

Illumina Human Comprehensive Panel

搭配 Illumina Complete Long
Read Prep with Enrichment,
Human 增加对蛋白编码基因
难测区域的覆盖度

简介

在进行人类全基因组测序 (WGS) 时, 可能存在一小部分基因区域, 仅使用短读长难以绘制。这些复杂区域包括高度同源或重复区域、复杂的结构变异、假基因和大片段插入缺失 (indel)。在这些情况下, 长读长测序可以作为标准短读长 WGS 数据的补充, 为复杂区域提供更高的覆盖度, 从而更加容易的检出相位变异和单倍型。

Illumina Complete Long Reads 技术采用标准的新一代测序 (NGS) 工作流程, 可以在因美纳测序仪上生成连续的长读长序列 (图 1)。Illumina Complete Long Read prep with Enrichment 和 Human Comprehensive Panel 提供了一种经济高效、可扩展的方法, 帮助研究人员解析复杂基因组区域。靶向长读长与高准确性的标准短读长 WGS 数据相结合, 以便绘制更多区域。

全面且优化的探针 panel

Illumina Human Comprehensive Panel 是一种经过优化的杂交捕获探针 panel, 主要针对采用长读长可提高绘制率的基因区域。该 panel 靶向超过 6,500 个蛋白编码基因 (表 1)¹ 的低绘制率区域, 以实现互补的长读长测序, 从而提高标准短读长 WGS 方法难以绘制的区域的分辨率。

表 1: Human Comprehensive Panel 参数^a

测序仪	NovaSeq X 系列 NovaSeq 6000 基因测序仪
目标区域大小	> 95 Mb
每个样本的测序产出 ^b	90–120 Gb
探针数量	约 40K
靶向基因的数量	> 6,500
样本类型	基因组 DNA
建议 DNA 起始量	50 ng
文库制备总时间	约 2 天
手动操作时间	约 6 小时
多重分析	多达 64 个样本
N50	6.3 kb
相区块 N50	15.6 kb
均一性	95%
侧翼 read 富集 (PRE) ^c	83%
定相的杂合子 SNV 百分比 ^d	98%

- a. 使用 50 ng HG002 基因组 DNA (Coriell, 货号 NA24385) 生成的数据。性能可能因 DNA 起始量和样本质量的不同而有所不同。
 b. 需要 2x150 bp 测序运行, 生成大约 30x 最终覆盖度的 Illumina Complete Long Reads。
 c. 覆盖均一性按 $> 0.2 * \text{平均值的百分比}$ 来计算。
 d. PRE 按 $100 * \text{计算} (\text{填补目标匹配 reads} / \text{总匹配 reads})$ 。
 e. SNV, 单核苷酸位点变异。

提取 DNA



文库制备和富集



测序



数据分析



无需专门的实验方案,
无需剪切,
无需片段长度筛选



Illumina Complete Long
Read Prep with
Enrichment 和 Human
Comprehensive Panel



NovaSeq X 系列或
NovaSeq 6000
基因测序仪



Illumina Complete Long
Read with Enrichment
DRAGEN 云端分析

图 1: 集成工作流程的一部分——使用可扩展、经过优化的文库制备与富集方案、成熟的因美纳测序化学技术和 DRAGEN 二级分析, 获取经济有效的靶向长读长 WGS 数据。需要来自同一样本的 $\geq 30 \times$ 标准短读长 WGS 数据用于分析。可以使用之前运行的样本中的 FASTQ 文件。

在开发该 panel 时，因美纳筛选了包括内含子和非翻译区 (UTR) 在内的所有 20,000 多个蛋白质编码基因，并靶向这些区域中所有难绘制碱基对的 70%。其中排除了仅通过短读长就能全面覆盖的基因。该 panel 已经过湿实验室测试和优化，旨在尽可能提高设计和测序的效率及性能。

高度可扩展的测序和分析

Illumina Complete Long Read Prep with Enrichment 工作流程可扩展性强，易实现自动化，可支持更多样本的全面 WGS (表 2)。在 NovaSeq™ X Plus 基因测序仪上，用户每年可以生成多达 15,000 个高准确度基因组。^{*} 为了降低分批处理要求，可使用 Novaseq X 1.5B 流动槽等低通量耗材对少量样本进行测序。Illumina Complete Long Read Prep with Enrichment 和 Human Comprehensive Panel 可用于扩展现有的 WGS 数据集，或作为反馈检测工具更广泛地检测变异。

^{*} Illumina Complete Long Read prep with Enrichment 和 Human Comprehensive Panel 与 NovaSeq X Plus 测序仪 (使用 25B 流动槽运行双流动槽) 联用时的潜在通量。

Illumina Complete Long Read Prep with Enrichment, Human 数据分析可在 Basespace™ Sequence Hub App 中进行或通过 Illumina Connected analytics 进行。数据会传输到云端，与来自同一样本的标准短读长 $\geq 30\times$ WGS 数据 (先前生成或并行生成) 一起分析。DRAGEN™ 流程将结果合并为一套输出文件，包括 DRAGEN 靶向检出程序³。

经济高效且覆盖度提升

与标准短读长 WGS 相比，Illumina Complete Long Read Prep with Enrichment 和 Human Comprehensive Panel 提高了变异检出的准确性以及整个目标区域的覆盖度 (图 2、图 3、表 3)。使用 Human Comprehensive Panel 的靶向长读长可提供出色的全基因组准确性，F1 评分 (SNV+ 插入缺失) 为 99.87%²，相比于单独使用短读长有很大的提高。对于蛋白质编码基因，其性能可与 Illumina Complete Long Read Prep, Human 和其他长读长全基因组技术相媲美，但成本更低，通量更高 (表 3)。

表 2: Human Comprehensive panel 和 Illumina Complete Long Read Prep with Enrichment, Human 产生 30 \times 最终覆盖度的建议样本通量^{a,b,c,d}

	SP	NovaSeq 6000				NovaSeq X		
		S1	S2	S4	1.5B	10B	25B	
300 循环试剂盒								
每个流动槽的样本数	2	4	10	24	4	24	64	
每个流动槽的产出	约 250 Gb	约 500 Gb	约 1.25 Tb	约 3 Tb	约 500 Gb	约 3 Tb	约 8 Tb	
运行时间	约 25 小时	约 25 小时	约 36 小时	约 44 小时	约 21 小时	约 25 小时	约 48 小时	

a. Human Comprehensive Panel 目标区域大小 > 95 Mb，每个样本需要 90-120 Gb 的测序输出。

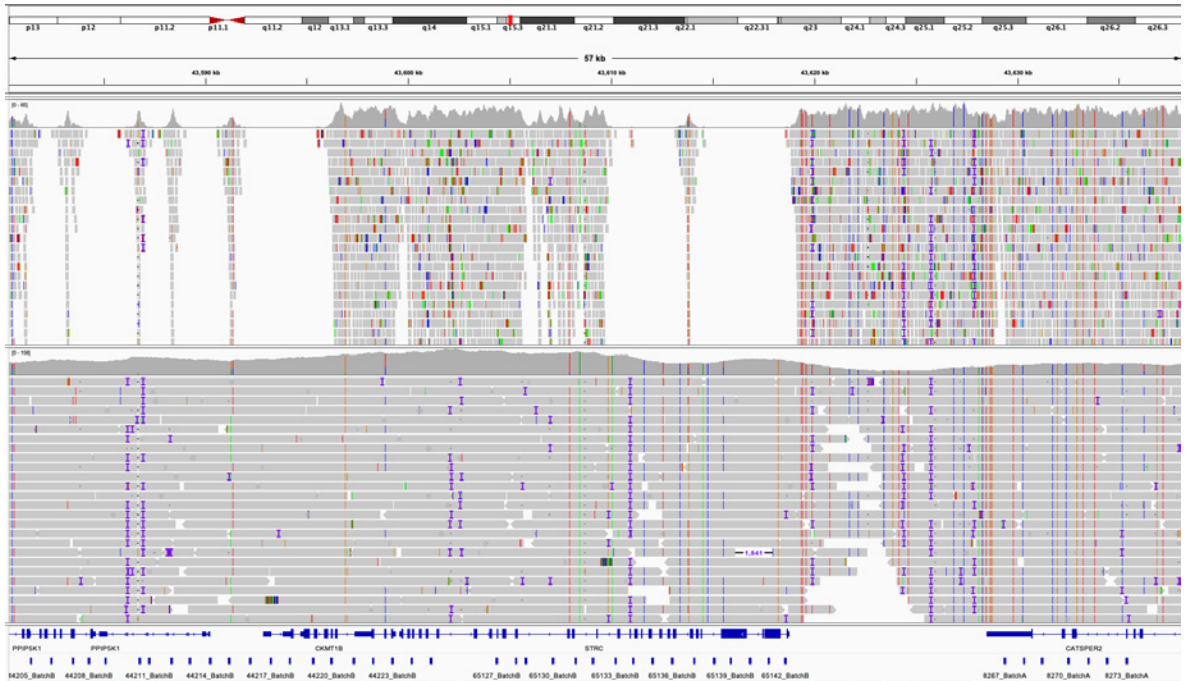
b. 需要 2 \times 150 bp 测序运行，生成大约 30 \times 最终覆盖度的 Illumina Complete Long Reads。

c. 需要来自同一样本的 30 \times 标准短读长人类基因组数据用于分析。推荐使用 [Illumina DNA PCR-Free Prep](#)，也兼容第三方 WGS 试剂盒。未标记的文库无需同时制备或测序，可以使用之前运行的样本中的 FASTQ 文件。

d. 在 NovaSeq 平台上对 Illumina Complete Long Read 文库执行测序有可能导致运行报告的 Q30 分值低于 NovaSeq 性能指标。这既不表示测序运行有性能问题，也不表示文库有问题。

A.

标准短读长
WGS



B.

标准短读长
WGS

Illumina
Complete Long
Read Prep with
Enrichment

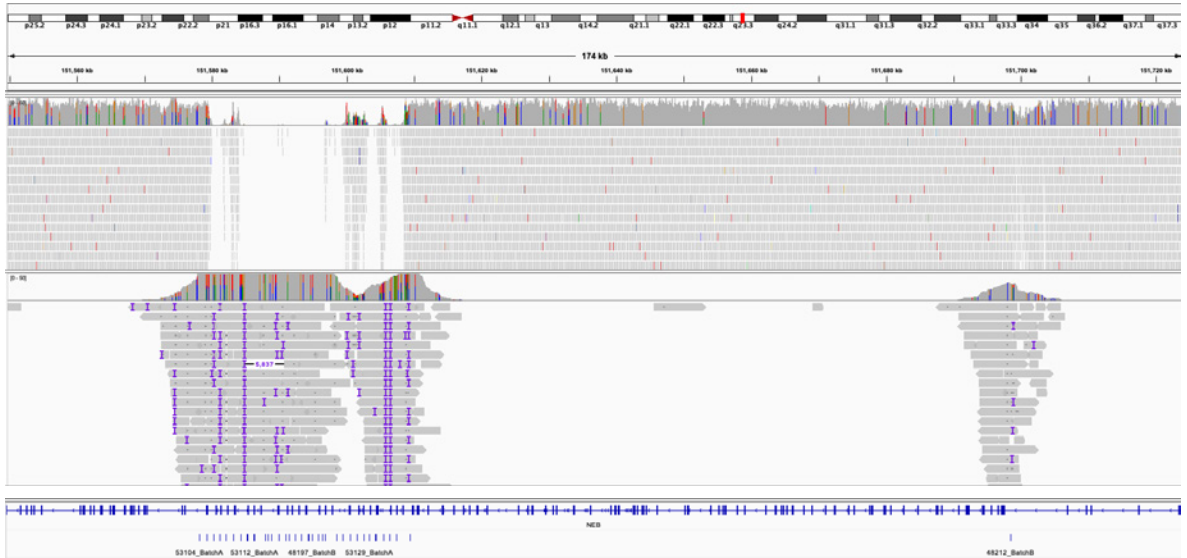


图 2： 使用 Human Comprehensive Panel 的靶向长读长解析难以绘制的区域——Illumina Complete Long Read Prep with Enrichment, Human 和 Human Comprehensive Panel 有助于提高对复杂基因区域的覆盖度，从而对标准短读长人类 WGS 进行补充。使用标准短读长 WGS 和 Illumina Complete Long Reads with Enrichment 获得的 Integrative Genomics Viewer (IGV) 图：(A) *STRC* 和 (B) *NEB*。

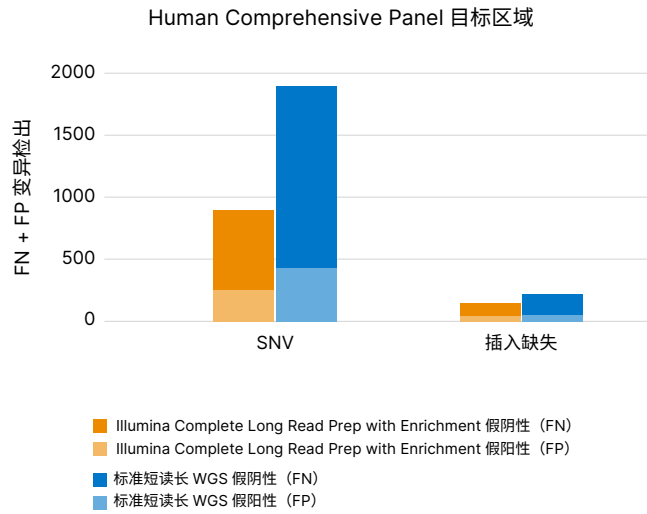


图 3: 通过靶向长读长提高变异检出准确性——与标准短读长 WGS (蓝色) 相比, 使用 Illumina Complete Long Read Prep with Enrichment 和 Human Comprehensive Panel (橙色) 在 Human Comprehensive Panel 靶向基因区域的假阴性 (FN) 和假阳性 (FP) SNV 和插入缺失变异检出情况。

总结

Illumina Human Comprehensive panel 能够对小部分基因区域实现高准确度长读长覆盖, 而仅使用短读长测序时难以覆盖这些基因区域。Illumina Complete Long Read Prep with Enrichment 和 Human Comprehensive Panel 提供了一种经过优化、经济高效的全基因组检测方法, 进一步完善了因美纳 WGS, 并专注于长读长测序, 通过完整的工作流程解决方案提供更高的价值。

了解更多

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

[Human Comprehensive Panel](#)

[长读长测序技术](#)

表 3: Human Comprehensive Panel 性能

Human Comprehensive Panel 目标区域的准确率 ^a			
	Illumina Complete Long Read Prep with Enrichment 和 Human Comprehensive Panel ^b	标准短读长 WGS ^c	
F1 评分 (SNV)	99.54%	99.04%	
F1 评分 (插入缺失)	99.47%	99.15%	
F1 评分 (SV)	80.43%	60.18%	
全基因组、所有基准区域的准确率 ^d			
F1 评分 (SNV+ 插入缺失)	Illumina Complete Long Read Prep with Enrichment 和 Human Comprehensive Panel ^b	Illumina Complete Long Read Prep, Human ^e	标准短读长 WGS ^c
	99.87%	99.90%	99.84%

a. 仅常染色体。SNV, 单核苷酸位点变异; SV, 结构变异。
b. Illumina Complete Long Read Prep with Enrichment, Human Comprehensive Panel + Illumina DNA PCR-Free Prep + NovaSeq 6000 基因测序仪 + DRAGEN v4.2。
c. Illumina DNA PCR-Free Prep + NovaSeq 6000 基因测序仪 + DRAGEN v4.2。
d. 使用 PrecisionFDA Truth Challenge V2 基准数据² 和 DRAGEN v4 的内部存档数据, 与所有参赛解决方案进行比较。
e. Illumina Complete Long Read Prep, Human + Illumina DNA PCR-Free Prep + NovaSeq 6000 基因测序仪 + DRAGEN v4.2。

参考文献

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-singlenucleotide-.html. Published 2021. Accessed August 30, 2023.
2. PrecisionFDA. Truth Challenge V2: Calling Variants from Short and Long Reads in Difficult-to-Map Regions. [Precision.fda.gov/challenges/10](https://precision.fda.gov/challenges/10). Accessed October 2, 2023.
3. Roessler K. Illumina Complete Long Reads software analysis workflow for human WGS. illumina.com/science/genomicsresearch/articles/complete-long-read-software-analysis.html. Published 2023. Accessed September 22, 2023.

订购信息

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

Illumina 中国

上海办公室 • 电话 (021) 6032-1066 • 传真 (021) 6090-6279
 北京办公室 • 电话 (010) 8441-6900 • 传真 (010) 8455-4855
 技术支持热线 400-066-5835 • chinasupport@illumina.com
 市场销售热线 400-066-5875 • china_info@illumina.com • www.illumina.com.cn

© 2024 Illumina, Inc. 保留所有权利。所有商标均为因美纳公司或其各自所有者的财产。
 关于具体的商标信息, 请访问 www.illumina.com.cn/company/legal.html。
 M-GL-02191 v1.0



illumina®