

Twist Exome 2.0 完美适配华大智造 MGISP自动化系统和DNBSEQ测序平台 赋能WES研究

本研究分别利用华大智造自动化系统MGISP-960 & MGISP-100以及测序平台MGISEQ-2000 & DNBSEQ-G99完成Twist WES文库构建和测序工作，并对相关性能进行综合性评估。

相关分析表明：Twist WES解决方案可完美搭配华大智造自动化系统以及DNBSEQ测序平台；实验过程建议采用八杂(8-plex)的杂交方案，生信分析建议截取6Gb的数据量。

推荐应用：全外显子组测序(WES)

推荐机型：MGISEQ-2000RS, DNBSEQ-G99ARS, DNBSEQ-T7RS(测序仪)
MGISP-960RS, MGISP-100RS(自动化样本制备系统)

• 为WES应用提供了一流的解决方案

Twist Exome 2.0探针有着较高的均一性, 较低的脱靶率, 使得测序效率更高, 数据质量更好。

• 完美适配DNBSEQ测序平台

华大智造通用文库转换试剂盒可使Twist WES 建库方案完美适配DNBSEQ测序平台。

• 数据产出高效且质量高

DNBSEQ测序技术具有高准确性, 低重复序列率以及低标签跳跃等重要特性。

• 可匹配自动化操作

华大智造MGISP自动化系统可为实验流程提供自动化解决方案, 能够极大地节约人工成本并提高效率。



背景介绍

大规模平行测序(MPS, massively parallel sequencing)技术的出现极大地改变了罕见的种系突变和体细胞突变的诊断过程,提高了诊断精度以及结果产出速度,降低了诊断成本¹。全外显子组测序(WES, whole exome sequencing)是一种被广泛开展的MPS应用,用于测序基因组中编码蛋白质的区域¹。外显子区域虽然约仅占全基因组的2%^{1,2},但已报道的致病变异中有85%位于外显子中。因此,与全基因组测序(WGS, Whole genome sequencing)相比,WES可以在测序数据量更少和分析成本更低的情况下,有效地覆盖95%以上的外显子区域²。WES可以通过单次测试揭示5000多种种系变异和体细胞变异³⁻⁸,肿瘤分子诊断实验室目前都采用WES作为常规的检测手段。

完美的WES解决方案应全面地覆盖蛋白质编码区,并且具有较高的均一性和中靶率。在各种各样的商业WES试剂盒中, Twist Exome 2.0是最受欢迎的WES靶向富集方案之一,其具有高均一性和低脱靶率等特点。Twist Exome 2.0探针为双链设计的DNA,可以靶定DNA的正义链和反义链并提高富集灵敏度。此外,该panel还涵盖额外的位于非编码区的与临床相关的致病性变异和可能致病的变异,使得种系变异和体细胞变异的检测更加容易。

华大智造自主研发的DNBSEQ测序技术具有高精度、高灵敏度、超低重复率以及低标签跳跃率等特点⁹。基于DNBSEQ测序技术推出的MGISEQ-2000、DNBSEQ-G99等一系列基因测序仪,可满足生命科学研究和临床应用中对MPS的各种需求。此外,华大智造还开发了高质量的MGISP自动化系统,可通过减少样品与样品之间的差异来提高文库一致性,该系统可极大地提升建库通量,减少文库制备的时间,进而提高测序结果的一致性,并降低实验室MPS运行成本⁸。

然而, Twist Exome 2.0与华大智造自动化系统和DNBSEQ测序平台的兼容性依旧未知,因此,本研究综合评估了基于华大智造自动化系统MGISP-960 & MGISP-100以及测序平台MGISEQ-2000 & DNBSEQ-G99, Twist Exome 2.0靶向富集panel的性能表现。

研究描述

为验证Twist的WES建库方案与华大智造相关产品的适配性，本研究利用自动化设备MGISP-960或MGISP-100完成文库构建，MGISEQ-2000或DNBSEQ-G99完成测序，成熟的WES生信分析流程完成数据分析后对相关表现进行了综合性评估。结果表明，Twist Exome 2.0在华大智造的自动化系统以及DNBSEQ测序平台上相关性能指标表现优异，且8杂(8-plex)是较优的杂交方案，每样本截取6Gb数据量更为经济高效。

研究方法

样本制备

本研究使用商业标准品 Human Genomic DNA (Takara PN: 636401) 作为性能评估实验的样本。

文库制备与测序

在文库制备阶段，每个样本投入50ng，使用Twist Library Preparation EF Kit 2.0 以及Twist Universal Adapter System完成杂交前的文库制备过程，其中Pre-PCR的循环数为8，具体操作流程可参考相关说明书。该过程在自动化系统MGISP-960或MGISP-100上完成。随后利用Twist Exome 2.0进行手工杂交，对于用于1-plex的样本，每个样本投入1500ng；对于用于8-plex的样本，投入样本的总质量为1500 ng (每个187.5 ng)。杂交后过程(包含洗涤、Post-PCR和纯化)通过自动化系统(MGISP-960或MGISP-100)或者手工完成。

Post-PCR (8个循环) 使用MGIEasy Universal Library Conversion kit (App-A)中的AC-primer完成扩增。后利用该试剂盒完成单链DNA环化，DNA纳米球(DNB)的制备。

后续测序工作在MGISEQ-2000或DNBSEQ-G99上按照双端150bp(PE150)的方案完成，均用App-C作为测序引物。利用DNBSEQ-G99完成1-plex文库的测序，利用MGISEQ-2000完成1-plex和8-plex文库的测序。本研究中，杂交过程均为手工进行，杂交前使用MGISP-960或MGISP-100完成操作；杂交后过程(包含洗脱、Post-PCR和纯化)使用MGISP-960或手工操作完成。2000-8plex(960, Manual)是指8-plex文库杂交前使用MGISP-960，杂交后过程为手工并在MGISEQ-2000上完成测序的实验组，其他名称含义依次类推。

数据分析

本研究在MegaBOLT生信分析加速器上完成相应的WES生信分析工作。简略流程如下：测序所得raw reads通过数据质控流程过滤得到clean reads，随后利用Burrows-Wheeler Aligner (BWA, <http://bio-bwa.sourceforge.net/>)将clean reads比对到人类参考基因组hg38上。比对结果被记录在SAM文件中，经过定位排序和标记去重后，被压缩成BAM文件，随后进行碱基质量值校正、BAM统计和VCF统计，最终生成可视化报告，结果报告中涵盖Duplication rate、Reads clean rate、Reads mapping rate、Q30、Average depth (rmdup)、Reads on target rate、Fold 80 base penalty、Target coverage (%)等关键数据指标。

样本采集	文库制备和测序	生信分析	结果分析
商业标准品Human Genomic DNA	 <p>Twist Library Preparation EF Kit 2.0 Twist Universal Adapter System Twist Exome 2.0 Twist Standard Hyb and Wash Kit v2</p>  <p>MGIEasy 通用文库转换试剂盒(App-A)</p>  <p>MGISP-960或MGISP-100</p>  <p>MGISEQ-2000或DNBSEQ-G99</p>	MegaBOLT 的 WES 分析流程	基于华大智造自动化系统和DNBSEQ测序平台对Twist Exome 2.0的系统性评估

结果

华大智造MGISP自动化系统可制备高质量的WES文库

本研究利用Twist Library Preparation EF Kit 2.0和Twist Exome 2.0来进行杂交前的文库制备和目标片段的富集，并采用自动化建库（MGISP-960 或MGISP-100）和手工建库两种方案作为对比以评估自动化系统的性能。图1展示的是自动化(9号样本)与手工(11号样本)建库所获文库对比情况。两样品的杂交前文库均由MGISP-960制备，杂交后

文库通过MGISP-960或者手工制备完成(图1A)。在相同50ng的投入量下，两种策略所构建文库均可满足后续DNB的制备要求(图1A)。同时，Agilent TapeStation仪器的分析结果表明杂交前与杂交后所获得的文库自动化与手工建库方案无明显差别、片段大小与预期相符、没有引物二聚体或其他非特异性峰且文库质量较高（图1B）。

A

样本编号	投入量	Index 名称	杂交前建库	杂交	杂交后建库	杂交前文库产量(ng)	杂交后文库产量(ng)
9	50 ng	257	MGISP-960	Manual	MGISP-960 (1-plex)	3800	495
11	50 ng	273	MGISP-960	Manual	Manual(1-plex)	3660	702

B

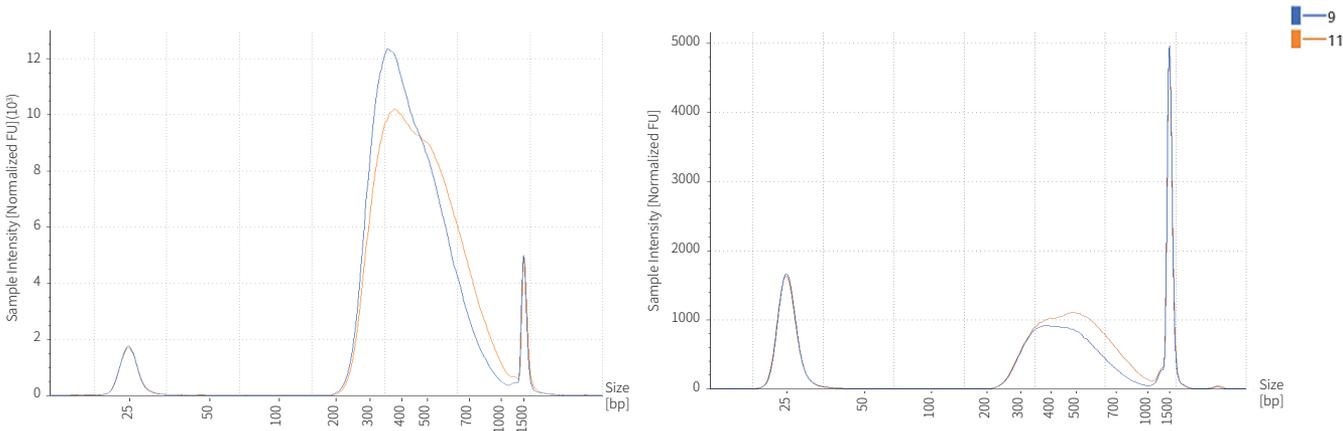


图1. 自动化与手工方案所构建文库的初步情况对比。(A) 自动化(9号样本)与手工(11号样本)建库所获文库情况对比; (B) 自动化(9号样本, 蓝色线)与手工(11号样本, 橙色线)建库所获文库电泳图以分析文库片段大小分布情况。左图为杂交前经过pre-PCR得到的文库, 右图为杂交后经过post-PCR得到文库。

6Gb是WES分析较为合适的数据量

为评估WES分析所需截取的测序数据量, 本研究从G99-1plex(960,960)样本的测序数据中分别独立截取2, 4, 6 和8 Gb的数据, 并比较这四组数据的测序质量。结果显示: Duplication rate, Average depth (rmdup)以及Target coverage这些指标随数据量增加都呈现递增的趋势 (图2 A,B,E)。但Reads on target rate以及Fold 80 base penalty相互之间无显著差别, Reads on target rate都达到了77.4%以上(图2 C), Fold 80 base penalty的值均小于1.5, 在截取数据量为2Gb时此值为1.41, 略高于

截取数据量为4, 6, 8 Gb的情况(图2 D)。据此表明测序数据量越多, 分析结果越准确。

由于Twist Exome 2.0探针的覆盖区域大小为36.5Mb, 在测序数据量为6Gb的情况下, 可以达到约150×的测序深度。在全外显子组测序中, 150×的测序深度已经能够得到较为准确的结果, 且6Gb的数据量下, 各项指标较好。因此基于测序成本和时间综合考虑, 推荐截取6GB的数据用于WES分析。

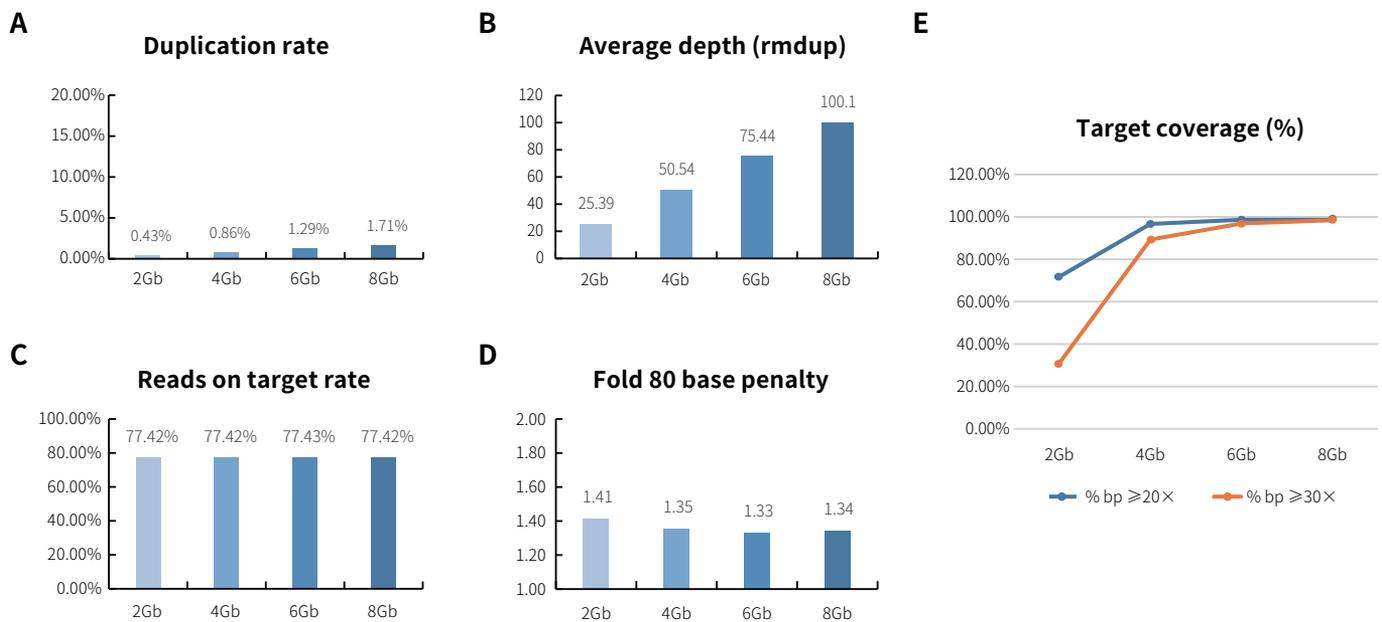


图2. 本研究选取G99-1plex (960,960)样本的测序结果, 分别截取2,4,6和8Gb的数据量以确定WES分析所需数据量。在截取数据量为2, 4, 6, 8 Gb时, 其Duplication rate呈递增趋势, 但数值均很低, 满足实验分析要求(A); 其Average depth (rmdup)也随着截取数据量的增加而增多, 表明截取的数据量越多, 测序深度越高(B); Reads on target rate基本无变化, 说明无论截取多少数据, Twist Exome 2.0对外显子有着较高的特异富集率(C); Fold 80 base penalty也基本无变化, 说明Twist Exome 2.0具有较好的覆盖均一性(D); Target coverage 在 $\geq 20\times$ 和 $\geq 30\times$ 的情况下均呈现递增的趋势, 说明截取的数据量越多, 目标覆盖度越高, 但在截取数据量为6 Gb和8 Gb的情况下, 基本无变化(E)。

WES实验流程中推荐8-plex的杂交方案

为确定杂交方案选取1-plex还是8-plex, 本研究对来自于2000-1plex(960,960)以及2000-8plex(960,960)的测序数据统一截取6Gb进行分析。2000-8plex(960,960)的Duplication rate为1.71%, 略高于2000-1plex(960,960); 但Reads on target rate、Target coverage(% bp $\geq 30\times$)、Average depth (rmdup)以及Fold 80 base penalty 的值分别为

80.18%、97.24%、79.56以及1.28(图3), 略优于2000-1plex(960,960)。2000-1plex(100,960)以及2000-8plex(100,960)的数据本文未展示。上述分析表明, 采用8-plex的杂交策略在节约成本、提高效率的同时仍可以获得较好的数据结果。因此在基于华大智造自动化系统与DNBSEQ测序平台开展WES研究时, 推荐采用8-plex的方案用于杂交。

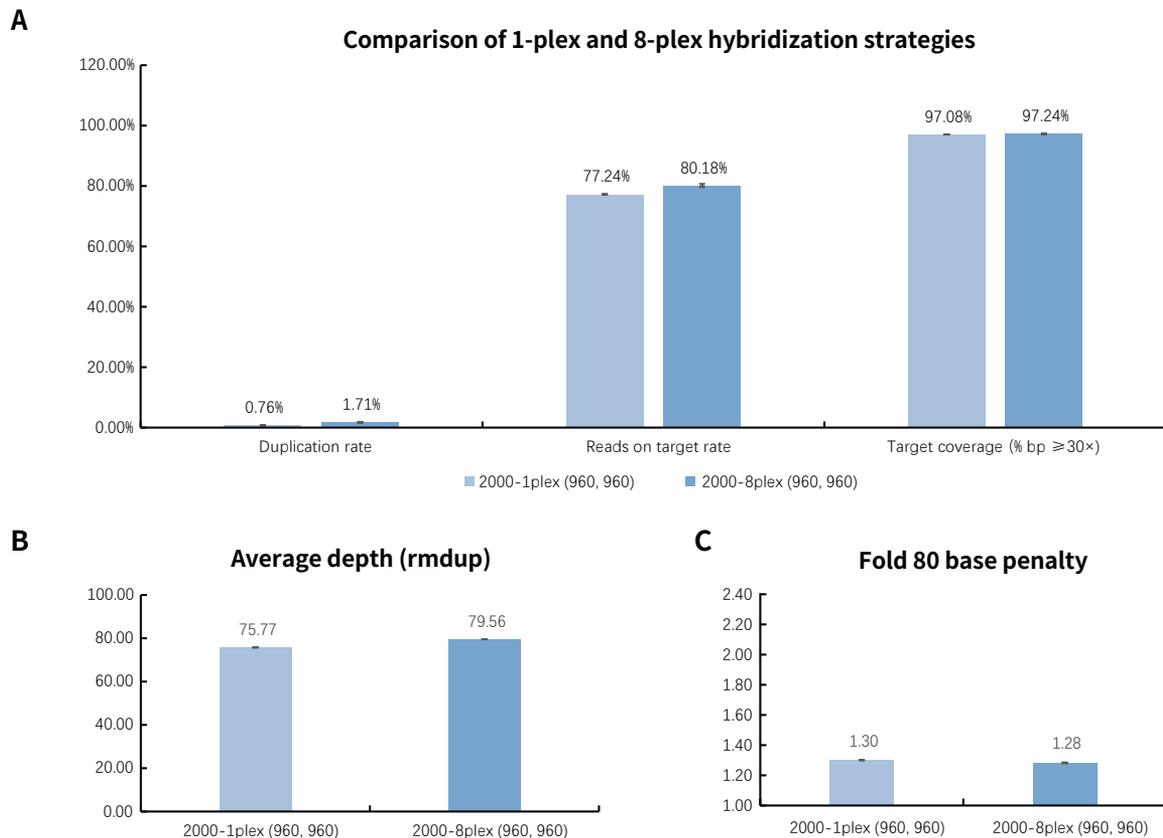


图3. 利用自动化系统进行文库制备时, 1-plex和8-plex的杂交策略比较。对2000-1plex(960,960)以及2000-8plex(960,960)这两组样本统一截取6Gb数据进行比较。对于Duplication rate, 8-plex方案的值略高于1-plex方案(A), 但对于Reads on target rate, Target coverage (% bp $\geq 30\times$), Average depth (rmdup)以及Fold 80 base penalty, 8-plex方案均略优于1-plex方案(A,B,C)。

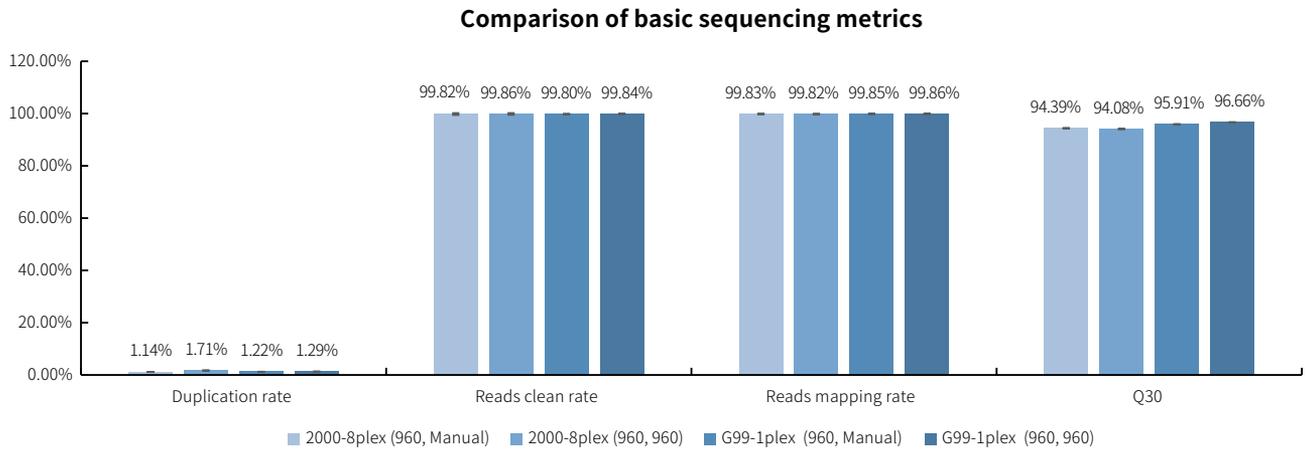
基于华大智造MGISP自动化系统所得WES文库的测序质量高

为了探究评估自动化建库与手工建库得到的测序数据质量是否有差别, 我们对来自于2000-8plex(960, Manual), 2000-8plex(960, 960), G99-1plex(960, Manual)以及G99-1plex(960, 960)的四组测序数据截取6Gb进行比较, 结果发现这四组数据的Duplication rate都低于2%, Reads clean rate以及Reads mapping rate都高达99%以上, Q30值都在95%左右(图4 A)。进一步的性能评估分析发现, 这四组测序数据的Average depth (rmdup)都高于65 \times , Reads on target rate都大于70%, Fold 80 base penalty都低于1.4, 并且自动化建库得到的

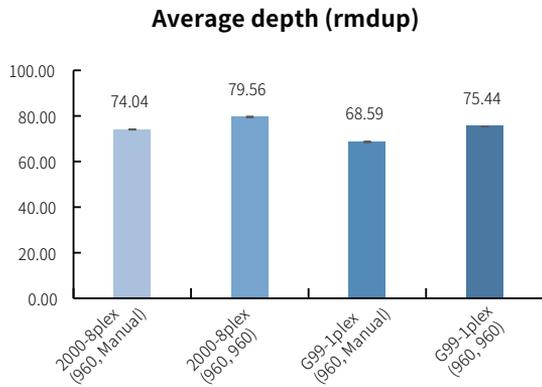
参数都略优于手工建库参数(图4 B, C, E)。同时, 这四组测序数据的Target coverage (% bp $\geq 30\times$)都高达96%以上, 足以获取可信度较高的变异检测结果(图4 D)。此外, 相关结果与Twist官网公布的Twist Exome 2.0相关性能指标无显著差别。

以上结果表明Twist建库方案搭配华大智造自动化系统与DNBSEQ测序平台开展WES研究, 得到的测序数据质量较高, 且与手工建库结果无显著差异, Twist建库方案可完美适配DNBSEQ平台。

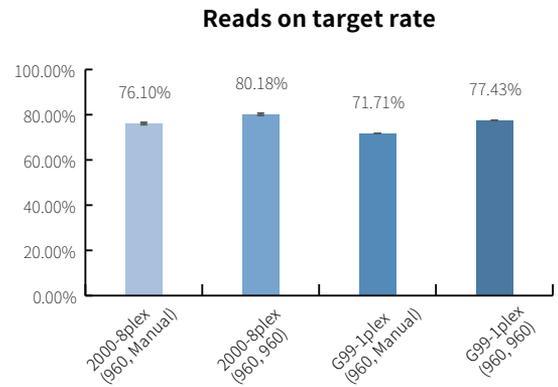
A



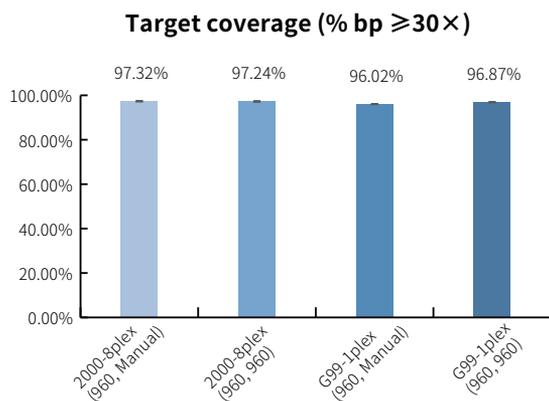
B



C



D



E

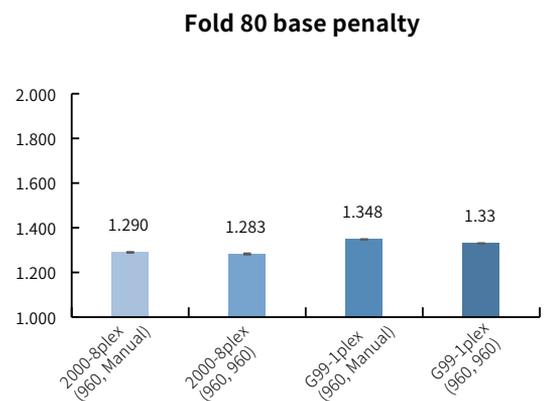


图4. 自动化建库与手工建库的关键测序质量指标比较。(A)手工建库和自动化建库的初级测序指标显示自动化建库与手工建库无差别；Duplication rate均低于2%，Reads clean rate以及Reads mapping rate均高于99%，Q30均在95%左右；(B, C, E)对于Twist 2.0性能评估参数Average depth (rmdup)，Reads on target rate，Fold 80 base penalty，自动化建库数值均略优于手工建库数值；(D)对于Target coverage (% bp $\geq 30\times$)，自动化建库与手工建库不相上下。

总结

Twist Exome 2.0在DNBSEQ平台的实测数据表明，Twist出品的WES解决方案可完美适配华大智造MGISP自动化系统以及DNBSEQ测序平台，在极大节省人工成本的同时，保证数据的高质量输出。本研究同时建议在使用该方案开展WES研究时，杂交过程中采用8杂(8-plex)的杂交方案，生信分析过程中截取6Gb数据进行后续分析。

MGISEQ-2000采用全新的载片系统，能够灵活支持多种不同的测序模式，并采用优化设计的光学及生化系统，满负荷PE150(FCL)测序仅需约56小时。DNBSEQ-G99作为全球同等通量测序仪中速度最快的机型之一，其单张载片通量为80M，满负荷PE150测序仅需12小时，将测序效率提升到极致，同时内置计算模块，测序生信一体化。

MGISEQ-2000 & DNBSEQ-G99搭配Twist方案，可全面助力全外显子组测序，支持科研、临床医学、司法、农业等多个领域的测序应用和数据分析。



自动化样本制备系统MGISP-960



基因测序仪MGISEQ-2000

参考文献

1. Rabbani, B., Tekin, M. & Mahdieh, N. The promise of whole-exome sequencing in medical genetics. *J Hum Genet* **59**, 5-15, doi:10.1038/jhg.2013.114 (2014).
2. Wang, Q., Shashikant, C. S., Jensen, M., Altman, N. S. & Girirajan, S. Novel metrics to measure coverage in whole exome sequencing datasets reveal local and global non-uniformity. *Sci Rep* **7**, 885, doi:10.1038/s41598-017-01005-x (2017).
3. Clark, M. M. et al. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom Med* **3**, 16, doi:10.1038/s41525-018-0053-8 (2018).
4. Wright, C. F., FitzPatrick, D. R. & Firth, H. V. Paediatric genomics: diagnosing rare disease in children. *Nat Rev Genet* **19**, 325, doi:10.1038/nrg.2018.12 (2018).
5. Lalonde, E. et al. Genomic Diagnosis for Pediatric Disorders: Revolution and Evolution. *Front Pediatr* **8**, 373, doi:10.3389/fped.2020.00373 (2020).
6. Yuan, J. et al. Novel technologies and emerging biomarkers for personalized cancer immunotherapy. *J Immunother Cancer* **4**, 3, doi:10.1186/s40425-016-0107-3 (2016).
7. Samstein, R. M. et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet* **51**, 202-206, doi:10.1038/s41588-018-0312-8 (2019).
8. Zimmerman Zuckerman, E. et al. Automation of hybridization and capture based next generation sequencing library preparation requires reduction of on-deck bead binding and heated wash temperatures. *SLAS Technol* **27**, 214-218, doi:10.1016/j.slast.2021.10.016 (2022).
9. Porreca, G. J. Genome sequencing on nanoballs. *Nat Biotechnol* **28**, 43-44, doi:10.1038/nbt0110-43 (2010).

推荐订购信息

产品类型	产品名称	产品货号
仪器	基因测序仪MGISEQ-2000RS	900-000035-00
	基因测序仪DNBSEQ-G99ARS	900-000560-00
	MGISP-100RS自动化样本制备系统	900-000070-00
	MGISP-960RS 自动化样本制备系统	900-000093-00
软件	MegaBOLT 生信分析加速器 (工作站式服务器)	970-000085-00
	Twist Library Preparation EF Kit 2.0 (96 RXN)	104207*
	Twist Universal Adapter System- TruSeq Compatible, 96 Samples Plate C	101310*
建库试剂	Twist Exome 2.0 (12RXN), Kit	104134*
	Twist Standard Hyb and Wash Kit v2 (12RXN)	104446*
	MGIEasy 通用文库转换试剂盒 (App-A) (16RXN)	1000004155
	MGISEQ-2000RS 高通量测序试剂套装 (App-A PE150)	1000014051
测序试剂	DNBSEQ-G99RS高通量测序试剂套装 (G99 SM App-C FCL PE150)	940-000413-00
	高通量测序引物试剂盒 (App-C)	1000027472

*相关产品可登录Twist官网进行查阅订购。

深圳华大智造科技股份有限公司

深圳市盐田区北山工业区综合楼11栋

☎ 4000-688-114

🌐 www.mgi-tech.com

✉ MGI-service@mgi-tech.com

股票简称: 华大智造

股票代码: 688114



仅供研究使用

版权声明: 本手册版权属于深圳华大智造科技股份有限公司所有,未经本公司书面许可,任何其他个人或组织不得以任何形式将本手册中的各项内容进行复制拷贝、编辑或翻译为其他语言。本手册中所有商标或标识均属于深圳华大智造科技股份有限公司及其提供者所有。

版本: 2023年7月版

撰稿: 张含菲 王其伟

责任编辑: 王其伟

审稿: 江遥