

使用 NovaSeq™ X 系列进行 RNA-Seq 测序

简化、集成的基因表达分析和
转录组分析工作流程



简介

多组学提供了一种综合方法，可帮助我们在多个生物学层面不断探索，获得新发现。通过将基因组学、转录组学、表观遗传学和蛋白质组学的数据相结合，研究人员可以更全面地了解正常发育、细胞反应和疾病相关的分子变化。

整合多组学，获得更全面的见解

多组学超越了单一组学，可揭示更深层次的生物学见解：

- **转录组学 + 基因组学**——对全基因组关联研究（GWAS）中发现的变异进行注释和优先排序，以了解疾病机制。
- **转录组学 + 表观遗传学**——研究基因调控与基因表达之间的联系，以了解控制目标表型的机制。
- **转录组学 + 蛋白质组学**——同时定量 RNA 和蛋白质表达，全面了解细胞功能和活动。

利用新一代测序（NGS）技术，RNA 测序（RNA-Seq）可为转录组学研究提供动力，以捕捉基因表达的变化并表征多种形式的 RNA，从而揭示其结构、变异和活性^{1,2}。

RNA-Seq 的优势

RNA-Seq 可提供特定时间点转录组的详细快照，与定量 PCR 相比具有诸多优势，包括：

- 无假设的实验设计，无需转录组的先验知识
- 强大的检测能力，利用全面的序列和变异信息检测已知和新型转录本
- 高通量，每次检测可定量数百至数千个区域
- 宽动态范围，基因表达的检测更加准确



图 1：NovaSeq X 和 NovaSeq X Plus 基因测序仪——旨在为执行数据密集型 RNA-Seq 应用提供所需的通量和准确性。

使用 NovaSeq X 系列进行 RNA-Seq 测序

因美纳在开发 RNA-Seq 解决方案方面拥有丰富成熟的经验，为研究人员开展转录组研究助力。因美纳 RNA 文库制备试剂盒通过简化的工作流程提供高质量数据，一个标准工作班次内即可完成全部实验。

NovaSeq X 系列为生产规模的 RNA-Seq 或多组学研究等数据密集型应用提供了超高的通量和准确性（图 1）。三种流动槽类型支持多种运行配置，提高了灵活性和可扩展性。

本应用说明介绍了如何使用 NovaSeq X 系列生成准确的 RNA-Seq 数据，数据质量媲美或优于 NovaSeq 6000 基因测序仪所生成的数据。

XLEAP-SBS™ 化学技术确保更出色的性能

NovaSeq X 系列使用 XLEAP-SBS 化学技术，该技术是我们成熟的因美纳边合成边测序（SBS）化学技术的升级版，速度更快、保真度更高、性能更可靠。NovaSeq X 系列试剂盒中的 XLEAP-SBS 试剂在性能和速度方面经过了优化，可在不牺牲数据质量的情况下大幅提高通量。与 NovaSeq 6000 基因测序仪相比，NovaSeq X 系列能以更短的运行时间显著提高通量（表 1）。

方法

NovaSeq X 系列支持 RNA-Seq 工作流程，该流程整合了 RNA 文库制备、测序以及简化的综合数据分析（图 2）。

文库制备

使用 Illumina Stranded Total RNA Prep with Ribo-Zero™ Plus（因美纳，货号 20040529），通过白血病细胞系 HL-60（Thermo Fisher Scientific，货号 AM7836）和 K562（BioChain，货号 R1255820-50）以及乳腺癌细胞系 RNA：MCF7（BioChain，货号 R1255830-50）制备总 RNA 文库。

表 1: 测序输出量相当，运行时间显著缩短

指标	NovaSeq X 系列 1.5B 流动槽	NovaSeq X 系列 10B 流动槽	NovaSeq X 系列 25B 流动槽	NovaSeq 6000 S4 流动槽
每次运行以 2×100 bp 输出 ^a	约 330-660 Gb	约 2-4 Tb	N/A	2-4 Tb
以 2×100 bp 运行的时间	约 20 小时	约 22 小时	N/A	约 36 小时
每次运行以 2×150 bp 输出	约 500-1000 Gb	约 3-6 Tb	约 8-16 Tb	3-6 Tb
以 2×150 bp 运行的时间	约 23 小时	约 25 小时	约 48 小时	约 44 小时
每个流动槽通过过滤的单端 read	1.6B	10B	26B	8-10B
总 RNA-Seq 文库	约 30	约 200	约 520	约 400

a. 输出量范围为一个或两个流动槽。



图 2: 简化、集成的 RNA-Seq 工作流程。

测序

采用 2 × 75 bp 运行配置，在 NovaSeq X Plus 基因测序仪上共对 682 个 RNA-Seq 文库进行了测序，具体如下。

- 使用 NovaSeq X Series 1.5B Reagent Kit（300 循环）（因美纳，货号 20104705）在 1.5B 流动槽上运行 61 个文库
- 使用 NovaSeq X Series 10B Reagent Kit（300 循环）（因美纳，货号 20085594）在 10B 流动槽上运行 141 个文库
- 使用 NovaSeq X Series 25B Reagent Kit（300 循环）（因美纳，货号 20104706）在 25B 流动槽上运行 480 个文库

为了进行比较，使用 NovaSeq 6000 S4 Reagent Kit v1.5（200 循环）（因美纳，货号 20028313）在 NovaSeq 6000 基因测序仪上对 60 个总 RNA 样本进行了测序，运行配置为 2 × 76 bp。

在 NovaSeq X Plus 基因测序仪上进行测序时使用了[定制的暗循环方案](#)，以减少测序过程中出现首碱基 T 悬垂（RNA 文库制备过程中连接测序接头时产生的副产物）的情况。暗测序循环只进行化学处理，省略了成像步骤，以改善测序指标。

机载或云端数据分析

将所有样本的测序数据取 10M read，使用 DRAGEN™ RNA pipeline v4.0.4 云端应用程序比较基因表达数据。将数据与基因组参考序列联盟的人类标准基因 GRCh38（hg38 组装序列）进行比对。同样使用 DRAGEN RNA pipeline v4.0.4 生成融合检出结果。使用 Broad 研究所的工具 [Morpheus](#) 生成热图。

虽然本应用说明中的分析是使用 DRAGEN RNA 流程进行的，但 RNA-Seq 分析也可以使用 [Partek Flow](#) 进行。该软件具有易于使用的界面、强大的统计算法和信息丰富的可视化效果。

结果

在 NovaSeq X Plus 基因测序仪上生成的总 RNA-Seq 数据超出了已公布的数据质量规范，也优于使用 NovaSeq 6000 基因测序仪生成的结果（[表 2](#)）。对于所有 NovaSeq X 系列流动槽中的总 RNA-Seq，两个平台的测序比对指标（包括重复率、已

表 2：总 RNA-Seq 的测序运行指标

指标	NovaSeq X 系列			NovaSeq 6000 基因测序仪
	1.5B 流动槽	10B 流动槽	25B 流动槽	S4 流动槽
运行配置	2 × 75 bp	2 × 75 bp	2 × 75 bp	2 × 76 bp
Read 1 碱基 ≥ Q30	94.00%	93.75%	94.98%	95.02%
Read 2 碱基 ≥ Q30	93.62%	93.64%	95.02%	92.77%
Read 1 错误率	0.09%	0.10%	0.14%	0.58%
Read 2 错误率	0.13%	0.15%	0.18%	1.07%
平均样本数	64	144	360	60
每个流动槽的多重分析	32	192	480	240
每次运行的通道数	4 个通道：2 次运行， 每次 2 个通道	6 个通道：3 次运行， 每次 2 个通道	6 个通道：3 次运行， 每次 2 个通道	1 次运行，1 个通道

来自单个流动槽的指标在具有不同通道数量的多个流动槽中运行的平均值。所有运行均符合已公布的产量规格。请注意，NovaSeq 6000 S4 流动槽和 NovaSeq X 10B 流动槽的每通道产量并不相同。

定位 read 百分比和唯一 read 百分比) 显示出出色的一致性 ($R^2 > 0.99$) (图 3 和图 4)。同样, 融合检出数据在两个平台上都表现出色 (图 5)。这些数据表明, 在 NovaSeq X 系列上进行全转录组测序产生的数据质量达到或超过了 NovaSeq 6000 基因测序仪的性能。

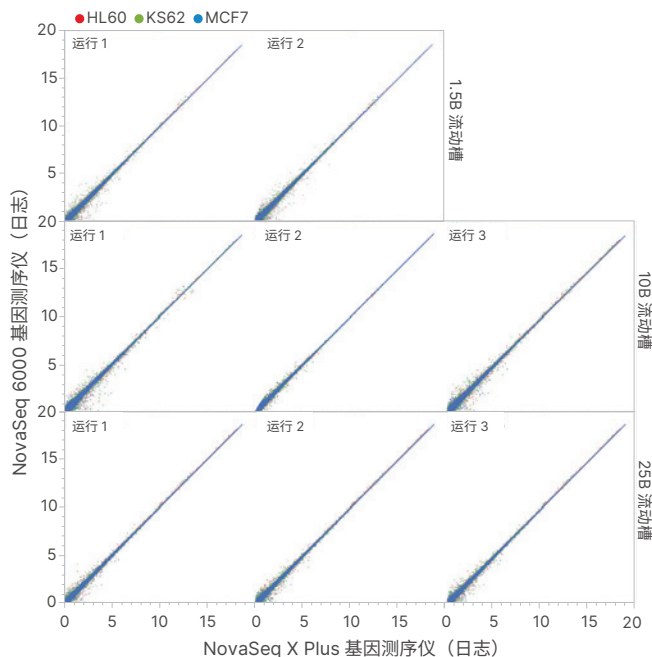


图 3: NovaSeq X Plus 和 NovaSeq 6000 基因测序仪之间的 RNA-Seq 数据一致性——测序数据显示，在所有 NovaSeq X 系列流动槽和所评估样本（HL60、KS62 和 MCF7，多次运行）的总 RNA-Seq 中，两个平台之间的一致性非常好， $R^2 > 0.99$ 。

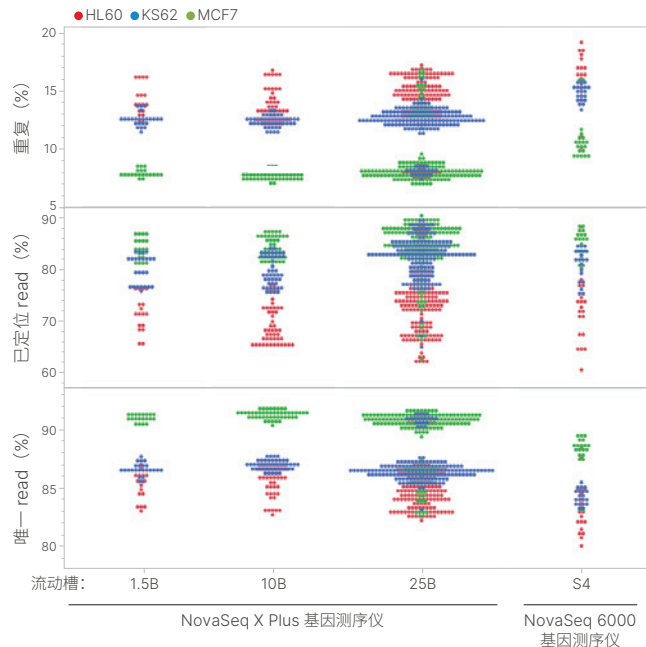


图 4: NovaSeq X Plus 和 NovaSeq 6000 基因测序仪上的 RNA-Seq 比对指标——包括重复率、已定位 read 百分比和唯一 read 百分比在内的测序比对指标显示，两个平台之间具有出色的一致性（在所有 NovaSeq X 系列流动槽和所评估样本（HL60、KS62 和 MCF7）中，总 RNA-Seq 的 $R^2 > 0.99$ ）。

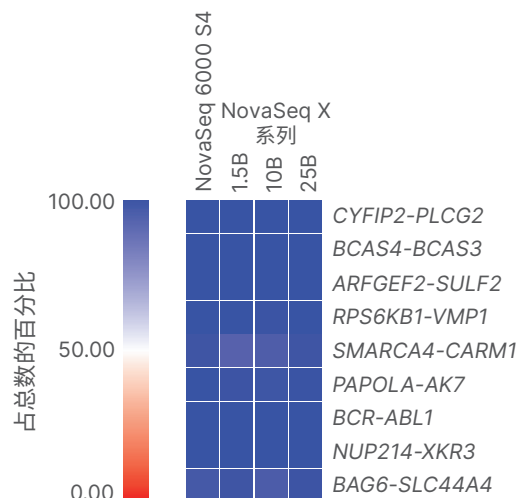


图 5: NovaSeq X Plus 和 NovaSeq 6000 基因测序仪之间的融合检出数据一致性——对于不同的流动槽，两种平台上的融合检出数据均表现出色，括号内标明。

总结

NovaSeq X 系列的 RNA-Seq 提供了简化的从 RNA 到结果的工作流程，集强大功能、速度和灵活性于一身，可实现高通量转录组分析。将 NovaSeq X Plus 基因测序仪的全转录组测序数据与使用 NovaSeq 6000 基因测序仪生成的数据进行了直接比较。结果表明，对于 NovaSeq X Plus 基因测序仪的所有三种流动槽配置，NovaSeq X 系列的性能均达到或优于 NovaSeq 6000 基因测序仪的性能，为从基因表达分析到全转录组发现研究的各种 RNA-Seq 项目提供了可扩展性。NovaSeq X 系列的转录组学分析可与其他组学研究相结合，从而获得深入的生物学见解。

了解更多

[NovaSeq X 系列](#)

[NovaSeq X 系列的数据一致性](#)

[多组学](#)

参考文献

1. Geraci F, Saha I, Bianchini M. [Editorial: RNA-Seq Analysis: Methods, Applications and Challenges](#). *Front Genet.* 2020;11:220. Published 2020 Mar 17. doi:10.3389/fgene.2020.00220.
2. Corchete LA, Rojas EA, Alonso-López D, De Las Rivas J, Gutiérrez NC, Burguillo FJ. [Systematic comparison and assessment of RNA-seq procedures for gene expression quantitative analysis](#). *Sci Rep.* 2020;10(1):19737. Published 2020 Nov 12. doi:10.1038/s41598-020-76881-x.

因美纳中国

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-02636 v1.0



illumina[®]