Previsione dell'oncogenicità con Illumina Connected Insights

Interpretazione automatizzata e altamente accurata delle varianti basata sui più recenti progressi delle linee guida oncologiche e sugli algoritmi di intelligenza artificiale di Illumina per la ricerca clinica

Previsione accurata dell'oncogenicità

Maggiore accuratezza della classificazione delle varianti rispetto ad altri classificatori e ai flussi di lavoro manuali

Interpretazione affidabile delle varianti

Elevata concordanza con le classificazioni cliniche consolidate e i criteri di oncogenicità assegnati dagli esperti

I migliori predittori *in silico*

Identificazione delle varianti che causano la malattia mediante algoritmi proprietari all'avanguardia

Introduzione

L'interpretazione delle varianti somatiche con metodi di sequenziamento di nuova generazione (NGS, Next-Generation Sequencing), come la mappatura genomica completa (CGP, Comprehensive Genomic Profiling) e gli studi di ricerca oncologica somatica, può essere complessa. Le linee guida per la classificazione stabilite da Association of Molecular Pathology (AMP), American Society of Clinical Oncology (ASCO) e College of American Pathologists (CAP), nonché gli standard di oncogenicità definiti rispettivamente da Clinical Genome Resource (ClinGen), Cancer Genomics Consortium (CGC) e Variant Interpretation for Cancer Consortium (VICC), forniscono un quadro di riferimento per la classificazione delle varianti somatiche in base all'utilità clinica e all'oncogenicità (ossia la capacità di contribuire alla formazione del tumore) rispettivamente. Tuttavia, gli attuali processi di interpretazione manuale sono soggetti a errori, richiedono molto tempo e spesso non sono coerenti nei risultati.[1] A causa di queste difficoltà, i laboratori di ricerca devono poter contare su soluzioni automatizzate in grado di applicare questi standard in modo più efficiente e coerente.

Illumina Connected Insights calcola automaticamente la previsione dell'oncogenicità e impiega PrimateAI-3D e SpliceAI nel proprio sistema di valutazione. Questi innovativi algoritmi proprietari consentono ai ricercatori di classificare le varianti in modo più accurato e di contare su risultati di ricerca più affidabili. Con PrimateAI-3D e SpliceAI, l'accuratezza della classificazione delle varianti è ancora più elevata, grazie all'identificazione più precisa delle varianti missenso patogene, che causano la malattia, rispetto a tutti gli altri classificatori testati.[2,3]

Per dimostrare l'accuratezza dell'interpretazione delle varianti oncologiche somatiche di Connected Insights, ne è stata valutata la capacità di previsione dell'oncogenicità tramite confronto con classificazioni assegnate dagli esperti e truth set consolidati.

Metodi

La concordanza della classificazione automatizzata dell'oncogenicità in Connected Insights è stata valutata tramite confronto con un truth set curato da esperti costituito da 94 varianti convalidate in base alle linee guida ClinGen/CGC/VICC. L'elevata concordanza indica l'allineamento con il giudizio degli esperti. L'accuratezza della classificazione in Connected Insights è stata valutata anche confrontando la precisione della previsione della patogenicità mediante PrimateAI-3D con quella degli algoritmi esistenti. L'accuratezza della previsione della patogenicità di PrimateAI-3D è stata misurata per 40 varianti missenso, confrontandola con quella di altri 19 predittori *in silico* (ad es., SIFT, PolyPhen-2). La concordanza è stata calcolata come percentuale di corrispondenze esatte alle classificazioni di riferimento. La forte concordanza di queste valutazioni dimostra la capacità di Connected Insights di emulare l'interpretazione manuale con elevata accuratezza.

Risultati

L'interpretazione automatizzata delle varianti con Connected Insights concorda con le classificazioni dell'oncogenicità curate dagli esperti

Nel set di convalida delle linee guida costituito da 94 varianti di tumore, Connected Insights ha riprodotto il 96% (90/94) delle etichette di oncogenicità assegnate dagli esperti. Il software ha identificato la maggior parte delle stesse etichette delle varianti: oncogena, probabilmente oncogena, variante di significato sconosciuto, probabilmente benigna o benigna (Tabella 1). In particolare, Connected Insights è stato in grado di incorporare evidenze aggiornate per alcune varianti evolutesi dopo la pubblicazione del truth set. In alcuni casi, Connected Insights ha trovato nuovi dati di supporto attingendo ai feed della knowledge base, con conseguente riclassificazione delle varianti come oncogene prima della riclassificazione post-pubblicazione da parte di fonti esterne (Tabella 2). Questo approccio evidenzia come l'automazione nella ricerca oncologica possa uguagliare e potenzialmente migliorare le interpretazioni integrando rapidamente nuove evidenze. Nel complesso, l'elevata concordanza indica che Connected Insights può raggiungere un'accuratezza paragonabile a quella degli esperti nell'applicazione dei criteri di oncogenicità ClinGen/CGC/VICC.

Previsione altamente accurata dell'oncogenicità delle varianti con PrimateAI-3D

PrimateAI-3D ha raggiunto il 100% di accuratezza nella previsione dell'oncogenicità rispetto a 40 varianti missenso, superando altri 19 strumenti di previsione *in silico* valutati e utilizzati come riferimenti (Tabella 3). In effetti, PrimateAI-3D è stato l'unico strumento a identificare tutte le varianti missenso patogene note nel riferimento, contrassegnando correttamente ogni variante benigna come tale (Tabella 3). Questi risultati confermano la leadership di Illumina nell'ambito della previsione della patogenicità delle varianti: gli algoritmi di intelligenza artificiale incorporati in Connected Insights dimostrano infatti una concordanza maggiore rispetto ad altri metodi ampiamente utilizzati.

Tabella 1: interpretazione altamente concordante delle varianti con Connected Insights

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

O, oncogena; LO (Likely Oncogenic), probabilmente oncogena; VUS (Variant of Unknown Significance), variante di significato sconosciuto; LB (Likely Benign), probabilmente benigna; B, benigna. La concordanza dell'interpretazione delle varianti è rappresentata dai colori: il verde indica che le classificazioni delle varianti con ClinGen e Connected Insights sono risultate allineate, il bianco indica che ClinGen e Connected Insights non hanno fornito classificazioni contrastanti e il blu indica che le classificazioni delle varianti con ClinGen e Connected Insights non sono risultate allineate.

Tabella 2: riclassificazione delle varianti come oncogene con Connected Insights

Variante	Descrizione	Link
PIK3CA N1068Kfs*5	Determina l'aumento della proliferazione cellulare in assenza di fattori di crescita	OncoKB
	Determina l'aumento della proliferazione cellulare, della migrazione e della capacità di trasformazione	CKB
RB1 S567L	Determina riduzione del legame a E2F, mancata induzione dell'arresto di G1 e mancata promozione della differenziazione	OncoKB
TP53 G334R	Determina compromissione della capacità di soppressione in un saggio di riduzione delle colonie, compromissione dell'attivazione di un sottogruppo di geni target *TP53* e ridotta stabilità termica	OncoKB

Connected Insights ha riclassificato come oncogene le varianti incluse in questa tabella prima delle basi di conoscenza degli esperti. Le varianti sono state classificate come oncogene secondo le linee guida Oncogenic Strong-2 (OS2), in base a studi funzionali *in vitro* o *in vivo* che ne hanno evidenziato l'effetto oncogeno.

Tabella 3: previsione altamente accurata dell'oncogenicità con PrimateAI-3D

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

O, oncogena; LO (Likely Oncogenic), probabilmente oncogena; LB (Likely Benign), probabilmente benigna; B, benigna. La concordanza dell'interpretazione delle varianti è rappresentata dai colori: il verde indica che le classificazioni delle varianti con ClinGen e Connected Insights sono risultate allineate e il bianco indica che ClinGen e Connected Insights non hanno fornito classificazioni contrastanti per alcuna variante.

Va notato che l'uso dei dati evolutivi dei primati e del contesto proteico 3D da parte di PrimateAI-3D potrebbe essere alla base delle sue prestazioni superiori, un vantaggio che si è tradotto in previsioni sensibilmente migliori rispetto agli algoritmi tradizionali su questo set di dati. L'integrazione di PrimateAI-3D e SpliceAI in Connected Insights consente ai ricercatori di trarre automaticamente vantaggio da questa accuratezza di alto livello e di poter contare su interpretazioni più affidabili delle varianti.

Riepilogo

Nel loro complesso, questi risultati dimostrano che Illumina Connected Insights offre elevata accuratezza analitica nell'interpretazione automatizzata delle varianti. Lo strumento riproduce molto fedelmente le classificazioni grazie ai migliori predittori di intelligenza artificiale della categoria e alle evidenze più attuali, aggiornando le interpretazioni con dati che i metodi manuali non sono in grado di integrare. La coerente applicazione di rigorose linee guida oncologiche e l'allineamento dimostrato ai riferimenti esterni evidenziano l'affidabilità di Connected Insights per l'uso nella ricerca clinica.

Ulteriori informazioni →

Richiedere una prova gratuita di Illumina Connected Insights

Bibliografia

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®

Numero verde 1.800.809.4566 (U.S.A.) | Tel. +1.858.202.4566
techsupport@illumina.com | www.illumina.com

© 2026 Illumina, Inc. Tutti i diritti riservati. Tutti i marchi di fabbrica sono di proprietà di Illumina, Inc. o dei rispettivi proprietari. Per informazioni specifiche sui marchi di fabbrica, visitare la pagina web www.illumina.com/company/legal.html.
M-GL-04087 ITA v1.0