

Takara Bio SMARTer建库方案搭配DNBSEQ平台赋能人类免疫图谱分析

SMARTer Human TCR a/b Profiling Kit v2和SMART-Seq Human BCR (with UMIs)在MGISEQ-2000测序仪上的性能评估

Takara基于其特有的SMART技术开发了多种免疫分析试剂盒。本研究选取了两个代表性试剂盒 SMARTer Human TCR a/b Profiling Kit v2和SMART-Seq Human BCR (with UMIs)以验证SMART技术与华大智造DNBSEQ测序平台的兼容性。综合的性能评估表明MGISEQ-2000与这些试剂盒完全适配，且其性能略优于Competitor A。

推荐应用：免疫图谱分析

推荐机型：MGISEQ-2000RS, DNBSEQ-G99ARS, DNBSEQ-T7RS

• 一流的免疫图谱解决方案

Takara Bio为人类TCR/BCR免疫组库分析提供灵敏可靠的分析方案

• 与DNBSEQ测序平台完美适配

MGIEasy通用文库转换试剂盒可使Takara Bio TCR/BCR建库试剂盒与DNBSEQ平台完美适配

• 高效且高质量的数据输出

DNBSEQ测序技术具有高准确性、低重复序列率、低标签跳跃率等特性，保证高质量的数据输出



背景介绍

T细胞受体（TCR）和B细胞受体（BCR）测序主要用于对临床患者的免疫组库进行分析进而评估患者的疾病状态。TCR和BCR测序虽主要用于肿瘤学研究中，但全方位的免疫分析数据可用于评估诸多疾病状态，了解不同淋巴细胞表达的免疫受体多样性对于免疫学研究和临床医学至关重要（图1）¹⁻³。

Takara Bio特有的SMART (Switching Mechanism At the 5' end of RNA Template)技术可以利用某种特定逆转录酶的template-switching特性在第一链cDNA合成过程中整合一个接头序列。基于这一特性，Takara Bio发明了一系列无缝、免连接的文库制备方法并将其应用于single-cell mRNA-seq、total RNA-seq、small RNA-seq、免疫分析以及ChIP-seq中。

本研究中选取了两款Takara Bio试剂盒——SMARTer Human TCR a/b profiling Kit V2和SMART-Seq Human BCR (with UMI)以全面评估Takara Bio SMART技术与华大智造DNBSEQ测序平台的兼容性。此两款试剂盒均采用SMART技术和基于5'-RACE的方法从TCR和BCR的V(D)J可变区捕获全长信息。它们是TCR和BCR分析的理想工具，让研究者们从大量样本（总RNA或纯化细胞）中深入了解TCR和BCR的多样性，期间还引入了独特分子标识符 (Unique molecular identifiers, UMI)以便在数据分析过程中进行PCR纠错和克隆型定量。

在本研究中，TCR和BCR文库皆采用PE300的测序策略在MGISEQ-2000和Competitor A测序仪上完成测序。随后将从两个平台生成的测序数据截取至相同的测序深度，对所有被鉴定的MGIs (molecular identifier groups)数目和克隆型数目进行比较，MGIs通过UMIs被分组。MGISEQ-2000与Competitor A的测序比较结果证明：Takara Bio 文库可以通过MGIEasy通用文库转换试剂盒进行转换，进而完美适配MGISEQ-2000平台。

材料与amp;方法

样本制备和RNA提取

本研究共收集两组样本：第一组样本包括13种来自不同人类细胞系和/或人外周血单核细胞（Peripheral blood mononucleated cell, PBMC）的细胞品系，使用Qiagen RNA提取试剂盒进行总RNA的提取；第二组样本包含4个从正常人类PBMCs中提取的总RNA样本（Human Blood, Peripheral Leukocytes Total RNA; Takara Bio, Cat. No. 636592）（表1）。

文库构建与测序

第1组样本由Takara Bio技术人员利用SMART-Seq Human BCR (with UMI)试剂盒制备成文库，第2组样本由华大智造技术人员利用SMART-Seq Human BCR (with UMI)试剂盒和SMARTer Human TCR a/b Profiling Kit v2进行文库制备，具体步骤请参考制造商说明书^{4,5}。最终得到纯化的双链DNA文库使用MGIEasy通用文库转换试剂盒(App-A)转换为单链环状DNA (ssCircDNA)，再将ssCircDNA制备为DNA纳米球 (DNB)，并上机至MGISEQ-2000测序仪FCS芯片上以PE300+8+8的测序策略在一条lane上完成测序步骤。

测序时，建议使用MGISEQ-2000(ECR6.0)、

MGISEQ-2000RS高通量快速测序试剂套装(FCS PE300)和高通量双末端测序引物试剂盒(App-A)进行测序。倘若MGISEQ-2000的ECR版本低于5.2，可使用MGISEQ-2000RS高通量快速测序试剂套装(FCS PE300)和高通量测序引物试剂盒(App-D)，并掺入20%-40%的华大智造标准文库或其他平衡文库以提高测序质量（本研究中掺入了40%的ATOPlex E450双标签平衡文库）。

生信分析

所有样本的FASTQ数据均使用Takara Bio公司的Cogent™ NGS Immune Profiler进行分析，用于分析使用SMART-Seq Human BCR (with UMI)或SMARTer Human TCR a/b Profiling Kit v2建库产生的数据。生信分析流程包括：数据前处理，基于UMI的分析，克隆型识别以及报告输出。生信分析输出的CSV文件被上传到基于Web的Cogent NGS Immune Viewer软件使数据可视化和并生成相关图片。具体分析流程可参考Takara Bio官网的Cogent™ NGS Immune Profiler v1.5 User Manual(<https://www.takarabio.com/documents/User%20Guides/Cogent%20NGS%20Immune%20Profiler%20v1.5%20User%20Manual.pdf>, P24~34)。

组别	样本编号	对应细胞系	文库类型	文库名称	
1 (Takara Bio)	U081	DAUDI cell line	BCR heavy chain	DAUDI-BH	
	U082	CA46 cell line	BCR heavy chain	CA46-BH	
	U083	DAUDI + CA46 cell line(ratio of 1:3)	BCR heavy chain	DAUDI/CA46-BH	
	U084	Human PBMC	BCR heavy chain	PBMC-BH	
	U085	Human PBMC + 10% of U083 cell line	BCR heavy chain	PBMC/DAUDI/CA46-10-BH	
	U086	Human PBMC + 1% of U083	BCR heavy chain	PBMC/DAUDI/CA46-1-BH	
	U089	DAUDI cell line	BCR light chain	DAUDI-BL	
	U090	CA46 cell line	BCR light chain	CA46-BL	
	U091	DAUDI + CA46 cell line (ratio of 1:3)	BCR light chain	DAUDI /CA46-BL	
	U092	Human PBMC cell line	BCR light chain	PBMC-BL	
	U093	Human PBMC + 10% of U091 cell line	BCR light chain	PBMC/DAUDI/CA46-10-BL	
	U094	Human PBMC + 1% of U091 cell line	BCR light chain	PBMC/DAUDI/CA46-1-BL	
	2 (MGI)	U029	Human PMBC	BCR heavy chain-02	PBMC-BH02
		U030	Human PMBC	BCR light chain-02	PBMC-BL02
U25		Human PMBC	TCR α/β -replicate 1	PBMC-Ta/ β -1	
U26		Human PMBC	TCR α/β -replicate 2	PBMC-Ta/ β -2	

表1. 由Takara Bio和華大智造技術人員分別推進實驗所涉及到的樣本及文庫信息。

结果

Takara Bio建库试剂盒搭配MGISEQ-2000测序仪可获得高质量测序数据

所有样本上机MGISEQ-2000获得了约10M条双端 (paired-end, PE) 测序reads, 分析结果发现测序结果的Q30值均高于95%, 显著高于公认标准 (85%), 这表明MGISEQ-2000的测序数据质量较高。可用于TCR或BCR克隆型鉴定的数据所占百

分比(Determined%), 都高于97%, 亦高于标准值 (93%) (表2)。Cogent NGS Immune Profiler 中使用MIGEC根据UMI序列将reads为不同的分子标识符组(Molecular Identifier Groups, MIG)。MIG 折叠后, 属于同一MIG的reads被折叠为单个 read, 用于下游克隆型的鉴定。如表2所示, MIG 折叠后得到的reads数(Collapsed)以及总MIGs数都符合预期。

文库名称	文库类型	总reads数	GC%	Q30%	Determined%	总MIGs数	Collapsed
DAUDI-BH	BCR	9,620,845	53.09	95.87	97.9	537,764	177,112
CA46-BH	BCR	8,987,764	59.93	97.12	97.8	362,526	125,210
DAUDI/CA46-BH	BCR	11,453,613	54.99	96.54	97.9	473,711	168,055
PBMC-BH	BCR	12,967,353	56.49	96.51	97.9	287,158	44,895
PBMC/DAUDI/CA46-10-BH	BCR	12,245,364	55.38	96.93	97.9	431,219	121,196
PBMC/DAUDI/CA46-1-BH	BCR	11,283,318	56.27	96.62	97.9	280,605	48,578
DAUDI-BL	BCR	9,408,528	49.37	96.86	98.3	420,193	148,720
CA46-BL	BCR	11,494,906	55.14	96.52	98.4	816,883	324,335
DAUDI /CA46 -BL	BCR	11,458,391	51.76	96.9	98.4	530,379	206,286
PBMC-BL	BCR	12,778,326	53.01	96.88	98.0	215,465	85,796
PBMC/DAUDI/CA46-10-BL	BCR	11,371,160	52.11	96.76	98.3	427,488	171,335
PBMC/DAUDI/CA46-1-BL	BCR	10,248,519	52.62	96.7	98.1	230,569	87,923
PBMC-BH02	BCR	9,171,213	55.75	96.52	99.2	449,988	108,118
PBMC-BL02	BCR	8,780,848	55.35	96.85	99.1	351,640	216,907
PBMC-Ta/ β -1	TCR	11,172,773	49.54	95.85	99.2	369,646	145,040
PBMC-Ta/ β -2	TCR	10,020,030	49.43	96.6	99.1	366,982	160,961

表2 TCR/BCR文库在MGISEQ-2000上的测序结果。Determined%: 确定为TCR或BCR克隆型的reads百分比; Total MIGs: 分子标识符组总数; Collapsed: MIG 折叠后的reads数。

MGISEQ-2000在TCR和BCR测序中展现出了准确的克隆型检测能力

为了获取更多的克隆型信息，研究者们计算了被鉴定为不同克隆型的reads数比例。结果发现，DAUDI-BH, CA46-BH以及 DAUDI/CA46-BH的IgM 含量都较高。类似的，DAUDI-BL, CA46-BL以及DAUDI/CA46 -BL都有着较高含量IgK。其他的BCR文库都包含PBMC，其中PBMC-BH02,PBMC-BH, PBMC/DAU-

DI/CA46-10-BH以及PBMC/DAUDI/CA46-1-BH都是BCR的重链，它们的含量都很类似，主要包含IgM, IgA以及IgG。PBMC-BL02, PBMC-BL, PBMC/DAUDI/-CA46-10-BL以及PBMC/DAUDI/CA46-1-BL都是BCR的轻链，它们的主要组成是IgK和IgL。最后PBMC-Tα/β-1 and PBMC-Tα/β-2 都是TCR，主要包含T细胞受体的alpha和beta基因序列(TRA以及TRB) (图1)。以上所有的克隆型鉴定结果均与预期相符。

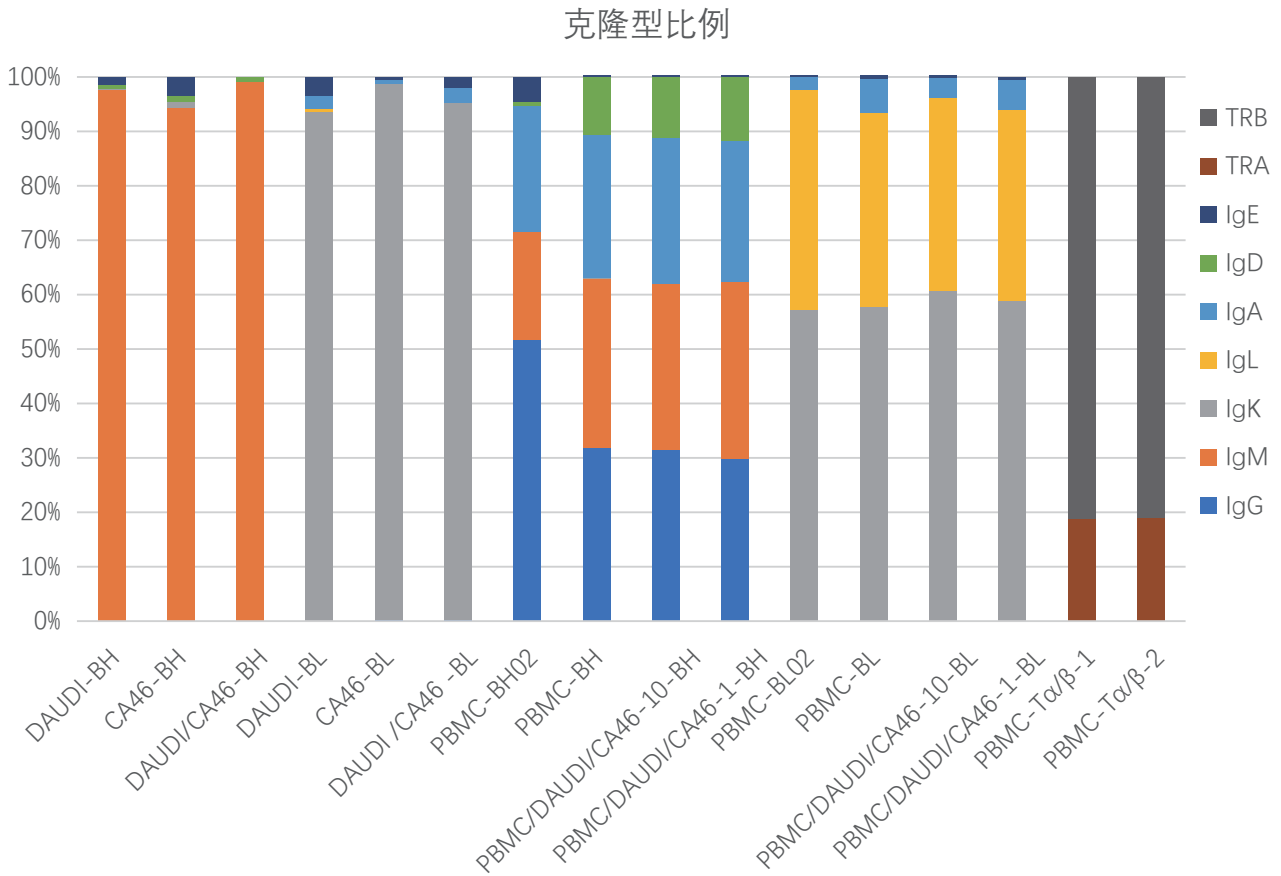


图1. TCR和BCR文库在MGISEQ-2000上测序得到的不同克隆型比例。克隆型比例: 检测到的特定克隆型barcodes比例,以上结果均源于MGISEQ-2000。

每个TCR或BCR都涵盖可变区和恒定区，可变区通过V(D)J重组进行组装。CDR3 (The third complementarity determining region)是V(D)J的核心区域，因此可以计算CDR3的长度分布以分析V(D)J重排的情况。图2的结果显示，大部分样本的CDR3长度在30 ~ 40个碱基左右，总体分布在30 ~ 80个碱基之间，符合之前研究的标准⁶。

克隆型分布分析显示，一些样本显示出相似比例的亚型(图1)。为了进一步探究这些样本的V(D)J重排情况，研究者们使用Cogent NGS Immune Profiler对每个CDR3基因的V(D)J组合进行了详尽的分析。由于D片段非常短且在重排过程中变化较大，因此研究V和J片段的重组更为可靠，V/J usage plot以及chord diagram如图3所示。图3A为DAUDI/CA46-BL、PBMC/DAUDI/CA46-10-BL以及PBMC-Tα/β-1的V/J usage plot，与图3 B、D、E中的chord diagram相对应，从chord diagram中可以看出，每个文库的V-J配对是不同的。例

如，DAUDI-BH和CA46-BH的细胞系不同，因此表现出明显不同的V-J配对，而DAUDI/CA46-BH如预期所示包含 DAUDI-BH和CA46-BH的V-J配对(图3B，左边)；在DAUDI-BL、CA46-BL和DAUDI/CA46-BL亦中观察到类似的现象(图3B，右边)。

在BCR重链组文库(PBMC-BH02、PBMC-BH、PBMC/DAUDI/CA46-10-BH和PBMC/DAUDI/CA46-1-BH)和轻链组文库(PBMC-BL02、PBMC-BL、PBMC/DAUDI/CA46-10-BL和PBMC/DAUDI/CA46-1-BL)中，一些样本有着较为复杂的V-J配对，这是由于PBMC中BCR多样性较高造成的(图3C,D)。一些样本如PBMC/DAUDI/CA46-10-BH有着较高比例的IGHJ4，IGHV3-74和IGHV5-51，这也与DAUDI/CA46-BH相符(图3B, C)。此外，TCR样本的两个技术性重复(PBMC-Tα/β-1和PBMC-Tα/β-2)具有几乎相同的V-J配对(图3E)，也与PBMC-BH02和PBMC-BL02有着截然不同的V-J配对形成鲜明对比(图3C, D)。

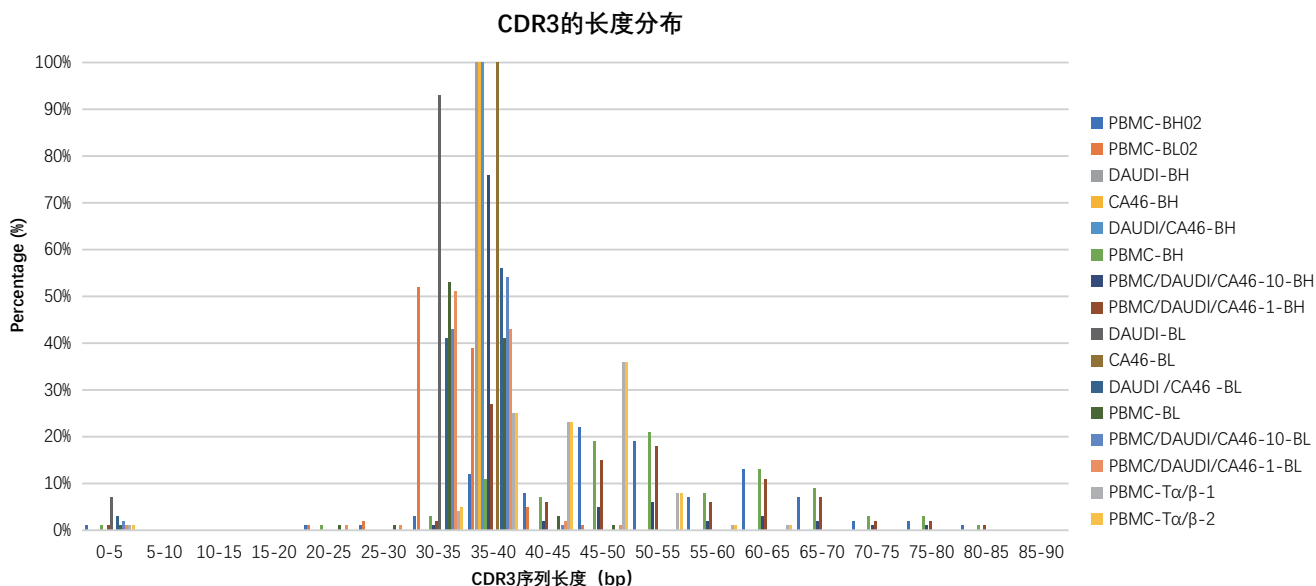


图2. CDR3序列的长度分布。长度分布由Cogent NGS Immune Profiler计算，长度超过90个碱基的CDR3序列不包括在其中,以上结果均源于MGISEQ-2000。

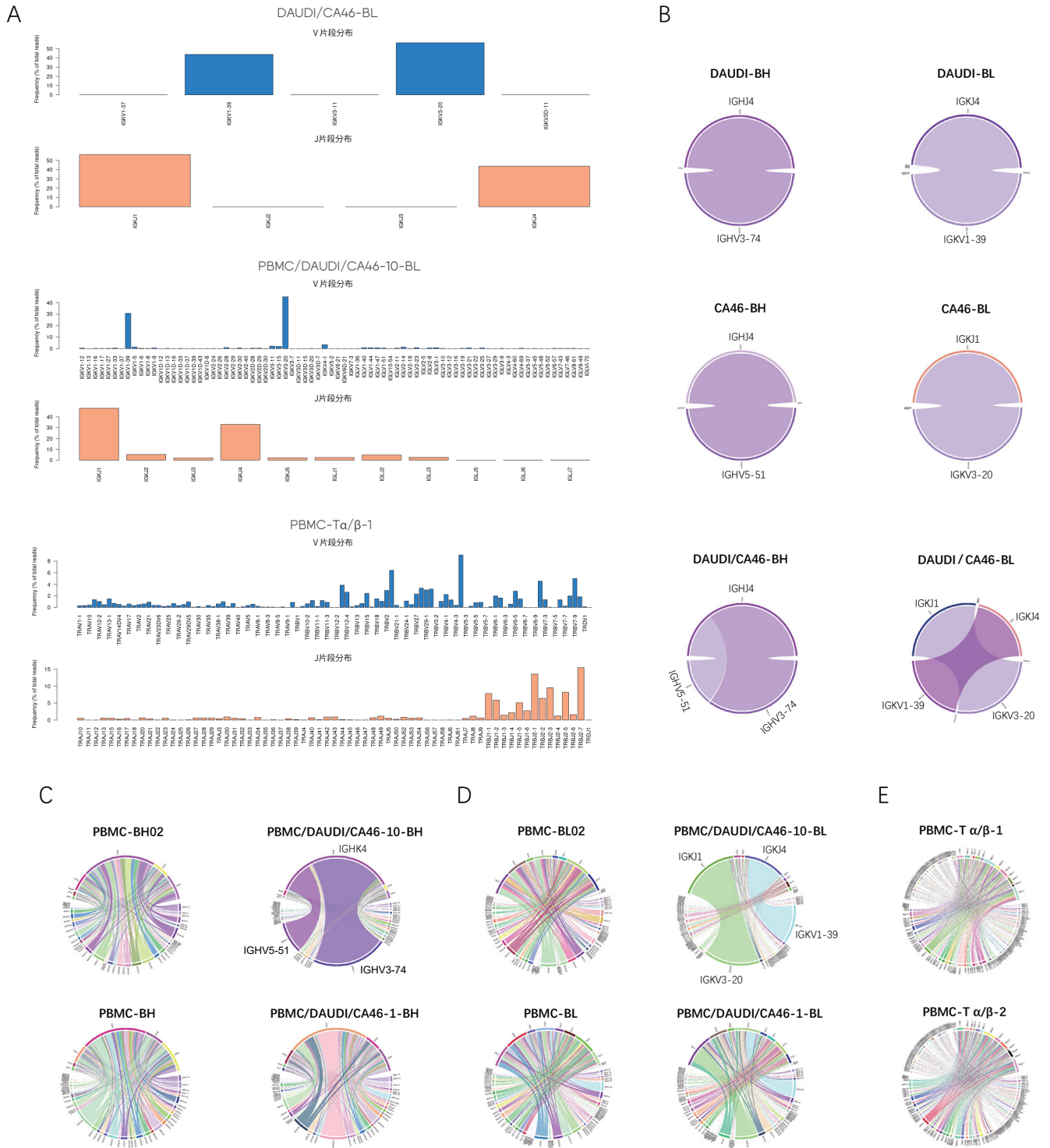


图3. V/J基因的usage plot以及V-J配对的chord diagram。(A) V/J基因的usage plot, 此处只展示了DAUDI/CA46-BL、PBMC/DAUDI/CA46-10-BL和PBMC-T α/β -1的样本数据, 这一数据可用于评估不同V/J基因片段的分布和共线性。(B, C, D, E) chord diagram直观地展示了克隆型分布, 并显示了V-J配对, Chord的宽度代表V和J片段配对的频率, 以上结果均源于MGISEQ-2000。

此外，研究人员还从PBMC和混合细胞系中确定了BCR或TCR比例排名前十的克隆型(表3)。

文库类型	排名	CDR3s的核苷酸序列	克隆型	所占比例 (%)
PBMCs 中的BCR	1	TGTCAGCAGCGTAGCAACAAGTCTTTT	IGK	2.12
	2	TGTGCGAGACAATACGTGGCCGTCTGG	IGG	0.58
	3	TGTGCGAAGCAGTACCTGGTTGGGCGGGGCTACTGG	IGG	0.47
	4	TGTCAGCAATATTACAGTACTCCGTGGACGTTC	IGK	0.44
	5	TGTCAGCAGCGTAGCAACTGGCCTCTCACTTTC	IGK	0.4
	6	TGTCAACAGTATGTTACTTATCCGTGGACGTTC	IGK	0.38
	7	TGTCAGCAGCGTAGCAAGAAGTCTTTT	IGK	0.32
	8	TGTCAGCAATATTATAGTACTCCGTACACTTTT	IGK	0.27
	9	TGTCAGCAGTATAATAACTGGCCTCGGACGTTC	IGK	0.23
	10	TGCTACAAGATTACAATTACCCTCGGACGTTC	IGK	0.22
混合细 胞系中 的BCR	1	TGTCAGCAGTATGGTAGTTCACCTCCGTGGACGTTC	IGK	47.58
	2	TGTCAACACAATTACAATTTCTCGTTCACCTTC	IGK	30.27
	3	GAGAACACTAGGAATTTACTCAGCCAGTG	IGK	1.67
	4	TGTCAATCAGCAGACAGCAGTGGTAGTTATGTCTTC	IGL	0.8
	5	GACATCCAGATGACCCAGTCTCCATCCTC	IGK	0.16
	6	TGTCAGCAGTATAATAACTGGCCTCCGTACACTTTT	IGK	0.08
	7	TGTCAGCAATATTATAGTACTCCGTACACTTTT	IGK	0.08
	8	TGTCAACAGAGTTACAGTACCCCGTACACTTTT	IGK	0.06
	9	TGTCAGCAATATCTTAAGTGGCTCACGATCACCTTC	IGK	0.06
	10	TGTCAGCAATCTTATAATCATCCTCGCACTTTT	IGK	0.06
PBMCs 中的TCR	1	TGTGTGGTGAACCCGCCGAACACAGGCTTTCAGAACTTGTATTT	TRA	4.14
	2	TGTGCCACCAGCAGAGATACGGGGTCTACGAGCAGTACTTC	TRB	1.51
	3	TGTGCCAGCAGCGGCCAGGGGGACCTGGGAGAGACCCAGTACTTC	TRB	0.74
	4	TGTGCCAGCAGCCCCGAGACAGGGAGTTCTACAATGAGCAGTTCTTC	TRB	0.71
	5	TGTGCCAGCAAGGTAGCGCTCAATACCTACGAGCAGTACTTC	TRB	0.67
	6	TGTGCCAGCAGTTTGGGGCAGGCCAATGAGCAGTTCTTC	TRB	0.66
	7	TGCGCCAGCAGCCAAGTTATGCGGGGCGGAACACTGAAGCTTTCTTT	TRB	0.49
	8	TGTGCCAGCAGTTTGGGCCTGCACTACGAGCAGTACTTC	TRB	0.49
	9	TGTGCCAGCAGTTACCAGGGCCATCAGCCCCAGCATTTT	TRB	0.45
	10	TGTGCCAGCAGCCCGACTACGTGGGAGGCCAGTACTTC	TRB	0.4

表3.在PBMC中的BCR(PBMC- BH02和PBMC- BL02)，混合细胞系中的BCR(PBMC/DAUDI/CA46-10-BL)以及PBMC中的TCR (PBMC-T α / β -1和PBMC-T α / β -2)里的比例排名前十的克隆型分布,以上结果均源于MGISEQ-2000。

MGISEQ-2000在PBMC和单克隆/混合细胞系文库测序中均显示出与Competitor A相当的测序质量

为了进一步验证MGISEQ-2000的测序性能，将同一文库在MGISEQ-2000和Competitor A的测序仪上测序得到的结果进行比较，所有样本都被截取至1M reads用于分析。与Competitor A相比，来自MGISEQ-2000测序数据的比对率(Aligned%)略高(图4A)，未确定的reads率(undetermined%)和

比对至chains failed linker-based correction的比例(flc%)略低(越低越好)(图4B, C)。比对至模板长度低于30bp的reads比例(short%)，两平台都为0%(此处未展示)。总之，MGISEQ-2000产生的数据质量略优于Competitor A。

进一步的分析表明，对于PBMC或单克隆细胞系中的总MIGs数检测，MGISEQ-2000的测序数据亦略优于Competitor A (图5A, B); PBMC中的克隆型计数也高于Competitor A (图5C)，再次证明MGISEQ-2000的性能略优于Competitor A平台。

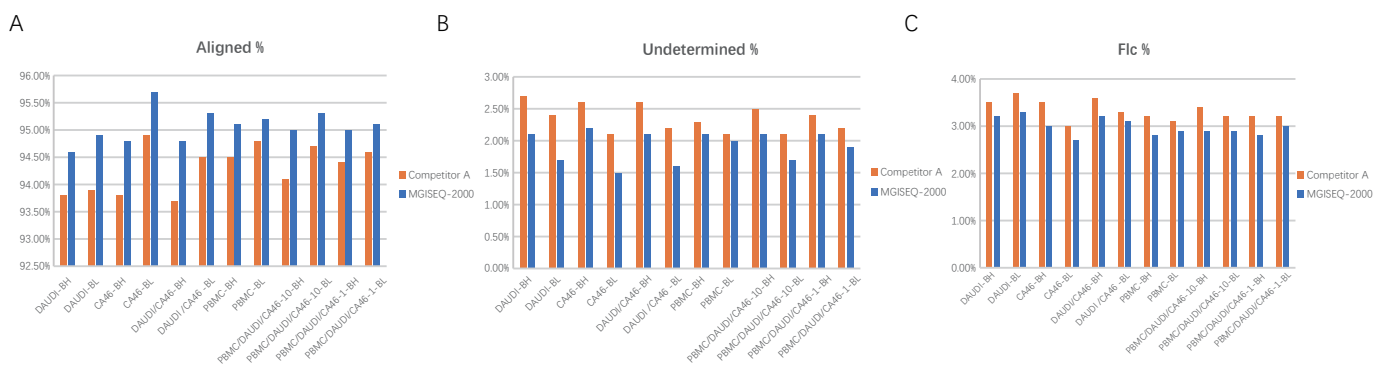


图4. MGISEQ-2000与Competitor A的测序参数比对。(A) Aligned %: 比对reads的百分比;(B) Undetermined %: 来自未确定链的reads百分比;(C) Flc %: 比对至chains failed linker-based correction的比例。

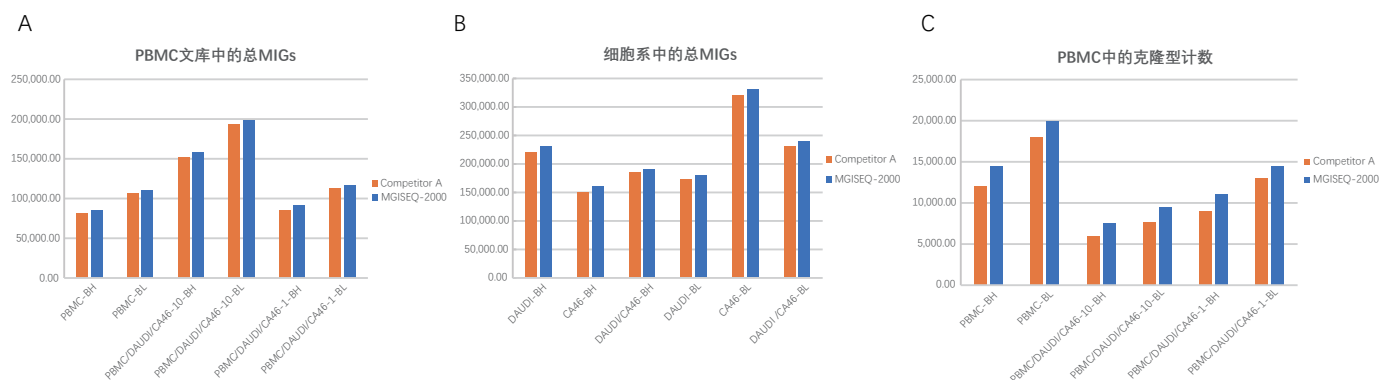


图5. MGISEQ-2000与Competitor A在MIGs总数和克隆型计数方面的比较。(A, B) PBMC和细胞系中的总MIGs;(C) PBMC样本中的克隆型计数。

总结

SMARTer Human TCR a/b Profiling Kit v2和SMART-Seq Human BCR (with UMIs)在MGISEQ-2000以及Competitor A测序平台上的实测结果表明，Takara SMART技术适配DNBSEQ平台。无论是PBMC还是单克隆细胞系，BCR还是TCR文库均能够获得较为的理想结果。TCR和BCR的多样性是通过V、D和J基因的重排来实现的，MGISEQ-2000平台搭配Takara TCR/BCR profiling kit鉴定TCR和BCR克隆型的能力十分优越，且其性能与Competitor A相当。

MGISEQ-2000采用全新的载片系统，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PE300 (FCS) 测序仅需约98小时。

MGISEQ-2000搭配Takara Bio免疫分析建库方案，有助于研究者们进一步了解V(D)J重排过程，识别疾病相关的TCR和BCR，为了解疾病状态，发现治疗靶点提供新思路。

参考文献

1. Friedensohn S, Khan TA, Reddy ST. Advanced Methodologies in High-Throughput Sequencing of Immune Repertoires. Trends in Biotechnology 2017. 35(3):203-14.
2. McDaniel JR, Brandon J DeKosky, Tanno H, Ellington AD & Georgiou G. Ultra-high-throughput sequencing of the immune receptor repertoire from millions of lymphocytes. Nature Protocols 2016. 11:429-42.
3. Hou D, Chen C, Seely EJ, Chen S and Song Y. High-Throughput Sequencing-Based Immune Repertoire Study during Infectious Disease. Front. Immunol. 2016. 7
4. SMART-Seq[®] Human BCR (with UMIs) User Manual: www.takarabio.com/documents/User%20Manual/SMART/SMART-Seq%20Human%20BCR%20%28with%20UMIs%29%20User%20Manual.pdf
5. SMARTer[®] Human TCR a/b Profiling Kit v2 User Manual: www.takarabio.com/assets/a/117462
6. Rock EP, Sibbald PR, Davis MM, Chien YH. CDR3 length in antigen-specific immune receptors. J Exp Med 1994. 179:323-328.



基因测序仪MGISEQ-2000

推荐订购信息

产品类型	产品名称	产品货号
仪器	MGISEQ-2000基因测序仪	900-000033-00
软件	Cogent NGS Immune Profiler	https://www.takara-bio.com/products/next-generation-sequencing/bioinformatics-tools/cogent-ngs-immune-profiler*
	Cogent NGS Immune Viewer	https://www.takara-bio.com/products/next-generation-sequencing/bioinformatics-tools/cogent-ngs-immune-viewer*
文库制备试剂	SMARTer Human TCR a/b Profiling Kit v2	634478*
	SMART-Seq Human BCR (with UMIs)	634777*
	MGIEasy通用文库转换试剂盒(App-A)	1000004155
	a) ATOplex E450双标签平衡文库试剂	940-000637-00**
测序试剂	b) MGISEQ-2000RS高通量快速测序试剂套装 (FCS PE300)	940-000039-00**
	c) 高通量双末端测序引物试剂盒 (App-A)	1000020832**
	d) 高通量测序引物试剂盒(App-D)	1000028550**

*相关产品可在Takara Bio官网查询订购

**当选择的MGISEQ-2000版本为ECR 6.0时使用b) + c), 或当选择的ECR版本低于5.2时使用a) + b) + d)。

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11月版

撰稿: 张含菲

责任编辑: 王其伟

审稿: 江遥