Illumina Connected Insights의 oncogenicity 예측 모델

임상 연구를 위한 최신 종양학 가이드라인 및 Illumina의 사유 AI 알고리즘을 기반으로 하는 높은 정확도의 자동 변이 해석

정확한 oncogenicity 예측

다른 변이 분류 도구나 수동 워크플로우보다 높은 분류 정확도 제공

신뢰할 수 있는 변이 해석

기존 등급 분류 및 전문가들이 지정한 oncogenicity 기준과 높은 일치도 유지

동급 최고 수준의 인실리코 예측 변수

혁신적인 사유 알고리즘으로 질병 원인 변이 식별

연구 전용입니다. 진단 절차에는 사용할 수 없습니다.

문서 번호: M-KR-00356 KOR
원본 문서: M-GL-04087 v1.0

소개

포괄적인 유전체 프로파일링(Comprehensive genomic profiling, CGP)과 같은 차세대 시퀀싱(next-generation sequencing, NGS) 방법과 체세포(somatic) 종양 연구를 통해 얻은 체세포 변이 데이터는 해석이 어려울 수 있습니다. 분자 병리학회(Association for Molecular Pathology, AMP), 미국 임상종양 학회(American Society of Clinical Oncology, ASCO) 및 미국 병리 학회(College of American Pathologists, CAP)의 tiering guidelines(등급 분류 가이드라인)와 Clinical Genome Resource(ClinGen), Cancer Genomics Consortium(CGC) 및 Variant Interpretation for Cancer Consortium(VICC)의 oncogenicity standards(발암성 기준)는 각각 clinical actionability(임상적 적용 가능성)와 oncogenicity(즉, 종양 형성을 유도하는 능력)에 따라 체세포 변이를 분류하는 체계를 제공합니다. 그러나 현재 사용되는 수동 변이 해석 절차는 오류가 발생하기 쉽고, 많은 시간이 소요되며, 결과에 일관성이 부족하다는 단점이 있습니다.[1] 따라서 이러한 어려움을 겪고 있는 연구 랩에는 앞서 언급한 기준을 보다 효율적이고 일관되게 적용해 줄 자동 변이 해석 솔루션이 필요합니다.

Illumina Connected Insights는 oncogenicity 예측을 자동 계산하고 채점 체계에 PrimateAI-3D와 SpliceAI를 활용합니다. 연구자는 이러한 혁신적인 사유 알고리즘을 이용해 더 높은 정확도로 변이를 분류하고 결과에 확신을 가질 수 있습니다. PrimateAI-3D와 SpliceAI는 테스트를 진행한 다른 모든 분류 도구보다 우수한 정확도로 질병을 유발하는 병원성 미스센스 변이(pathogenic missense variant)를 식별하므로, 변이 분류 정확도를 한층 더 높여 줍니다.[2,3]

Connected Insights의 체세포 종양 변이 해석 정확도를 확인하기 위해 이 솔루션의 oncogenicity 예측 능력을 평가 전문가들이 지정한 분류 기준과 이미 존재하는 진리 집합(truth set)과 비교했습니다.

방법

Connected Insights가 제공하는 자동 oncogenicity 분류의 일치도(concordance)를 측정하기 위해 ClinGen/CGC/VICC 가이드라인에 따라 검증된 전문가 큐레이션을 거친 94가지 변이로 구성된 진리 집합과 비교한 결과, 높은 일치도가 확인되어 전문가 판단에 부합하는 것으로 나타났습니다. 추가적으로 기존 알고리즘 대비 PrimateAI-3D pathogenicity(병원성) 예측 정밀도를 테스트하여 Connected Insights 내 분류 정확도를 확인했습니다. 40가지 미스센스 변이에 대한 PrimateAI-3D의 pathogenicity 예측 정확도를 측정했으며, 다른 19가지 인실리코*(in silico*, 컴퓨터 기반) 예측 변수(예: SIFT, PolyPhen-2)를 벤치마킹했습니다. 일치도는 참조 분류(reference classification)와 정확하게 일치하는 결과를 백분율로 산출하여 계산했습니다. 이렇듯 여러 테스트에 걸쳐 높은 일치도가 관찰되어, Connected Insights가 수동 변이 해석을 높은 정확도로 재현할 수 있음이 입증되었습니다.

결과

Connected Insights의 자동 변이 해석과 전문가 oncogenicity 분류 간 일치도

Connected Insights는 94가지 종양 변이로 구성된 가이드라인 검증 데이터 집합에서 전문가들이 지정한 oncogenicity 라벨을 96%(90/94) 재현했습니다. 또한 이 소프트웨어는 대부분의 동일한 종양 라벨(oncogenic, likely oncogenic, variant of unknown significance, likely benign 또는 benign)을 식별했습니다(표 1). 특히 Connected Insights는 해당 진리 집합이 발표된 후 새롭게 발견된 특정 변이에 대한 최신 근거를 포함할 수 있었습니다. 외부 지식 소스가 이미 발표된 변이를 다시 분류하기에 앞서 Connected Insights가 지식 베이스(knowledge base) 피드에서 새로운 근거 데이터를 발견한 후 변이를 'oncogenic'으로 재분류한 사례도 일부 있었습니다(표 2). 이 방법은 종양 연구 시 자동화가 어떻게 신속한 새로운 근거를 통합하여 변이 해석을 일치시키고 잠재적으로 개선할 수 있는지를 보여줍니다. 전반적으로 높은 일치도는 곧 Connected Insights가 ClinGen/CGC/VICC oncogenicity 기준을 적용함에 있어 전문가 수준의 정확도를 제공할 수 있음을 의미합니다.

PrimateAI-3D의 정확도 높은 변이 oncogenicity 예측 결과

PrimateAI-3D는 40가지 미스센스 변이에 대해 100% oncogenicity 예측 정확도를 달성했으며, 벤치마크로 비교 확인한 19가지 인실리코 예측 변수보다 우수한 성능을 보였습니다(표 3). 실제로 PrimateAI-3D는 유일하게 벤치마크에 포함된 알려진 병원성 미스센스 변이를 모두 식별하였고, 모든 'benign' 변이를 정확하게 배정했습니다(표 3). 이러한 결과를 토대로 Connected Insights의 임베디드 AI 알고리즘이 현재 널리 사용되고 있는 다른 방법보다 높은 일치도를 제공하고 Illumina의 변이 pathogenicity 예측 기술이 업계를 선도함을 알 수 있습니다. Connected Insights의 우수한 성능은 PrimateAI-3D의 영장류 진화 데이터와 3D 단백질 구조 사용에 기인할 수 있는데, 이러한 장점은 이 데이터 세트에서 다른 알고리즘보다 눈에 띄게 더 정확한 예측을 가능하게 해 줍니다. 연구자는 Connected Insights에 통합된 PrimateAI-3D와 SpliceAI를 사용하여 자동적으로 최고 수준의 정확도를 누리고 변이 해석의 신뢰도를 높일 수 있습니다.

표 1: Connected Insights의 일치도 높은 변이 해석 결과

		Connected Insights		
		O + LO	VUS	LB + B
ClinGen	O + LO	43	3	0
	VUS	1	32	0
	LB + B	0	0	15

O = oncogenic(발암성), LO = likely oncogenic(준발암성), VUS = variant of unknown significance(미분류 변이), LB = likely benign(준양성), B = benign(양성). 변이 해석 일치도가 각기 다른 색상으로 표시되어 있는데, 초록색은 ClinGen과 Connected Insights 간에 변이 분류가 일치함을, 흰색은 ClinGen과 Connected Insights 간에 상충되는 분류를 가진 변이가 없음을, 보라색은 ClinGen과 Connected Insights 간에 변이 분류가 불일치함을 나타냄

표 2: Connected Insights로 변이를 'oncogenic'으로 재분류한 케이스

변이	설명	링크
PIK3CA N1068Kfs*5	성장 인자 부재에도 세포 증식이 증가함	OncoKB
	세포 증식, 이동 및 형질전환 능력이 증가함	CKB
RB1 S567L	E2F와의 결합이 감소하고 G1기 정지(G1 arrest) 유도와 분화 촉진에 실패함	OncoKB
TP53 G334R	집락 감소 분석에서 억제 능력이 저하되고 *TP53* 표적 유전자 하위집합의 활성화가 저하되며 열 안정성이 감소함	OncoKB

이 표에 포함된 변이는 Connected Insights가 전문가 지식 베이스보다 먼저 'oncogenic'으로 재분류함. 해당 변이는 oncogenic strong-2(OS2) 가이드라인에 따라 'oncogenic'으로 분류되었으며, 변이의 oncogenicity 영향은 이미 잘 알려진 체외*(in vitro)* 또는 생체 내*(in vivo)* 기능 연구가 뒷받침함

표 3: PrimateAI-3D의 정확도 높은 변이 oncogenicity 예측

		PrimateAI-3D	
		O + LO	LB + B
ClinGen	O + LO	33	0
	LB + B	0	7

O = oncogenic(발암성), LO = likely oncogenic(준발암성), LB = likely benign(준양성), B = benign(양성). 변이 해석 일치도가 각기 다른 색상으로 표시되어 있는데, 초록색은 ClinGen과 Connected Insights 간에 변이 분류가 일치함을, 흰색은 ClinGen과 Connected Insights 간에 상충되는 분류를 가진 변이가 없음을 나타냄

요약

이러한 결과를 모두 종합하면, Illumina Connected Insights가 자동 변이 해석 시 높은 분석 정확도를 제공함을 재확인할 수 있습니다. Connected Insights는 동급 최고 수준의 AI 예측 변수와 최신 근거를 사용하여 높은 정확도로 분류를 재현하며, 수동 변이 해석 방법으로는 놓칠 수 있는 최신 데이터로 기존 해석을 업데이트해 줍니다. Connected Insights는 엄격한 종양학 연구 가이드라인을 일관성 있게 적용하고 외부 벤치마크에 부합하는 결과를 제공함으로써 임상 연구에 적합한 신뢰도를 보여줍니다.

상세 정보

Illumina Connected Insights 무료 체험판을 신청하실 수 있습니다.

문서 번호: M-KR-00356 KOR
원본 문서: M-GL-04087 v1.0

연구 전용입니다. 진단 절차에는 사용할 수 없습니다.

참고 문헌

1. Austin-Tse CA, Jobanputra V, Perry DL, et al. Best practices for the interpretation and reporting of clinical whole genome sequencing. *NPJ Genom Med*. 2022;7:27 doi.org/10.1038/s41525-022-00295-z
2. Gao H, Hamp T, Ede J, et al. The landscape of tolerated genetic variation in humans and primates. *Science*. 2023;380(6648):eabn8153. doi:10.1126/science.abn8197
3. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell*. 2019;176(3):535-548.e24. doi:10.1016/j.cell.2018.12.015

illumina®

무료 전화(한국) 080-234-5300
techsupport@illumina.com | www.illumina.com

© 2026 Illumina, Inc. All rights reserved.
모든 상표는 Illumina, Inc. 또는 각 소유주의 자산입니다.
특정 상표 정보는 www.illumina.com/company/legal.html을 참조하십시오.
M-KR-00356 KOR