

Illumina Human Comprehensive Panel

Cobertura aprimorada de regiões desafiadoras em genes codificadores de proteínas com o Illumina Complete Long Read Prep with Enrichment, Human



Introdução

Ao realizar o sequenciamento do genoma completo (WGS) humano, pode haver uma pequena fração de regiões gênicas que são difíceis de mapear com leituras curtas isoladamente. As regiões desafiadoras incluem regiões altamente homólogas ou repetitivas, variantes estruturais complexas, pseudogenes e grandes deleções de inserção (indels). Nesses casos, o sequenciamento de leitura longa pode complementar os dados WGS de leitura curta padrão para ajudar a fornecer cobertura aprimorada para resolver regiões desafiadoras e facilitar o faseamento de variantes e a identificação de haplótipos.

A tecnologia Illumina Complete Long Reads usa um fluxo de trabalho padrão de sequenciamento de última geração (NGS) para gerar seqüências de leitura longa contíguas em sistemas de sequenciamento Illumina (Figura 1). O Illumina Complete Long Read Prep with Enrichment e o Human Comprehensive Panel oferecem uma abordagem econômica e escalável para ajudar os pesquisadores a resolver regiões desafiadoras conhecidas do genoma. Leituras longas direcionadas são analisadas com dados WGS padrão de alta precisão para permitir o mapeamento de mais regiões.

Painel de sondas abrangente e otimizado

O Illumina Human Comprehensive Panel é um painel de sonda de captura híbrida otimizado—projetado para abordar regiões gênicas que podem se beneficiar da maior capacidade de mapeamento de leituras mais longas. O painel tem como alvo regiões de baixa mapeabilidade em > 6.500 genes codificadores de proteínas (Tabela 1)¹ para permitir leituras longas complementares que melhoram a resolução de áreas que são difíceis de mapear usando métodos padrão de leitura curta de WGS.

Tabela 1: Parâmetros do Human Comprehensive Panel¹

Sistema	NovaSeq X Series ou NovaSeq 6000 System
Tamanho da região alvo	> 95 Mb
Saída de sequenciamento por amostra ^b	90–120 Gb
Nº de sondas	~40 mil
Nº de genes alvo	> 6.500
Tipo de amostra	DNA genômico
Faixa de entrada de DNA recomendada	50 ng
Tempo total de preparação de biblioteca	~2 dias
Tempo de trabalho	~6 h
Multiplexação	Até 64 amostras
N50	6,3 kb
Bloco de fase N50	15,6 kb
Uniformidade	95%
Enriquecimento de leitura “padded” (PRE) ^c	83%
% de SNV heterozigótica em fases ^d	98%

- Dados gerados usando 50 ng de DNA genômico de HG002 (Coriell, nº de catálogo NA24385). O desempenho pode variar com a entrada de DNA e a qualidade da amostra.
- Requer uma corrida de sequenciamento de 2 × 150 bp, gerando aproximadamente 30× de cobertura final do Illumina Complete Long Reads.
- Uniformidade da cobertura calculada como % > 0,2 * média.
- PRE calculado como 100 * (leituras alinhadas de alvo “padded”/leituras alinhadas totais).
- Variantes de nucleotídeo único (SNV)



Figura 1: Parte de um fluxo de trabalho integrado: acesse dados WGS de leitura longa direcionados e econômicos usando uma preparação de biblioteca escalável e otimizada com protocolo de enriquecimento, química comprovada de sequenciamento Illumina e análise secundária DRAGEN. Requer ≥ 30× dados padrão de leitura curta de WGS da mesma amostra para análise. Os arquivos FASTQ da amostra de uma corrida anterior podem ser usados.

Ao desenvolver esse painel, a Illumina considerou o conjunto completo de mais de 20.000 genes codificadores de proteínas, incluindo íntrons e regiões não traduzidas (UTRs), e almejava 70% de todos os pares de bases difíceis de mapear nessas regiões. Os genes que são amplamente cobertos por leituras curtas isoladamente foram excluídos. Este painel foi testado em laboratório úmido e otimizado para maximizar a eficiência e o desempenho do projeto e do sequenciamento.

Sequenciamento e análise altamente escaláveis

O fluxo de trabalho do Illumina Complete Long Read Prep with Enrichment é altamente escalonável e fácil de automatizar para oferecer suporte a WGS abrangente para um número maior de amostras (Tabela 2). Com o NovaSeq™ X Plus System, os usuários podem gerar até 15.000 genomas de alta precisão por ano.* Para reduzir os requisitos de lotes, sequencie números de amostras menores em materiais de consumo de menor produtividade, como a lâmina de fluxo NovaSeq X 1.5B. O Illumina Complete Long Read Prep with Enrichment e o Human Comprehensive Panel podem ser usados para aumentar os conjuntos de dados WGS existentes como uma ferramenta reflexa para detecção mais ampla de variantes.

* Rendimento potencial quando o Illumina Complete Long Read Prep with Enrichment e o Human Comprehensive Panel são usados com o NovaSeq X Plus System, corrida de lâmina de fluxo duplo com lâminas de fluxo 25B.

A análise de dados do Illumina Complete Long Read Prep with Enrichment, Human está disponível como um aplicativo BaseSpace™ Sequence Hub ou por meio do Illumina Connected Analytics. Os dados são transmitidos para a nuvem e analisados juntamente com dados padrão de leitura curta $\geq 30\times$ WGS da mesma amostra (gerados anteriormente ou em paralelo). O pipeline do DRAGEN™ mescla resultados em um único conjunto de arquivos de saída, incluindo chamadores direcionados ao DRAGEN.³

Cobertura aprimorada e econômica

O Illumina Complete Long Read Prep with Enrichment e o Human Comprehensive Panel melhoram a precisão da identificação de variantes e melhoram a cobertura nas regiões-alvo, em comparação com o WGS de leitura curta padrão (Figura 2, Figura 3, Tabela 3). Leituras longas direcionadas usando o Human Comprehensive Panel fornecem precisão abrangente do genoma completo com uma pontuação F1 (SNVs + indels) de 99,87%,² uma melhora substancial em relação a leituras curtas isoladamente. Isso fornece desempenho comparável ao do Illumina Complete Long Read Prep, genomas humanos e outros genomas inteiros de leitura longa em genes codificadores de proteínas, mas com menor custo e maior rendimento (Tabela 3).

Tabela 2: Rendimento recomendado da amostra para gerar cobertura final de 30 \times para o Human Comprehensive Panel e Illumina Complete Long Read Prep with Enrichment, Human^{a,b,c,d}

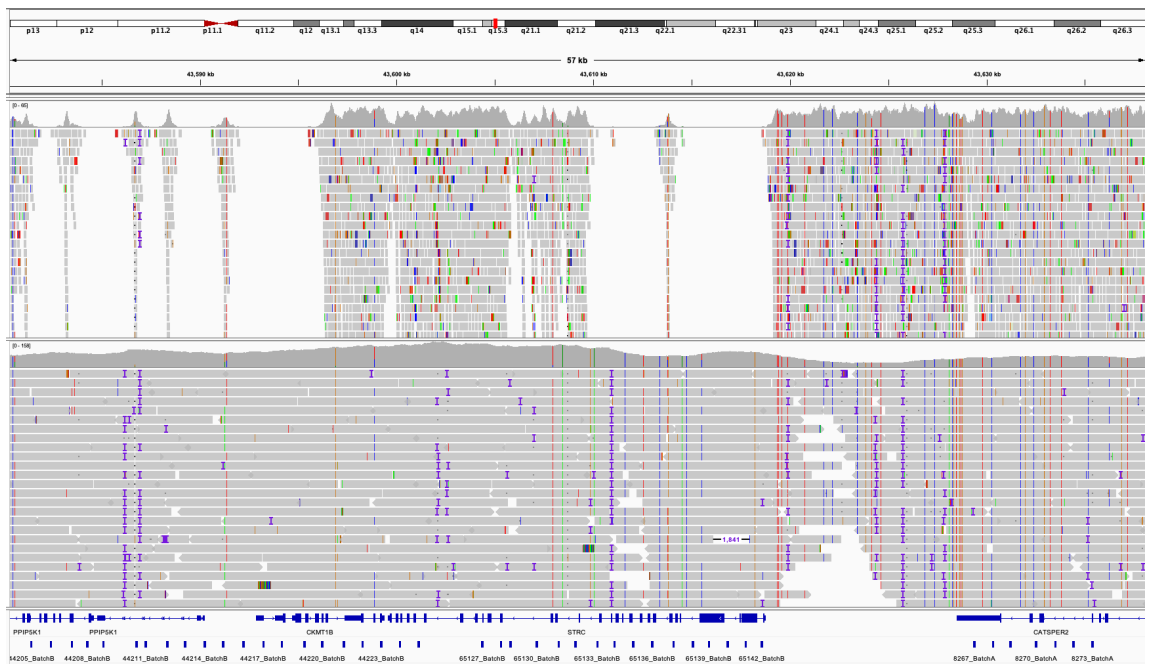
	NovaSeq 6000				NovaSeq X		
Kits de reagentes de 300 ciclos	SP	S1	S2	S4	1.5B	10B	25B
Amostras por lâmina de fluxo	2	4	10	24	4	24	64
Saída por lâmina de fluxo	~250 Gb	~500 Gb	~1,25 Tb	~3 Tb	~500 Gb	~3 Tb	~8 Tb
Tempo de corrida	~25 h	~25 h	~36 h	~44 h	~21 h	~25 h	~48 h

- O tamanho da região-alvo do Human Comprehensive Panel é > 95 Mb e requer uma saída de sequenciamento de 90 a 120 Gb por amostra.
- Requer uma corrida de sequenciamento de 2 \times 150 bp, gerando aproximadamente 30 \times de cobertura final do Illumina Complete Long Reads.
- Requer 30 \times dados padrão de leitura curta do genoma completo humano da mesma amostra para análise. Recomenda-se o Illumina DNA PCR-Free Prep. Kits WGS de terceiros também são compatíveis. A biblioteca não marcada não precisa ser preparada nem sequenciada em paralelo; podem ser usados arquivos FASTQ da amostra de uma corrida anterior.
- Sequenciar bibliotecas de leitura longa completas da Illumina nas plataformas NovaSeq pode fazer com que a pontuação Q30 relatada em uma corrida fique abaixo da especificação NovaSeq. Isso não indica um problema de desempenho com a corrida do sequenciamento nem com a biblioteca.

A.

WGS de leitura curta padrão

Illumina Complete Long Read Prep with Enrichment



B.

WGS de leitura curta padrão

Illumina Complete Long Read Prep with Enrichment

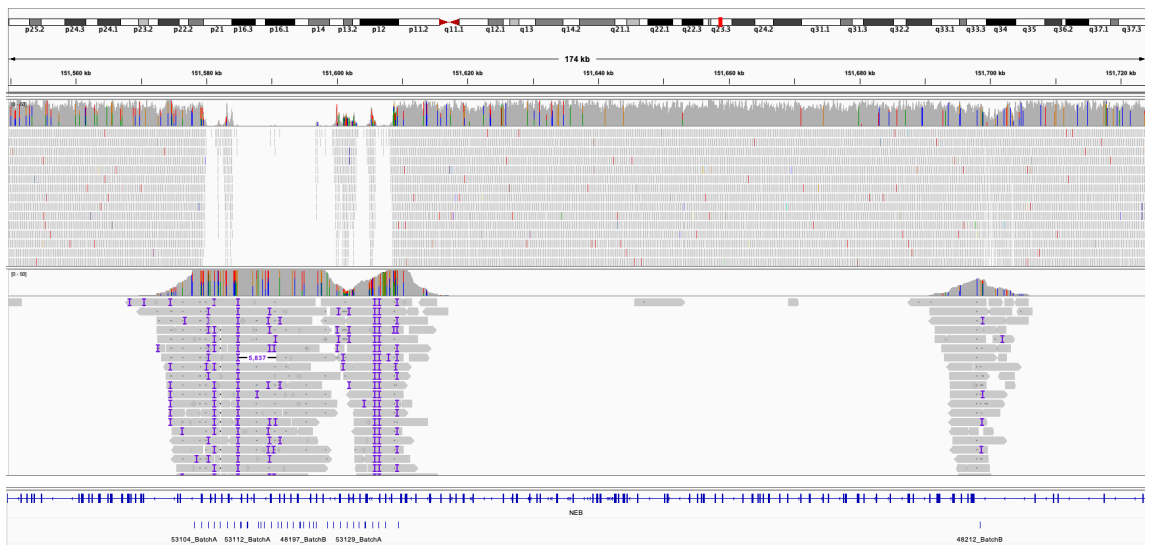


Figura 2: Resolva regiões difíceis de mapear usando leituras longas direcionadas com o Human Comprehensive Panel: o Illumina Complete Long Read Prep with Enrichment, Human e o Human Comprehensive Panel, ajudam a melhorar a cobertura em regiões gênicas desafiadoras para complementar o WGS humano de leitura curta padrão. Gráficos do Integrative Genomics Viewer (IGV) de (A) *STRC* e (B) *NEB* usando WGS de leitura curta padrão e o Illumina Complete Long Reads with Enrichment.

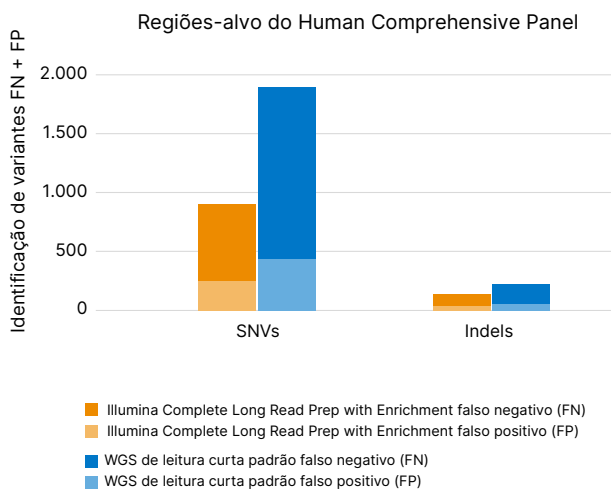


Figura 3: Leituras longas direcionadas para melhorar a precisão da identificação de variantes: identificação de variantes falso negativas (FN) mais SNV falso positivas (FP) e variantes indel em regiões gênicas direcionadas pelo Human Comprehensive Panel, usando Illumina Complete Long Read Prep com Enrichment e Human Comprehensive Panel (laranja) em comparação com WGS de leitura curta padrão (azul).

Resumo

O Illumina Human Comprehensive Panel permite uma cobertura de leitura longa de alta precisão da pequena porção de regiões gênicas que são difíceis de mapear com leituras curtas isoladamente. O Illumina Complete Long Read Prep with Enrichment e o Human Comprehensive Panel oferecem um ensaio de genoma completo otimizado e econômico que complementa o Illumina WGS e se concentra em leituras longas, onde fornecem maior valor com uma solução de fluxo de trabalho completa.

Saiba mais

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

[Human Comprehensive Panel](#)

[Tecnologia de sequenciamento de leitura longa](#)

Tabela 3: Desempenho do Human Comprehensive Panel

Precisão para regiões-alvo do Human Comprehensive Panel ^a		
	Ilumina Complete Long Read Prep with Enrichment e Human Comprehensive Panel ^b	Standard short-read WGS ^c
Pontuação F1 (SNVs)	99,54%	99,04%
Pontuação F1 (Indels)	99,47%	99,15%
Pontuação F1 (SVs)	80,43%	60,18%

Precisão para genoma completo, todas as regiões de referência ^e pontuação F1 (SNVs + indels)			
	Ilumina Complete Long Read Prep with Enrichment e Human Comprehensive Panel ^b	Ilumina Complete Long Read Prep, Human ^e	Standard short-read WGS ^c
	99,87%	99,90%	99,84%

a. Somente autossomos. SNV, variantes de nucleotídeo único; SV, variantes estruturais.
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. Conforme medido em relação a todas as soluções participantes usando os dados internos do PrecisionFDA Truth Challenge v2 Benchmark Data,² em arquivo para o DRAGEN v4.
e. Illumina Complete Long Read Prep, Human + Illumina DNA PCR-Free Prep + NovaSeq 6000 System + DRAGEN v4.2.

Referências

- 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](https://www.illumina.com/science/genomics-research/articles/identifying-genomic-regions-with-high-quality-single-nucleotide-.html). Publicado em 2021. Acessado em 30 de agosto de 2023.
- PrecisionFDA. Truth Challenge V2: Calling Variants from Short and Long Reads in Difficult-to-Map Regions. precision.fda.gov/challenges/10. Acessado em 2 de outubro de 2023.
- Roessler K. Illumina Complete Long Reads software analysis workflow for human WGS. [illumina.com/science/genomics-research/articles/complete-long-read-software-analysis.html](https://www.illumina.com/science/genomics-research/articles/complete-long-read-software-analysis.html). Publicado em 2023. Acessado em 22 de setembro de 2023.

Informações para pedido

Produto	Nº do catálogo
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



1 800-809-4566, ligação gratuita (EUA) | tel. +1 (858) 202-4566
techsupport@illumina.com | www.illumina.com

© 2024 Illumina, Inc. Todos os direitos reservados. Todas as marcas comerciais pertencem à Illumina, Inc. ou aos respectivos proprietários. Para obter informações específicas sobre marcas comerciais, consulte www.illumina.com/company/legal.html.
M-GL-02191 PTB v1.0