력하기 위하여 Summary 버튼을 클릭하면 통계 정보 출력을 위한 대화상자가 실행된다 (도 8 참조). 통계 정보 출력 창에서 SAVE 버튼을 클릭하면 출력 화면을 파일로 저장할 수 있으며, Tab으로 구분된 tsv 파일로 저장된다.

표 4

항목	의미
Gene Name	Gene의 이름
Number of SNPs	해당 Gene에서 추출된 모든 SNP 수
Number of dbSNPs	해당 Gene에서 추출된 모든 SNP 중 dbSNP에 등록된 SNP 수
Number of Noble SNPs	해당 Gene에서 추출된 모든 SNP 중 dbSNP에 등록되지 않은 SNP 수
Number of Indels	해당 Gene에서 추출된 모든 Indel 수
Number of dbSNP_Indels	해당 Gene에서 추출된 모든 Indel 중 dbSNP에 등록된 Indel 수

Number of Noble_Indels	해당 Gene에서 추출된 모든 Indel 중 dbSNP에 등록되지 않은 Indel 수
Number of multiallelic sites	해당 Gene에서 추출된 모든 multiallelic sites 수
Number of multiallelic SNP sites	해당 Gene에서 추출된 모든 multiallelic sites 중 SNP 수

[0078] 구체적으로, 환자의 게놈 데이터와 유전자 그룹이 업로드되면 분석된 출력 데이터가 도 8과 같이 요약 테이블로 표시된다. 도 8은 선택된 유전자 그룹의 각 유전자에 대한 SNP 데이터를 보여준다. SNP 데이터에는 dbSNP 데이터베이스에 등록된 SNP뿐만 아니라 희귀 변형 버전 및 dbSNP 버전일 수 있는 새로운 SNP도 포함될 수 있다.

[0079] 도 9와 같이 유전자 변형이 있는 유전자 셋은 Exome Browser의 유전자 필터 패널을 통해 선택될 수 있다. 이 필터 패널에서 대립유전자 빈도 값을 조정하여 유전적 변이가 더 큰 유전자의 범위를 좁힐 수도 있다. 유전자 필터링 외에도 VCF 필터 패널(도 10, 상단 패널)을 통해 보다 자세한 필터링 옵션이 사용될 수 있다. 유전자 필터 패널에서 특정 유전자가 선택되면 해당 유전자의 모든 변이체 정보가 Exome Browser의 exon variant call 패널에 표시된다(도 10, 하단 패널). 이 패널은 위치, dbSNP 참조 번호(rsID), 참조 게놈과 비교한 서열 변경 및 변이 유형과 같은 상세한 변이 정보를 제공한다. 위의 VCF 필터 패널에서 시퀀스 변이 유형을 클릭하고 대립 유전자 빈도를 조정하여 추가 필터링을 수행할 수 있다.

[0081] (3) Gene/Transcript View Panel

[0082] Gene/Transcript View 영역에는 사용자가 선택한 Phenotype/Gene_Group에 해당하는 Gene set이 Transcript 정보와 함께 출력된다(도 9 하단 참조). 해당 Gene/Transcript 영역에 변이 추출이 보고되어 있는 경우, 해당하는 Gene/Transcript 이름의 배경이 녹색으로 출력되며, 해당 Gene/Transcript 영역에 변이 추출이 보고되어 있지 않은 경우, 해당하는 Gene/Transcript 이름의 배경이 흰색으로 출력된다. 여기에서 사용되는 변이 정보는 사용자가 입력한 VCF 파일을 기반으로 추출된다. Gene/Transcript View 영역에 출력되는 모든 Gene은 HGNC에서 승인한 Official Symbol을 사용한다. 이때 동의어(Synonym)나 이전에 승인된 기호(Previous symbol)는 지원하지 않는다. Gene/Transcript View 영역에 출력된 임의의 Gene을 더블클릭하면 해당 Gene에 대한 Transcript 정보가 지식 노트로 출력된다. 또한 지식 노트에 출력된 임의의 Transcript를 더블클릭 하면 Browser의 Exon Variant Call 영역과 Information 영역이 활성화되어, 해당 Transcript에 대한 엑손/인트론 구조, 변이 정보 등을 상세히 분석할 수 있는 환경이 제공된다.

[0084] (4) Gene Filter Panel

[0085] Gene/Transcript View Panel에 출력되는 Gene/Transcript set에 필터링 조건을 부가할 수 있다(도 9 상단 참조). <표 5>에 보이는 두 가지의 Filtering 조건이 지원되며, 각각의 필터링 조건은 체크 버튼을 이용하여 활성화된다. “View only genes with variations” 항목을 이용하여 Gene/Transcript View 영역에 변이가 보고되어 있는 Gene/Transcript만을 출력하도록 설정할 수 있다. 또한 “Allele Frequency” 항목을 이용하여 해당 조건을 만족하는 변이를 갖는 Gene/Transcript만을 출력하도록 제약 조건을 부가할 수 있다. Allele Frequency 필터 항목을 체크하지 않았을 때 기본값으로 Allele Frequency가 1 이하인 것들을 모두 출력한다. 반면 해당 항목을 체크하고 Allele Frequency 값을 설정하는 경우, 해당 조건에 따라 출력되는 Gene/Transcript set이 제한되어 출력된다. 단, 이 둘 필터링 조건을 부가한 후, Search 버튼을 클릭하여야 하며, 그 이후 필터링 결과가 Gene/Transcript View 영역에 반영되어 나타난다.

표 5

필터 종류	기능
View only genes with variations	Gene/Transcript View 영역에 변이가 있는 Gene/Transcript만 출력
Allele Frequency	설정된 Allele Frequency 조건을 만족하는 변이가 있는 Gene/Transcript만 출력

[0087] (5) Exon Variant Call Panel

[0088] Exon Variant Call Panel 영역은 Gene/Transcript View 영역에서 사용자가 선택한 Transcript에서 추출된 변이

정보를 출력한다 (도 10의 하단 (1) 패널 참조). 변이 정보는 사용자가 입력한 VCF 파일을 기반으로 생성된다. VCF 파일에서 추출된 변이 중 rsID를 가지고 있는 변이는 dbSNP 데이터베이스와 연동되어, dbSNP 정보의 일부가 함께 출력된다. rsID가 없는 경우에는 VCF 파일에 포함된 정보만이 출력된다.

[0089] 임의의 변이를 선택하여 마우스로 클릭하면 선택된 변이 위치를 포함하는 엑손 영역이 Main View Panel에 출력되며, 해당하는 엑손 영역의 정보가 Information Panel에 출력된다. 또한 변이를 선택하여 마우스로 더블 클릭하면 해당 변이 위치를 중심으로 하는 Genome Browser가 호출되어, 해당 영역의 세부 분석에 사용할 수 있다.

[0090] 상단에 있는 VCF Format 버튼을 클릭하면 VCF Format 정보 확인 창이 실행되며, VCF 파일 생성시 입력한 Filter 정보와 함께, FORMAT 정보, INFO 정보가 출력된다 (도 15 참조). 또한 Report 버튼을 클릭하면 Exon Variant Call Panel 패널에 출력된 결과를 저장하기 위한 대화상자가 생성되며 대화상자 내의 저장 버튼을 클릭하여 해당 결과를 tsv 파일로 저장할 수 있다.

[0092] (6) VCF Filter Panel

[0093] VCF Filter Panel은 Exon Variant Call 영역에 출력된 변이 정보를 필터링하기 위하여 사용된다 (도 10의 상단 (2) 패널 참조). 변이 종류에 따라 변이를 선택적으로 출력 가능하며, 또한 allele frequency 값을 지정하여 지정된 값 이하의 값을 갖는 변이만을 선택적으로 출력할 수 있다. 필터링 조건을 모두 설정한 후 Reload 버튼을 클릭하면 Exon Variant Call 영역에 필터링 된 결과가 출력된다.

[0095] (7) Information Panel

[0096] Information Panel에는 현재 브라우징을 수행하고 있는 유전자, 트랜스크립트의 이름, 위치 정보 등이 출력된다 (도 12 참조). Gene/Transcript View 영역에서 선택한 Gene명/Transcript명의 정보가 출력되며, Exon Variant Call Panel에서 선택한 변이가 포함된 엑손의 위치 정보 등이 출력된다. 또한 Main View Panel에서의 마우스 포인터가 있는 위치 정보가 함께 출력된다.

[0098] (8) Main Exon View Panel

[0099] Exome Browser의 최종 출력은 유전자 구조, 참조 서열에 따른 변이체 및 읽기 정렬이 표시되는 Main Exon View 패널에 표시된다(도 13).

[0101] 구체적으로, Exon Variant Call Panel에서 임의의 변이를 클릭하여 선택하면, Main Exon View Panel에 해당 변이를 포함하는 엑손 영역에 관한 상세 정보가 출력된다 (도 13 참조). Main View Panel 영역은 3개 영역으로 구분되어 있으며, 각 영역의 주요 기능은 다음과 같다 (표 6 참조).

표 6

번호	영역	기능
①	Transcript structure display	해당 변이가 포함된 Transcript의 exon-intron 구조 표시
②	Variants along the reference sequence	Reference sequence 상의 변이 위치 표시
③	Read alignment viewer	해당 변이가 포함된 Exon 영역에 대한 리드 정렬 결과 출력

[0103] 유전자 구조 영역에서 서열 변이가 있거나 없는 엑손은 각각 파란색 또는 녹색 상자로 표시될 수 있다. Read Alignment Viewer는 SNP(도 18), 삭제(deletion)(도 19) 및 삽입(insertion)(도 20)과 같은 다양한 유형의 서열 변형을 보여주었다. 따라서 Read Alignment Viewer에서 특정 시퀀스 변형이 식별될 수 있다.

[0105] ① Transcript 구조 표현

[0106] 해당 변이가 포함된 Transcript의 exon-intron 구조를 보인다. Exon 영역과 Intron 영역은 사각형 박스와 실선으로 각각 표시되며, Exon 영역을 나타내는 사각형 박스는 녹색, 파란색, 빨간색의 3가지 색으로 구별되어 출력된다 (도 14 참조). 변이가 없는 Exon 영역은 녹색으로 출력되며, 변이가 있는 Exon 영역은 파란색으로 출력되고, 현재 Read alignment viewer에 출력되고 있는 (해당 변이가 포함된) Exon 영역은 빨간색으로 출력된다. 또한 Exon 영역을 나타내는 사각형 박스 위치에 마우스 포인터를 Overlay하면 해당 Exon의 일련번호, 위치 등의 정보가 출력된다 (도 15 참조). Transcript의 exon-intron 구조에서 마우스 클릭에 의하여 임의의 Exon 영역으로 이동 가능하며, Exon 영역 이동 시, “Variants along the reference sequence” 패널과 “Read alignment viewer” 패널이 해당 Exon 영역 정보를 표시하도록 동기화된다.

[0107] ② Variants along the reference sequence

[0108] 해당 엑손 영역에서 추출된 변이가 reference sequence 하단에 함께 출력된다 (도 16 참조). Reference sequence의 각 염기는 색깔로 구분되어 표시되며 (표 7 참조), reference sequence 하단에 보이는 변이도 색깔에 의하여 염기 종류가 구별되어 표현된다. 마우스 커서를 reference sequence의 임의의 위치에 가져가면 유전체 공간 상의 정확한 위치가 Information Panel의 Position 정보에 출력된다.

표 7

색	염기
Lawn Green	A
Blue	C
Sandy Brown	G
Red	T
Gray	N

[0111] ③ Read alignment viewer 해당 Exon 영역에 대한 리드 정렬 결과를 다양한 정보와 함께 출력한다 (도 17 참조). 주요 사양은 다음과 같다.

[0112] ▶ 출력되는 유전체 영역의 범위: 기본적으로 출력되는 유전체 영역은 해당 Exon 영역과 해당 Exon의 앞 뒤 인트론 영역의 일부를 포함한다. 도 17에서 보이는 검은색 실선의 바깥쪽 영역은 해당 Exon 영역에 인접한 Intron 영역을 나타내며, 디폴트 값으로 Exon 영역의 앞뒤 30 base가 기본 출력 영역에 포함된다. 단, Viewer Control Panel에서 출력되는 intron 영역의 범위를 조절 가능하다. 정렬된 read sequence들의 coverage depth가 깊어 리드 정렬결과가 한 화면에 출력이 되지 않은 경우, 마우스 휠 혹은 스크롤 바를 이용하여 상하 화면의 리드 정렬결과를 확인할 수 있다. 또한 정렬된 임의의 리드 시퀀스 상에서 마우스 더블클릭을 수행하면 Genome Browser가 호출되어, 해당 위치를 중심으로 하는 유전체 영역에 대한 리드 정렬 결과를 세부적으로 분석할 수 있다.

[0113] ▶ 리드 정렬 결과의 표현: 리드 정렬결과에서 read sequence의 염기가 reference sequence의 염기와 일치하는 경우, 리드 sequence의 각 염기는 LightGray 색으로 표시되며, 일치하지 않는 경우 read sequence의 염기는 염기의 종류에 따라 다른 색깔로 구분되어 표시된다 (표 7 참조). Paired-end 리드의 각 리드는 방향성을 가지고 있으며, 정방향의 리드는 왼쪽에서 시작하여 오른쪽에서 끝나며, 역방향의 리드는 오른쪽에서 시작하여 왼쪽에서 끝난다. 정렬되어 있는 리드들의 방향성을 표시하기 위하여 각 리드의 단말 염기에 화살표를 부가하여 각 리드의 방향성을 표현한다. 또한 정렬된 paired-end 리드의 insert size를 구분하여 표시하기 위하여, insert size 값이 최대 설정값 (디폴트: 400)보다 크면 해당 리드들을 빨간색으로 표시하고 최소 설정값 (디폴트: 100)보다 작으면 파란색으로 표시하며, 최대값과 최소값 사이의 경우, LightGray로 표시한다. Insert size의 최대/최소 값은 Exon View Control Panel에서 설정/변경이 가능하다.

[0114] ▶ 변이 추출 결과의 출력: 추출된 SNP, insertion, deletion 변이 정보는 Read alignment viewer에서 다음과 같은 방식으로 출력된다

[0115] ▶▶ SNPs: 리드 정렬결과에서 read sequence의 염기가 reference sequence의 염기와 일치하지 않는 경우 read sequence의 염기는 LightGray가 아닌 염기의 종류에 따른 색깔로 구분되어 표시된다 (표 7 참조). 따라서 리드 정렬결과에 SNP 변이가 존재하는 경우, 해당 SNP 위치에 LightGray가 아닌 다른 색깔로 표시된 염기의 열을 시각적으로 확인 가능하다. 다음 도 18의 빨간색 박스로 표시된 부분이 SNP 변이의 출현 예를 나타내며, 마우스

커서를 해당 위치에 가져가면 유전체 공간 상의 정확한 위치가 Information Panel의 Position 정보에 출력된다.

▶▶ Deletions: reference sequence와 비교하여 리드 정렬 결과에서 Deletion 변이가 발견되는 경우, 해당하는 read sequence의 염기 서열 부분을 검정 가로 막대로 표현하여 나타낸다. 따라서 리드 정렬결과에 Deletion 변이가 존재하는 경우, 해당 변이 위치에 LightGray가 아닌 검은색 막대로 표시된 염기 부분 서열을 시각적으로 확인 가능하다. 다음 도 19의 빨간색 박스로 표시된 부분이 Deletion 변이의 출현 예를 나타내며, 마우스 커서를 해당 위치에 가져가면 유전체 공간 상의 정확한 위치가 Information Panel의 Position 정보로 출력된다.

▶▶ Insertions: reference sequence와 비교하여 리드 정렬 결과에서 Insertion 변이가 발견되는 경우, 해당하는 read sequence의 염기 위치에 보라색 box를 삽입하여 표시한다. 다음 도 20에서 마우스 포인터가 가리키는 부분에 Insertion 변이의 출현 예를 보인다. 그러나 Insertion 변이는 염기 문자들 사이에 출현 위치가 표현되므로 Read alignment viewer 상에서 시각적으로 구별이 어려울 수 있다. 이 경우, 해당 위치에서 마우스를 더블 클릭하여 Genome Browser를 호출하면, Insertions 변이 영역을 더욱 세밀히 확인할 수 있다 (도 21 참조).

(9) Exon View Control Panel

Main Exon View Panel을 제어하기 위하여 사용된다 (도 22 참조). 기능은 크게 네 가지로 구분할 수 있으며, 그 기능을 <표 8>에 보인다.

표 8

영역	기능
Exon Number 1 Exon Move << < > >>	Exon 영역 이동
Exon Move Controller Expanded area 30 Base	기본 출력 영역에 포함되는 Intron 영역의 범위 지정
Intron Area Select Controller Insert Size High 400 Insert Size Low 100 Base	Paired-end read의 insert size 지정
Insert size Controller Capture	Panel 캡처 기능
Capture 버튼	

① Exon Move Controller Exon Move Controller에서 Exon Number를 지정하는 텍스트 박스에 Exon Number를 입력한 후 Exon Move 버튼을 클릭하면 해당 Exon 영역으로 이동한다. 여기에서 Exon Number는 해당 Transcript에 존재하는 엑손을 유전체 위치에 따라 순차적으로 인덱싱한 숫자를 나타낸다. 또는 << < > >> 버튼을 이용하여 Exon 영역을 이동할 수 있다. << >> 버튼은 Exon Number를 5개 단위로 이동시키며, < > 버튼은 Exon Number를 1개 단위로 이동시킨다.

② Intron Area Select Controller

Read alignment viewer에 표시되는 Exon의 앞 뒤 인트론 영역의 크기를 조정하는데 사용된다. 초기값은 30으로 설정되어 있으며, Exon 영역의 앞뒤 30 base의 Intron 영역이 출력된다. 조절 가능한 범위는 0부터 200까지 정수의 값을 사용할 수 있다.

③ Insert size controller

Paired-end 리드의 insert size에 따라 리드 정렬 결과를 색깔로 구별하여 출력하기 위하여 insert size 값의 최소/최대 값을 조정하는데 사용된다. 초기값은 최소/최대 값으로 100/400이 설정되어 있다. 조절 가능한 범위는 0부터 99,999까지 정수의 값을 사용할 수 있다.

④ Capture 버튼

현재 출력되고 있는 Main Exon View panel과 Information Panel 내용을 캡처하여 이미지로 저장한다. 저장되는 파일은 jpg 포맷으로 제공된다 (도 23 참조).

[0130] *이하 도면들을 통해서, Genome Browser에 대해 순차적으로 설명한다.

[0132] Genome Browser는 Exome Browser에서 호출되어 리드 정렬을 기반으로 하는 변이 추출 결과를 세밀히 브라우징 하는데 사용된다. Genome Browser는 호출되기 전에 Exome Browser를 통하여 리드 정렬 파일(BAM 파일)이 로드 되어야 하며, 변이 추출 파일(VCF 파일)은 선택적으로 입력될 수 있다. Genome Browser는 다음 세 가지 방식에 의하여 호출되어 실행된다.

[0133] 기본적으로 Exome Browser의 Menu Bar에서 Genome Browser 메뉴를 선택하면 Genome Browser가 호출되어 실행된다 (도 24 참조).

[0134] 다음, Exome Browser의 Exon Variant Call Panel에서 임의의 변이 레코드를 선택하여 더블클릭하면 Genome Browser가 호출된다 (도 25 참조). Genome Browser는 해당 변이 위치를 전달받아 호출되므로, 기본적으로 해당 위치를 중심으로 하는 앞뒤 유전체 영역이 브라우저에 로드 된다. 다음 도 26은 도 25에서 선택된 변이 레코드를 더블클릭하여 호출된 Genome Browser의 실행 화면을 나타내며, 해당 변이 위치를 중심으로 하는 리드 정렬 결과를 보다 상세히 검색할 수 있는 환경이 제공된다.

[0135] 또한 Exome Browser의 Main Exon View Panel에 있는 Variants along the reference sequence 영역과 Read alignment viewer 영역에서 마우스 더블클릭을 하면 Genome Browser가 호출된다 (도 27 참조). Genome Browser는 마우스 클릭을 수행한 유전체 위치를 전달받아 호출되므로 기본적으로 해당 위치를 중심으로 하는 앞뒤 유전체 영역이 브라우저에 로드 된다.

[0136] Genome Browser의 메인 화면은 총 5개의 영역으로 구성된다 (도 28 참조). 5개 영역의 각 기능을 <표 9>에 보인다.

표 9

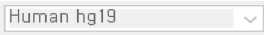
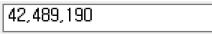
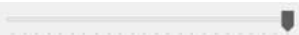
번호	영역	기능
(1)	Control Panel	표준 참조 서열 상의 유전체 위치 및 브라우징을 위한 해상도 설정
(2)	Cytoband Panel	cytoband representation
(3)	Coverage graph Panel	read depth representation
(4)	Genetic structure Panel	유전자 구조 표현
(5)	Main Genome View Panel	리드 정렬 결과 및 reference sequence 상의 변이 위치 표시

[0138] 단일 뉴클레오타이드 분해능에서 유전적 변이를 분석하기 위해 위에서 사용한 개인 게놈 시퀀싱 데이터를 사용하여 다양한 유형의 서열 변이가 모니터링된 결과, 도 28과 같이 개인 게놈의 특정 유전자에서 SNP 변이체, deletion(결실) 및 insertion(삽입)이 확인되었다. 동의 돌연변이가 코딩 아미노산을 변경하지 않는 ATR 유전자의 C에서 T로의 단일 뉴클레오타이드 변화는 모든 서열 판독에서 확인되었다(도 28). MST1L 유전자에서 서열 판독의 절반에서 5개 뉴클레오타이드의 결실이 확인되었다(도 43). 이것은 두 쌍의 염색체 중 하나가 프레임 시프트와 함께 돌연변이된 MST1L 유전자를 가지고 있음을 의미한다. 또한, PRDM2 유전자에 추가 CTC 서열의 삽입이 확인되었다(도 44). 시퀀스 리드의 약 절반에 이 삽입이 포함되어 있기 때문에 두 쌍의 염색체 중 하나에 추가 CTC 삽입이 있는 돌연변이된 PRDM2 유전자가 있는 것으로 추정될 수 있다. 이렇듯, PhenGenVar 게놈 브라우저를 사용한 육안 검사가 변이 호출 검증을 위한 효과적이고 강력한 도구임이 확인되었으며, 위양성 횡수를 줄이고 진정한 유전 변이 발견을 확인하는 데 도움을 주는 점이 검증되었다.

[0140] (1) Control Panel

[0141] Control Panel은 BAM/VCF 파일 분석을 위한 데이터/환경 설정 패널로써 4개의 영역으로 구성된다 (도 29 참조). 4개 영역의 각 기능은 <표 10>과 같다.

표 10

번호	영역	기능
①	Reference sequence combo box 	표준 참조 서열 표시
②	Chromosome combo box 	Chromosome 선택
③	Search text box  	Gene Name 또는 Position 입력
④	Track Bar 	해상도 설정

① Reference sequence combo boxGenome Browser를 호출한 Exome browser에서 선정한 reference sequence가 기본으로 표시되며, 변경할 수 없다. 본 브라우저에서는 Reference sequence 버전으로서 Human hg_19와 Human hg_38을 사용할 수 있다.

② Chromosome combo box

Genome Browser를 호출한 Exome browser에 설정되어 있는 Chromosome 번호가 기본으로 표시된다. 단, 이 값은 변경할 수 있다. Exome browser에서 Chromosome 번호가 설정되지 않은 상태에서 호출되면 디폴트 값으로 Chromosome 1번으로 설정된다.

③ Search text box

Chromosome 상의 유전체 위치를 설정하는데 사용되며, 입력된 위치는 각 패널의 중심 위치로 설정된다. 위치 정보는 임의로 변경 가능하며, Gene Name을 직접 입력하여 해당 유전자의 위치를 패널의 중심 위치로 설정 가능하다. 단, 브라우저 실행 중에는 각 패널의 선두 위치를 나타낸다.

③ TrackBar

각 패널 화면의 해상도를 높이거나 낮추는데 사용된다. TrackBar를 이용하여 해상도를 변경시키면 Genome Browser의 모든 패널의 화면은 중앙 위치를 기준으로 확대/축소된다. 해상도는 다음의 수식에 의하여 계산된다.

수학식 1

$$S = \frac{ChrLength}{2^n}$$

여기에서 n의 값은 해상도 level을 나타내며, n은 0에서 22까지의 값을 가질 수 있다. S는 해상도 레벨 n의 값에 따라 화면에 출력되는 nucleotide bps 단위의 유전체 영역의 최대 길이를 나타내며, ChrLength는 염색체의 최대 길이를 나타낸다. 예를 들어 Read alignment viewer에서는 Chromosome 1번 유전체 영역에 대하여 최소 해상도에서 121,704bp 영역이 한 패널에 출력되며, 최대 해상도에서 59bp 영역이 한 패널에 출력된다.

한편, 일 실시 예에 따르면, 전자 장치는 각 해상도에 대하여 변이에 해당하는 유전자를 가장 많이 보여줄 수 있도록 중앙 위치로 다시 선택할 수 있다. 구체적으로, 전자 장치는 중앙 위치를 시퀀스 상에서 좌우로 일정 수치 범위 내에서 변경한 결과 변이에 해당하는 유전자가 가장 많이 보일 수 있는 지점을 선택할 수 있다.

여기서, 일정 수치는, 현재의 해상도에서 표시되는 bp 영역의 수를 10으로 나눈 값에 해당할 수 있으며, 일정 수치 범위 내에서 한 단계씩(ex. 현재의 해상도에서 표시되는 bp의 수를 100으로 나눈 값) 중앙 위치를 이동하는 각각의 경우들 중 변이에 해당하는 유전자 요소가 가장 많이 보일 수 있는 지점이 중앙 위치로 선택될 수 있다.

[0155] (2) Cytoband Panel

[0156] 현재 브라우저를 수행하고 있는 Chromosome의 Cytoband를 표시한다. Cytoband 표현 상의 임의의 위치를 마우스로 클릭하면, 해당 위치가 모든 패널의 중앙 위치로 지정되어 브라우저를 수행하게 된다. 또한 현재 패널에서 브라우저를 수행하고 있는 유전체 영역을 Cytoband 상에 붉은색 박스로 나타낸다 (도 30 참조).

[0158] (3) Coverage graph Panel

[0159] 리드 정렬 파일(BAM 파일)로부터 각 유전체 위치에 대한 read depth를 동적으로 계산하여 계산된 결과를 coverage graph로 출력한다. Control Panel의 TrackBar를 이용하여 패널에 표시되는 Coverage graph의 해상도를 임의로 조절할 수 있다 (도 31 참조). 또한 Coverage graph 상의 임의의 위치에 마우스 포인터를 가져가면, 해당 위치의 position 정보와 read depth가 수치로 표시된다 (도 32 참조).

[0161] (4) Genetic structure Panel

[0162] 해당 유전체 영역에 존재하는 유전자의 구조를 표현하며, 유전자명과 등록된 트랜스크립트들의 exon-intron 세부 구조를 동시에 나타낸다. Exom Browser의 Exon Variant Call Panel에서 임의의 변이 레코드를 더블클릭하여 Genome Browser가 호출된 경우, 해당 변이가 존재하는 유전자(트랜스크립트) 정보가 Genome Browser에게 전달되며, 전달받은 유전자명/트랜스크립트명은 빨간색으로 구별하여 표시한다. 단 그 외의 경우, 유전자명/트랜스크립트명은 모두 검은색으로 표현된다 (도 33 참조). 브라우저의 해상도가 낮게 설정되면 이와 같은 유전자/트랜스크립트의 Exon-Intron 구조가 출력되며, 해상도가 높게 설정되어 화면이 확대되면, 해당 유전자/트랜스크립트의 아미노산 시퀀스가 출력된다 (도 34 참조). 아미노산 시퀀스는 파란색과 하늘색이 반복되는 형태의 사각형으로 표시된다. 그리고 엑손 영역은 굵은 선으로 표시되며, 인트론 영역은 얇은 선으로 표시된다 (도 34 참조). 또한 하나의 유전자에 다수의 Transcript가 보고되어 있어 정해진 패널 영역에 함께 출력이 되지 않는 경우, 마우스 휠을 사용하여 현재 보이는 패널 영역의 위, 아래 영역으로 이동하며 이 들을 검색하는 것이 가능하다.

[0164] (5) Main Genome View Panel

[0165] 리드 정렬 및 변이 추출 결과가 다양한 정보와 함께 출력된다 (도 35 참조). Main View Panel 영역은 2개 영역으로 구분되어 있으며, 각 영역의 주요 기능은 다음과 같다 (표 11 참조).

표 11

번호	영역	기능
①	Variants along the reference sequence	Reference sequence 상의 변이 위치 표시
②	Read alignment viewer	리드 정렬 결과 출력

[0167] ① Variants along the reference sequence 영역변이 추출 파일(VCF 파일)에 보고된 변이가 reference sequence 하단에 함께 출력된다. Reference sequence의 각 염기는 색깔로 구분되어 표시되며 (표 12 참조), reference sequence 하단에 보이는 변이도 색깔에 의하여 염기 종류가 구별되어 표현된다. reference sequence의 임의의 위치에 마우스오버 하면 유전체 위치가 브라우저 하단의 position 정보에 출력된다. 이 영역의 정보는 최소로 축소된 경우 컬러 막대로 출력되고, 최대로 확대된 경우 컬러 글씨로 출력된다 (그림 36-(a), 36-(b) 참조).

표 12

색	염기
Lawn Green	A (adenine)
Blue	C (cytosine)
Sandy Brown	G (guanine)
Red	T (thymine)
Gray	N

- [0169] ② Read alignment view 영역Reference Sequence에 정렬된 각 리드의 세부 정보, 매핑 정보를 다양한 출력 방식으로 나타내며, 리드 정렬 결과 분석에 의하여 추출되는 변이 정보를 추가 정보와 함께 나타낸다. 기본으로 출력되는 브라우저 영역에서는 마우스의 드래그를 이용하여 표시 영역을 이동할 수 있다. 왼쪽으로 마우스를 드래그하면 현재 출력되고 있는 리드 정렬 결과의 이전 위치로 이동 가능하며, 화면을 오른쪽으로 드래그하면 현재 출력된 리드 정렬 결과의 이후 위치로 이동 가능하다. 또한 출력되는 리드들의 coverage depth가 깊어 리드 정렬 결과가 한 화면에 출력되지 않을 경우 마우스 휠을 이용하여 상/하 이동이 가능하다. Read alignment viewer의 주요 기능은 다음과 같다.
- [0170] ▶ 리드 재정렬 결과 출력: 브라우저에 출력되는 리드 정렬 결과는 기본적으로 입력된 BAM 파일의 리드 정렬 정보를 기반으로 한다. 본 브라우저에서는 추가적으로 BAM 파일의 CIGAR 정보를 이용하여 리드들을 reference sequence에 바르게 정렬시키는 리드 재정렬 작업을 수행한다. 또한 정렬되어 표시되는 리드 사에 마우스오버 하면 각 리드의 상세 정보가 출력되어 나타난다.
- [0171] ▶ 리드 정렬 결과의 확대 및 축소 표현: Control Panel의 TrackBar를 이용하여 해상도를 조절하여 리드 정렬 결과를 확대 및 축소 표현할 수 있다. 해상도가 높아지면, read sequence와 reference sequence의 각 염기는 염기의 종류 (표 5 참조)에 따라 다른 색깔로 구분되어 표시된다 (도 37 참조). 그러나 해상도가 낮아지면 리드 정렬결과에서 read sequence의 염기가 reference sequence의 염기와 일치하는 경우, 리드 sequence의 각 염기는 LightGray 색으로 표시되며, 일치하지 않는 경우 read sequence의 염기만이 염기의 종류에 따라 다른 색깔로 구분되어 표시된다 (도 38 참조).
- [0172] ▶ 리드 시퀀스의 정렬 방향 표시: Paired-end 리드의 각 리드는 방향성을 가지고 있으며, 정방향의 리드는 왼쪽에서 시작하여 오른쪽에서 끝나며, 역방향의 리드는 오른쪽에서 시작하여 왼쪽에서 끝난다. 정렬되어 있는 리드들의 방향성을 표시하기 위하여 각 리드의 단말 염기에 화살표를 부가하여 각 리드의 방향성을 표현하였다 (그림 39 참조).
- [0173] ▶ 리드의 Insert Size 표시: 정렬된 paired-end 리드의 insert size를 구분하여 표시하기 위하여 read 시퀀스의 배경색에 컬러를 달리하여 표시한다. Insert size 값이 최대 설정값 (디폴트: 400)보다 크면 해당 리드 시퀀스의 배경색을 빨간색으로 표시하고, 최소 설정값 (디폴트: 100)보다 작으면 해당 리드 시퀀스의 배경색을 파랑색으로 표시하며, 최대값과 최소값 사이의 경우, 회색으로 표시한다 (도 40 참조). Insert size의 최대/최소 값은 Exome Browser의 Exon View Control Panel에서 설정/변경이 가능하다.
- [0174] ▶ 변이 추출 결과의 출력: 추출된 SNP, insertion, deletion 변이 정보는 Read alignment viewer에서 다음과 같은 방식으로 출력된다.
- [0175] ▶▶ SNP: 리드 정렬결과에서 read sequence의 염기가 reference sequence의 염기와 일치하지 않는 경우 염기의 종류에 따른 다른 색깔 (표 5 참조)로 구분되어 표시된다. 따라서 리드 정렬결과에 SNP 변이가 존재하는 경우, 해당 SNP 위치에 레퍼런스 시퀀스와 다른 색깔로 표시된 염기의 열을 시각적으로 확인 가능하다 (도 41 참조). 도 41의 빨간색 박스로 표시된 부분이 SNP 변이의 출현 예를 나타내며, 마우스 커서를 해당 위치에 가져가면 유전체 위치가 브라우저 왼쪽하단의 Position 정보에 출력된다. 그러나, 사용자가 입력한 변이 정보 파일 (VCF 파일)이 로드되어 있어, 해당 SNP 정보가 포함되어 있는 경우, Variants along the reference sequence 영역에 추출된 변이가 reference sequence 하단에 함께 출력된다 (도 42 참조).
- [0176] ▶▶ Deletions: Reference sequence와 비교하여 리드 정렬 결과에서 Deletion 변이가 발견되는 경우, 해당하는 read sequence의 염기 서열 부분을 검정 가로 막대로 표현하여 나타낸다 (도 43 참조). 따라서 리드 정렬결과에 Deletion 변이가 존재하는 경우, 해당 변이 위치에 검은색 막대로 표시된 염기 부분 서열을 시각적으로 확인 가능하다. 다음 그림 20의 빨간색 박스로 표시된 부분이 Deletion 변이의 출현 예를 나타내며, 마우스 커서를 해당 위치에 가져가면, 유전체 위치가 브라우저 왼쪽 하단의 Position 정보에 출력된다.
- [0177] ▶▶ Insertions
- [0178] reference sequence와 비교하여 리드 정렬 결과에서 Insertion 변이가 발견되는 경우, 해당하는 read sequence의 염기 위치에 보라색 box를 삽입하여 표시한다 (도 44 참조). 도 44에서 빨간색 박스로 표시된 부분에 Insertion 변이들이 출현하고 있는 것을 볼 수 있다. 또한 Insertion 변이가 발견된 위치에 마우스 커서를 가져가면, 삽입된 염기 수 및 종류 등 삽입된 변이의 상세 정보가 출력된다.

- [0179] 도 45는 본 개시의 일 실시 예에 따른 전자 장치의 구성을 설명하기 위한 구성이다.
- [0180] 전자 장치(100)는 데스크탑 PC, 노트북 PC, 태블릿 PC, 스마트폰 등 다양한 단말 기기에 해당할 수 있다. 이 경우, 전자 장치(100)는 적어도 하나의 서버로부터 상술한 다양한 실시 예에 따른 동작을 수행하기 위한 컴퓨터 프로그램 내지는 소프트웨어를 다운로드하여 실행할 수 있다.
- [0181] 또한, 전자 장치(100)는 하나 이상의 단말 기기와 통신 가능한 서버로 구현될 수도 있다. 이 경우, 전자 장치(100)는 하나 이상의 사용자 단말과 통신을 수행할 수 있다. 구체적으로, 서버인 전자 장치(100)는 애플리케이션 또는 웹 페이지의 형태로 사용자 단말을 통해 서비스를 제공하는 방식으로 상술한 다양한 실시 예에 따른 동작을 수행할 수 있다.
- [0182] 도 45를 참조하면, 전자 장치(100)는 메모리(110), 프로세서(120)를 포함한다.
- [0183] 메모리(110)는 전자 장치(100)의 구성요소들의 전반적인 동작을 제어하기 위한 운영체제(OS: Operating System) 및 전자 장치(100)의 구성요소와 관련된 적어도 하나의 인스트럭션 또는 데이터를 저장하기 위한 구성이다.
- [0184] 메모리(110)는 ROM, 플래시 메모리 등의 비휘발성 메모리를 포함할 수 있으며, DRAM 등으로 구성된 휘발성 메모리를 포함할 수 있다. 또한, 메모리(110)는 하드 디스크, SSD(Solid state drive) 등을 포함할 수도 있다.
- [0185] 프로세서(120)는 전자 장치(100)를 전반적으로 제어하기 위한 구성이다. 구체적으로, 프로세서(120)는 메모리(110)와 연결되는 한편 메모리(110)에 저장된 적어도 하나의 인스트럭션을 실행함으로써 본 개시의 다양한 실시 예들에 따른 동작을 수행할 수 있다.
- [0186] 프로세서(120)는 CPU, AP, DSP(Digital Signal Processor) 등과 같은 범용 프로세서, GPU, VPU(Vision Processing Unit) 등과 같은 그래픽 전용 프로세서 또는 NPU와 같은 인공지능 전용 프로세서 등을 포함할 수 있다. 인공지능 전용 프로세서는, 특정 인공지능 모델의 훈련 내지는 이용에 특화된 하드웨어 구조로 설계될 수 있다.
- [0187] 도 45를 참조하면, 전자 장치(100)는 통신부(130), 사용자 입력부(140), 디스플레이(150) 등을 더 포함할 수 있다.
- [0188] 통신부(130)는 다양한 유무선 통신방식으로 적어도 하나의 외부 장치와 통신을 수행하기 위한 회로, 모듈, 칩 등을 포함할 수 있다. 로봇(100)의 이동성 확보를 위해 통신부(130)는 무선 통신을 수행하기 위한 다양한 회로/모듈/칩을 구비할 수 있다.
- [0189] 통신부(130)는 다양한 유무선 통신방식으로 적어도 하나의 외부 장치와 통신을 수행하기 위한 회로, 모듈, 칩 등을 포함할 수 있다. 통신부(130)는 다양한 네트워크를 통해 외부 장치와 연결될 수 있다.
- [0190] 네트워크는 영역 또는 규모에 따라 개인 통신망(PAN; Personal Area Network), 근거리 통신망(LAN; Local Area Network), 광역 통신망(WAN; Wide Area Network) 등일 수 있으며, 네트워크의 개방성에 따라 인트라넷(Intranet), 엑스트라넷(Extranet), 또는 인터넷(Internet) 등일 수 있다.
- [0191] 통신부(130)는 LTE(long-term evolution), LTE-A(LTE Advance), 5G(5th Generation) 이동통신, CDMA(code division multiple access), WCDMA(wideband CDMA), UMTS(universal mobile telecommunications system), WiBro(Wireless Broadband), GSM(Global System for Mobile Communications), DMA(Time Division Multiple Access), WiFi(Wi-Fi), WiFi Direct, Bluetooth, NFC(near field communication), Zigbee 등 다양한 무선 통신 방식을 통해 외부 장치들과 연결될 수 있다.
- [0192] 또한, 통신부(150)는 이더넷(Ethernet), 광 네트워크(optical network), USB(Universal Serial Bus), 선더볼트(ThunderBolt) 등의 유선 통신 방식을 통해 외부 장치들과 연결될 수도 있다.
- [0193] 통신부(130)를 통해, 전자 장치(100)는 컴퓨터 프로그램을 전송/수신할 수 있으며, reference sequence 또는 read sequence로 구성된 생체 데이터를 수집할 수도 있다.
- [0194] 사용자 입력부(140)는 다양한 사용자 명령 내지는 정보를 입력 받기 위한 구성이다. 사용자 입력부(160)는 적어도 하나의 버튼, 터치 패드, 마이크, 카메라 등을 포함할 수 있다.
- [0195] 디스플레이(150)는 유전자와 관련된 다양한 데이터를 시각적으로 출력하기 위한 구성이다. 디스플레이(150)는 LCD(Liquid Crystal Display), LED(Light Emitting Diodes), OLED(Organic Light Emitting Diodes), TOLED(Transparent OLED), Micro LED 등으로 구현될 수 있다. 또한, 디스플레이(110)는 평면 디스플레이 외에

곡면 디스플레이, 플렉서블 디스플레이, 폴더블 디스플레이 등으로 구현될 수도 있다.

- [0196] 전자 장치(100)는 디스플레이(150)를 통해 상술한 다양한 실시 예에 따른 메뉴, 패널 등 다양한 UI(User Interface)를 제공할 수 있다.
- [0197] 본 개시의 일 실시 예에 따른 전자 장치(100)는 일 표현형에 대한 특정 인간(: 제1 인간)의 유전자 셋의 변이 정보를 획득하고, 획득된 변이 정보를 동일한 표현형의 다른 인간들의 유전자 셋과 비교할 수 있다. 이때, 전자 장치(100)는 변이의 (유전자 내) 위치 또는 종류(ex. SNP, insertion, deletion)를 비교할 수 있다.
- [0198] 여기서, 전자 장치(100)는, 비교 결과에 따라, 상술한 제1 인간의 변이 정보 중 (표현형이 동일한) 다른 인간들의 유전자 셋의 변이 정보와 다른 변이 정보를 추출할 수 있다. 그리고, 상술한 다른 변이 정보가 매칭된 유전자 영역을 모니터링하기 위한 사용자 입력이 수신되는 경우(: 화면에 해당 유전자 영역이 표시됨), 전자 장치(100)는 상술한 다른 변이 정보가 현재 제공되는 표현형과의 연관성이 낮음을 알리는 메시지를 제공할 수 있다. 이 경우, 표현형과 관련된 변이 영역을 찾기 위한 사용자 입력을 통해 다른 유전자 영역이 출력될 수 있다.
- [0199] 또한, 전자 장치(100)는 상술한 다른 변이 정보를 가지면서도 표현형이 동일한 또다른 제2 인간의 유전자 셋을 검색할 수 있으며, 제1 인간 및 제2 인간의 복수의 표현형 중 공통되는 적어도 하나의 표현형에 대한 정보를 제공할 수 있다. 이 경우, 다양한 변이 정보와 표현형 간의 연관관계가 발견될 가능성이 생기는 장점이 있다.
- [0200] 한편, 일 실시 예로, 전자 장치(100)는 중복되는 표현형의 수에 따라 복수의 인원의 유전자 셋을 그룹핑 할 수 있다.
- [0201] 예를 들어, 기설정된 복수의 표현형 중 타겟에 해당하는 일 표현형이 사용자 입력에 따라 선택된 경우, 전자 장치(100)는 타겟에 해당하는 표현형을 가지는 복수의 인원의 유전자 셋을 그룹핑 할 수 있다.
- [0202] 이때, 복수의 인원의 유전자 셋을 대상으로 하여, 전자 장치(100)는 기설정된 복수의 표현형 중에서 타겟인 표현형 외에 중복되는 다른 표현형이 없는 제1 그룹, 타겟인 표현형 외에 중복되는 다른 표현형이 하나인 제2 그룹, 타겟인 표현형 외에 중복되는 다른 표현형이 두 개인 제3 그룹, 타겟인 표현형 외에 중복되는 다른 표현형이 세 개인 제4 그룹 등을 순차적으로 획득할 수 있다. 여기서, 각 그룹은 두 명 이상의 인원을 포함한다(특정 수만큼의 표현형이 서로 중복되는 인원들이 존재하지 않을 수도 있으므로, 해당 그룹이 존재하지 않는 경우도 물론 가능하다).
- [0203] 여기서, 제2 그룹, 제3 그룹, 제4 그룹 등은 각각 복수 개 획득될 수 있음은 물론이다. 예를 들어, 타겟 표현형이 A인 경우, 표현형 A와 b가 겹치는 인원들을 포함하는 제2-1 그룹과 표현형 A와 c가 겹치는 인원들을 포함하는 제2-2 그룹이 각각 획득되는 등 복수의 제2 그룹이 획득될 수 있다.
- [0204] 여기서, 전자 장치(100)는 중복되는 표현형이 비교적 적은 제2 그룹 및 제3 그룹 중 그룹에 포함되는 인원의 수가 가장 많으면서도 인원의 수가 '일정 수'(ex. 상술한 복수의 인원 중 일정 비율) 이상인 적어도 하나의 대표 연관 그룹을 선택할 수 있다. 예를 들어, 상술한 제2-1 그룹의 인원 수가 가장 많은 경우 제2-1 그룹이 대표 연관 그룹으로 식별될 수 있으며, 전자 장치(100)는 제2-1 그룹의 중복되는 표현형인 A와 b를 식별할 수 있다.
- [0205] 또한, 전자 장치(100)는 중복되는 표현형이 가장 많은 n개에 해당하는 최다 중복 그룹을 식별하고, 최다 중복 그룹에 속하는 복수의 인원 간에 서로 중복되는 표현형들을 식별할 수 있다. 예를 들어, 5개의 표현형(ex. A, b, C, E, f)이 중복되는 인원들을 포함하는 최대 중복 그룹이 식별될 수 있다.
- [0206] 여기서, 전자 장치(100)는 대표 연관 그룹 상에서 중복되는 표현형들(ex. A, b) 각각을 최다 중복 그룹 상에서 중복되는 표현형들(ex. A, b, C, E, f) 각각과 비교하고, 타겟 표현형(ex. A) 외에 일치하는 표현형이 존재하는 경우 해당 표현형(ex. b)을 타겟 표현형에 대한 연관 표현형으로 식별할 수도 있다. 그리고, 전자 장치(100)는 타겟 표현형과 연관 표현형 간에 연관 가능성이 있음을 알리는 정보를 출력할 수 있다.
- [0207] 이렇듯, 전자 장치(100)는 사용자 입력에 따라 선택된 타겟 표현형과 비교적 연관성이 깊은 적어도 하나의 다른 연관 표현형을 찾아내는 과정을 자동으로 수행할 수 있으며, 서로 연관되는 표현형들의 발견에 기여할 수 있게 된다.
- [0208] 한편, 타겟 표현형과 연관 표현형 간에 연관 가능성이 있음을 알리는 정보가 출력된 이후, 전자 장치(100)는 사용자의 피드백을 수신할 수 있다. 만약, 타겟 표현형과 연관 표현형 간의 연관 가능성을 부정하는 피드백이 수신되는 경우, 전자 장치(100)는 향후 적어도 하나의 타겟 표현형이 선택되었을 때 대표 연관 그룹을 선정하는 기준이 되는 상술한 '일정 수'를 더 높게 업데이트할 수 있다. 그 결과, 실제 연관 가능성을 찾기 위한 전자 장

치(100)의 동작 기준 및 성과가 점진적으로 업그레이드될 수 있다.

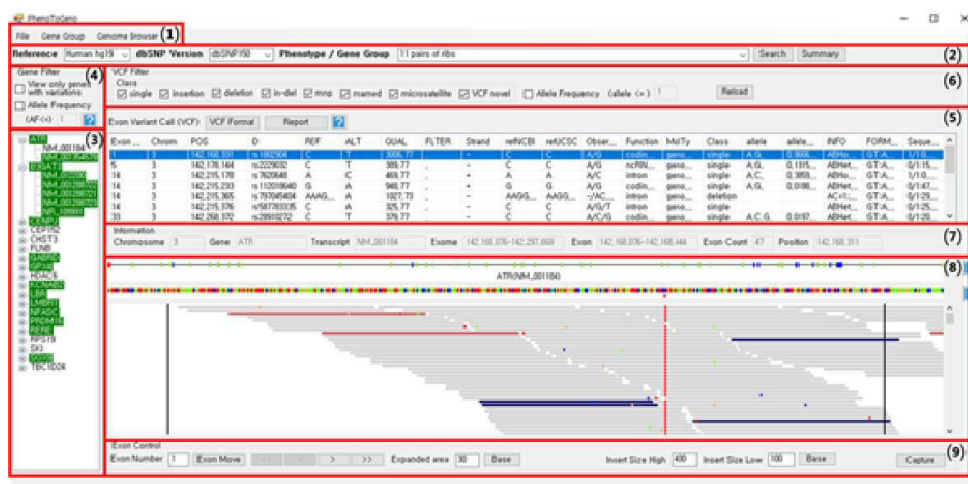
- [0209] 한편, 이상에서 설명된 다양한 실시 예들은 서로 저촉되지 않는 한 복수의 실시 예가 결합되어 구현될 수 있다.
- [0210] 한편, 이상에서 설명된 다양한 실시 예들은 소프트웨어(Software), 하드웨어(Hardware) 또는 이들의 조합된 것을 이용하여 컴퓨터(Computer) 또는 이와 유사한 장치로 읽을 수 있는 기록 매체 내에서 구현될 수 있다.
- [0211] 하드웨어적인 구현에 의하면, 본 개시에서 설명되는 실시 예들은 ASICs(Application Specific Integrated Circuits), DSPs(Digital Signal Processors), DSPDs(Digital Signal Processing Devices), PLDs(Programmable Logic Devices), FPGAs(Field Programmable Gate Arrays), 프로세서(Processors), 제어기(Controllers), 마이크로 컨트롤러(Micro-controllers), 마이크로 프로세서(Microprocessors), 기타 기능 수행을 위한 전기적인 유닛(Unit) 중 적어도 하나를 이용하여 구현될 수 있다.
- [0212] 일부의 경우에 본 명세서에서 설명되는 실시 예들이 프로세서 자체로 구현될 수 있다. 소프트웨어적인 구현에 의하면, 본 명세서에서 설명되는 절차 및 기능과 같은 실시 예들은 별도의 소프트웨어 모듈들로 구현될 수 있다. 상술한 소프트웨어 모듈들 각각은 본 명세서에서 설명되는 하나 이상의 기능 및 작동을 수행할 수 있다.
- [0213] 한편, 상술한 본 개시의 다양한 실시 예들에 따른 각 장치에서의 처리동작을 수행하기 위한 컴퓨터 명령어(Computer Instructions) 또는 컴퓨터 프로그램은 비일시적 컴퓨터 판독 가능 매체(Non-transitory Computer-readable Medium)에 저장될 수 있다. 이러한 비일시적 컴퓨터 판독 가능 매체에 저장된 컴퓨터 명령어 또는 컴퓨터 프로그램은 전자 장치 등 특정 기기의 프로세서에 의해 실행되었을 때 상술한 다양한 실시 예에 따른 전자 장치의 동작을 상술한 특정 기기가 수행하도록 한다.
- [0214] 비일시적 컴퓨터 판독 가능 매체란 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(Reading)이 가능한 매체를 의미한다. 비일시적 컴퓨터 판독 가능 매체의 구체적인 예로는, CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM 등이 있을 수 있다.
- [0215] 이상에서는 본 개시의 바람직한 실시 예에 대하여 도시하고 설명하였지만, 본 개시는 상술한 특성의 실시 예에 한정되지 아니하며, 청구범위에서 청구하는 본 개시의 요지를 벗어남이 없이 당해 개시에 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 개시의 기술적 사상이나 전망으로부터 개별적으로 이해되어서는 안 될 것이다.

부호의 설명

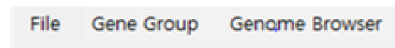
- [0216] 100: 전자 장치 110: 메모리
120: 프로세서

도면

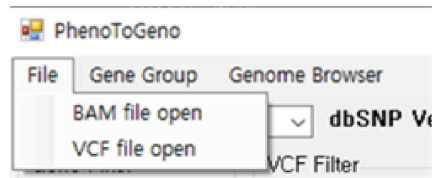
도면1



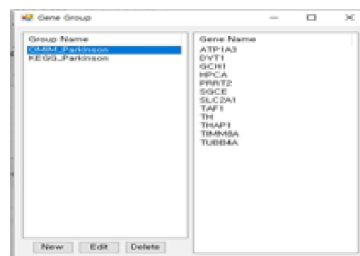
도면2



도면3



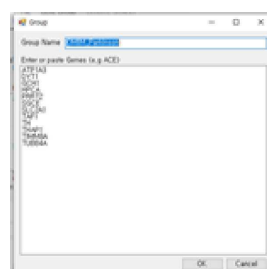
도면4



도면5



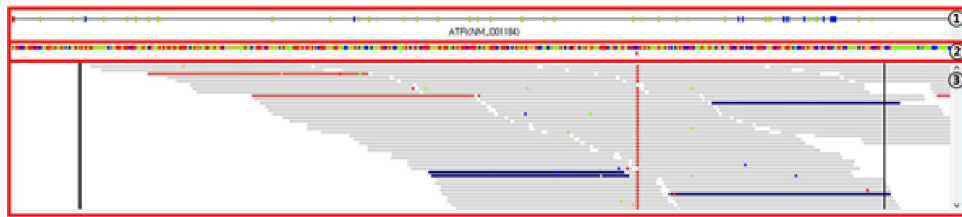
도면6



도면7



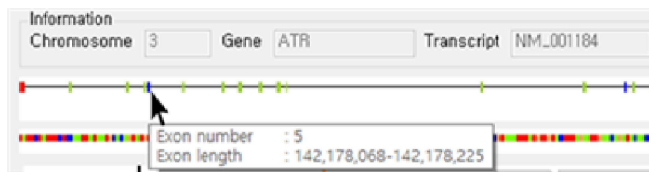
도면13



도면14



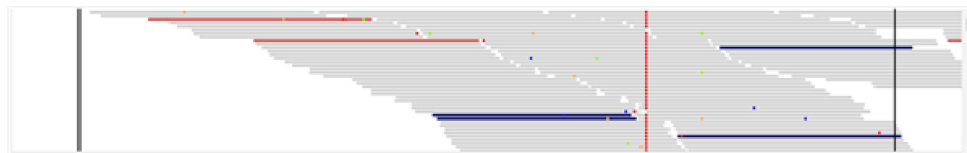
도면15



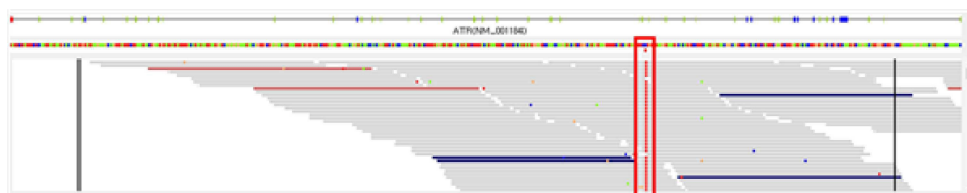
도면16



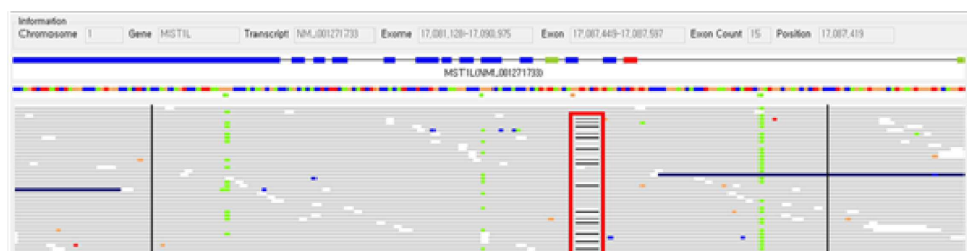
도면17



도면18



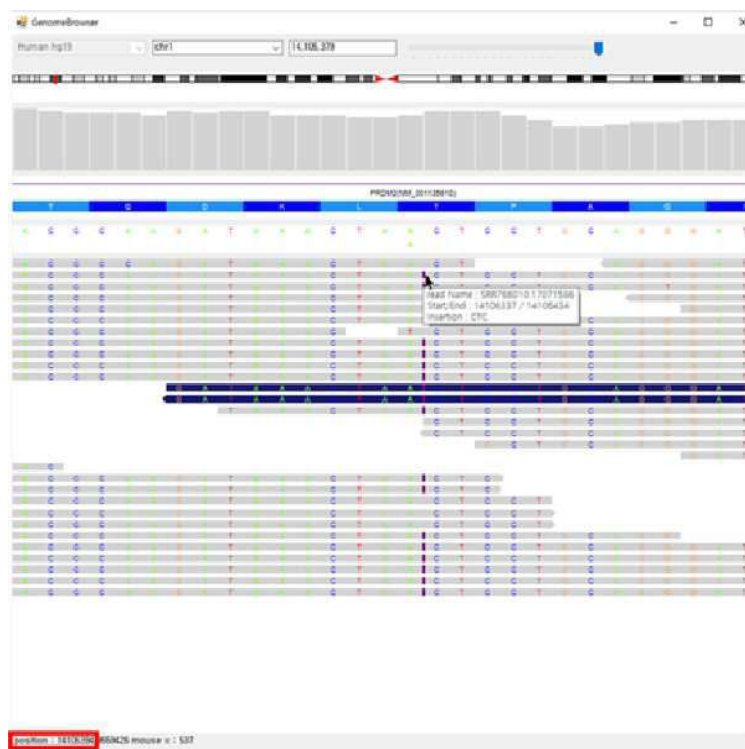
도면19



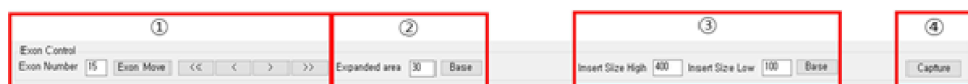
도면20



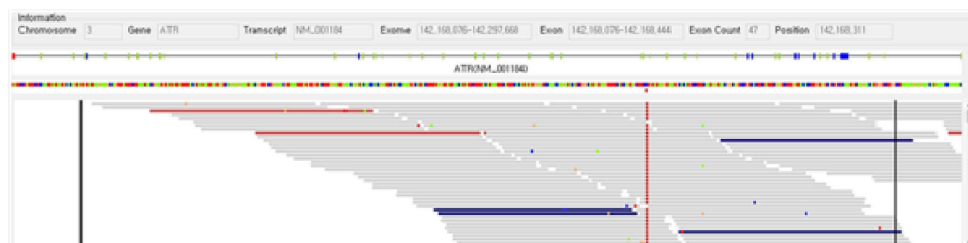
도면21



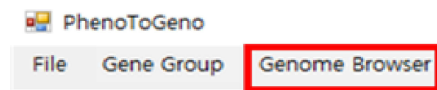
도면22



도면23



도면24



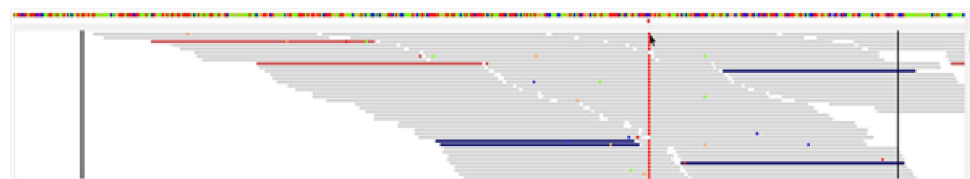
도면25

Exon Variant Call (VCF) VCF Format Report															
Exon ...	Chrom	POS	ID	REF	ALT	QUAL	FILTER	Strand	refNCBI	refUCSC	Obs...	Function	MolTip	Class	allele...
1	1	1,417,165	rs14069979	C	T	75.77		+	C	C	C/T	intron	geno...	single	C.T.
6	1	1,417,590	rs20130793	C	T	22.79		+	C	C	C/T	intron	geno...	single	C.T.
5	1	1,417,696	rs819970	C	G	1710.77		+	C	C	C/G	intron	geno...	single	C.G.
7	1	1,418,004	rs819972	T	C	734.07		+	T	T	C/T	intron	geno...	single	C.T.
8	1	1,421,569	rs37138022	G	A	1301.79		+	CCCC	CCCTC	~CC...	intron	geno...	deletion	AC+1...
10	1	1,421,531	rs819976	C	A	957.77		+	C	C	A/C/T	codon	geno...	single	A.C.T.
11	1	1,421,916	rs201951488	T	C	1836.77		+	T	T	C/T	intron	geno...	single	C.T.

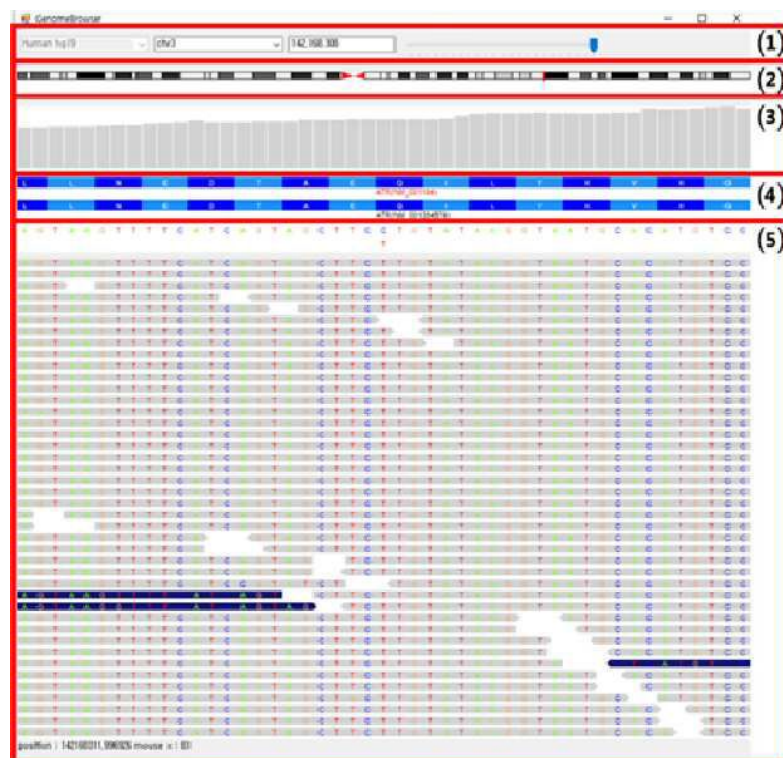
도면26



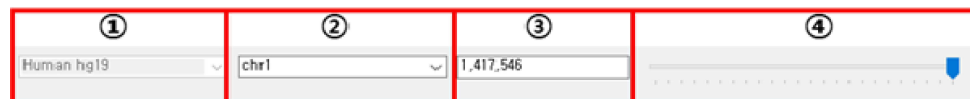
도면27



도면 28



도면29



도면30



도면31

도면32

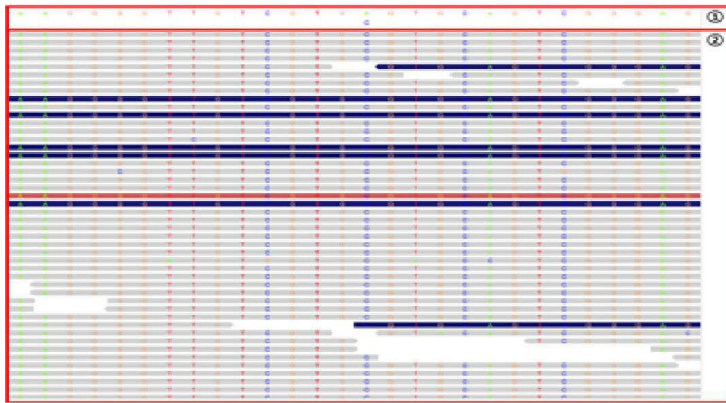
도면33



도면34



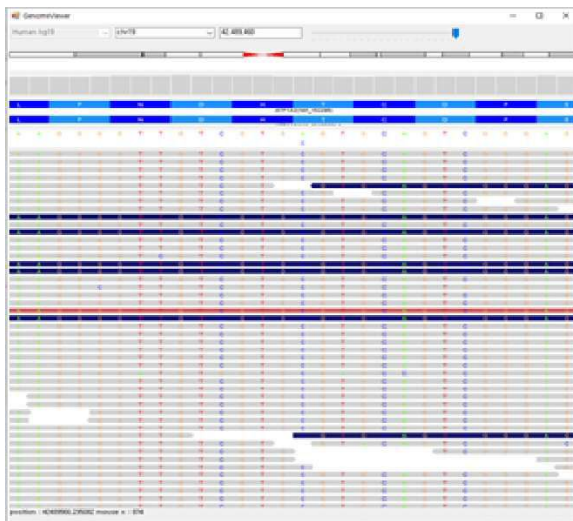
도면35



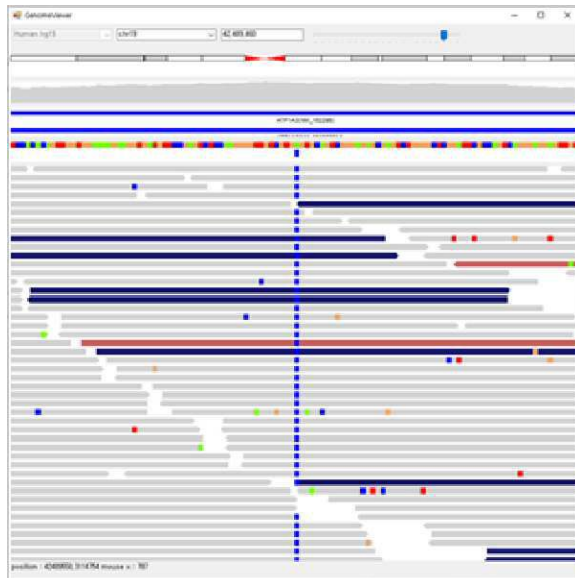
도면36



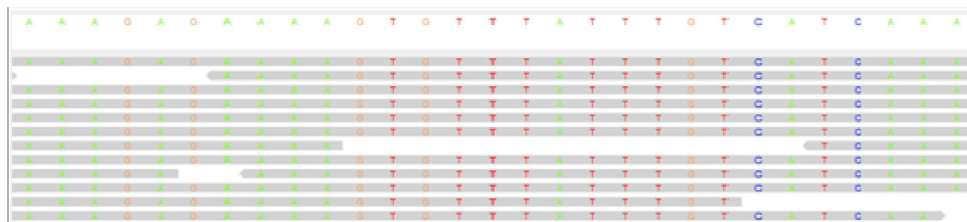
도면37



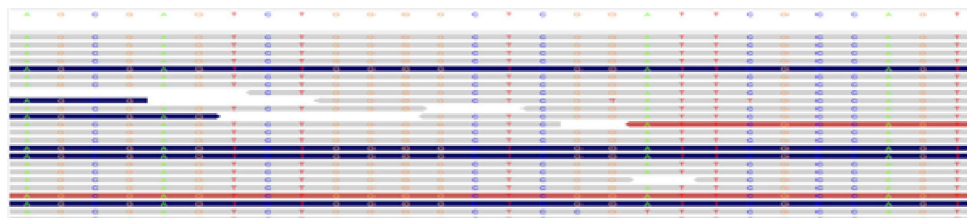
도면38



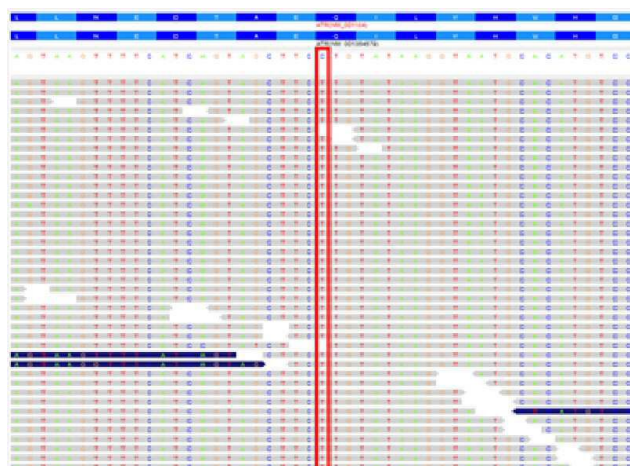
도면39



도면40



도면41



도면42

도면42는 유전체 시퀀싱 데이터의 시각화입니다. 상단에는 'chr10:100000000-100000000'과 'chr10:100000000-100000000'라는 좌표 정보가 표시되어 있습니다. 주된 영역은 수백 개의 수직 줄로 구성되어 있으며, 각 줄은 다양한 색(파란, 빨간, 초록, 노란)의 블록으로 채워져 있습니다. 이는 DNA 시퀀스의 변이 또는 특정 유전자 영역을 나타낼 수 있습니다. 오른쪽에는 'chr10:100000000-100000000'이라는 레이블이 붙은 한 줄이 강조되어 있습니다.

도면43

도면43은 유전체 시퀀싱 데이터의 시각화입니다. 상단에는 'chr10:100000000-100000000'과 'chr10:100000000-100000000'라는 좌표 정보가 표시되어 있습니다. 주된 영역은 수백 개의 수직 줄로 구성되어 있으며, 각 줄은 다양한 색(파란, 빨간, 초록, 노란)의 블록으로 채워져 있습니다. 이는 DNA 시퀀스의 변이 또는 특정 유전자 영역을 나타낼 수 있습니다. 오른쪽에는 'chr10:100000000-100000000'이라는 레이블이 붙은 한 줄이 강조되어 있습니다.

도면44

도면44는 유전체 시퀀싱 데이터의 시각화입니다. 상단에는 'chr10:100000000-100000000'과 'chr10:100000000-100000000'라는 좌표 정보가 표시되어 있습니다. 주된 영역은 수백 개의 수직 줄로 구성되어 있으며, 각 줄은 다양한 색(파란, 빨간, 초록, 노란)의 블록으로 채워져 있습니다. 이는 DNA 시퀀스의 변이 또는 특정 유전자 영역을 나타낼 수 있습니다. 오른쪽에는 'chr10:100000000-100000000'이라는 레이블이 붙은 한 줄이 강조되어 있습니다.

도면45

