

Illumina Complete Long Read Prep with Enrichment, Human

Letture lunghe flessibili,
convenienti e accurate sul sistema
di sequenziamento Illumina

- Letture lunghe mirate abbinate a letture brevi per informazioni complementari e avanzate sul genoma umano
- Preparazione scalabile e solida delle librerie con risultati affidabili con un ridotto input di DNA
- Flusso di lavoro di preparazione delle librerie di due giorni compatibile con l'automazione senza la necessità di apparecchiature specializzate



Introduzione

La comprovata chimica di sequenziamento mediante sintesi (SBS, sequencing by synthesis) di Illumina combinata con la premiata analisi secondaria DRAGEN™ fornisce dati di sequenziamento dell'intero genoma (WGS, whole-genome sequencing) con un'accuratezza eccezionale.^{1,2} Tuttavia, piccole porzioni del genoma, incluse le regioni altamente omologhe o ripetitive, sono difficili da mappare solo con letture brevi. In questi casi, il sequenziamento a lettura lunga può integrare i dati WGS standard a lettura breve. L'utilizzo combinato di letture brevi e lunghe ad alta precisione può contribuire a fornire una risoluzione e una mappatura migliorate delle regioni storicamente difficili.

In passato, il sequenziamento a lettura lunga non era compatibile con molti tipi di campioni a causa della rigorosa qualità del DNA e degli elevati requisiti di input.³⁻⁶ Grazie alla tecnologia Illumina Complete Long Reads ora il sequenziamento a lettura lunga è accessibile ai laboratori di genomica consentendo sia letture brevi sia letture lunghe complementari sullo stesso sistema con una singola pipeline di analisi. Questo saggio ad alte prestazioni utilizza un flusso di lavoro standard di sequenziamento di nuova generazione (NGS, next-generation sequencing) di Illumina per generare sequenze di lettura lunga contigue e richiede solo 10 ng di input di DNA senza estrazioni, frammentazione (shearing) o selezione delle dimensioni specializzati (Figura 1).

Illumina Complete Long Read Prep with Enrichment, Human crea una soluzione economica e flessibile per il sequenziamento mirato a lettura lunga. L'efficiente preparazione delle librerie di due giorni con arricchimento è di facile automazione e semplice da scalare per studi a elevata processività.

Le letture lunghe mirate sono utilizzabili per occuparsi di quelle regioni notoriamente a bassa mappabilità mediante sequenziamento standard a lettura breve. In alternativa, le letture lunghe mirate possono essere applicate su interi geni o regioni per consentire il sequenziamento a fasi fino a centinaia di kilobasi per risolvere gli aplotipi. Utilizzare Illumina Complete Long Read Prep with Enrichment, Human per aumentare i set di dati WGS esistenti come strumento reflex per un rilevamento più ampio delle varianti.

Letture lunghe mirate di alta qualità dalla piattaforma NovaSeq™

L'uso di Illumina Complete Long Read Prep with Enrichment, Human è stato convalidato con il NovaSeq X Plus Sequencing System, il NovaSeq X Sequencing System, il NovaSeq 6000 Sequencing System e il NovaSeq 6000Dx Sequencing System (modalità RUO). Il saggio è flessibile e consente di ottenere risultati coerenti su campioni di qualità variabile, oltre a richiedere il 90% in meno di input di DNA rispetto ad altre soluzioni di sequenziamento a lettura lunga. Illumina raccomanda 50 ng di DNA di input, ma vengono restituiti risultati validi con appena 10 ng. La tecnologia Complete Long Read di Illumina è resistente ai comuni inibitori e contaminanti ed è efficace con il DNA di sangue, saliva o tessuto.^{7,8} Ciò consente di ottenere informazioni complete da più tipi di campioni rispetto ad altre soluzioni a lettura lunga.



Figura 1: flusso di lavoro Illumina Complete Long Reads with Enrichment. Consente di accedere ai dati WGS a lettura lunga mirati e convenienti utilizzando un protocollo scalabile e ottimizzato per la preparazione delle librerie, 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.

L'Illumina Complete Long Reads unisce un saggio di preparazione delle librerie proprietario, la comprovata chimica di sequenziamento SBS di Illumina e un'efficace analisi secondaria DRAGEN per generare dati altamente accurati di letture lunghe (Figura 2). I frammenti lunghi di DNA a singola molecola vengono marcati con pattern unici (o "landmark"), quindi amplificati, arricchiti e sequenziati. L'aggiunta di una fase di arricchimento mediante ibridazione-cattura con pannelli sonda mirati consente al sequenziamento di concentrarsi sulle regioni che beneficiano maggiormente delle informazioni a lettura lunga. I landmark consentono la differenziazione di regioni ripetitive o difficili da mappare per generare letture lunghe con un N50 di 5-7 kb.⁷ I dati di lettura lunga sono combinati con una libreria WGS standard non marcata per produrre letture lunghe contigue che sono una rappresentazione completa e accurata dei frammenti originali mirati a singola molecola.

Letture lunghe dove servono

Le letture lunghe mirate stratificate su dati WGS a lettura breve e ad alta precisione aiutano i ricercatori a destinare le risorse economiche per il sequenziamento alla risoluzione di regioni notoriamente difficili del genoma. Le letture lunghe mirate offrono elevata flessibilità con la scelta di più pannelli predefiniti o pannelli personalizzati abilitati dal software Illumina DesignStudio™, uno strumento di progettazione saggi gratuito e intuitivo (Tabella 1, Tabella 2). Scegliere il pannello sonda di arricchimento che meglio soddisfa le proprie esigenze di ricerca: per migliorare le regioni a bassa copertura o le regioni di fase del genoma.

Illumina Human Comprehensive Panel

L'ampio Illumina Human Comprehensive Panel si rivolge alla piccola porzione di regioni geniche che trarrebbe vantaggio da letture lunghe, mirando a punti a bassa copertura su più di 6.500 geni codificanti le proteine.⁹ Nello sviluppo di questo pannello, Illumina ha valutato l'intero set di oltre 20.000 geni codificanti le proteine per mirare alle regioni a bassa mappabilità. Sono stati esclusi i geni coperti in modo esaustivo solo da letture brevi. Human Comprehensive Panel offre una copertura migliorata e capacità avanzate di identificazione di varianti nelle regioni target (Figura 3A). Questo pannello è ottimizzato per l'uso con Illumina Complete Long Read Prep with Enrichment, Human ed è disponibile prefabbricato o come pannello pre-progettato che può essere personalizzato nello strumento DesignStudio.

Pannelli pre-progettati con il software DesignStudio

A un costo inferiore e con un'opzione di processività più elevata, i ricercatori possono scegliere pannelli mirati per l'arricchimento a lettura lunga (Tabella 1).

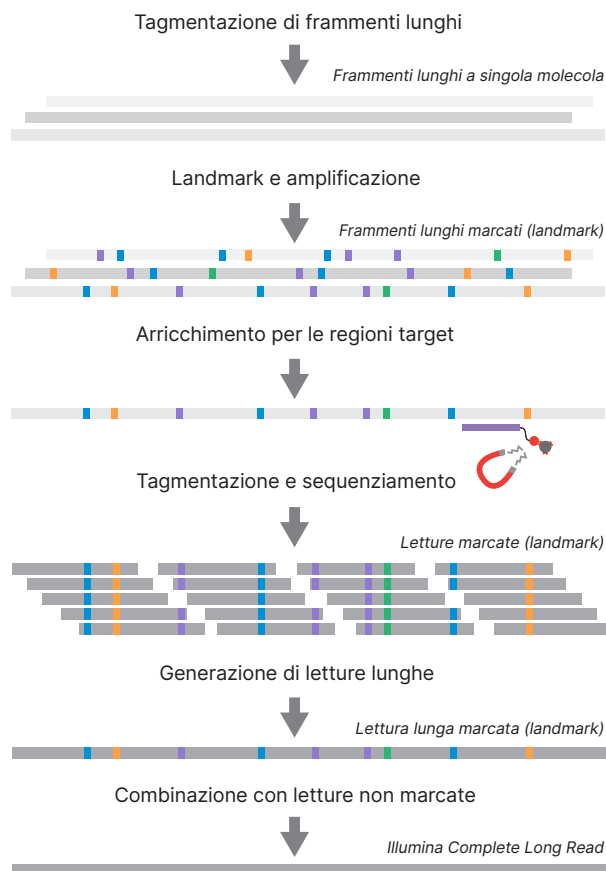


Figura 2: funzionamento del saggio Illumina Complete Long Reads. Il saggio utilizza la tagmentazione per generare frammenti lunghi di DNA (più di 10 kb), eliminando la frammentazione (shearing) o la selezione della dimensione. I frammenti lunghi vengono marcati (landmark) a livello di singola molecola per acquisire e preservare le informazioni sulla lettura lunga all'interno del frammento (senza barcoding o adattatori complessi). I frammenti lunghi marcati (landmark) sono arricchiti per le regioni target utilizzando un pannello sonda con ibridazione-cattura. Dopodiché, i frammenti lunghi arricchiti vengono sottoposti a tagmentazione ancora una volta in una libreria pronta per il sequenziamento. L'analisi genera letture lunghe e combina i dati con una libreria WGS standard non marcata (dallo stesso campione, sequenziata separatamente) per produrre Illumina Complete Long Read altamente accurate.

Tabella 1: processività del campione raccomandata per generare una copertura finale 30× per Illumina Complete Long Read Prep with Enrichment, Human^{a,b,c}

	Illumina Human Comprehensive Panel	Pannello grande	Pannello medio	Pannello piccolo		
Dimensione regione target ^d	>95 Mb	Circa 20 Mb	Circa 10 Mb	Circa 5 Mb		
Dati raccomandati per campione ^a	90-120 Gb	30-60 Gb	15-30 Gb	7,5-15 Gb		
Kit di reagenti da 300 cicli	N. di campioni per cella a flusso ^e				Output per cella a flusso	Durata corsa
NovaSeq 6000 SP Reagents	2	8	16	32	Circa 250 Gb	Circa 25 ore
NovaSeq 6000 S1 Reagents	4	16	32	66	Circa 500 Gb	Circa 25 ore
NovaSeq 6000 S2 Reagents	10	40	82	166	Circa 1,25 Tb	Circa 36 ore
NovaSeq 6000 S4 Reagents	24	100	200	400 ^f	Circa 3 Tb	Circa 44 ore
NovaSeq X Series 1.5B Reagents	4	16	32	66	Circa 500 Gb	Circa 21 ore
NovaSeq X Series 10B Reagents	24	100	200	400 ^f	Circa 3 Tb	Circa 25 ore
NovaSeq X Series 25B Reagents	64	266	532 ^f	1.066 ^f	Circa 8 Tb	Circa 48 ore

a. Richiede una corsa di sequenziamento di 2 × 150 bp e 5-10 milioni di letture paired-end (circa 1,5-3 Gb di dati) per regione target Mb, generando circa 30× di copertura finale di Illumina Complete Long Read. I requisiti relativi ai dati del pannello personalizzato per campione rappresentano solo il punto di partenza raccomandato. È possibile ottimizzare i dati assegnati in base alle prestazioni del pannello.

b. Per l'analisi richiede almeno 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.

c. 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.

d. La dimensione della regione target corrisponde alla somma delle lunghezze della posizione della sonda "padded", unite dove si sovrappongono.

e. Esempio di numeri di processività del campione calcolati per 5 milioni di letture paired-end (1,5 Gb di dati) per regione target Mb.

f. È disponibile un massimo di 384 doppi indici unici. Per la NovaSeq X Series, il caricamento indipendente delle corsie consente il multiplex di più campioni. Utilizzare il flusso di lavoro NovaSeq 6000 Xp per il caricamento indipendente delle corsie sul NovaSeq 6000 System.

Diversi pannelli pre-progettati sono disponibili nel software DesignStudio e possono essere personalizzati (Tabella 2). Questi pannelli mirano a geni rilevanti dal punto di vista medico (CMRG, challenging medically relevant genes),¹⁰ a geni comunemente coperti dai saggi di analisi farmacogenetica (PGx, pharmacogenetic),¹¹⁻¹³ ai geni presenti nell'elenco dei risultati secondari dell'American College of Medical Genetics and Genomics (ACMG) (ACMG SF v3.1)¹⁴ o alla regione completa con complesso maggiore di istocompatibilità (MHC, major histocompatibility complex).¹⁵

Pannelli personalizzati

Gli utenti possono utilizzare le proprie conoscenze per creare un pannello unico dedicato a un sottogruppo di geni. Lo strumento DesignStudio supporta la progettazione di pannelli personalizzati con un algoritmo ottimizzato per l'arricchimento di frammenti lunghi. Anche i pannelli di oligonucleotidi di terze parti sono compatibili con Illumina Complete Long Read Prep with Enrichment, Human.

Flusso di lavoro scalabile e a elevata processività

Il flusso di lavoro Illumina Complete Long Read Prep with Enrichment, Human è facile da automatizzare, richiede solo apparecchiature standard da laboratorio ed è altamente scalabile per supportare un WGS completo per più campioni (Figura 1). Il semplice protocollo di preparazione delle librerie richiede circa 6,5 ore il primo giorno (con circa 3 ore di interventi manuali), seguito da una reazione di ibridazione notturna e 5,5 ore il secondo giorno (con circa 3 ore di interventi manuali).

Utilizzare le celle a flusso NovaSeq X Series 10B o 25B per una maggiore processività per coorti più grandi e costi ottimizzati per pannelli più grandi (Tabella 1). Con il NovaSeq X Plus System, ogni anno gli utenti possono generare con elevata accuratezza fino a 15.000 genomi migliorati. Utilizzare materiali di consumo a bassa processività come la cella a flusso NovaSeq X Series 1.5B per piccoli pannelli e per ridurre i requisiti di batch.

Tabella 2: pannelli sonda con ibridazione-cattura progettati per Illumina Complete Long Read Prep with Enrichment, Human

Pannello	Geni o regioni target	Dimensione pannello
Illumina Human Comprehensive Panel ^a	Macchie a bassa copertura su più di 6.500 geni codificanti le proteine ⁹	>95 Mb
Pannello CMRG ^b	391 geni clinicamente rilevanti notoriamente difficili da risolvere con letture brevi ¹⁰	22,5 Mb
Pannello PGx ^b	98 geni comunemente coperti dai saggi di analisi farmacogenetica ¹¹⁻¹³	8,1 Mb
Pannello ACMG ^b	78 geni unici dall'elenco dei risultati secondari dell'ACMG (ACMG SF v3.1) ¹⁴	7 Mb
Pannello MHC ^b	La regione MHC completa (più di 140 geni) nel gruppo GRCh38.p14 ¹⁵	4,9 Mb

a. Pannello prefabbricato ottimizzato.
b. Pannelli pre-progettati e personalizzati disponibili nel software DesignStudio. CMRG, geni rilevanti dal punto di vista medico; PGx, farmacogenomica; ACMG, American College of Medical Genetics and Genomics; MHC, complesso maggiore di istocompatibilità.

Analisi esaustiva e semplificata

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. La singola pipeline DRAGEN analizza sia le letture brevi sia quelle lunghe per fornire risultati WGS completi. I risultati vengono uniti in un unico set di file di output, inclusi gli identificatori mirati DRAGEN.¹⁶

Accesso a WGS altamente accurato con dati a lettura lunga arricchiti

I dati di Illumina Complete Long Read Prep with Enrichment, Human mostrano una copertura e un'accuratezza migliorate in regioni difficili mirate rispetto al solo WGS standard a lettura breve (Figura 3, Figura 4). Targeted Illumina Complete Long Reads risolve con successo anche gli aplotipi di geni completi e può frazionare grandi blocchi di regioni altamente polimorfiche come l'MHC (Figura 5).

Riepilogo

Illumina Complete Long Read Prep with Enrichment, Human è una soluzione flessibile e conveniente che integra il comprovato WGS Illumina e si concentra sulle letture lunghe laddove offrono il massimo valore. L'Illumina Complete Long Read consente ai laboratori di genomica di accedere facilmente al WGS esaustivo poiché sullo stesso strumento possono essere analizzate sia letture lunghe sia letture brevi con una singola pipeline di analisi DRAGEN. Il flusso di lavoro semplificato e di uso comune offre il saggio dell'intero genoma più scalabile e accurato disponibile sul mercato.

Maggiori informazioni

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

[Human Comprehensive Panel](#)

[Tecnologia di sequenziamento delle letture lunghe](#)

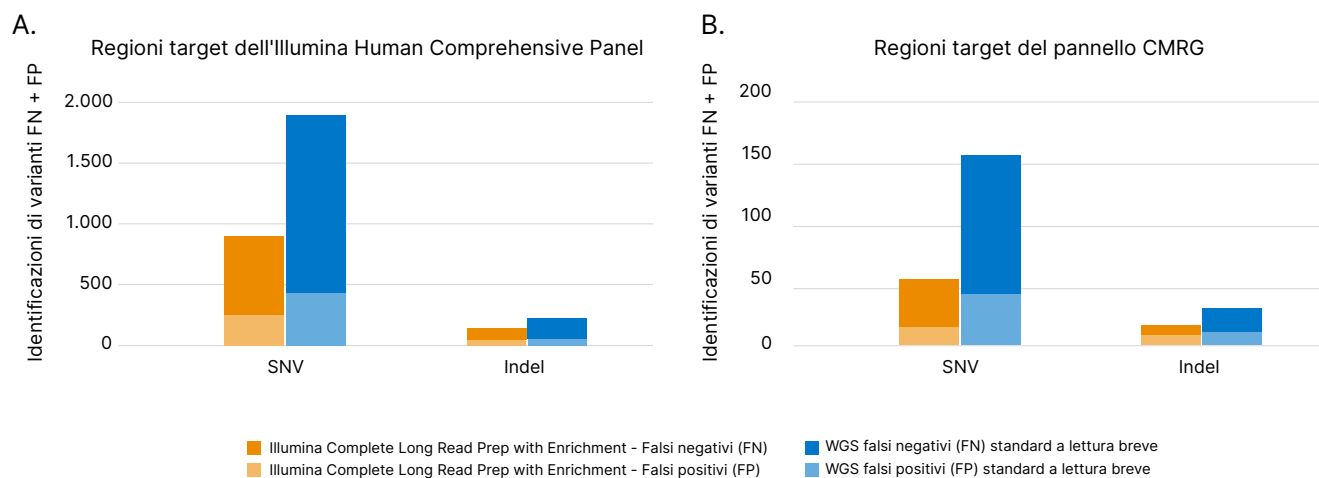


Figura 3: letture lunghe mirate per migliorare l'accuratezza dell'identificazione di varianti nelle regioni difficili. Identificazioni di varianti false negative (FN) e false positive (FP) per le varianti di singolo nucleotide (SNV, single nucleotide variant) e delezione/ inserzione (indel) nelle regioni geniche HG002 coperte (A) dall'Illumina Human Comprehensive Panel o (B) dal pannello CMRG, utilizzando Illumina Complete Long Read Prep with Enrichment (arancione) rispetto al WGS standard a lettura breve (blu).

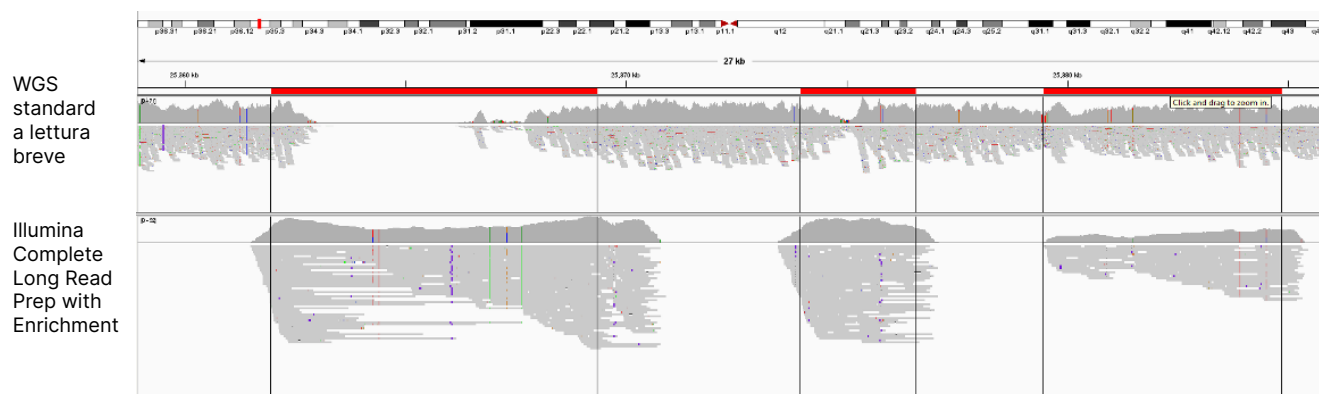
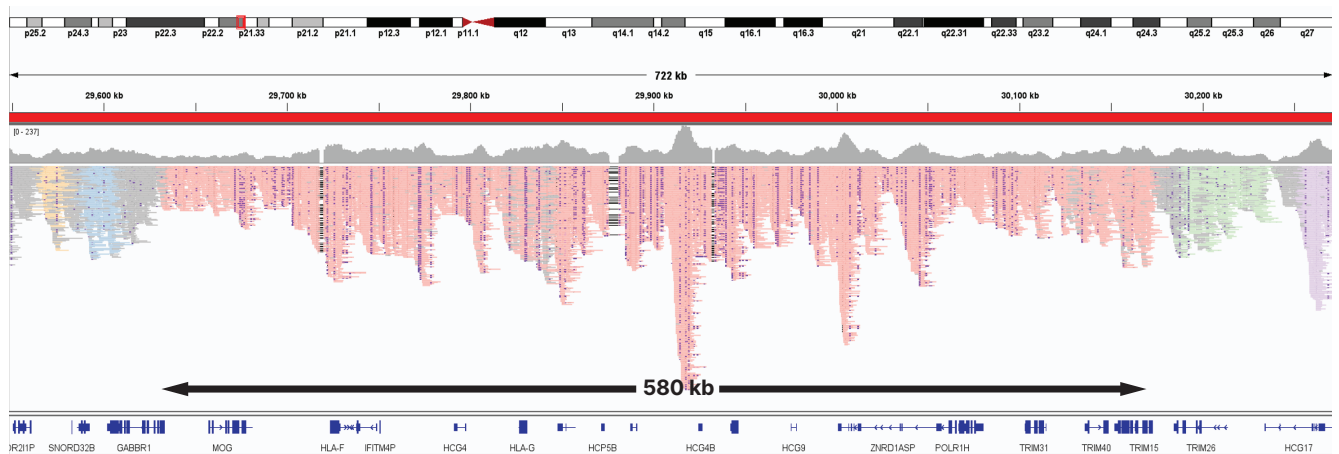


Figura 4: raggiungimento di una copertura conveniente dell'intero genoma umano in regioni difficili. Illumina Complete Long Read Prep with Enrichment, Human aiuta a migliorare la copertura in regioni geniche difficili come quelle del gene *RHCE* per integrare il WGS umano standard a lettura breve. Grafico Integrative Genomics Viewer (IGV) di *RHCE* sequenziato utilizzando WGS standard a lettura breve (in alto) e Illumina Complete Long Read Prep with Enrichment (in basso). Regioni target indicate in rosso.

A.



B.

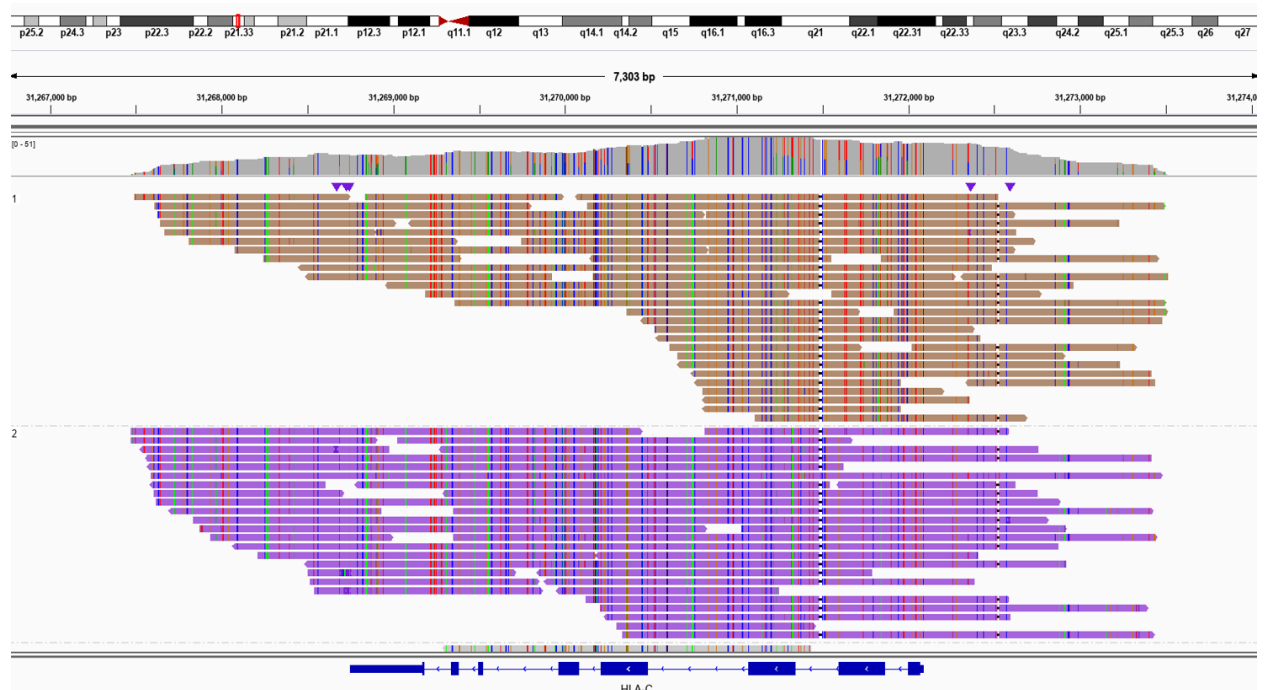


Figura 5: le letture lunghe mirate aiutano a risolvere gli aplotipi nei geni polimorfici. Grafici IGV del sequenziamento a lettura lunga con Illumina Complete Long Read Prep with Enrichment, Human. (A) Determinazione delle fasi su una regione di 722 kb nel locus MHC. Una regione da 580 kb (rosa) è incapsulata in un blocco di fase. (B) Si effettua la determinazione completa delle fasi del gene *HLA-C*. Le letture sono separate per aplotipo.

Bibliografia

1. Mehio R, Ruehle M, Catreux S, et al. DRAGEN Wins at Precision- FDA Truth Challenge V2 Showcase Accuracy Gains from Alt-aware Mapping and Graph Reference Genomes. illumina.com/science/genomics-research/dragen-wins-precisionfda-challenge-showcase-accuracy-gains.html. Pubblicato nel 2020. Consultato il 12 gennaio 2023.
2. Illumina. Accuracy improvements in germline small variant calling with the DRAGEN Bio-IT Platform. illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/dragen-v4-accuracy-app-note-m-gl-01016/dragen-v4-accuracy-app-note-m-gl-01016.pdf. Pubblicato nel 2022. Consultato il 1° febbraio 2023.
3. Pacific Biosciences. Preparing DNA for PacBio HiFi sequencing—Extraction and quality control. pacb.com/literature/technical-note-preparing-dna-for-pacbio-hifi-sequencing-extraction-and-quality-control/. Pubblicato nel 2022. Consultato il 12 gennaio 2023.
4. Pacific Biosciences. Preparing whole genome and metagenome libraries using SMRTbell prep kit 3.0. pacb.com/wp-content/uploads/Procedure-checklist-Preparing-whole-genome-and-metagenome-libraries-using-SMRTbell-prep-kit-3.0.pdf. Pubblicato nel 2022. Consultato il 5 ottobre 2023.
5. Oxford Nanopore Technologies. Ligation Sequencing Kit. <https://store.nanoporetech.com/us/productDetail/?id=ligation-sequencing-kit-v14>. Consultato il 5 ottobre 2023.
6. Pacific Biosciences. Low Yield Troubleshooting Guide. pacb.com/wp-content/uploads/Guide-Low-Yield-Troubleshooting.pdf. Pubblicato nel 2018. Consultato il 12 gennaio 2023.
7. Illumina. Scheda tecnica di Illumina Complete Long Read Prep, Human. illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/illumina-long-read-prep-human-data-sheet-m-gl-01420/illumina-long-read-prep-data-sheet-m-gl-01420.pdf. Pubblicato nel 2022. Consultato il 22 settembre 2023.
8. Illumina. Comprehensive whole-genome sequencing with Illumina Complete Long Read Prep, Human technical note. illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/illumina-long-read-prep-human-tech-note-m-gl-01421/ilmn-long-read-hu-tech-note-m-gl-01421.pdf. Pubblicato nel 2022. Consultato il 22 settembre 2023.
9. 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.
10. Wagner J, Olson ND, Harris L, et al. Curated variation benchmarks for challenging medically relevant autosomal genes. *Nat Biotechnol.* 2022;40(5):672-680. doi:10.1038/s41587-021-01158-1.
11. PharmGKB. VIP: Very Important Pharmacogenes. pharmgkb.org/vips. Consultato il 22 settembre 2023.
12. National Library of Medicine. GTR: Genetic Testing Registry. Precision HealthPGx Panel (25 Genes). ncbi.nlm.nih.gov/gtr/tests/593428/. Aggiornato il 29 novembre 2022. Consultato il 22 settembre 2023.
13. Pratt VM, Everts RE, Aggarwal P, et al. Characterization of 137 Genomic DNA Reference Materials for 28 Pharmacogenetic Genes: A GeT-RM Collaborative Project. *J Mol Diagn.* 2016;18(1):109-123. doi:10.1016/j.jmoldx.2015.08.005.
14. Miller DT, Lee K, Abul-Husn NS, et al. ACMG SF v3.1 list for reporting of secondary findings in clinical exome and genome sequencing: A policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med.* 2022;24(7):1407-1414. doi:10.1016/j.gim.2022.04.006.
15. Kulski JK, Suzuki S, Shiina T. Human leukocyte antigen super-locus: nexus of genomic supergenes, SNPs, indels, transcripts, and haplotypes. *Hum Genome Var.* 2022;9(1):49. doi:10.1038/s41439-022-00226-5.
16. 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 Custom Enrichment Panel v2 (32 µl, 120 bp)	20073953
Illumina Custom Enrichment Panel v2 (384 µl, 120 bp)	20073952
Illumina Custom Enrichment Panel v2 (1536 µl, 120 bp)	20111339

Informazioni per gli ordini

Prodotto	N. di catalogo
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-02188 ITA v1.0