

IDT xGen 16S Amplicon Panel v2在MGISEQ-200平台的实测数据展示

IDT产品匹配华大智造DNBSEQ测序平台赋能微生物群落鉴定

本研究基于IDT xGen 16S Amplicon Panel v2建库流程和MGISEQ-200测序平台，精准解析及再现了ZymoBIOMICS™ Microbial Community DNA Standard（D6305）标准品中微生物在种属水平上的组成和丰度情况，搭配IDT开发的16S SNAP APP分析工具，使得数据有效利用率接近100%，菌种分型检测结果更加准确。

推荐应用：细菌种属鉴定

推荐机型：MGISEQ-200，MGISEQ-2000和DNBSEQ-G99ARS

• 微生物鉴定精确度高

IDT xGen 16S Amplicon Panel v2目标区域设计覆盖16S rRNA基因区域的V1-V9区域，让微生物的鉴定更加精确。

• IDT方案可完美适配DNBSEQ测序平台

华大智造通用文库转换方案可使该试剂盒完美适配DNBSEQ测序平台。

• 测序数据质量高

DNBSEQ测序技术具有高准确性，低重复序列率、低标签跳跃等重要特性，MGISEQ-200平台以其灵活、小巧和快速的特点成为微生物测序的经典机型。

• 可匹配自动化操作

华大智造可为实验流程提供自动化解决方案，能够极大的节约人工成本并提高效率。



背景介绍

16S rRNA是原核核糖体30S亚单位的RNA成分，其编码基因为16S rRNA基因，该基因区域因进化速度缓慢而导致序列高度保守，常用于不同种类的细菌和古细菌的系统发育研究^{1,2}。细菌16S基因包含9个高变区(V1-V9)，每个高变区的长度约为30到100个碱基对，涉及小核糖体亚基的二级结构。这些高变区因序列高度特异能够被用于不同种属细菌鉴定³。

在研究中，可通过在保守区设计通用引物扩增高变区并开展16S rRNA基因测序以达到鉴定菌种的目的。目前，16S rRNA基因测序通常用于复杂生物混合物中的微生物的鉴定、分类和定量，如环境样本和肠道样本等^{4,5}。传统工作流程一般采用对9个高变区中的一到两个区域测序实现物种鉴定，所提供的菌种信息及检测准确性有限。IDT xGen 16S Amplicon Panel v2采用xGen扩增子技术，通过重叠扩增子的设计，实现了对16S rRNA基因区域9个高变区的单管扩增覆盖。除此之外，相对于覆盖V3、V4区域的设计，重叠式的“tilling”设计允许以更短的读长得到更多有用的菌种信息（图1上）。该方案推荐采用PE150读长进行测序，其间使用23对引物，可以获得平均大小为425bp的扩增子，覆盖16S rRNA全部的V1-V9区域，适用于医学、环境、农业、法医等场景下的环境微生物群落检测和鉴定（图1下）。

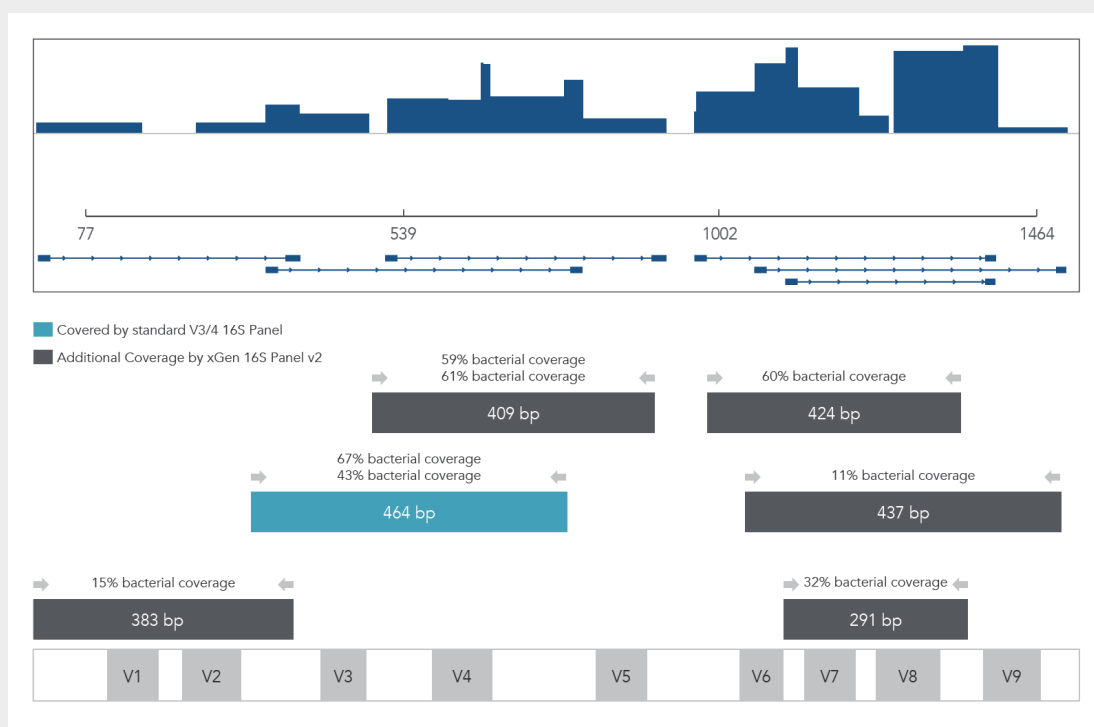


图1.IGV(Integrative Genomics Viewer)可视化展示，IDT xGen 16S Amplicon Panel v2覆盖16S rRNA的V1-V9区域(上)；相对于V3、V4(蓝色)扩增子设计，IDT xGen 16S Amplicon Panel v2使用6个扩增子、23对引物的设计(下)。

研究描述

华大智造推出的测序平台因其核心DNBSEQ测序技术从而具有高准确性、低重复序列率、低标签跳跃等重要特性。为了验证IDT推出的用于微生物鉴定的xGen 16S Amplicon Panel v2在DNBSEQ测序平台的性能表现，本研究基于商用ZymoBIOMICS™ Microbial Community DNA Standard标准品开展测试。测试结果验证发现IDT方案搭载DNBSEQ测序平台，在低至10pg的投入量、100k reads的测序量条件下，也可得到稳定的菌种鉴定结果；而在搭配16S SNAP APP生信分析软件后，可有效地提高数据的利用率。

测试方法

DNA样本获取

为测试IDT xGen 16S Amplicon Panel v2 在DNBSEQ测序平台上的性能表现，本研究基于ZymoBIOMICS™ Microbial Community DNA Standard（D6305）DNA标准品开展性能评估。该标准品由8种不同微生物DNA按特定比例混合而成（表1）。其中，Avg.GC指基因组DNA的GC平均值,Gram strain指革兰氏菌阴阳性，16S gDNA Abun. 指

Species	物种组成	Avg. GC (%)	Gram Stain	16S gDNA Abun. (%)
<i>Pseudomonas aeruginosa</i>	绿脓杆菌	66.2	-	4.2
<i>Escherichia coli</i>	大肠杆菌	56.8	-	10.1
<i>Salmonella enterica</i>	肠道沙门氏菌	52.2	-	10.4
<i>Lactobacillus fermentum</i>	发酵乳杆菌	52.8	+	18.4
<i>Enterococcus faecalis</i>	粪链球菌	37.5	+	9.9
<i>Staphylococcus aureus</i>	金黄色葡萄球菌	32.7	+	15.5
<i>Listeria monocytogenes</i>	李斯特菌	38	+	14.1
<i>Bacillus subtilis</i>	枯草芽孢杆菌	43.8	+	17.4
<i>Saccharomyces cerevisiae</i>	酿酒酵母	38.4	Yeast	NA
<i>Cryptococcus neoformans</i>	新型隐球菌	48.2	Yeast	NA

表1. D6305中的微生物组分

16S gDNA的丰度。

为检测该方案的性能情况，本研究将该标准品梯度稀释并取不同起始投入量（10ng、5ng、1ng和10pg）用于建库，测序及生物信息学分析。

文库制备与测序

本研究在华大智造客户体验中心（武汉）开展。每种DNA投入起始量设置三个实验重复（10ng为两个重复），利用IDT xGen 16S Amplicon Panel v2建库试剂盒通过两轮PCR反应（循环数分别为18和7）完成16S rRNA基因可变区域的富集，具体操作流程可以参考该试剂盒的说明书进行。随后利用MGIEasy通用文库转换试剂盒（App-A）通过PCR完成接头转换，将文库制备成适配DNBSEQ测序平台的测序文库，并完成DNA纳米球（DNB）的制备。在MGISEQ-200平台上完成双端150bp(PE150)的测序，保证每个样本的数据量为15M reads左右。

生物信息学分析

测序结果主要是利用IDT开发的16S SNAP APP分析工具完成。它是一种利用16S rRNA可变区分析微生物群落结构的自动化分析工作流，并且适用于双端测序，工作流的开始是利用Cutadapt工具对数据进行引物去除，随后对数据进行质量过滤、去噪、双端测序reads融合，再利用DADA2去除嵌合reads，确定扩增子变异序列(ASVs)，ASV通过VSEARCH删除重复序列，进而产生独特序列的测序子集，用于后续利用BASTn比对客户端16S数据库(RDP 11.5,https://rdp.cme.msu.edu)。不同于其他分析工具，16S SNAP APP工具会将下机得到的来自不同菌种不同区域的单独reads，组装成更长序列的“template”，使用组装后、

来自于不同V1-V9的“template” 比对16S数据库，可增加检测结果的准确性及数据的有效利用率。最终，使用RDP Classifier 对测序结果进行

归类进而产生一个标准格式的丰度表用于后续分析使用（分析流程见图2）。

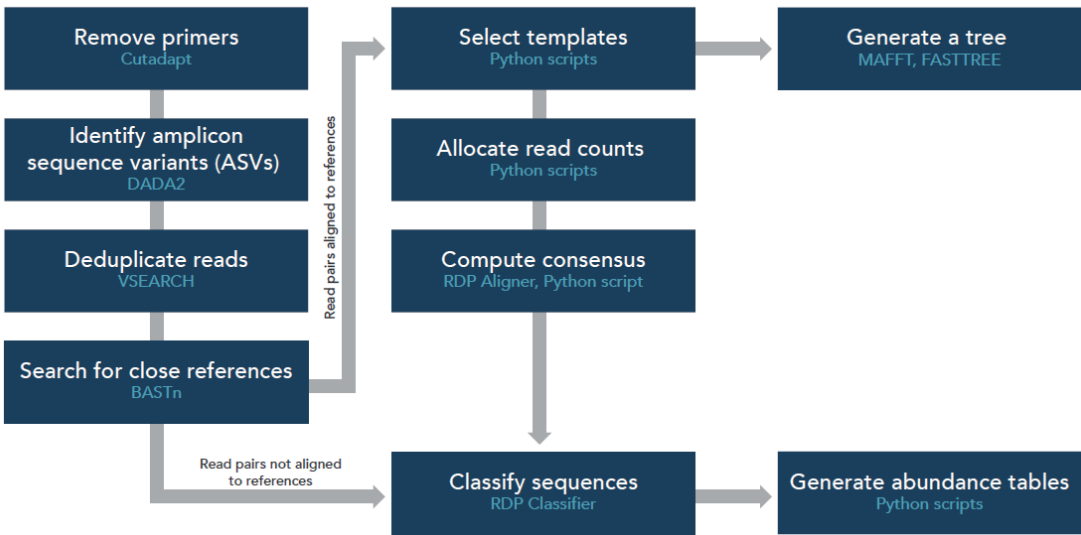


图2. 16S SNAP APP的分析工作流程

样本采集	文库制备和测序	生信分析	结果分析
ZymoBIOMICS™ Microbial Community DNA Standard (D6305) DNA标准品	 IDT xGen 16S Amplicon Panel v2 MGIEasy 通用文库转 换试剂盒 (App-A)  MGISEQ-200基因 测序仪	16S SNAP APP	分析标准品中各菌种 的实验值并全方位评 估该方案

结果

高质量的文库制备

本研究利用IDT xGen 16S Amplicon Panel v2建库试剂盒对不同投入起始量（10pg、1ng、5ng和10ng）的DNA标准品（D6305）进行文库制备，结果显示，在不同的样本投入量下，得到的文库的

浓度和产量均能满足后续的检测需求（图3A）。Agilent 2100电泳图显示投入起始量为1ng时，文库的扩增子大小也与预期相符，且没有引物二聚体等非特异峰，得到的文库质量较高（图3B）。

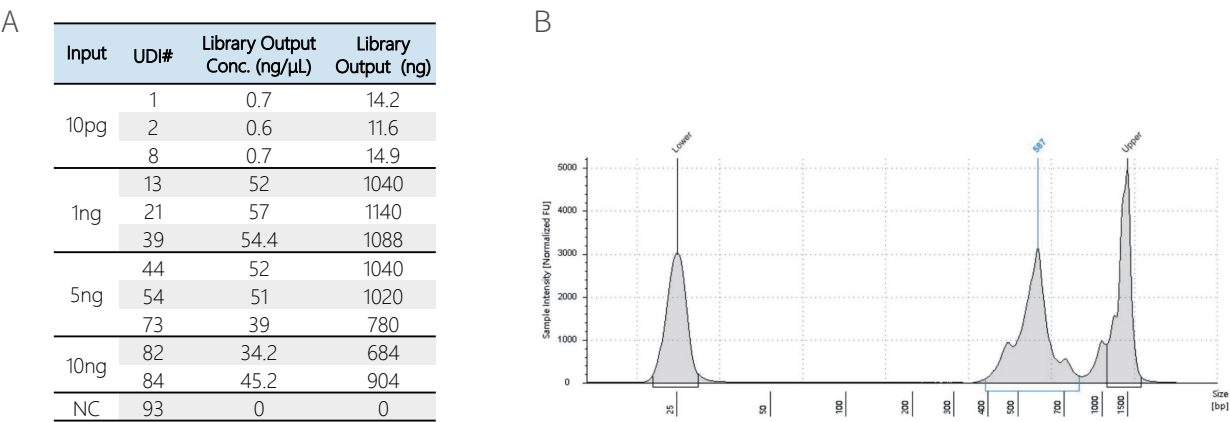


图3. 不同样本投入起始量的建库结果及文库质量情况。图A是在不同的样本投入量下（10 pg、1 ng、5 ng为三次重复，10 ng为两次重复）得到的文库浓度和产量，图B显示的是样本起始量为1ng时，文库Agilent 2100电泳图。

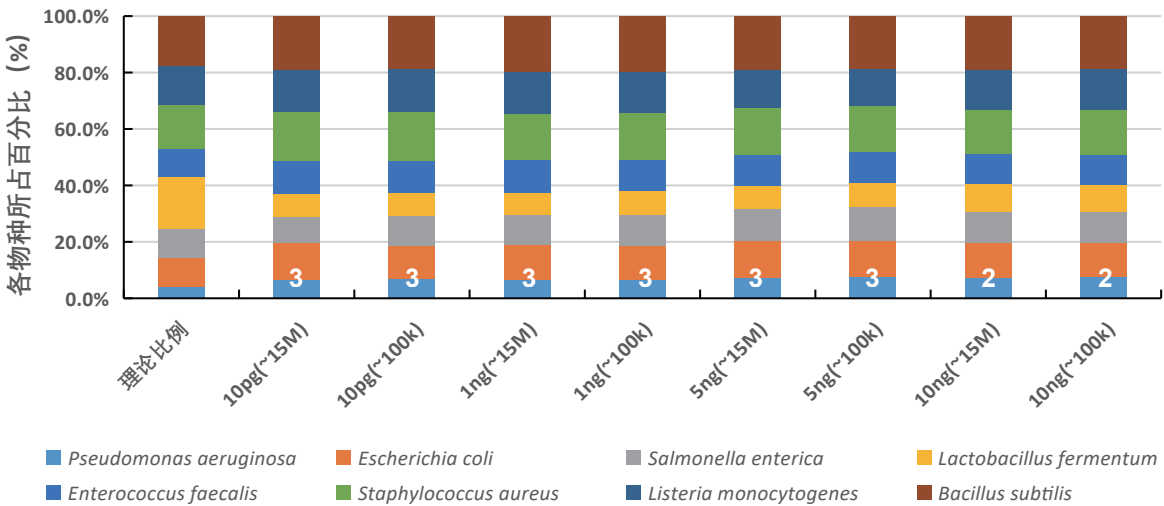
稳定的性能指标

文库在完成测序后利用IDT自研的16S SNAP APP分析工具对数据进行分析。实测数据显示，基于IDT xGen 16S v2建库流程和MGISEQ-200测序平台，不同样品起始量的检测结果在精确度和物种相对丰度上表现一致，与预期相符（表2）。如在10 pg的投入量下，绿脓杆菌、大肠杆菌、肠道沙门氏菌、发酵乳杆菌、粪链球菌、金黄色葡萄球、李斯特菌、枯草芽孢杆菌等8种细菌的相对丰度分别为6.6%±0.1%、12.9%±0.8%、9.5%±0.9%、

8.1%±0.2%、11.6%±0.2%、17.4%±0.2%、15.1%±0.3%、18.8%±1.0%，与理论比例基本相一致，其他投入量结果与此相似（表2）。通过截取不同的数据量（~15M reads和~100k reads）进行数据分析，发现在数据量低至~100k reads的情况下，仍能得到较好的菌种相对丰度检测结果，与~15M reads数据量时无显著差异，且结果大致与理论值相符（图4）。

物种组成	物种组成 (中文名称)	理论 比例	10pg (AVE±SEM)	1ng (AVE±SEM)	10pg (AVE±SEM)	1ng (AVE±SEM)
<i>Pseudomonas aeruginosa</i>	绿脓杆菌	4.2%	6.6%±0.1%	6.7%±0.4%	6.6%±0.1%	6.7%±0.4%
<i>Escherichia coli</i>	大肠杆菌	10.1%	12.9%±0.8%	12.5%±0.8%	12.9%±0.8%	12.5%±0.8%
<i>Salmonella enterica</i>	肠道沙门氏菌	10.4%	9.5%±0.9%	10.3%±0.8%	9.5%±0.9%	10.3%±0.8%
<i>Lactobacillus fermentum</i>	发酵乳杆菌	18.4%	8.1%±0.2%	8.1%±0.6%	8.1%±0.2%	8.1%±0.6%
<i>Enterococcus faecalis</i>	粪链球菌	9.9%	11.6%±0.2%	11.4%±0.3%	11.6%±0.2%	11.4%±0.3%
<i>Staphylococcus aureus</i>	金黄色葡萄球菌	15.5%	17.4%±0.2%	16.3%±0.4%	17.4%±0.2%	16.3%±0.4%
<i>Listeria monocytogenes</i>	李斯特菌	14.1%	15.1%±0.3%	15%±0.8%	15.1%±0.3%	15%±0.8%
<i>Bacillus subtilis</i>	枯草芽孢杆菌	17.4%	18.8%±1.0%	19.6%±0.5%	18.8%±1.0%	19.6%±0.5%

表2. 不同样本投入量的检测结果



高效的数据利用率

IDT 16S SNAP APP软件分析，与常规16S rRNA所使用的分析工具不同，其能进一步将测序单独 reads进行相关联，以产生更长的序列用于后续菌种分类，提高了分型精准度，还提高了数据有效利用率。本研究基于IDT xGen 16S v2建库流程和

MGISEQ-200平台，搭配IDT的16S SNAP APP分析流程发现：在10 pg、1ng、5ng、10 ng的投入量下，数据的有效利用率分别是99.90%、99.91%、99.61%以及99.39%，数据有效利用率近100%（图5）。

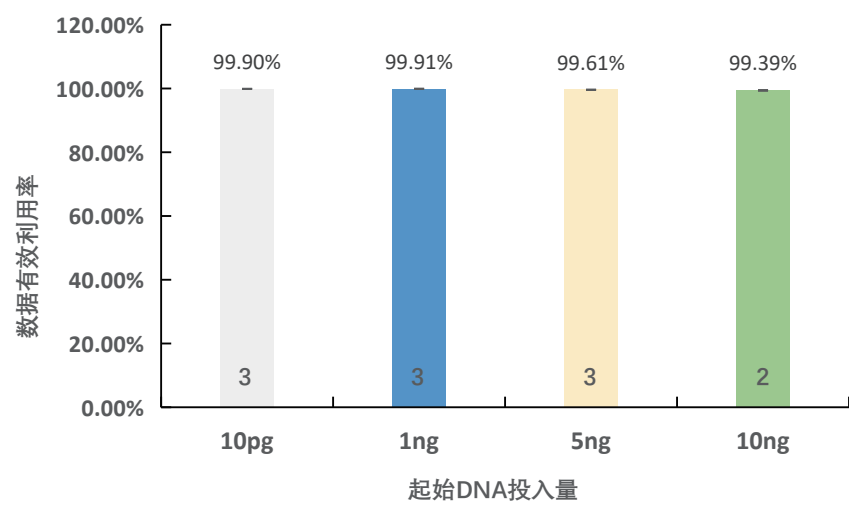


图5. 16S SNAP APP对数据的有效利用率。横坐标代表起始投入量，纵坐标代表数据的有效利用率，10 pg、1 ng、5 ng的投入量分别进行了三次重复，10 ng的投入量进行了两次重复。

总结

IDT相关方案在MGISEQ-200平台上的实测结果表明，xGen 16S Amplicon Panel v2适配MGISEQ-200平台，不同的起始量、数据量都可以得到稳定的菌种检出结果，搭配16S SNAP APP，可以有效地提高数据的利用率。

MGISEQ-200平台基于DNBSEQ™专利技术，可以稳定获得高质量的测序数据，大小双载片的设计满足灵活的测序需求。500M reads大载片在需要高深度测序的应用场景下可以大大降低测序数据量，降低测序成本。按照每样本100k reads的数据量，MGISEQ-200的100M reads小载片理论上能够满足1000个样本的16S rDNA测序，MGISEQ-200以其灵活、小巧和快速的特点成为微生物测序的经典机型。



基因测序仪MGISEQ-200

参考文献

1. Woese, C. R. & Fox, G. E. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A* **74**, 5088-5090, doi:10.1073/pnas.74.11.5088 (1977).
2. Weisburg, W. G., Barns, S. M., Pelletier, D. A. & Lane, D. J. 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* **173**, 697-703, doi:10.1128/jb.173.2.697-703.1991 (1991).
3. Pereira, F. *et al.* Identification of species by multiplex analysis of variable-length sequences. *Nucleic Acids Res* **38**, e203, doi:10.1093/nar/gkq865 (2010).
4. Schmidt, T. M., DeLong, E. F. & Pace, N. R. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J Bacteriol* **173**, 4371-4378, doi:10.1128/jb.173.14.4371-4378.1991 (1991).
5. Ley, R. E. *et al.* Obesity alters gut microbial ecology. *Proc Natl Acad Sci U S A* **102**, 11070-11075, doi:10.1073/pnas.0504978102 (2005).

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-200RS	900-000350-00
	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	900-000100-00
软件	MegaBOLT生信分析加速器(工作站式服务器)	970-000085-00
	16S SNAP APP	https://github.com/swiftbio-sciences/16S-SNAPP-py3
建库试剂	IDT xGen™ Amplicon Core (96RXN)	10009827
	IDT xGen™ 16S Amplicon Panel v2 (96RXN)	10009828
	IDT xGen™ Amplicon UDI Primers Set 1	10009846
	MGIEasy 通用文库转换试剂盒 (App-A) (16RXN)	1000004155
测序试剂	MGISEQ-200RS 高通量测序试剂套装(FCL PE150)	1000019843
	高通量双末端测序引物试剂盒(App-A)	1000020832
	高通量条形码引物3试剂盒(App-A)	1000014047

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1月版

撰稿：王其伟 张含菲

责任编辑：王其伟

审稿：江遥