

TruSeq™ DNA Exome

Una soluzione per la preparazione delle librerie e l'arricchimento degli esomi efficace in termini di costi con eccezionale accuratezza.

Punti principali

- Qualità dei dati TruSeq dimostrata**
 La frammentazione (shearing) meccanica e la chimica di arricchimento TruSeq forniscono una copertura uniforme e $\geq 80\%$ di letture di sequenziamento sul target
- Sequenziamento dell'esoma efficace in termini di costi**
 Il raggruppamento in pool delle librerie pre-arricchimento e la copertura ottimale consentono il sequenziamento dell'esoma a costi contenuti
- Risultati affidabili e accurati**
 L'analisi dei dati premendo un pulsante identifica le varianti esoniche con sicurezza
- Soluzione integrata di flusso di lavoro**
 Il flusso di lavoro completo ottimizza il sequenziamento dell'esoma dalla preparazione delle librerie all'analisi dei dati

Introduzione

La comunità scientifica ha riconosciuto il sequenziamento dell'esoma come metodo potente per la scoperta di potenziali varianti causative per malattie genetiche.¹⁻³ TruSeq DNA Exome fornisce una soluzione per il sequenziamento dell'esoma a costi contenuti, per permettere ai ricercatori di sequenziare più esomi per studio e accelerare la ricerca. Il kit unisce la comprovata tecnologia TruSeq all'eccezionale accuratezza, anche per i campioni difficili. Come parte di un flusso di lavoro integrato che include preparazione delle librerie, arricchimento dell'esoma, sequenziamento e analisi dei dati, TruSeq DNA Exome fornisce identificazioni accurate delle varianti e permette una più profonda comprensione delle mutazioni codificanti.

Qualità dei dati TruSeq dimostrata

La capacità di ottenere identificazioni di varianti con elevata sicurezza dipende dall'accuratezza del sequenziamento ma anche dall'elevata qualità della preparazione e dell'arricchimento delle librerie. TruSeq DNA Exome è compatibile con le serie MiSeq™, NextSeq™, HiSeq™ e NovaSeq™ di sistemi di sequenziamento (Tabella 1). Questi sistemi utilizzano il sequenziamento per sintesi chimica (SBS) per generare più del 90% dei dati di sequenziamento al mondo.* Il sequenziamento per sintesi chimica (SBS) Illumina fornisce un'elevata percentuale di basi sequenziate con un punteggio qualitativo superiore a Q30. Unendo TruSeq DNA Exome con la chimica SBS, i ricercatori possono identificare il numero più elevato di varianti vere codificanti e ridurre al minimo le identificazioni false positive e false negative.

Contenuto esonico mirato

TruSeq DNA Exome è ottimizzato per fornire una copertura uniforme e specifica di 45 Mb di contenuto esonico. Il gruppo di sonde è stato progettato per arricchire 214.405 esoni (Tabella 2). Questa progettazione mirata, accoppiata con arricchimento specifico e uniforme, consente il sequenziamento completo dell'esoma e l'identificazione affidabile di varianti codificanti vere.

Tabella 1: Confronto del rendimento con TruSeq DNA Exome^a

Sistema di sequenziamento	n. di esomi per corsa a 50x	n. di esomi per corsa a 100x
Serie MiSeq	1	N/A
Serie NextSeq		
Cella a flusso a output medio	3	2
Cella a flusso a output elevato	12	6
Serie HiSeq		
Sistema HiSeq 2500, modalità Rapid Run (Corsa rapida) (doppia cella a flusso)	24	12
Sistema HiSeq 2500, modalità High Output (Output elevato) (doppia cella a flusso)	156	78
Sistema HiSeq 3000	96	48
Sistema HiSeq 4000 (doppia cella a flusso)	192	96

a. Il numero di esomi stimati sequenziati per corsa è calcolato con una copertura media di 50x e 100x, rispettivamente. Illumina raccomanda una lunghezza di lettura di 2×75 bp su tutti i sequenziatori quando si utilizza TruSeq DNA Exome.

Tabella 2: Contenuto esonico con TruSeq DNA Exome e Nextera™ Exome

Specifiche della copertura	TruSeq DNA Exome o Nextera Exome
Dimensione target	45 Mb
N. di esoni target	214.405
Contenuto target	esoni codificanti
Percentuale di copertura dell'esoma (per database)	
RefSeq ^a	99,45%
CCDS ^b	98,83%
ENSEMBL ^c	99,68%
GENCODE v19 ^d	99,68%

- RefSeq - database delle sequenze di riferimento NCBI. www.ncbi.nlm.nih.gov/refseq/. Consultato l'11 febbraio, 2015.
- CCDS - database della consensus CDS (CCDS). www.ncbi.nlm.nih.gov/projects/CCDS/CcidsBrowse.cgi. Consultato l'11 febbraio, 2015.
- ENSEMBL - Ensembl Genome Browser. www.ensembl.org/index.html. Consultato l'11 febbraio, 2015.
- GENCODE - Progetto GENCODE: enciclopedia dei geni e delle varianti geniche. www.gencodegenes.org/. Consultato l'11 febbraio, 2015.

* Calcoli dei dati in archivio. Illumina, Inc., 2015.

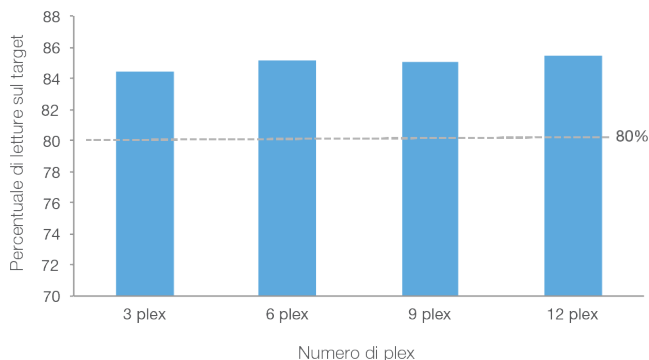


Figura 1: Arricchimento sul target: TruSeq DNA Exome fornisce il $\geq 80\%$ delle letture di sequenziamento sul target per un sequenziamento dell'esoma efficiente ed efficace in termini di costi.

Sequenziamento dell'esoma efficace

TruSeq DNA Exome supporta il raggruppamento in pool di 12 plex, per permettere ai ricercatori di massimizzare il rendimento del sequenziamento e identificare le varianti in tempi più brevi sequenziando fino a 12 librerie per corsia sulla cella a flusso. TruSeq DNA Exome fornisce $\geq 80\%$ delle letture del sequenziamento sul target (Figura 1) e una buona uniformità di copertura per risultati ad elevata affidabilità. Permette inoltre il sequenziamento di più esomi per corsia, per permettere ai ricercatori di massimizzare i budget disponibili.

Chimica di arricchimento efficiente

TruSeq DNA Exome è ottimizzato per offrire prestazioni coerenti per molti diversi tipi di campione. Il flusso di lavoro per la preparazione delle librerie (Figura 2) inizia con la frammentazione meccanica, generando dimensioni di frammenti uniformi per ottenere la massima riproducibilità tra le librerie (Figura 2A). Questa frammentazione (shearing) meccanica eseguita mediante l'ultrasonificazione Covaris, o un metodo simile, supporta l'utilizzo di campioni compromessi, come i campioni che contengono frammenti brevi di DNA. I frammenti di DNA con estremità tronche vengono generati utilizzando una combinazione di reazioni di riempimento e di attività esonucleasica, seguita dalla selezione della dimensione con microsfere di immobilizzazione reversibile in fase solida (SPRI) AMPure (Beckman Coulter). Gli adattatori che contengono il complemento completo di siti di ibridazione dei primer di sequenziamento per letture unidirezionali, paired-end e con indicizzazione sono legati ai frammenti (Figura 2B). I prodotti sottoposti a ligazione sono amplificati mediante PCR (Figura 2C).

Quindi le librerie vengono raggruppate in pool e denaturate (Figura 2D). Le sonde biotinilate vengono ibridate sulle regioni target (Figura 2E e 2F), che vengono arricchite mediante microsfere coniugate con la streptavidina. Dopo un'altra reazione della PCR (Figura 2G), i frammenti vengono eluiti dalle microsfere e sono pronti per il sequenziamento (Figura 2H). Questo flusso di lavoro ottimizzato genera una dimensione di inserto target di circa ~150 bp e può essere completato in meno di < 2,5 giorni. TruSeq DNA Exome arricchisce in modo efficiente il contenuto dell'esoma per il sequenziamento,

fornendo elevata uniformità di copertura con più dell'85% di basi coperte a una profondità di 10x (Figura 3).

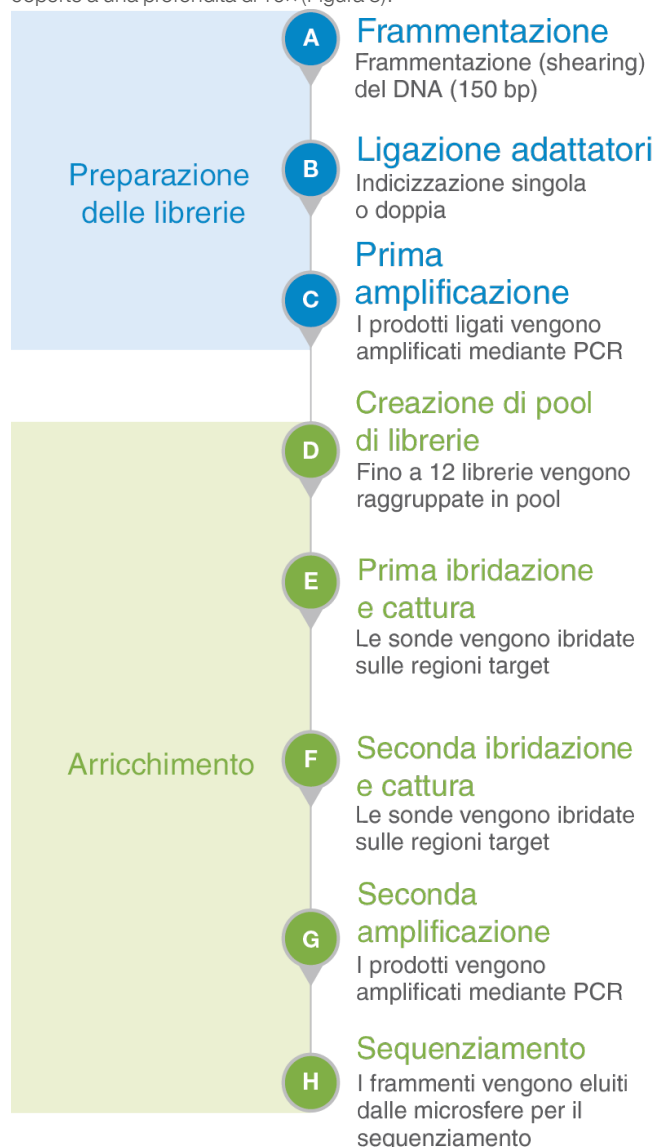


Figura 2: Flusso di lavoro TruSeq DNA Exome: TruSeq DNA Exome unisce la preparazione delle librerie con l'arricchimento dell'esoma e può essere completato in meno di 2,5 giorni.

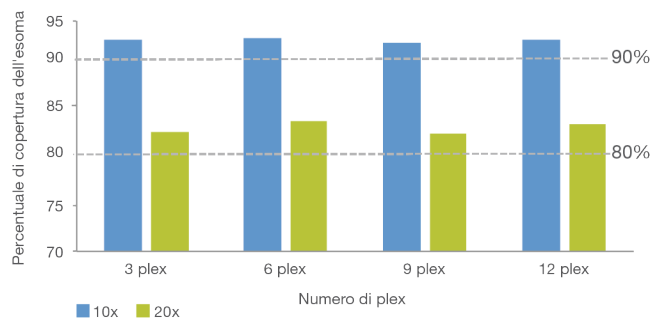


Figura 3: Elevata uniformità di copertura: TruSeq DNA Exome fornisce una copertura uniforme, con più dell'85% coperte a 10x di profondità.

Risultati affidabili e accurati

TruSeq DNA Exome fornisce copertura target eccellente su un ampio intervallo di profondità di letture (Figura 4). La riproducibilità del saggio TruSeq DNA Exome, l'elevata uniformità di copertura e la chimica SBS forniscono identificazioni delle varianti altamente accurate. Più del 99,65% delle identificazioni delle varianti ottenute in seguito a sequenziamento TruSeq DNA Exome e Illumina coincide con i dati di riferimento standard contenuti nel database del National Institute of Standards and Technology (NIST) (Figura 5).^{4,5}

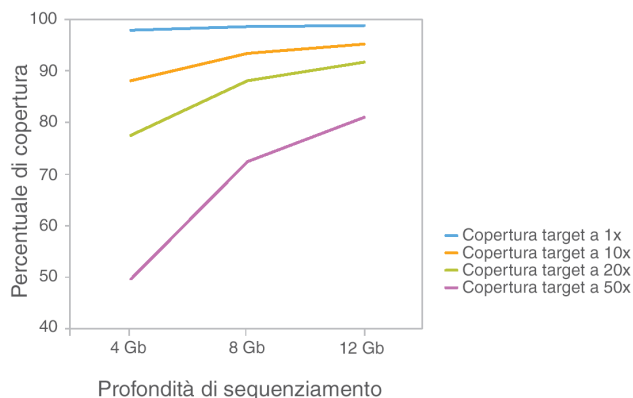


Figura 4: Efficacia della copertura a varie profondità: TruSeq DNA Exome garantisce una copertura eccezionale a diverse profondità di sequenziamento, con una copertura di > 80% dei target fino a una profondità di 20x.

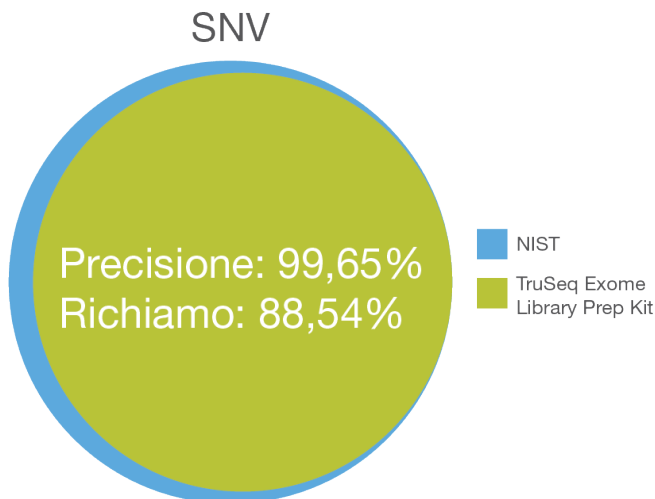


Figura 5: Elevata correlazione con il database NIST: le identificazioni delle varianti ottenute con TruSeq DNA Exome dimostrano l'elevata concordanza con i dati di riferimento standard. Il campione di DNA NA12878 del Centre de' Etude du Polymorphism Humain (CEPH) è stato sequenziato a una profondità di copertura di 100x. Sono riportate le identificazioni delle varianti a singolo nucleotide (single nucleotide variant, SNV). **La precisione** è la probabilità che una variante identificata sia accurata. **Il richiamo** è la probabilità di identificare una variante convalidata.

Flusso di lavoro di sequenziamento integrato

TruSeq DNA Exome fa parte di una soluzione compatta e comprovata che guida i ricercatori dalla preparazione delle librerie all'analisi dei dati (Figura 7). Il kit unisce la preparazione delle librerie e l'arricchimento dell'esoma, eliminando la necessità di acquistare indici, microsferi per la purificazione dei campioni e altro materiale ausiliario. Tutti i componenti di TruSeq DNA Exome sono progettati, ottimizzati e convalidati assieme analiticamente, eliminando la necessità di valutare componenti multipli e disparati. Gli esperti scienziati Illumina forniscono una singola fonte di supporto tecnico e sul campo per ciascuna fase del flusso di lavoro. Unendosi alla comunità Illumina, i ricercatori possono sfruttare l'esperienza del team di supporto Illumina e collaborare con un'ampia rete di scienziati che utilizzano la tecnologia Illumina.

I dati di sequenziamento vengono trasferiti automaticamente dai sistemi Illumina a BaseSpace® Sequence Hub, l'ambiente di calcolo genomico Illumina. BaseSpace Sequence Hub rimuove molte delle complessità create da un tipico flusso di lavoro di analisi, semplificando l'analisi dei dati e l'interpretazione biologica. BaseSpace Sequence Hub offre un ecosistema consolidato di strumenti integrati per l'analisi dei dati progettati per i biologi. Con le applicazioni BaseSpace, gli strumenti di analisi preferiti dagli esperti sono raccolti in un'interfaccia utente intuitiva e di facile utilizzo, per permettere a qualsiasi ricercatore di accedere a un gruppo di strumenti di analisi affidabile senza la necessità di avere precedente esperienza in bioinformatica (Figura 6). I ricercatori possono scegliere di analizzare i dati dell'esoma utilizzando l'applicazione BWA Enrichment, che utilizza il metodo BWA/GATK standard nel settore, oppure l'applicazione Isaac™ Enrichment, che utilizza un gruppo di strumenti Illumina veloce e accurato.⁶

Per i biologi che studiano la base genetica della malattia, l'applicazione VariantStudio permette l'identificazione e l'interpretazione funzionale di varianti di singolo nucleotide (single nucleotide variant, SNV), inserzioni e delezioni (Indel) associate alla malattia. I ricercatori possono filtrare e isolare rapidamente varianti consequenziali per arricchire i dati di sequenziamento con contesto biologico. I risultati significativi vengono esportati in report concisi. L'applicazione VariantStudio permette ai ricercatori di esplorare il significato biologico in pochi semplici passaggi.

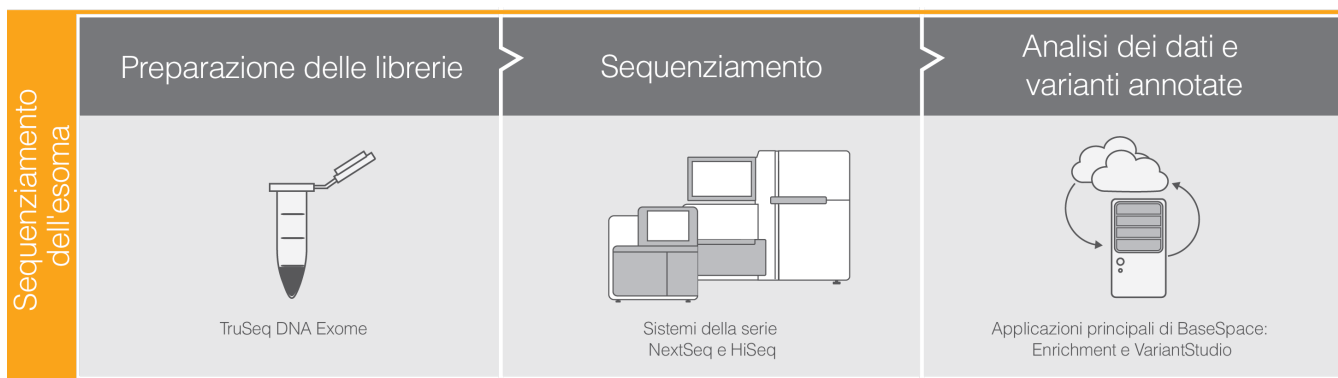


Figura 7: Flusso di lavoro del sequenziamento dell'esoma: TruSeq DNA Exome fa parte di un flusso di lavoro del sequenziamento dell'esoma integrato che include la preparazione delle librerie, il sequenziamento e l'analisi dei dati.

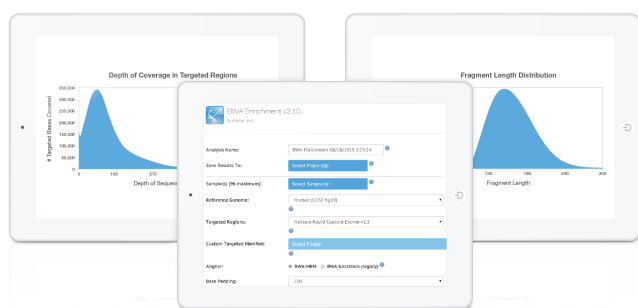


Figura 6: Analisi semplificata dei dati con le applicazioni BaseSpace: i dati di sequenziamento di TruSeq DNA Exome possono essere caricati facilmente e senza rischi su BaseSpace Sequence Hub e analizzati con l'applicazione BWA Enrichment. I risultati sono forniti in formati di facile lettura.

Confronto delle prestazioni per il sequenziamento dell'esoma

Ilumina offre due soluzioni integrate di flusso di lavoro per il sequenziamento dell'esoma. Sono disponibili inoltre flussi di lavoro che combinano la preparazione delle librerie Ilumina con TruSeq DNA Exome o Nextera DNA Exome, con successivo arricchimento dell'esoma con xGen® Universal Blockers, xGen Lockdown Reagents e xGen Exome Research Panel v1.0, disponibili grazie a IDT (Tabella 3).

Tabella 3: Confronto delle prestazioni per il flusso di lavoro dell'esoma

Metrica	TruSeq-xGen ^a	Nextera-xGen ^a	TruSeq Exome	Nextera Exome
Input di DNA	100 ng	50 ng	100 ng	50 ng
Tipi di campioni	DNA	DNA	DNA e FFPE	DNA
Durata interventi manuali	5 ore	2 ore	6 ore	3 ore
Durata totale del saggio	2,5 giorni	2 giorni	2,5 giorni	2 giorni
Durata dell'ibridazione	4 ore	4 ore	16 ore	2 ore
% sul target	> 91%	> 92%	> 80%	> 75%
% di copertura a 20x ^b	> 95%	> 85%	> 90%	> 85%

- a. Le specifiche del flusso di lavoro di arricchimento dell'esoma Ilumina-IDT si basano su dati preliminari registrati su BaseSpace Sequence Hub.
- b. La percentuale di copertura a 20x è stata determinata per i kit TruSeq-xGen e Nextera-xGen con un sequenziamento di 3,5 Gb. La percentuale di copertura a 20x è stata determinata per i kit TruSeq DNA Exome e Nextera DNA Exome con un sequenziamento di 8 Gb.

Riepilogo

TruSeq DNA Exome offre un metodo ottimizzato ed efficace in termini di costi per identificare e comprendere le varianti codificanti con un'accuratezza dei dati eccellente. L'integrazione con un flusso di lavoro completo composto da tecnologia di sequenziamento leader nel settore e da strumenti di analisi di facile utilizzo permette ai ricercatori di accedere a una singola fonte per tutte le necessità di sequenziamento dell'esoma.

Maggiori informazioni

Per maggiori informazioni sul sequenziamento dell'esoma, visitate la pagina Web www.illumina.com/techniques/sequencing/dna-sequencing/targeted-resequencing/exome-sequencing.html.

Ordering Information

Product	Catalog No.
TruSeq Exome Kit (24 samples)	20020614
TruSeq Exome Kit (96 samples)	20020615
IDT for Ilumina – TruSeq DNA UD Indexes (24 indexes, 96 samples)	20020590
IDT for Ilumina – TruSeq DNA UD Indexes (96 indexes, 96 samples)	20022370

Bibliografia

1. Litchfield K, Summersgill B, Yost S, et al. Whole-exome sequencing reveals the mutational spectrum of testicular germ cell tumours. *Nat Commun.* 2015;6:5973.
2. Srivastava S, Cohen JS, Vernon H, et al. Clinical whole exome sequencing in child neurology practice. *Ann Neurol.* 2014;76:473–483.
3. Worthey EA, Mayer AN, Syverson GD, et al. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genet Med.* 2011;13:255–262.
4. Standard Reference Data (www.nist.gov/srd). Consultato l'11 febbraio 2015.
5. Genome in a Bottle Consortium | Advances in Biological and Medical Measurement Science (sites.stanford.edu/abms/giab). Consultato il 20 febbraio 2015.
6. Racz C, Petrovski R, Saunders CT, et al. Isaac: ultrafast whole-genome secondary analysis on Illumina sequencing platforms. *Bioinformatics.* 2013;29:2041–2043.