

Illumina Human Comprehensive Panel

Copertura avanzata delle
regioni difficili nei geni
codificanti le proteine con
Illumina Complete Long Read
Prep with Enrichment, Human

illumina[®]

Introduzione

Quando si esegue il sequenziamento dell'intero genoma (WGS, whole-genome sequencing) umano, potrebbe essere presente una ridotta porzione di regioni geniche difficili da mappare solo con letture brevi. Le regioni difficili includono quelle altamente omologhe o ripetitive, varianti strutturali complesse, pseudogeni e ampie inserzioni-delezioni (indel). In questi casi, il sequenziamento a lettura lunga può integrare i dati WGS standard a lettura breve per contribuire a fornire una copertura migliorata atta a risolvere regioni difficili e a semplificare la determinazione delle fasi delle varianti e l'identificazione degli aplotipi.

La tecnologia Illumina Complete Long Reads utilizza un flusso di lavoro standard di sequenziamento di nuova generazione (NGS, next-generation sequencing) per generare sequenze a lettura lunga contigue sui sistemi di sequenziamento Illumina (Figura 1). Illumina Complete Long Read Prep with Enrichment e Human Comprehensive Panel offrono un approccio conveniente e scalabile per aiutare i ricercatori a risolvere le regioni del genoma notoriamente difficili. Le letture lunghe mirate vengono analizzate con dati WGS standard ad alta precisione per consentire la mappatura di più regioni.

Pannello sonda completo e ottimizzato

L'Illumina Human Comprehensive Panel è un pannello sonda con ibridazione-cattura ottimizzato e progettato per le regioni geniche che possono trarre vantaggio da una maggiore mappabilità delle letture più lunghe. Il pannello si occupa di regioni a bassa compatibilità con più di 6.500 geni codificanti le proteine (Tabella 1)¹ per consentire letture lunghe complementari che migliorano la risoluzione delle aree difficili da mappare utilizzando metodi WGS standard a lettura breve.

Tabella 1: parametri dello Human Comprehensive Panel^a

Sistema	NovaSeq X Series NovaSeq 6000 System
Dimensione regione target	>95 Mb
Output di sequenziamento per campione ^b	90-120 Gb
N. di sonde	Circa 40.000
N. di geni mirati	>6.500
Tipo di campione	DNA genomico
Input di DNA raccomandato	50 ng
Durata totale della preparazione delle librerie	Circa 2 giorni
Interventi manuali	Circa 6 ore
Multiplex	Fino a 64 campioni
N50	6,3 kb
Blocco di fase N50	15,6 kb
Uniformità	95%
Arricchimento della lettura "padded" (PRE) ^c	83%
% di SNV eterozigote a fasi ^d	98%

- Dati generati utilizzando 50 ng di DNA genomico HG002 (Coriell, n. di catalogo NA24385). Le prestazioni possono variare in base all'input di DNA e alla qualità del campione.
- Richiede 2 × corse di sequenziamento da 150 bp e genera una copertura finale di circa 30× dell'Illumina Complete Long Reads.
- Uniformità della copertura calcolata come % >0,2 * media.
- PRE calcolato come 100 * (letture allineate target "padded"/letture allineate totali).
- SNV, variante a singolo nucleotide.



Figura 1: parte di un flusso di lavoro integrato. Consente di accedere ai dati WGS a lettura lunga mirati e convenienti utilizzando un protocollo scalabile e ottimizzato per la preparazione delle librerie con arricchimento, la comprovata chimica di sequenziamento Illumina e l'analisi secondaria DRAGEN. Per l'analisi richiede almeno 30× di dati WGS standard a lettura breve dallo stesso campione. È possibile utilizzare i file FASTQ di un campione precedentemente analizzato.

Nello sviluppo di questo pannello, Illumina ha preso in considerazione l'intero gruppo di oltre 20.000 geni codificanti le proteine, inclusi gli introni e le regioni non tradotte (UTR, untranslated region), e ha scelto come target il 70% di tutte le coppie di basi difficili da mappare in queste regioni. Sono stati esclusi i geni coperti in modo esaustivo solo da letture brevi. Questo pannello è stato testato e ottimizzato per laboratorio umido (wet-lab) al fine di ottimizzare l'efficienza e le prestazioni di progettazione e sequenziamento.

Sequenziamento e analisi altamente scalabili

Il flusso di lavoro Illumina Complete Long Read Prep with Enrichment è altamente scalabile e facile da automatizzare al fine di supportare il WGS completo per un numero maggiore di campioni (Tabella 2). Con il NovaSeq™ X Plus System, gli utenti possono generare con elevata accuratezza fino a 15.000 genomi l'anno.* Per ridurre i requisiti di batch, sequenziare meno campioni su materiali di consumo a bassa processività come la cella a flusso NovaSeq X 1.5B. Illumina Complete Long Read Prep with Enrichment e Human Comprehensive Panel possono essere utilizzati per aumentare i set di dati WGS esistenti come strumento reflex per un rilevamento più ampio delle varianti.

* Potenziale processività quando Illumina Complete Long Read Prep with Enrichment e Human Comprehensive Panel vengono utilizzati con il NovaSeq X Plus System, in corse con doppia cella a flusso con le celle a flusso 25B.

L'analisi dei dati per Illumina Complete Long Read Prep with Enrichment, Human è disponibile come app BaseSpace™ Sequence Hub o tramite Illumina Connected Analytics. I dati vengono trasmessi in streaming sul cloud e analizzati insieme ad almeno 30× di dati WGS standard a lettura breve dallo stesso campione (generati in precedenza o in parallelo). La pipeline DRAGEN™ unisce i risultati in un unico set di file di output che include gli identificatori mirati DRAGEN.³

Copertura migliorata e conveniente

Illumina Complete Long Read Prep with Enrichment e Human Comprehensive Panel migliorano l'accuratezza dell'identificazione di varianti e la copertura per tutte le regioni target rispetto al WGS standard a lettura breve (Figura 2, Figura 3, Tabella 3). Le letture lunghe mirate che impiegano lo Human Comprehensive Panel forniscono un'accuratezza completa dell'intero genoma con un punteggio F1 (SNV + indel) pari al 99,87%,² il che rappresenta un miglioramento sostanziale rispetto alle sole letture brevi. Ciò assicura prestazioni paragonabili a quelle offerte da Illumina Complete Long Read Prep, Human e da altri genomi interi a lettura lunga nei geni codificanti le proteine, ma con costi inferiori e una maggiore processività (Tabella 3).

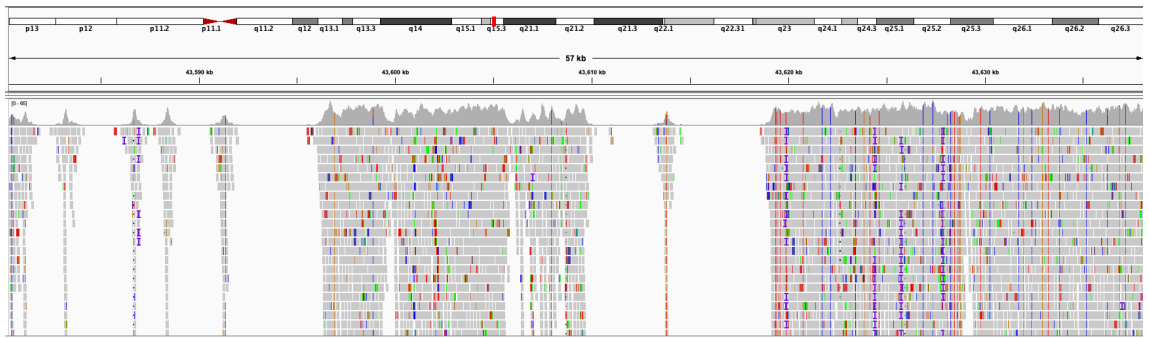
Tabella 2: processività del campione raccomandata per generare una copertura finale 30× per Human Comprehensive Panel e Illumina Complete Long Read Prep with Enrichment, Human^{a,b,c,d}

	NovaSeq 6000				NovaSeq X		
	SP	S1	S2	S4	1.5B	10B	25B
Kit di reagenti da 300 cicli							
Campioni per cella a flusso	2	4	10	24	4	24	64
Output per cella a flusso	Circa 250 Gb	Circa 500 Gb	Circa 1,25 Tb	Circa 3 Tb	Circa 500 Gb	Circa 3 Tb	Circa 8 Tb
Durata corsa	Circa 25 ore	Circa 25 ore	Circa 36 ore	Circa 44 ore	Circa 21 ore	Circa 25 ore	Circa 48 ore

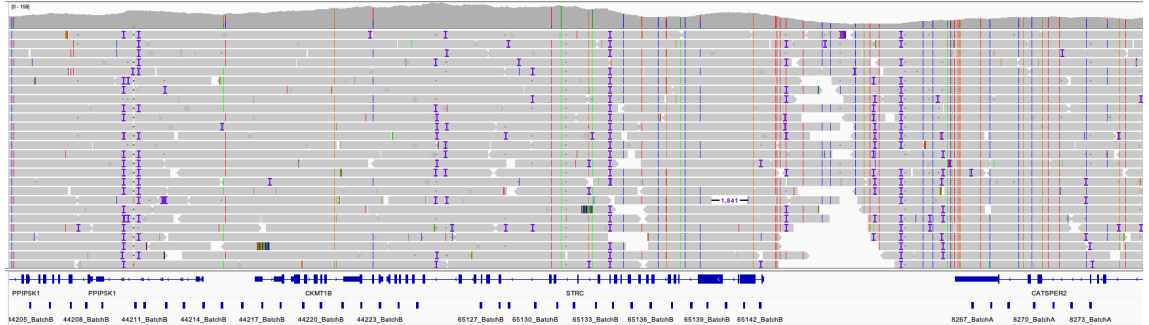
- a. La dimensione della regione target dello Human Comprehensive Panel è maggiore di 95 Mb e richiede un output di sequenziamento di 90-120 Gb per campione.
 b. Richiede 2 × corse di sequenziamento da 150 bp e genera una copertura finale di circa 30× dell'Illumina Complete Long Reads.
 c. Per l'analisi richiede 30× di dati del genoma intero umano di letture brevi standard dallo stesso campione. Si raccomanda l'Illumina DNA PCR-Free Prep. Sono compatibili anche i kit WGS di terze parti. Le librerie non marcate non devono essere preparate o sequenziate in parallelo; è possibile utilizzare i file FASTQ di un campione precedentemente analizzato.
 d. Se il sequenziamento delle librerie Illumina Complete Long Read viene eseguito sulle piattaforme NovaSeq, il punteggio Q30 riportato per una corsa potrebbe risultare al di sotto delle specifiche del NovaSeq. Questo non indica un problema relativo alle prestazioni con la corsa di sequenziamento o con la libreria.

A.

WGS standard
a lettura breve

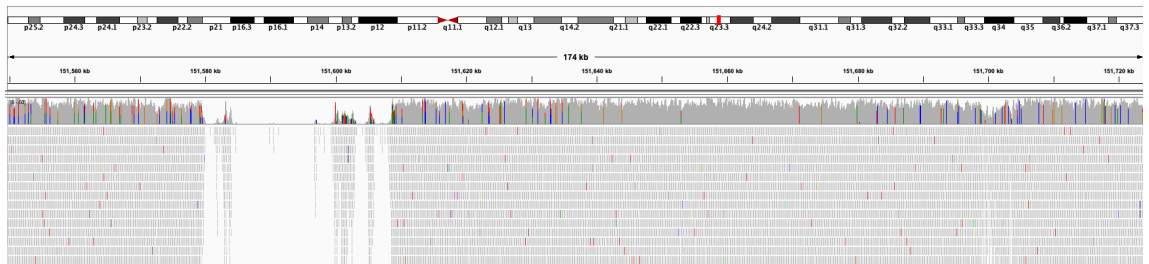


Illumina
Complete Long
Read Prep with
Enrichment



B.

WGS standard
a lettura breve



Illumina
Complete Long
Read Prep with
Enrichment

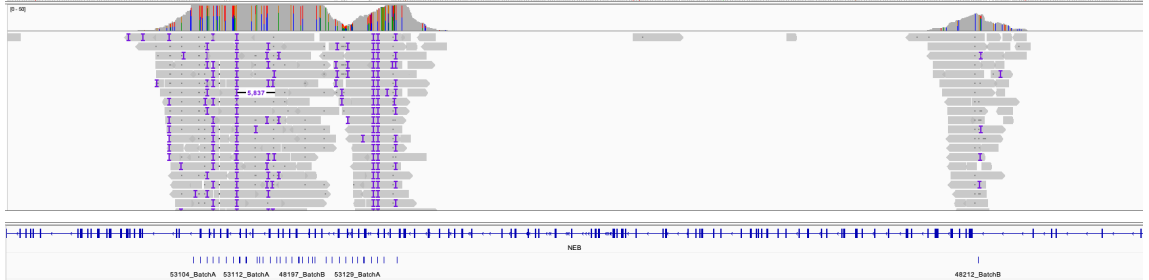


Figura 2: risoluzione delle regioni difficili da mappare utilizzando letture lunghe mirate con Human Comprehensive Panel. Illumina Complete Long Read Prep with Enrichment, Human e Human Comprehensive Panel aiutano a migliorare la copertura in regioni geniche difficili per completare il WGS umano standard a lettura breve. Grafici Integrative Genomics Viewer (IGV) di (A) *STRC* e (B) *NEB* utilizzando il WGS standard a lettura breve e l'Illumina Complete Long Reads with Enrichment.

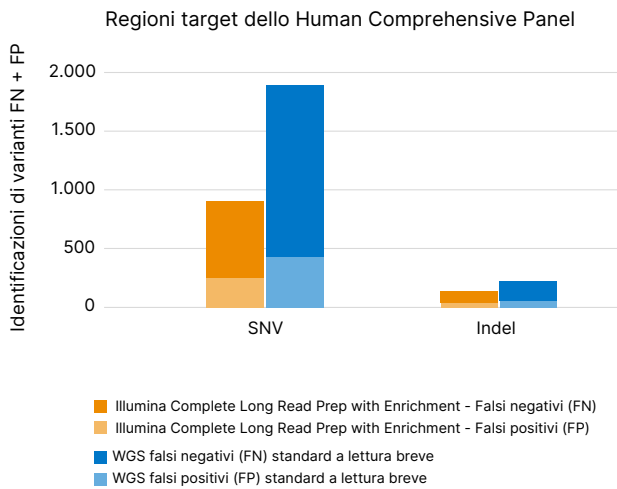


Figura 3: letture lunghe mirate per migliorare l'accuratezza dell'identificazione di varianti. Identificazioni di varianti false negative (FN) e false positive (FP) per le varianti SNV e indel nelle regioni geniche mirate dallo Human Comprehensive Panel utilizzando Illumina Complete Long Read Prep with Enrichment e Human Comprehensive Panel (arancione) rispetto al WGS standard a lettura breve (blu).

Riepilogo

L'Illumina Human Comprehensive Panel consente una copertura a lettura lunga ad alta precisione della piccola porzione di regioni geniche difficili da mappare solo con letture brevi. Illumina Complete Long Read Prep with Enrichment e Human Comprehensive Panel offrono un saggio ottimizzato e conveniente dell'intero genoma che integra l'Illumina WGS e si occupa delle letture lunghe laddove assicurano il massimo valore con una soluzione completa per il flusso di lavoro.

Maggiori informazioni

[Illumina Complete Long Read Prep with Enrichment, Human Comprehensive Panel](#)

[Tecnologia di sequenziamento delle letture lunghe](#)

Tabella 3: prestazioni dello Human Comprehensive Panel

Accuratezza per le regioni target dello Human Comprehensive Panel ^a			
	Illumina Complete Long Read Prep with Enrichment e Human Comprehensive Panel ^b	WGS standard a lettura breve ^c	
Punteggio F1 (SNV)	99,54%	99,04%	
Punteggio F1 (indel)	99,47%	99,15%	
Punteggio F1 (SV)	80,43%	60,18%	
Accuratezza per l'intero genoma, tutte le regioni di riferimento ^d Punteggio F1 (SNV + indel)			
	Illumina Complete Long Read Prep with Enrichment e Human Comprehensive Panel ^b	Illumina Complete Long Read Prep, Human ^e	WGS standard a lettura breve ^c
	99,87%	99,90%	99,84%

a. Solo autosomi. SNV, varianti a singolo nucleotide; SV, varianti strutturali.
b. Illumina Complete Long Read Prep with Enrichment, Human Comprehensive Panel + Illumina DNA PCR-Free Prep + NovaSeq 6000 System + DRAGEN v4.2.
c. Illumina DNA PCR-Free Prep + NovaSeq 6000 System + DRAGEN v4.2.
d. Secondo le misurazioni rispetto a tutte le soluzioni partecipanti utilizzando PrecisionFDA Truth Challenge v2 Benchmark Data,² dati interni in archivio per DRAGEN v4.
e. Illumina Complete Long Read Prep, Human + Illumina DNA PCR-Free Prep + NovaSeq 6000 System + DRAGEN v4.2.

Bibliografia

1. Bekritsky MA, Bekritsky MA, Colombo C, Eberle MA. Identifying genomic regions with high quality single nucleotide variant calling. illumina.com/science/genomics-research/articles/identifying-genomic-regions-with-high-quality-single-nucleotide-.html. Pubblicato nel 2021. Consultato il 30 agosto 2023.
2. PrecisionFDA. Truth Challenge V2: Calling Variants from Short and Long Reads in Difficult-to-Map Regions. precision.fda.gov/challenges/10. Consultato il 2 ottobre 2023.
3. Roessler K. Illumina Complete Long Reads software analysis workflow for human WGS. illumina.com/science/genomics-research/articles/complete-long-read-software-analysis.html. Pubblicato nel 2023. Consultato il 22 settembre 2023.

Informazioni per gli ordini

Prodotto	N. di catalogo
Illumina Complete Long Read Prep with Enrichment, Human (24 samples)	20113832
Illumina Complete Long Read Prep with Enrichment, Human (96 samples)	20113833
Illumina Complete Long Read Prep with Enrichment, Human Comprehensive Panel (24 samples)	20113834
Illumina Complete Long Read Prep with Enrichment, Human Comprehensive Panel (96 samples)	20113835
Illumina Human Comprehensive Panel (24 samples)	20113836
Illumina Human Comprehensive Panel (96 samples)	20113837
Illumina Unique Dual Indexes, LT (48 indexes, 48 samples)	20098166
Illumina DNA/RNA UD Indexes Set A, Tagmentation (96 indexes, 96 samples)	20091654
Illumina DNA/RNA UD Indexes Set B, Tagmentation (96 indexes, 96 samples)	20091656
Illumina DNA/RNA UD Indexes Set C, Tagmentation (96 indexes, 96 samples)	20091658
Illumina DNA/RNA UD Indexes Set D, Tagmentation (96 indexes, 96 samples)	20091660
Illumina Analytics - 1 iCredit	20042038



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

© 2024 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-02191 ITA v1.0