

灵活、高质量 活跃你的日处理能力



MGISEQ-2000系列 基因测序仪



双载片、双机型, 给你更多选择

实力版

MGISEQ-2000RS

MGISEQ-2000

灵活、稳定、资质完备
中高通量基因测序主力

极速版

MGISEQ-2000RS FAST

快速上手、快速交付
中低通量基因测序首选

产品介绍

MGISEQ-2000作为一款全能型桌面测序仪，能够为用户提供全面、灵活、高效的测序选择。除了标准的高通量测序试剂(Standard-MPS)，基于MGISEQ-2000平台研发的CoolMPS高通量测序试剂为追求更高质量的用户提供更多选择，具有稳定的高强度信号和随机的低测序错误率，在科学和临床应用，如肿瘤的低频突变检测中表现出优异的性能。

MGISEQ-2000RS FAST定位于中小规模测序项目的快速交付，相比于MGISEQ-2000，MGISEQ-2000RS FAST则更加注重于快速交付和“轻”成本，与MGISEQ-2000形成了绝好的高低搭配。

MGISEQ-2000系列测序仪应用场景广泛，涵盖科学研究、临床医学、疾病防控、环境、农业等多种应用领域，助力医疗和科研领域高通量测序系统的全面普及。

选择多

• 两种载片规格

大载片FCL 1800M，小载片FCS 550M

• 双载片平台

一台可当两台用，通量覆盖550M~3600M reads/run

质量高

• DNBSEQ™技术优势

扩增错误零累积，低扩增偏向性，低标签跳跃

• CoolMPS™技术优势

独有的CoolMPS测序试剂让肿瘤低频突变更准确

速度快

• 测序和数据处理时间短

FCS SE100从测序到生成Fastq文件仅需13小时

应用广

• 适用范围广

既适合科研用户平台建设
亦可满足临床样本快速交付

采用全新的双载片系统，在双载片独立运行平台的基础上，支持不同规格的载片，实现更全面更灵活的测序体验。



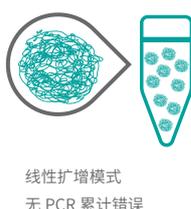
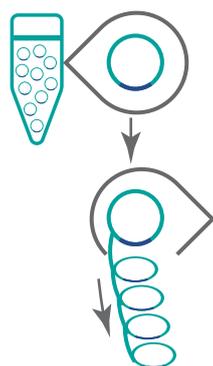
载片类型	试剂类型	MGISEQ-2000RS FAST	MGISEQ-2000RS	MGISEQ-2000
FCS	StandardMPS	●	●	●
FCL			●	●
	CoolMPS		●	●

◎ 技术原理

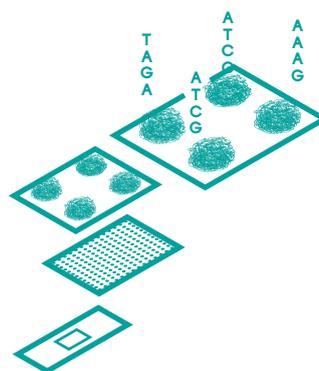
自主可控的「DNBSEQ™」核心技术

DNA纳米球测序技术——线性扩增无错误累积

↑ 高准确性 ↓ 低重复序列率 ↓ 低标签跳跃率



线性扩增模式
无 PCR 累计错误



纳米级阵列位点设计
测序精度高且不产生信号干扰

CoolMPS 测序基于 DNBSEQ 技术原理，保留了 DNBSEQ 技术特有的高准确性，低重复序列率和低标签跳跃率。

CoolMPS测序化学——天然碱基让碱基识别更清晰

CoolMPS 是华大智造推出的一项创新性技术，是全球首个基于抗体的大规模平行测序方法。CoolMPS 高通量测序试剂套装是基于该方法的测序试剂，适配华大智造 DNBSEQ 测序平台，兼容常规文库制备方法。

CoolMPS 使用的 cold dNTPs（未在碱基上标记荧光），在 DNA 聚合酶的作用下，被聚合到测序链上，通过可与其特异结合的荧光标记抗体来实现无损碱基识别。由于这些 cold dNTPs 使用了天然碱基，可有效避免传统测序方法 (StandardMPS) 中的 DNA “疤痕” 积聚对后续读取准确性的影响，荧光信号更强，聚合反应效率更高，测序准确性更高，测序潜力更大。

已有用户数据表明，在 MGISEQ-2000 平台采用 CoolMPS 高通量测序试剂套装的下机数据质量更高，错误率更低。

CoolMPS测序原理

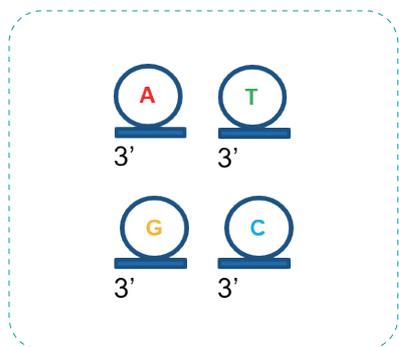


图 1 四种 cold dNTPs

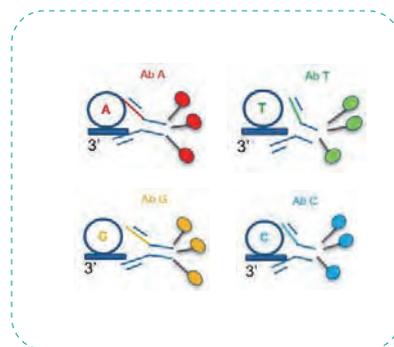


图 2 四种荧光标记的抗体与对应的 cold dNTPs 特异结合

在 CoolMPS 测序过程中，将四种带可逆阻断的 cold dNTPs 引入测序载片。通过 DNA 聚合酶将 cold dNTPs 聚合到锚定有引物的 DNB 测序链的每个拷贝中。

A, T, G, C 特异性抗体交叉反应几乎为零，每个抗体都有一个特定的染料（标签）分子附着在其上。

在 CoolMPS 测序过程中，cold dNTPs 聚合后，将特定荧光染料标记的四种抗体混合液引入测序载片中。例如，T 特异性抗体（如图 2 中 Ab T 所示）与 DNB 中被聚合的所有 cold-T 结合。

CoolMPS测序优势

· 信号更强

每个抗体分子可以标记上多个荧光分子，相同拷贝数的 DNB，用 CoolMPS 可以获得更高的信号和更高的信噪比，从而使测序更准确。图 3 比较了 StandardMPS 和 CoolMPS 信号强度。CoolMPS 信号更强，在 PE100 的测序模式下，CoolMPS 无论是起始信号还是在 200 循环时均更高。

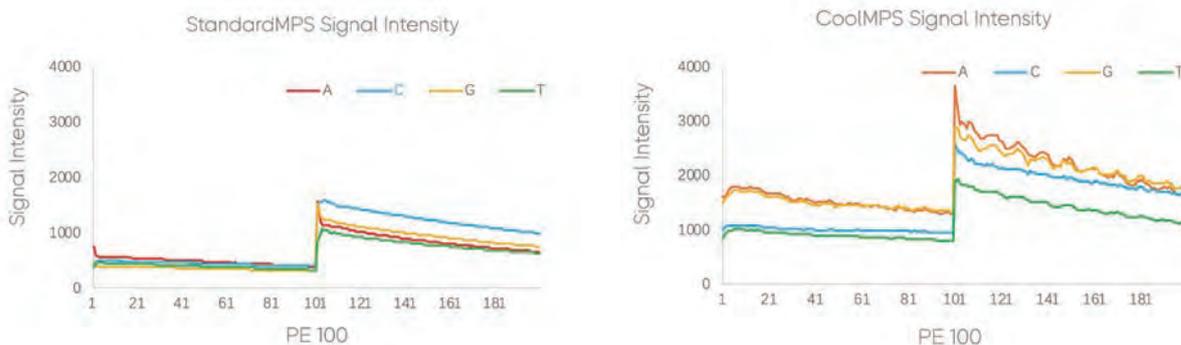


图 3 测序时 StandardMPS 和 CoolMPS 的信号强度及随循环数的变化趋势

· 无荧光猝灭，信号更稳定

传统高通量测序方法会在聚合的碱基上留下疤痕，有时会对待测碱基上的荧光基团有淬灭效应，从而影响测序的准确性。CoolMPS 方法中新合成链的碱基全为天然碱基，不存在因为碱基上的疤痕而产生的荧光淬灭效应，从而可以获得更高的准确性和更长的测序读长。

图 5 显示了传统测序和 CoolMPS 测序中，前两个碱基对被测碱基信号的影响。传统高通量测序化学中，A、C、G 和 T 的信号强度由于疤痕的存在而受到部分抑制，这种信号抑制可能会导致测序错误。而在 CoolMPS 中，由于不存在任何疤痕，不管前面两个碱基是什么，其 A、C、G、T 的信号保持相当稳定。

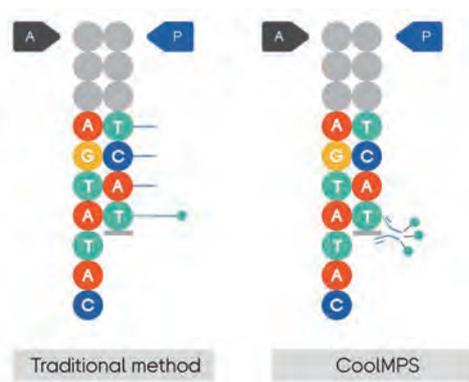


图 4 CoolMPS 与传统高通量测序方法的比较：无疤痕

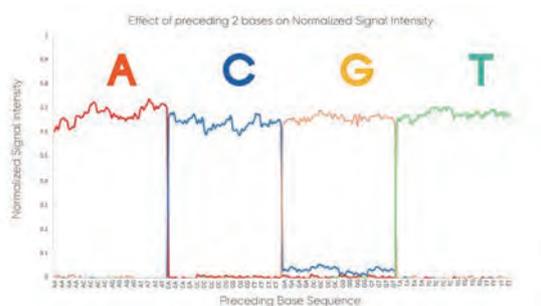


图 5 CoolMPS 测序中，前两个碱基对被测碱基信号的影响

· 测序错误更随机

图 6 展示了 CoolMPS 的测序错误类型与前一个碱基的关系：无论前两个碱基是什么，CoolMPS 的测序错误都非常低，且没有明显的错误偏好。



图 6 CoolMPS 前两个碱基对被测碱基错误类型的影响

Runon 和 Lag 显著降低

聚合酶结合及新加的都是天然碱基，DNA 聚合酶在生物体内的底物是天然 dNTPs。虽然经过突变体的改造，DNA 聚合酶也能加碱基上有庞大荧光基团标记的 dNTPs，但其反应效率远远不如加碱基上没有标记的 dNTPs。此外每一轮反应接近 100% 的完成率也是进行长读长测序的必要条件，否则信号会呈现指数下降，产生测序错误。

因此，CoolMPS 的反应更完全，即 DNA 聚合天然碱基的效率更高，测序读长更长。

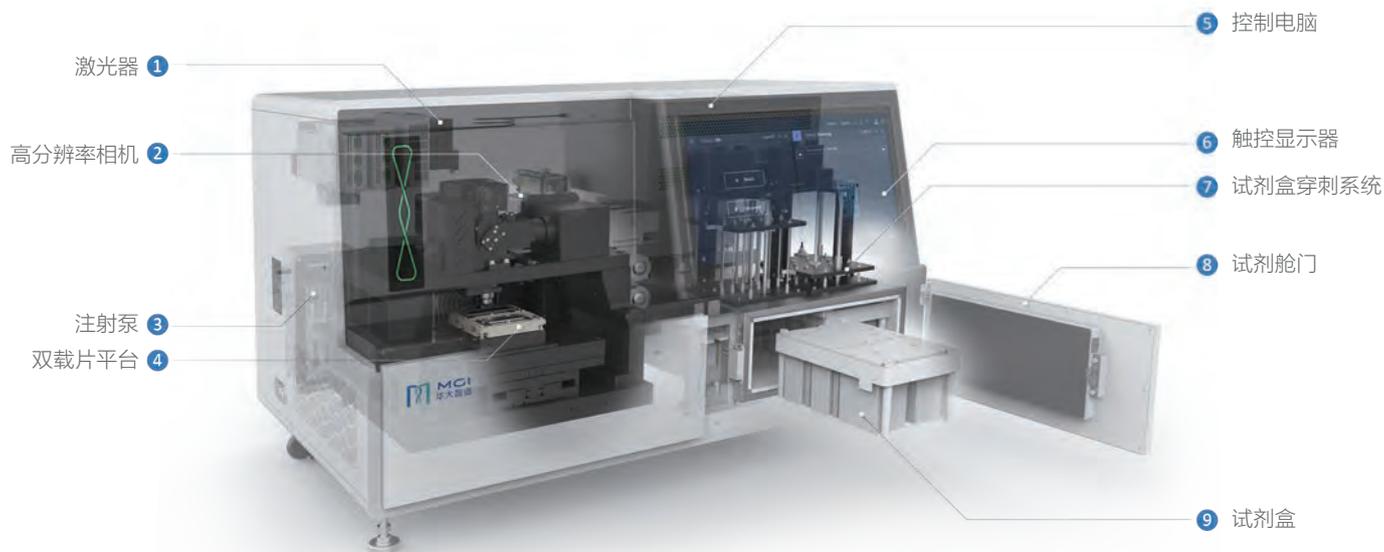
测序时，常用 lag 及 runon 指标来反映反应是否完全。lag 是指测序进行到 N 位置时，其某些拷贝才反应到 N-1 位置的比例；而 runon 则指某些拷贝已反应到 N+1 位置上的比例。我们测定了用于 CoolMPS 的测序酶在加 100% 的有荧光标记碱基的 dNTPs 与加无修饰碱基的 cold dNTPs 的 lag 与 runon，其结果如图 7 所示，加 cold dNTPs 的平均 lag 值 (0.2%) 比加荧光标记 dNTPs 的平均 lag 值 (0.7%) 低很多。同时，加 cold dNTPs 的平均 runon 值 (0.06%) 比加荧光标记 dNTPs 的平均 Runon 值 (0.24%) 低很多。CoolMPS 技术的低 lag 及低 runon 可能使得测序读长更长。



图 7 DNA 聚合酶对 Cold dNTP 与 Labeled dNTPs 的聚合的 lag 和 runon 的效率比较

◎ 硬件平台

采用优化设计的光学及生化系统，可在较短时间内完成完整的测序过程，且平台自带初级数据分析软件，可以自动完成图像分析，并产生有质量打分的碱基序列实现精简流畅的测序体验。



◎ 性能参数

载片类型	试剂类型	有效 Reads 数 *	读长	数据产出	运行时间 **	Q30***	
FCS	StandardMPS	550M	SE100	55G	13H	>85%	
			PE100	110G	26H	>85%	
			PE150	165G	37H	>75%	
FCL		1500-1800M	SE50	75-90G	14H	>85%	
			SE100	150-180G	25H	>85%	
			PE100	300-360G	48H	>85%	
			PE150	450-540G	66H	>75%	
			SE400	600-720G	109H	>70%	
			PE200	600-720G	107H	>75%	
		CoolMPS	1500-2000M	SE50	75-100G	17H	>90%
				SE100	150-200G	30H	>90%
				PE100	300-400G	58H	>90%

* 有效 Reads 数的最大值根据特定标准文库运行所得，实际应用文库受样本类型、建库方式会有所波动。

** 运行时间根据双载片进行统计，包括：样本加载、测序、碱基识别和数据处理时间。

*** 高于 Q30 的碱基百分比是特定标准文库通过整个运行平均所得。实际应用表现受样本类型，文库质量，插入片段长度等因素影响。

◎ 适配应用

MGISEQ-2000 支持 6 种读长，从生育检测和病原快检（SE50、SE100），到肿瘤检测、转录组、WGBS、WES、WGS（PE100、PE150），再到个体识别（SE400）以及动植物基因组测序（PE200），全面地覆盖了科研领域和临床领域的应用需求。

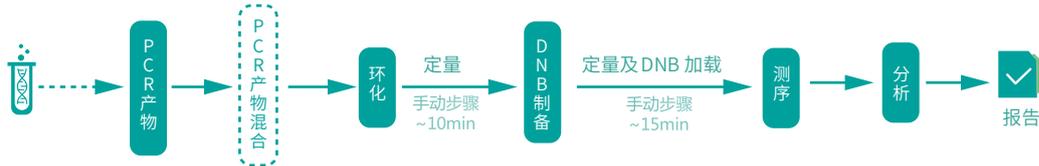
应用类型	推荐读长	MGISEQ-2000 平台单次运行可支持的推荐样本数 *				
		1 张 FCS	2 张 FCS	1 张 FCL	1 张 FCL+1 张 FCS	2 张 FCL
		550M reads	1100M reads	1800M reads	2350M reads	3600M reads
无创产前筛查 5M reads/ 样本	SE50	85 样本	170 样本	275 样本	360 样本	550 样本
病原微生物快检 25M reads/ 样本	SE50/SE100	17 样本	34 样本	55 样本	72 样本	110 样本
单细胞 RNA-Seq 5000 个细胞, 100K reads/ 细胞, 100 Gb/ 样本	PE100	1 样本	2 样本	4 样本	5 样本	8 样本
肿瘤小 panel 1 Gb/ 样本		85 样本	170 样本	275 样本	360 样本	550 样本
肿瘤大 panel 5 Gb/ 样本		17 样本	34 样本	55 样本	72 样本	110 样本
转录组 40M reads/ 样本	PE150	11 样本	22 样本	35 样本	46 样本	70 样本
全外显子组测序 100× 平均测序深度, 15 Gb/ 样本		8 样本	16 样本	28 样本	36 样本	56 样本
全基因组测序 30× 平均测序深度, 100 Gb/ 样本		1 样本	2 样本	4 样本	5 样本	8 样本
个体识别 1M reads/ 样本	SE400	/	/	1500 样本	/	3000 样本

* 综合考虑 Pooling 偏差及其应用场景后估算得出的样本个数，仅供参考，用户可根据实际需求调整。

解决方案

全能建库测序分析方案

基因测序仪MGISEQ-2000RS 2006A 全程自动化,全场景适用



- 从提取到报告全流程自动化操作, 简单易懂, 提升效率
- 可靠的建库、测序和分析流程, 助力低成本获取高准确度测序分析结果
- 提供多种测序、分析模式选择, 可拓展多种应用场景

注: 代表自动化步骤
 代表即可自动化又可手动的步骤

灵活测序分析方案

基因测序仪MGISEQ-2000RS 2006B 超强分析

流程自动化

从文库到分析报告, 通过 Zlims 控制全流程自动化

使用独立性好

既可作为生信加速器, 也可作为分析服务器, 与测序仪运行互不干扰

解决方案丰富

可集成多种 MGI 或第三方软件

分析兼容性强

不限制数据来源, 均可使用 MegaBOLT 分析

维护更简单

外置服务器, 方便独立维护



流畅建库测序方案

基因测序仪MGISEQ-2000RS 2006C解放双手

提取建库测序自动化

减少 90% 以上的手工操作时间

手动操作时间短

从样本到报告全程手动操作时间不超过 30 分钟

可拓展性强

支持定制化样本制备流程开发



从样本到文库 自动化样本提取与建库

MGISP-100 是全球第一台获得 NMPA 认证的自动化样本制备系统, 该工作站搭载 8 通道固定间距移液器, 单次运行最大支持 16 个样本, 采用自动化的流程设计, 全面提升实验室整体工作效率。

- 一机整合: 包含自动化建库 + 提取, 自动化程度高达 90% 以上
- 功能全面: 可自动化完成一系列实验操作, 包括核酸提取、磁珠纯化、酶反应、PCR 反应、片段筛选、杂交捕获以及 DNB 制备
- 开放平台: 根据不同需求支持定制化开发, 70+ 验证脚本, 强大的拓展性



自动化样本制备系统
MGISP-100

从文库到数据 精准的测序结果

MGISEQ-2000 是一款准确、全面、灵活的桌面型基因测序仪, 适用于科研用户平台建设, 亦可满足临床样本快速交付, 可极大的活跃你的日处理能力。

- DNBSEQ™技术: 高准确性、低重复序列率、低标签跳跃
- 更高测序质量: 独有的 CoolMPS 测序试剂进一步提升测序准确性
- 灵活运转: 高性价比的大载片 (FCL) 和运行快速的小载片 (FCS), 支持单 / 双载片, 五种测序模式可选, 通量覆盖 550M ~ 3600M reads/run, 读长覆盖 SE50、SE100、SE400、PE100、PE150 和 PE200



基因测序仪
MGISEQ-2000

从数据到报告 高效的分析方案

MegaBOLT 生信分析加速器, 专注于加速高通量测序数据分析, 是兼顾设备、样本、应用、数据管理的一体化工作站, 并整合实验室信息管理系统 ZLIMS-Lite, 部署灵活, 使用便捷, 并可拓展支持更多增值应用。

解决方案所包含的应用*:

Germline/Somatic WGS/WES分析

- 完成下机数据至 SNP、InDel 的变异检测, 含 MegaBOLT-DV 深度学习超高精度模块

病原微生物快检PFI

- Meta 样本的病原快检应用

多重新冠分析 MGI-SARS-CoV-2

- 鉴定 SARS-CoV-2 序列, 并报告测序样本的感染状态

* 更多应用将持续上线 注: 仅供科研使用



生信分析加速器
MegaBOLT

快速
省时
20 倍以上加速

集成
全自动一站式
测序 + 分析
丰富的应用

易用
操作友好的
web 端设计

经济
分析费用节省低至
十分之一

应用案例

◎ 全基因组测序

案例1: 人重测序

样本来源: 1025 个中原汉族人 DNA 样本

建库方法: MGIEasy PCR-Free DNA 文库制备试剂套装

测序策略: MGISEQ-2000 PE150

表 1-1 测序数据质量

		Min	Median	Mean	Max	High quality	PASS
Total Reads		601727956	726056164	726494436	952285662	/	/
Mean Reads Length		150	150	150	150	/	/
Reads*	R1	100%	100%	100%	100%	=100%	=100%
	R2	100%	100%	100%	100%	=100%	=100%
Q30	R1	87.21%	90.43%	90.34%	92.91%	>=85%	>=80%
	R2	84.22%	89.79%	89.56%	92.00 %	>=85%	>=80%

*passed filter

表 1-2 测序数据分析关键指标

	Min	Median	Mean	Max	High quality	PASS
Properly Paired	96.38%	98.27%	98.26%	98.88%	>=95%	>=90%
Raw Depth (GRCh38)	29.23	35.27	35.29	46.25	>=30	>=10
Mapping Rate	97.81%	99.99%	99.99%	100.00%	>=99%	>=95%
Duplication	0.25%	0.88%	0.99%	3.22%	/	/
Mean Insert Size	262.38	329.40	332.64	382.18	/	/
Insert Size SD	51.23	71.40	71.75	82.80	/	/

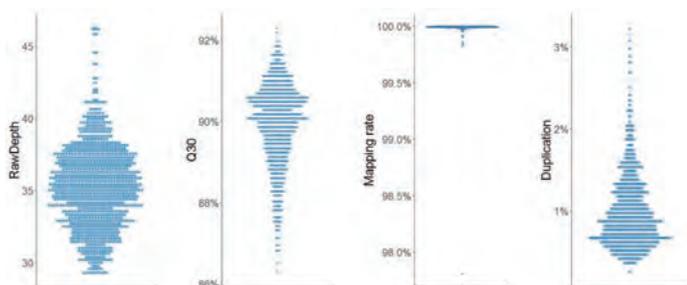


图 1 总体测序质量优

结论:

- 超过98%的样本Raw Depth都达到了30X以上,最小的也有29.226X
 - 全部样本碱基Q30的比例均达到了85%以上
 - 除了有一个样本的mapping rate在97%,其他的均大于99.5%
 - 超过95%的样本duplication比例小于2%,最大的也在4%以下
- 基于 MGISEQ-2000 的全基因组测序整体测序质量优秀,raw data、Q30和比对率均处于高水平,且Duplication比例低,可产出准确可靠的全基因组测序的数据。

案例2: 动植物重测序

Hi-C 技术是以整个细胞核为研究对象，利用高通量测序技术，结合生物信息学方法，研究全基因组范围内整个染色质 DNA 在空间位置上的交互关系。

样本来源：鱼

建库方法：Hi-C 文库进行 App-A 文库转化

测序策略：MGISEQ-2000 PE200, N 平台 PE150

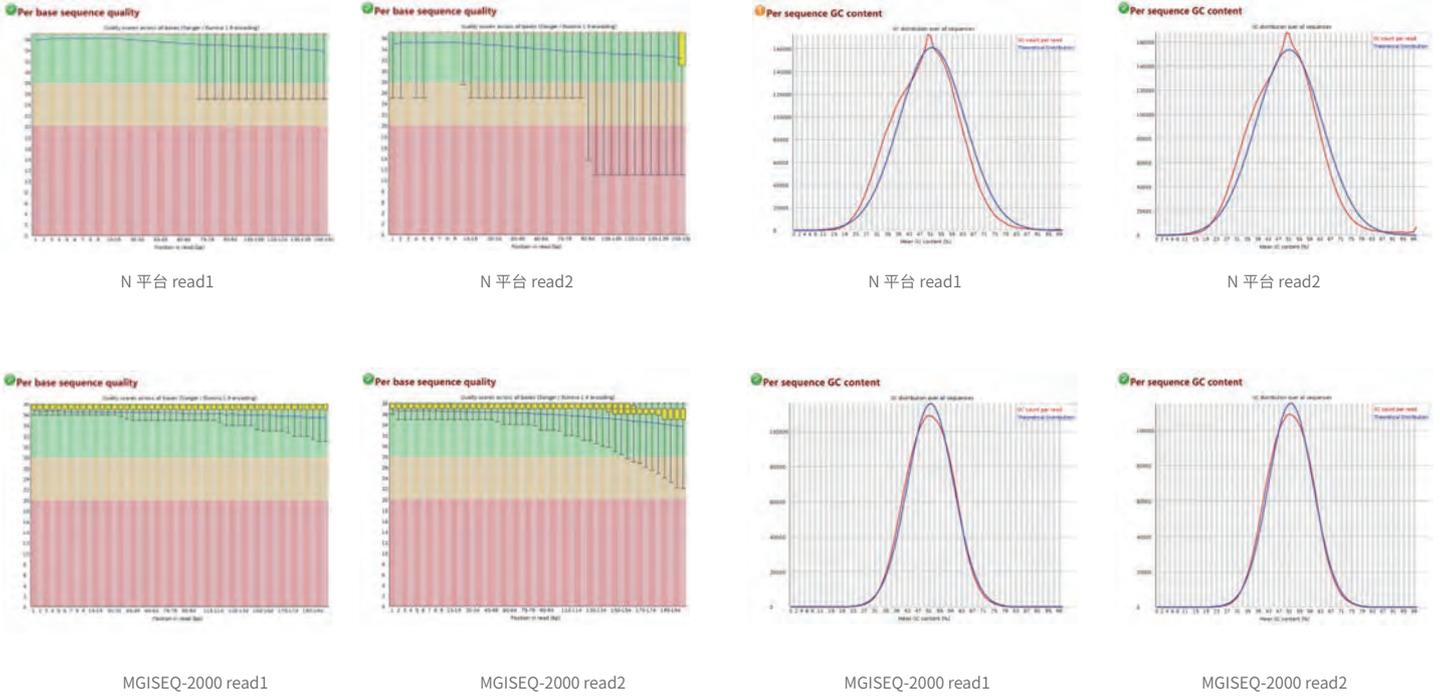


图 2-1 碱基测序质量
MGISEQ-2000 测序质量优于 N 平台

图 2-2 序列平均 GC 分布
MGISEQ-2000 GC 分布更平滑

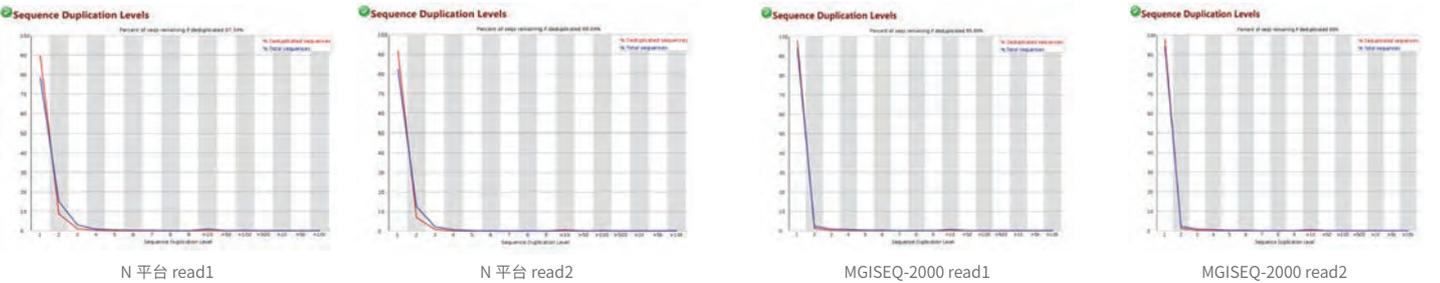


图 2-3 Sequence duplication levels
MGISEQ-2000 duplication rate 更低

案例3: stLFR

stLFR(single tube Long Fragment Read) 是由华大智造研发的无分割共标签长片段读取技术, 搭配 DNBSEQ 系列测序平台, 可通过短读长测序获取长片段 DNA 信息, 从而实现一次测序获得高准确度的 SNP/InDel/CNV/SV 变异检测结果, 以及单倍体分型结果, 为大规模 WGS 提供新的解决方案。

文章: Huang, Chuanfeng, et al. "An integrated Asian human SNV and indel benchmark established using multiple sequencing methods." Scientific reports 10.1 (2020): 1-11.

样本来源: 中国食品药品检定研究院提供的来自中国北京的汉族男性志愿者的细胞系 (XHEC-C-2019-086, HJ)

建库方法: 常规 MPS 建库、stLFR 建库、HiFi CCS 建库

测序策略: BGISEQ-500; DNBSEQ-G400* (stLFR); N 平台; P 平台

* MGISEQ-2000 在欧盟区域及部分海外国家产品名称切换为 DNBSEQ-G400

30X 测序深度的情况下, 各个平台对于基因组中大约 44.41Mb 的盲区覆盖度的比较结果中, 常规 WGS 数据覆盖度低于 2%, 无法有效覆盖盲区, 但 stLFR 与 CCS 覆盖率分别是 73.3% 及 68.53%, stLFR WGS 数据表现非常优秀。

DNBSEQ-G400 stLFR 数据在盲区的覆盖度优于常规短读长平台数据。

表 3 CCS 和 stLFR 在盲区的覆盖度

	CCS	stLFR
Length(bp)	30403236	32519955
Percentage(%)	68.53	73.3

从 SNV 和 InDel 的比较结果来看, stLFR 与 CCS 的检测结果优于短读长平台。从下图可以看出, stLFR WGS 特有的 SNV 和 InDel 数量超过 20 万个, 数量远超常规 WGS。

DNBSEQ-G400 stLFR 数据在 SNV 和 InDel 的检测能力优于常规短读长数据。

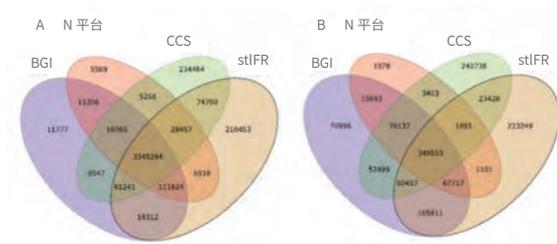


图 3-1 各个平台检测到的 SNV 和 InDel 变异结果展示

在盲区中的基因及其变异, 有很多和人类疾病相关, 举例如下:

NBPF4: 神经母细胞瘤断点基因家族 (NBPF) 的一员, 该家族由几十个相近重复的基因组成, 主要位于人类 1 号染色体的重复片段中。各平台的检测结果显示, 只有 stLFR 和 PB CCS 测序完整检出。

DNBSEQ-G400 stLFR 数据检测疾病相关基因能力优于常规短读长数据。

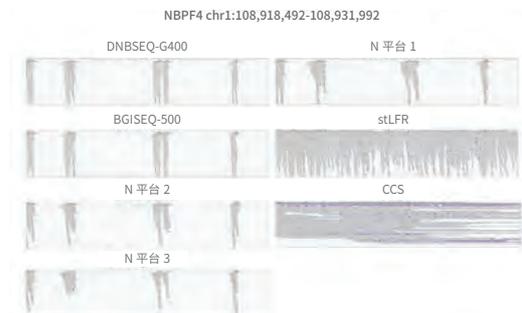


图 3-2 不同平台在 NBPF4 基因区的覆盖度和测序深度的 IGV 结果图

◎ 外显子测序

案例4: IDT探针全外测序

样本来源: NA12878

捕获探针: IDT 全外探针

建库方法: MGIEasy 外显子组通用文库制备试剂套装

测序策略: MGISEQ-2000 PE150

表 4 测序数据分析结果

Sample	1	2	3	4	5	6	7	8	9	10	11	12
杂交基数	12 杂											
探针	IDT 全外											
Q20_clean	95.73%	95.67%	95.71%	96.00%	95.73%	95.57%	95.74%	95.86%	94.83%	94.82%	94.89%	95.18%
Q30_clean	86.72%	86.52%	86.67%	87.34%	86.66%	86.29%	86.72%	87.00%	84.46%	84.45%	84.58%	85.31%
GC_clean	51.76%	52.08%	52.16%	52.24%	52.05%	52.12%	52.11%	52.32%	51.40%	49.98%	49.59%	51.08%
Mapping_Rate	98.99%	99.04%	99.00%	99.08%	99.04%	99.00%	99.02%	99.04%	98.78%	98.73%	98.85%	98.84%
PE_Mapping_Rate	98.28%	98.34%	98.29%	98.44%	98.34%	98.26%	98.32%	98.36%	97.88%	97.84%	98.02%	98.02%
Duplication_Rate	2.91%	1.95%	1.87%	2.41%	1.99%	2.23%	2.34%	2.58%	2.69%	1.95%	2.79%	2.36%
Mismatch_Rate	0.66%	0.66%	0.66%	0.63%	0.65%	0.67%	0.65%	0.64%	0.76%	0.76%	0.75%	0.73%
Capture_Rate_on_Reads	74.08%	74.80%	74.93%	75.61%	74.88%	74.93%	75.25%	75.60%	72.38%	70.59%	71.38%	72.59%
Capture_Rate_on_Bases	51.37%	52.00%	52.10%	52.60%	52.04%	52.10%	52.32%	52.64%	50.10%	48.61%	49.12%	50.26%
Average_depth(rmdup)	101.57	101.73	99.39	99.71	99.25	100.49	100.64	99.18	100.42	99.52	99.65	99.72
Coverage(>=1X)	99.02%	99.02%	99.02%	99.02%	99.02%	99.02%	99.02%	99.02%	99.06%	99.06%	99.06%	99.05%
Coverage(>=20X)	98.22%	98.13%	98.04%	98.10%	98.10%	98.10%	98.13%	97.99%	98.19%	98.44%	98.52%	98.36%

结论:

MGISEQ-2000 平台可以产出高质量的 WES 测序数据, 对比率和捕获率高, 错误率和 duplication 低, 测序覆盖度好。

案例5: CoolMPS 全外测序

对基因组 DNA 样本 Y1、Y2 样本采取 NadPrep for MGI_UDI + IDT Exome v2 的捕获测序方案，然后将同一捕获文库在 N 平台及 MGISEQ-2000 平台分别进行测评。

捕获方法：IDT Exome v2

建库方法：纳昂达 NadPrep for MGI_UDI

测序平台：MGISEQ-2000, N 平台

测序试剂：CoolMPS 高通量测序试剂套装 (MGISEQ-2000RS FCL PE100)，N 平台测序试剂将测序数据直接比对到 UCSC 的 hg38 参考基因组后，分析去重后覆盖深度，如图 5-1 所示。

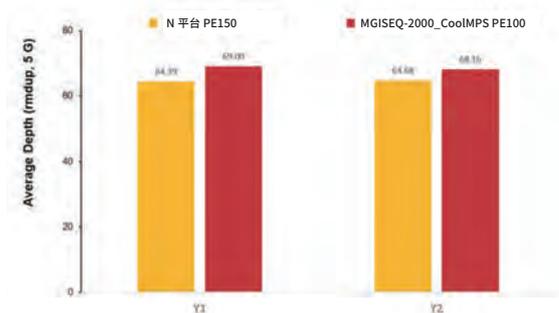


图 5-1 IDT Exome v2 在 N 平台及 MGISEQ-2000 平台的平均覆盖深度

分别用 NadPrep for MGI 对 Y1、Y2 建库后进行杂交捕获，同一杂交文库分别采用 MGISEQ-2000 CoolMPS (PE100, ECR4.0) 和 N 平台 (PE150) 进行测序，取 5 Gb 数据进行分析。

采用 CoolMPS 测序试剂套装的 MGISEQ-2000 的有效覆盖深度上相对更高。

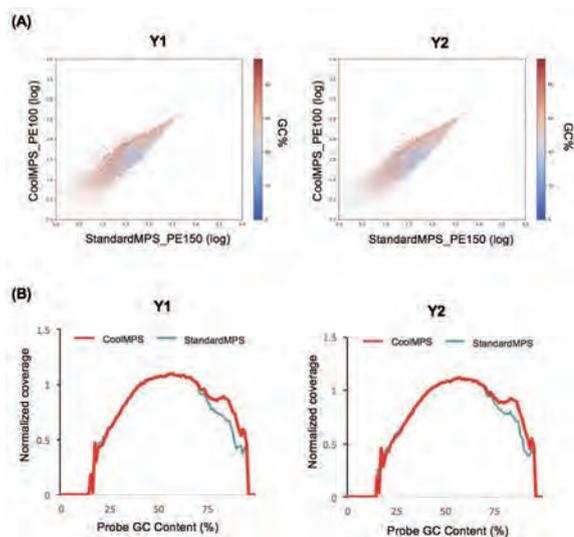


图 5-2 IDT Exome v2 在 MGISEQ-2000 平台 CoolMPS 和 StandardMPS 测序试剂套装下的 GC 偏好性

A. 相同探针区域在测序数据中覆盖深度的比较。颜色代表探针 GC 含量。

B. 不同 GC 含量区域覆盖。

采用 CoolMPS 测序试剂套装的 MGISEQ-2000 在 GC-rich 区域的表现得到了一定程度的整体提升。

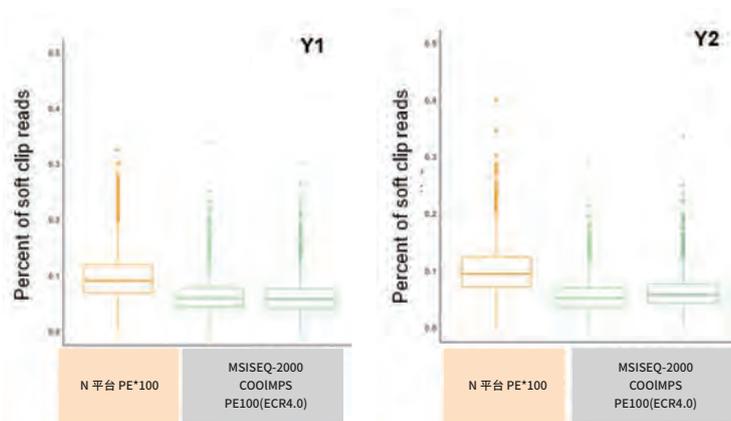


图 5-3 不同测序平台与测序试剂微卫星位点的 Soft clip reads 占比。微卫星筛选标准如下：A/T >= 15 bp。数据比对后的 bam 中的 Soft clip reads 相对微卫星区域总 reads 数的比率。PE*100 原始 PE150 取前 100 碱基。

与 N 平台相比，MGISEQ-2000 重复碱基读取能力相对更有优势。

◎ RNA测序

案例6: Bulk RNA-Seq

样本来源: UHRR

建库方法: MGIEasy 系列文库在 MGISEQ-2000 平台测序

N 系列文库分别采用 A 和 B 品牌同类建库产品构建, 在 N 平台测序

测序策略: MEISEQ-2000 PE150

实验方法: 相同样品 UHRR, 起始量 300ng, 分别在 MGISEQ-2000 和 N 平台测序 PE150, 数据量统一到 10G 后进行分析

测序质量

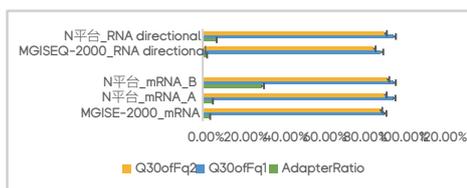


图 6-1 MGISEQ-2000 平台与 N 平台测序质量比较

比对率

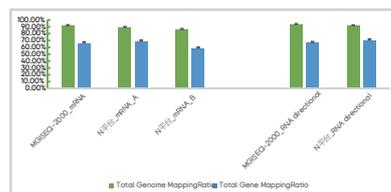


图 6-2 MGISEQ-2000 平台与 N 平台比对率比较

基因检出数

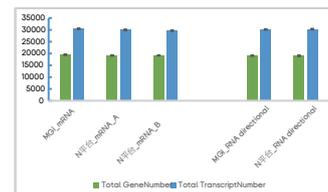


图 6-3 MGISEQ-2000 平台与 N 平台基因检出数比较

相关性

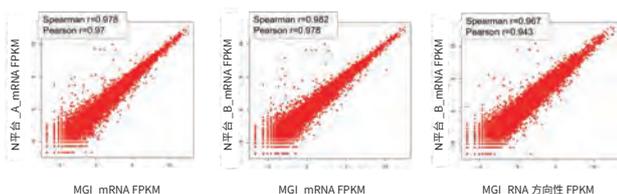


图 6-4 MGISEQ-2000 与 N 平台相关性比较

基因覆盖情况

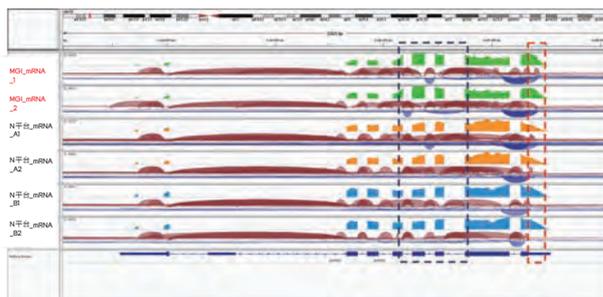


图 6-5 管家基因 GAPDH 覆盖情况

对于管家基因 GAPDH, MGI 试剂盒具有更好的基因覆盖度, 尤其是 3' 端; 有更多覆盖了 Junction 区域的序列。

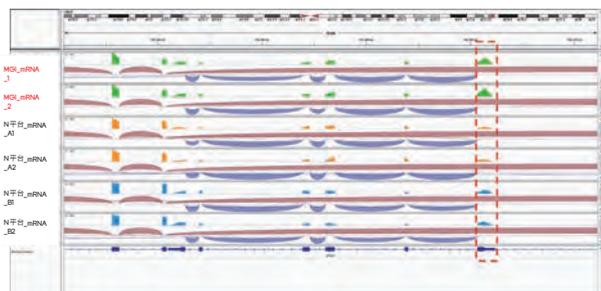


图 6-6 中等表达量的基因, 如 GFM1 覆盖情况

对于中等表达量的基因, 如 GFM1, MGI 试剂盒与竞品基因覆盖度基本一致, 3' 端覆盖度优于竞品。

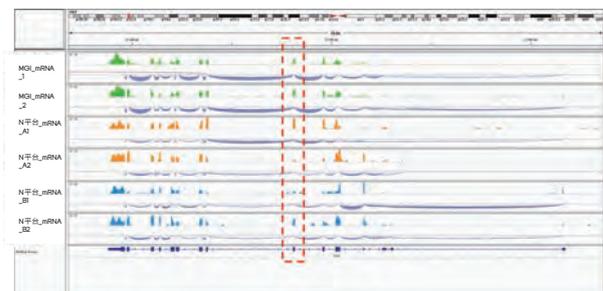


图 6-7 低表达量的基因, 如 FGR 覆盖情况

对于低表达量的基因, 如 FGR, MGI 试剂盒具有更均匀的基因覆盖度。

结论:

- 两个平台测序质量基本一致, Q30 约为 90%, MGI 平台接头污染率更低;
- 两个平台基因组和基因比对率基本一致, 基因和转录本检出数基本一致;
- 两个平台在基因定量相关性上, mRNA 的相关性 >0.97, RNA 方向性的相关性 >0.94;
- 与其他平台比较, MGISEQ-2000 具有优异的基因覆盖度, 尤其 3' 端覆盖度更好。

案例7: DNBelab C4 RNA测序

DNBelab C 系列细胞组学整体解决方案，基于独特的 DNBelab C 系列单细胞文库制备技术和强大的 DNBSEQ 测序技术，由 DNBelab C4 便携式单细胞系统、DNBelab C 系列单细胞 RNA 文库制备套装，搭配华大智造 DNBSEQ 系列测序平台，以及相关的单细胞分析软件构成，可实现便携式、即时化、一站式单细胞组学研究全流程。

样本来源：人 293T 细胞系：鼠 3T3 细胞系 =1:1

建库方法：DNBelab C 系列 RNA 文库制备套装

测序策略：MGISEQ-2000 PE100

表 7-1 细胞数据结果

Estimated Number of Cells	2,544
Estimated Number of Human Cells	1,309
Estimated Number of Mouse Cells	1,155
Fraction Reads in Cells	79.70%
Fraction Reads in Human Cells	80%
Fraction Reads in Mouse Cells	79.30%
Mean Reads per Cell	40,882
Mean Reads per Human Cell	41,485
Mean Reads per Mouse Cell	39,518
Median UMI Counts per Human Cell	28,915
Median UMI Counts per Mouse Cell	26,411
Median Genes per Human Cell	6,168
Median Genes per Mouse Cell	5,395

表 7-2 测序数据结果

Number of Reads	172,491,759
Reads Pass QC	148,377,085
Reads with Valid Barcodes	148,377,085
Filtered Reads with Failed Barcodes	23,040,453
Filtered Reads with Low Quality	1,074,221
Filtered Reads with Unknown Sample Barcodes	0
Q30 Bases in Cell Barcode	89.30%
Q30 Bases in Sample Barcode	0.00%
Q30 Bases in UMI	86.50%
Q30 Bases in RNA Read	79.80%

表 7-3 比对结果

Reads Mapped Confidently to Genome	95.20%
Reads Mapped Confidently to Gene	95.10%
Reads Mapped Confidently to Exonic Regions	66.30%
Reads Mapped Confidently to Intronic Regions	2.80%
Reads Mapped Antisense to Gene	5.70%

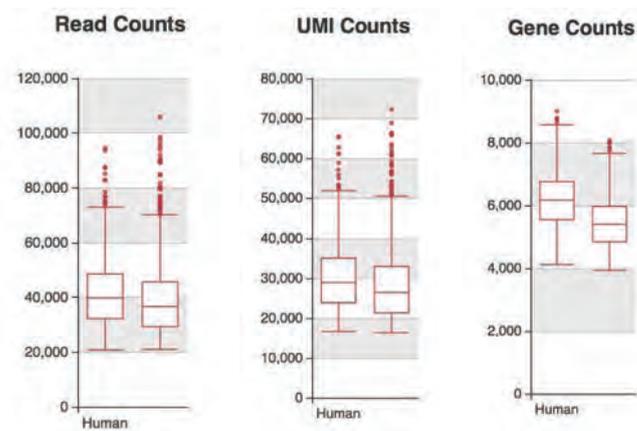


图 7 数据分布情况

结论:

华大智造 DNBelab C4 单细胞建库产品在 MGISEQ-2000 平台数据量及质量表现符合预期。

案例8: 10x空间转录组

样本来源: 小鼠组织

建库方法: 10x genomics 空间转录组

测序策略: MGISEQ-2000 PE100

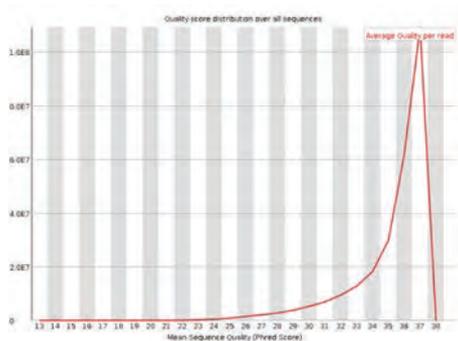


图 8-1 整体测序质量分布

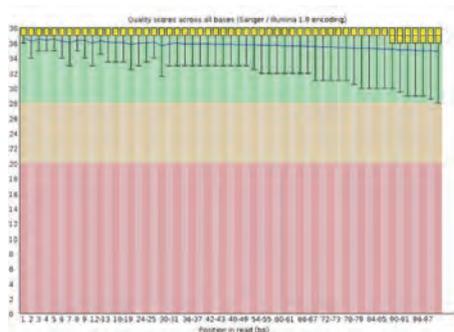


图 8-2 碱基测序质量分布

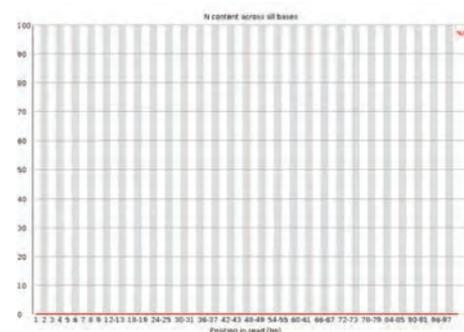


图 8-3 N 碱基的含量

结论:

测序质量峰值及每条 reads 的不同位置上的碱基的测序质量大部分超过 Q30，即绝大部分 reads 的准确度超 99.9%。

测序质量稳定，准确性在测序后期也没有明显下降。

测到的 N 碱基含量很低，说明测序的可信度高。

基于 MGISEQ-2000 的单细胞转录组测序整体测序质量优秀，准确性强，可信度高。

◎ 甲基化测序

案例9: gDNA 甲基化测序

样本来源: NA12878

建库方法: MGIEasy 全基因组甲基化文库制备试剂盒

测序策略: MGISEQ-2000 vs N 平台

表 9-1 MGISEQ-2000 与 N 平台测序结果比较

#Sample	MGI_CoolMPS_PE100	MGI_PE150	N 平台_PE150
Sequencer	MGISEQ-2000	MGISEQ-2000	N 平台
Average_MappedReads_Length(bp)	100	150	150
Clean_Rate(%)	99.29	95.73	99.41
Clean_reads	1025773790	733333344	733333334
Mapped_reads	937684291	666856392	601991927
Mapping_rate(%)	91.41	90.93	82.09
Dup_rate(%)	14.38	11.22	16.82
Unique_Mapped_reads	893465925	645354142	581230354
Unique_Mapped_bases(bp)	89346592500	96803121300	87184553100
Unique_Mapping_rate(%)	87.1	88	79.26
Pair_Mapped_reads	810688002	620914692	507611396
Pair_Mapped_rate(%)	79.03	84.67	69.22
Single_Mapped_reads	126996289	45941700	94380531
Single_Mapped_rate(%)	12.38	6.26	12.87
Mismatch_bases(bp)	669511867	753366881	813154657
Mismatch_rate(%)	0.71	0.75	0.9
BS_conversion_rate(%)	99.29	99.4	99.56

表 9-2 MGISEQ-2000 与 N 平台数据分析结果比较

#Sample	MGI_CoolMPS_PE100	MGI_PE150	N 平台_PE150
Sequencer	MGISEQ-2000	MGISEQ-2000	N 平台
Average_MappedReads_Length(bp)	100	150	150
Clean_reads	1025773790	733333344	740000000
Initial_bases_in_genome(bp)	2861392204	2861392204	2861392204
Mapped_bases(bp)	75844315400	84777839100	71761247100
Covered_bases(bp)	2798579563	2814517888	2815568908
Coverage(%)	97.805	98.362	98.399
Average_sequencing_depth	26.51	29.63	25.08
Fraction_covered_with_at_least_20x(%)	76.29	79.541	61.768
Fraction_covered_with_at_least_10x(%)	94.014	92.634	91.702
Fraction_covered_with_at_least_4x(%)	96.913	96.753	97.095

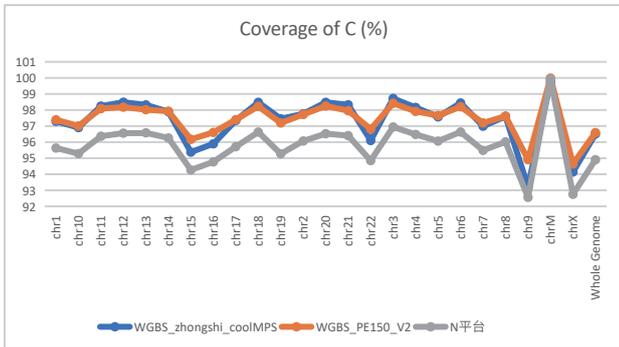


图 9-1 MGISEQ-2000 与 N 平台 C 位点相关的甲基化覆盖度比较

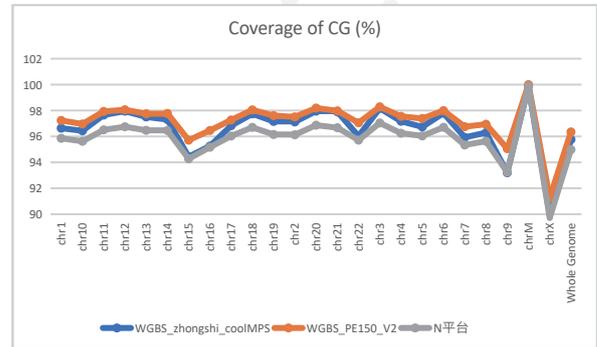


图 9-2 MGISEQ-2000 与 N 平台 CG 位点相关的甲基化覆盖度

结论:

MGISEQ-2000 平台 WGBS 表现出更高的比对率、更低的重复序列率及错误率；
MGISEQ-2000 平台 WGBS 表现出更高的与 C/CG 位点相关的甲基化覆盖度。

◎ 生育方向

案例10: 遗传性疾病检测——脆性X染色体综合征

发病率: 1/1000~1/1500, 仅次于先天愚型

发病机理: 致病基因为 FMR1, 含有 (CGC)_n 三核苷酸重复序列, 在遗传过程中, (CGC)_n 重复异常, 会导致患病; 女性携带者的重复数越大, 越容易生出患儿

样本来源: FMR1 基因标准品 gDNA

建库方法: MGIEasy PCR-Free DNA 文库制备试剂套装

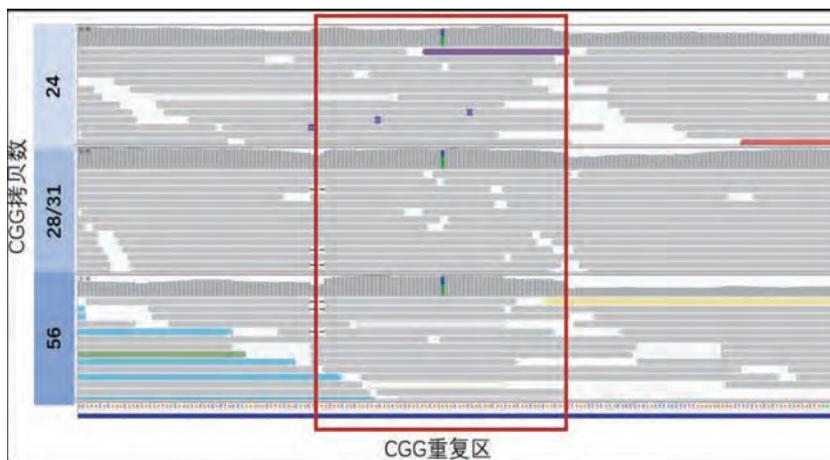


图 10 检出的 CGG 拷贝数

结论:

MGISEQ-2000 平台 PCR-free 建库检出的 CGG 拷贝数的结果与理论值一致。

◎ 病原方向

案例11: 新冠检测

文章: Hu, Fengyu, et al. "A compromised specific humoral immune response against the SARS-CoV-2 receptor-binding domain is related to viral persistence and periodic shedding in the gastrointestinal tract." Cellular & Molecular Immunology (2020): 1-7.

样本来源: 广州市第八人民医院收治的中国 COVID-19 患者

建库方法: 华大智造 ATOPlex 多重 PCR 平台

测序平台: MGISEQ-2000

分析软件: nCoV Finder pipeline (基因组组装)

nCoV Variant detection pipeline (变异检测)

SARS-CoV-2 Multi-PCR v1.0 (变异检测)

表 11 部分复检阳性患者样本 SARS-CoV-2 测序结果

Sample representative	Patient	Sample type	Ct value (N)	Ct value (orf1ab)	Clean reads	SARS-CoV-2 reads	SARS-CoV-2 reads (RPM)	sgmRNA reads	sgmRNA reads / clean reads (RPM)	Coverage (>=100x)	Assembly length (bp)
yes	Patient 01	Throat swab	35	38	32,510,228	217,194	6,681	8	0	39.33%	11,762
yes	Patient 02	Anal swab	35	37	26,340,150	555,489	21,089	75	3	91.51%	27,363
yes	Patient 03	Anal swab	31	33	20,391,376	6,812,954	334,110	496	24	99.11%	29,638
yes	Patient 04	Anal swab	38	40	27,205,700	8,422	310	0	0	2.31%	692
yes	Patient 05	Anal swab	36	38	60,091,632	51,188	852	0	0	21.66%	6,478
yes	Patient 06	Anal swab	34	37	86,500,624	10,674,143	123,400	802	9	99.44%	29,737
yes	Patient 07	Throat swab	40	negative	40,103,674	125,579	3,131	5	0	48.54%	14,514
yes	Patient 08	Anal swab	33	34	21,685,592	15,425,408	711,321	1,238	57	99.15%	29,649
yes	Patient 10	Throat swab	36	negative	27,268,380	2,426	89	0	0	1.42%	424
yes	Patient 11	Anal swab	35	41	58,483,382	484,586	8,286	18	0	77.91%	23,297
yes	Patient 12	Throat swab	37	40	25,781,818	817,946	31,726	31	1	95.66%	28,606
yes	Patient 13	Anal swab	38	40	22,708,342	22,542	993	0	0	10.20%	3,049
yes	Patient 14	Anal swab	33	36	12,136,122	579,353	47,738	21	2	91.23%	27,280
yes	Patient 15	Throat swab	35	37	21,996,818	127,984	5,818	2	0	37.83%	11,313
yes	Patient 16	Anal swab	37	40	27,757,050	474,841	17,107	7	0	83.48%	24,964
yes	Patient 18	Anal swab	36	38	23,167,418	239,350	10,331	27	1	41.72%	12,476
	PC1&	NA	32	NA	28,343,628	9,216,440	325,168	340	12	99.12%	29,640
	PC2&	NA	29	NA	23,862,664	22,548,264	944,918	2,301	96	99.16%	29,651
1st admission	Patient 03 §	Throat swab	40	negative	33,419,450	4,560	136	2	0	2.18%	652
1st admission	Patient 08 §	Throat swab	24	25	37,759,960	34,009,562	900,678	2,652	70	99.68%	29,796

&: PC1 and PC2. positive control, SARS-CoV-2 viral RNA extracted from cell culture.

§: samples from the first admssion

结论:

华大智造新冠 ATOPlex 多重 PCR 在 MGISEQ-2000 平台可以产出高质量的病毒序列数据, 搭配华大基于新冠检测的生信分析软件, 能够完成病毒基因组的准确组装和变异检测, 协助 SARS-CoV-2 在人群中反弹机制的进一步探索, 助力病毒预防 and 有效疫苗的研发。

◎ 肿瘤方向

案例12: CoolMPS 检测低频突变

样本来源: 菁良泛肿瘤 800 gDNA
 文库类型: WES (IDT 探针)
 测序策略: N 平台 PE150, MGISEQ-2000 CoolMPS PE100
 测序深度: 700X (去重后)

表 12 不同平台特定位点突变检出情况

Gene	AA	STD_FREQ	N 平台	MGI_1	MGI_2
EGFR	V769_D770insASV	3%	1.73%	1.27%	2.32%
EGFR	L858R	1%	1.44%	0.62%	1%
EGFR	T790M	2%	2.00%	0.53%	1.49%
EGFR	E746_A750del	2%	NA	NA	1.03%
KRAS	G12D	2%	1.29%	1.93%	1.62%
KRAS	G13D	4%	3.67%	4.66%	2.60%
KRAS	A146T	1%	NA	2.02%	NA
NRAS	Q61K	1%	NA	1.26%	1.02%
KIT	D816V	2%	2.21%	2.61%	2.56%
FLT3	I836del	2%	2.16%	1.51%	1.75%
EGFR	G719S	4%	4.23%	2.86%	2.97%
BRAF	V600E	7%	4.34%	4.72%	5.47%
PIK3CA	H1047R	7%	7.76%	4.91%	10.14%

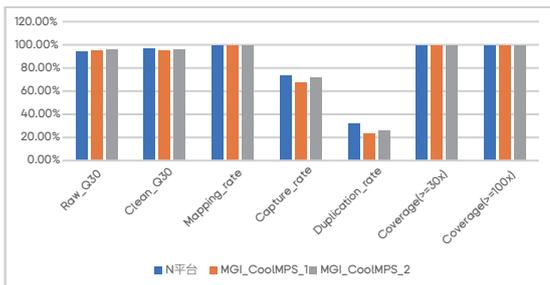


图 12 不同平台的质控数据对比

结论:

两个平台的 Q30, Mapping rate, Coverage 等表现基本一致, 但 MGISEQ-2000 平台 CoolMPS 试剂 Duplication rate 更低, 表现更好; 在 ddPCR 验证的 13 个位点中, 预设突变率在 1%-7% 的低频突变标准品中, N 平台检出率为 76.9% (10/13), MGISEQ-2000 CoolMPS 试剂检出率为 92.3%(12/13)。

案例13: 在结直肠癌中的应用

文章: Zheng, Kai, et al. "A novel NGS-based microsatellite instability (MSI) status classifier with 9 loci for colorectal cancer patients." Journal of Translational Medicine 18 (2020): 1-9.

样本来源: 64 例原发性结直肠癌患者样本, 包括 14 例 MSI-High 病例和 50 例微卫星稳定 (MSS) 病例
 建库方法: MGIEasy universal DNA library kit (MGI) 建库, xGen Hybridization and Wash Kit (IDT) 杂交捕获
 测序平台: MGISEQ-2000 PE100

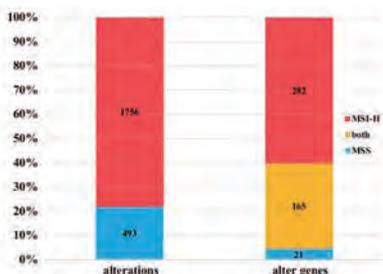


图 13-1 64 例结直肠癌病例中发生变化的位点与基因的示意图

在 64 例结直肠癌病例中, 有 468 个基因共 2249 个位点发生改变。突变在 MSI-H 样品中明显富集, 在 14 个 MSI-H 样品中发现 78% (447 个基因中 1756 个突变), 在 50 个 MSS 样品中发现 22% (186 个基因中 493 个突变)。MSI-H 病例的平均突变数为 125 (范围 63-302), MSS 为 10 (范围 1-26)。其中 60.2% (282/468) 的基因仅在 MSI-H 病例中被检测到, 而在 MSS 中只有 4.5% (21/468)。

Gene	Hotspot	Number of variant cases		
		Total	MSI-H	MSS
KRAS	pG12V/S/D/A	15	3	12
	pG13D	7	5	2
	pG61H	2	0	2
NRAS	pA146T	4	3	1
	pQ61R/H/V/L	3	0	3
BRAF	pV600E	4	4	0
TCF7L2	pE384fs	10	10	0
RNF43	pG939fs	7	7	0
TGFBR2	pE125fs	6	6	0
DNX1	pK125fs	3	3	0
CARD11	pR555fs	3	3	0
ICDRL1	pS1579fs	4	4	0
JAK3	pQ239fs	2	2	0
BCOR	pQ1156fs	2	2	0

pG12V/S/D/A, pG13D and pA146T of KRAS were found in both MSI-H and MSS cases, while BRAF pV600E was only found in a MSI-H case. pE384fs of TCF7L2 were highly mutated in MSI-H samples.
 R: fragment shift.

图 13-2 结直肠癌中的热点突变

在 85 个基因中检测到 90 个频发突变, 其中最高频的为 TCF7L2 中的 p.E384fs, RNF43 中的 p.G659fs 和 TGFBR2 中的 p.E125fs。这三个突变都位于单核苷酸重复区, 并分别在 10、7 和 6 例 MSI-H 病例中检测到, 这表明微卫星位点在 MSI-H 病例中通常是不稳定的。热点突变 p.G12V/S/D/A, p.G13D 和 p.A146T 在 MSI-H 和 MS 病例中均有检测到, 而 BRAF p.V600E 仅在 4 例 MSI-H 病例中检测到, 说明 BRAF p.V600E 可能与 MSI 有关。

结论: 基于华大智造 MGISEQ-2000 测序平台建立了一个新的基于高通量测序的 MSI 分类器 USCI-MSI, 包含 9 个微卫星位点用于检测 CRC 病例中的 MSI 状态。该方法具有 100% 的灵敏度和特异性, 在低纯度的肿瘤标本中表现良好。高 MSI 的肿瘤可能含有丰富的可以引起免疫反应的新抗原, 因此, 确定 MSI 状态可为确定哪些患者可能受益于免疫治疗提供了机会。

附录

◎ 参数规格

MGISEQ-2000基因测序仪

	机 型	预期用途
产品型号 *	MGISEQ-2000	IVD
	MGISEQ-2000RS	RUO
	MGISEQ-2000RS FAST	RUO
尺寸	1086 × 756 × 710 mm	
重量	200kg	
电 源	电源种类	100-240 V, 50/60 Hz
	运行功率	1200 VA
操作环境 **	环境温度	19°C-25°C
	相对湿度	20% RH-80% RH, 无冷凝
	气压范围	70 kPa-106 kPa
	防水等级	IPX0
控制电脑配置 ***	CPU	Intel Xeon E5 10Core * 2 2.2GHz
	内存	256 GB RAM
	机械硬盘	16 Tb
	固态硬盘	480 Gb
	操作系统	Windows 10 Enterprise

* 仅作型号区分

** 仅供室内使用；测序载片需要在 -25°C~-15°C的温度下储藏及运输，无需液态介质

*** 支持计算机配置、系统版本升级等

◎ 订购信息

MGISEQ-2000系列产品

货号	产品名称	货号	产品名称
900-000035-00	基因测序仪 MGISEQ-2000RS	1000018234	CoolMPS 高通量测序试剂套装 (MGISEQ-2000RS FCL SE50)
900-000208-00	基因测序仪 MGISEQ-2000RS FAST	1000018236	CoolMPS 高通量测序试剂套装 (MGISEQ-2000RS FCL SE100)
900-000033-00	基因测序仪 MGISEQ-2000	1000018238	CoolMPS 高通量测序试剂套装 (MGISEQ-2000RS FCL PE100)
1000012551	MGISEQ-2000RS 高通量测序试剂套装 (SE50)	1000019479	CoolMPS 高通量测序试剂套装 (MGISEQ-2000RS Small RNA FCL SE50)
1000012552	MGISEQ-2000RS 高通量测序试剂套装 (SE100)	1000019163	测序反应通用试剂盒 (联合探针锚定聚合测序法) 型号: G400 SM FCS SE100
1000012554	MGISEQ-2000RS 高通量测序试剂套装 (PE100)	1000019164	测序反应通用试剂盒 (联合探针锚定聚合测序法) 型号: G400 SM FCS PE100
1000012555	MGISEQ-2000RS 高通量测序试剂套装 (PE150)	1000019165	测序反应通用试剂盒 (联合探针锚定聚合测序法) 型号: G400 SM FCS PE150
1000013857	MGISEQ-2000RS 高通量测序试剂套装 (SE400)	1000012545	测序反应通用试剂盒 (联合探针锚定聚合测序法) 型号: G400 SM FCL SE50
1000013858	MGISEQ-2000RS 高通量测序试剂套装 (PE200)	1000012546	测序反应通用试剂盒 (联合探针锚定聚合测序法) 型号: G400 SM FCL SE100
1000011719	MGISEQ-2000RS 高通量快速测序试剂套装 (FCS SE100)	1000012548	测序反应通用试剂盒 (联合探针锚定聚合测序法) 型号: G400 SM FCL PE100
1000013155	MGISEQ-2000RS 高通量快速测序试剂套装 (FCS PE100)	1000012549	测序反应通用试剂盒 (联合探针锚定聚合测序法) 型号: G400 SM FCL PE150
1000011718	MGISEQ-2000RS 高通量快速测序试剂套装 (FCS PE150)	1000022460	测序反应通用试剂盒 (联合探针锚定聚合测序法) 型号: G400 CM FCL SE50
1000011545	MGISEQ-2000RS 高通量测序试剂套装 (stLFR) (PE100)	1000022461	测序反应通用试剂盒 (联合探针锚定聚合测序法) 型号: G400 CM FCL SE100
1000006138	MGISEQ-2000RS 高通量测序试剂套装 (SE50) (Small RNA)	1000022462	测序反应通用试剂盒 (联合探针锚定聚合测序法) 型号: G400 CM FCL PE100