

使用改良的 Agilent SureSelect XT HS2 靶向序列捕获流程分析模式生物的 DNA 甲基化状态

方法和分析综述

作者

Dr. Matthieu Egloff 医学博士，
普瓦提埃大学，
INSERM 1084, LNEC, F-86000
Poitiers, France

Quentin Riche-Piotax 博士，
普瓦提埃大区医疗中心，
Service de Génétique, F-86000
Poitiers, France

François Lozach
安捷伦科技有限公司

摘要

通过二代测序 (NGS) 检测和分析甲基化可能具有挑战性，湿实验室操作和分析流程通常较长且复杂。本应用简报旨在详细介绍该方法，帮助研究人员将 Agilent SureSelect XT HS2 杂交捕获试剂盒用于该特定 NGS 应用。此外，本文还为甲基化分析提供了一些指南。本文档描述了对小鼠血液和大脑提取的 DNA 的目标区域捕获的完整流程。在这一应用实例中，重亚硫酸盐转化文库揭示了靶向捕获能力和获得数据的精密度。

前言

脊椎动物 DNA 的甲基化通常在胞嘧啶 (C) 核苷酸上连接一个甲基 (CH₃)，主要发生于脊椎动物的 GC 序列背景信息中。这些甲基化位点被称为 CpG (p 代表磷酸二酯键连接)，以便与 CG 碱基对区分开来。

甲基化模式可以通过复制传代，从而在子细胞中维持甲基化状态。甲基化状态与基因表达的其他表观遗传修饰因子（如组蛋白修饰）共同作用调节转录。例如，启动子位点的 CpG 甲基化可以通过阻止与用于转录起始的蛋白结合来抑制转录，也可以位于特异性修饰酶结合位点以产生完整的基因表达调控机制^[1]。

因此，CpG 通常位于启动子中和启动子附近，其甲基化影响基因表达。然而，在某些情况下，研究 CpG 组之外的差异甲基化区域 (DMR) 也很重要^[2]。

在本应用简报中，采用改良的 Agilent SureSelect XT HS2 靶向序列捕获方案研究小鼠大脑和血细胞中的 DNA 甲基化状态。该动物模型提供了相同的个体间遗传背景和受控环境，有助于识别与暴露于丙戊酸 (VPA) 的研究药物特异性相关的标记。VPA 是一种抗癫痫药物，常用于通过子宫内给药生成 ASD（孤独症谱系障碍）的环境动物模型。在这项研究中，我们评估了野生型小鼠和子宫内暴露于 VPA 的小鼠之间的甲基化差异。这代表了使用 Agilent SureSelect 甲基化测序靶向序列捕获系统，并可以获得良好结果的一个示例。

本应用简报简短描述了本研究中使用的 SureSelect 定制 Panel，随后详细描述了该方案，最后一节专门描述了分析流程。

定制 Panel

SureSelect 3 Mb Panel 的靶标区域对应于 1407 个基因组区域，位于 1029 个基因的已知或预测启动子中，已知这些基因与 ASD 有关。为建立这一靶标列表，本文交叉引用了两个来源：来自 Simons 基金会孤独症研究倡议 (SFARI) 的孤独症相关基因列表 (<https://gene.sfari.org/>) 和真核启动子数据库 (<https://epd.expasy.org/epd/>)。设计策略使用计算机模拟探针性能标准，将探针置于 CpG 位点或 CpG 位点附近。与 SureSelect 全甲基化组 (80 Mb) Panel 相比，3 Mb Panel 除了专门关注 ASD 靶标区域外，主要优势是能够对每个靶标进行深度测序。

分步方案

在本节中，使用快速启动方案[使用捕获后混合的 SureSelect XT HS2 DNA](#) 作为参考。如适用，请参考已发表的方案。根据已发表的流程对该方案的部分章节进行了修改，所有差异详见下文。快速启动方案的修改部分在图 8 中以红色字体突出显示。如果没有特别说明，步骤与原始方案相同。运行该方案需要 EZ DNA-Methylation Lightening Kit (Zymo Research)。

样品采集：处死时心内采集 500 µL 血液，立即提取 DNA。心内灌注 50 mL 0.9% NaCl 后，取大脑并储存在 -80 °C 下。使用 DNEasy blood&Tissue Kit (Qiagen) 从血液和大脑中提取 DNA。

定制 Panel：详见以上描述。

DNA 片段化：按照快速启动方案[使用捕获后混合的 SureSelect XT HS2 DNA](#) 所述进行，选项 2：Covaris 剪切；第 2 页。

文库制备：按照快速启动方案[使用捕获后混合的 SureSelect XT HS2 DNA](#) 所述进行。**重要提示：**您需要遵循第 1、2、3、4 步（第 3 页），但用 SureSelect 甲基化测序 XT 试剂盒中的 5 µL SureSelect 甲基化测序甲基化接头瓶（**绿盖**）替换 SureSelect XT HS2 接头寡核苷酸混合物（白盖）。注：根据 DNA 的初始量，可能需要优化接头稀释量。

使用 Agilent TapeStation 高灵敏度 DNA 试剂盒 (HSD1000) 评估文库质量。预期峰分布在 200–300 bp 的范围内。通过积分峰下面积测定文库的浓度。

杂交/捕获：

1. 杂交设置

使用 SureSelect 甲基化测序杂交试剂盒 1 盒与 2 盒进行过夜杂交。

- ☐ 连接和纯化的文库须使用 SpeedVac 干燥至 4 µL
- ☐ 制备杂交缓冲液混合物（表 1）

表 1. 制备杂交缓冲液

试剂	单次反应 (µL)	8 次反应，含冗余 (µL)	17 次反应，含冗余 (µL)
SureSelect Hyb 1（橙盖或瓶）	6.63	60	112.7
SureSelect Hyb 2（红盖）	0.27	2.5	4.59
SureSelect Hyb 3（黄盖或瓶）	2.65	24	45
SureSelect Hyb 4（黑盖或瓶）	3.45	31	58.65
	13 µL	117.5 µL	220.94 µL

- ☐ 制备 SureSelect Block 混合物（表 2）

表 2. 制备 SureSelect Block 混合物

试剂	单次反应 (µL)	8 次反应，含冗余 (µL)
SureSelect Indexing Block 1（绿盖）	2.5	22.5
SureSelect Block 2（蓝盖）	2.5	22.5
SureSelect Methyl-Seq Block 3（棕色瓶盖）	0.6	5.4
	5.6 µL	50.4 µL

- ☐ 向每个未扩增文库中加入 5.6 µL 以上述方法制备的 SureSelect Block 混合物。涡旋混合 5 s，然后离心
- ☐ 盖上微孔，然后将密封的孔板或联管转移至热循环仪中，并运行表 3 所示的程序

表 3. 杂交前 DNA + Block 混合物的热循环仪程序

步骤	温度	时间
第 1 步	95 °C	5 min
第 2 步	65 °C	保持（至少 5 min）

- ☐ 根据捕获文库的大小，制备适当的 SureSelect RNase Block 混合物稀释液（表 4）

表 4. 根据捕获文库大小稀释 SureSelect RNase Block 混合物

捕获文库大小	RNase Block 稀释液 (RNase Block:水)	每次杂交反应所需的 RNase Block 稀释液体积
> 3 Mb	25% (1:3)	2 µL
< 3 Mb	10% (1:9)	5 µL

- 根据捕获文库的大小，制备捕获文库杂交混合物（表 5 和表 6）。在室温下短暂保存混合物，待用

表 5. 制备用于 > 3 Mb 捕获文库的捕获文库杂交混合物

试剂	单次反应 (μL)	8 次反应，含冗余 (μL)
杂交缓冲液混合物	13	117
25% RNase Block 溶液	2	18
SureSelect 捕获文库	5	45
	20 μL	180 μL

表 6. 制备用于 > 3 Mb 捕获文库的捕获文库杂交混合物

试剂	单次反应 (μL)	8 次反应，含冗余 (μL)
杂交缓冲液混合物	13	117
10% RNase Block 溶液	5	45
SureSelect 捕获文库	2	18
	20 μL	180 μL

- 将文库 + RNase Block 混合物孔板或联管置于 65 °C，同时向每个样品孔中加入 20 μL 捕获文库杂交混合物。用移液器吹打 8-10 次，混匀
- 将杂交混合物在 65 °C（热盖 105 °C）下孵育 16 小时（盖好盖子，避免蒸发）

2. 制备链霉亲和素磁珠

请参阅快速启动方案使用捕获后混合的 SureSelect XT HS2 DNA，杂交/捕获第 2 步；第 5 页。

3. 捕获杂交文库

- 杂交后检查样品体积 (> 25 μL)
- 将杂交混合物转移到室温下的磁珠溶液中。用移液器吹打混合，使磁珠完全重悬
- 以 1400 rpm 室温孵育 30 min
- 置于磁性分离器上。弃去上清液
- 使用 200 μL SureSelect 清洗液 1 重悬磁珠，然后在室温下孵育 15 min
- 置于磁性分离器上。弃去上清液

- 使用 200 μL（65 °C 预热）SureSelect 清洗液 2 清洗磁珠，置于热循环仪（热盖）上 65 °C 下孵育 10 min。重复 3 次

- 新鲜配制 0.1 M NaOH，用于洗脱捕获的文库
- 向磁珠结合样品中加入 20 μL 新鲜配制的 0.1 M NaOH 溶液，并在涡旋混合器上混合 5 s，重悬磁珠
- 室温孵育样品 20 min
- 在磁性分离器上收集洗脱混合物中的磁珠，并将每个孔中的上清液转移到新孔板或联管的孔中

重亚硫酸盐转化：

1. 杂交设置

- 向 PCR 管中的 20 μL DNA 样品中加入 130 μL **Lightning 转化试剂**。混合，然后短暂离心，确保管盖或管壁无液滴

注：如果 DNA 体积小于 20 μL，则用无核酸酶水定容至 20 μL

注：如果样品 > 20 μL，则必须使用多个转化反应进行处理。可以重复第 3 步至第 5 步，对每个重复反应使用相同的过滤柱进行清洗

- 将 PCR 管置于热循环仪中，运行以下步骤：
 1. 98 °C 循环 8 min
 2. 54 °C 循环 60 min
 3. 4 °C 下储存长达 20 小时
- 向 Zymo-Spin™ IC 柱中加入 600 μL **M 结合缓冲液**，并将柱置于提供的收集管中

注：插入柱的收集管容量为 800 μL

必要时清空收集管，以防流通液污染柱内容物。

- ☐ 将样品（来自第 2 步）上样至含有 M 结合缓冲液的 Zymo-Spin™ IC 柱中。盖上管盖，上下翻转柱数次进行混合
- ☐ 全速 (> 10000 × g) 离心 30 s。弃去流通液
- ☐ 向柱中加入 100 μL **M 清洗缓冲液**。全速离心 30 s
- ☐ 向柱中加入 200 μL **L 脱硫缓冲液**，并在室温 (20–30 °C) 下静置 17 min。孵育后，全速离心 30 s
- ☐ 向柱中加入 200 μL **M 清洗缓冲液**。全速离心 30 s。重复该清洗步骤
- ☐ 将柱置于 1.5 mL 微量离心管中，并直接向柱基质中加入 20 μL H₂O。全速离心 30 s 以洗脱 DNA

2. PCR 扩增

首先使用 SureSelect 甲基化测序文库制备试剂盒进行 PCR。

- ☐ 制备 PCR #1 反应混合物（表 7）

表 7. 制备 PCR #1 反应混合物

试剂	单次反应 (μL)	8 次反应，含冗余 (μL)
无核酸酶水	30	270
SureSelect 甲基化测序 PCR 预混液	50	450
甲基化测序 PCR1 引物 F	1	9
甲基化测序 PCR1 引物 R	1	9
	82 μL	738 μL

- ☐ 向各 20 μL DNA 样品中加入 82 μL PCR #1 混合物，混匀
- ☐ 在热循环仪中根据 PCR #1 循环曲线进行孵育（表 8）

表 8. 重亚硫酸盐转化文库的 PCR #1 循环程序

PCR #1 循环曲线（反应体积 100 μL）			
阶段	循环次数	温度	时间
1	1	95 °C	2 min
		95 °C	30 s
2	根据设计大小确定	60 °C	30 s
		72 °C	30 s
3	1	72 °C	7 min
4	1	4 °C	保持

3. Ampure 纯化

- ☐ 涡旋混合 AMPure XP 磁珠
- ☐ 将 DNA 样品与 100 μL AMPure XP 磁珠在室温下孵育 5 min
- ☐ 用新鲜配制的 70% 乙醇 (200 μL) 清洗 2 次
- ☐ 用 25 μL 无核酸酶水洗脱（涡旋混合，孵育 5 min）

4. 标签 PCR

SureSelect XT HS2 靶向序列捕获试剂盒 ILM 杂交模块盒 2 和用于 ILM（捕获前 PCR）的 SureSelect XT HS2 标签引物对。

- ☐ 制备 PCR #2 反应混合物（表 9）

表 9. 制备 PCR #2 反应混合物

试剂	单次反应 (μL)	8 次反应，含冗余 (μL)
无核酸酶水	9	81
5X Herculase II 反应缓冲液含 dNTP（透明）	10	90
Herculase II Fusion DNA 聚合酶（红色）	1	9
	20 μL	180 μL

- ☐ 向 PCR 反应中的每个样品孔中加入 5 μL 适当的 SureSelect XT HS2 标签引物对
- ☐ 向 PCR 反应中的每个样品孔中加入 20 μL PCR #2 混合物。充分混合
- ☐ 在热循环仪中根据 PCR #2 循环曲线进行孵育（表 10）

表 10. 用于标签扩增的 PCR #2 循环程序

PCR #2 循环曲线（反应体积为 50 μL）			
阶段	循环次数	温度	时间
1	1	98 °C	2 min
		98 °C	30 s
2	8	60 °C	30 s
		72 °C	1 min
3	1	72 °C	5 min
4	1	4 °C	保持

5. 最终 Ampure 纯化

- ☐ 涡旋混合 AMPure XP 磁珠
- ☐ 将扩增文库上清液与 50 μL AMPure XP 磁珠在室温下孵育 5 min
- ☐ 用新鲜配制的 70% 乙醇 (200 μL) 清洗 2 次
- ☐ 在 37 °C 的加热模块上干燥样品约 2 min 或在室温下干燥约 10 min
- ☐ 用 25 μL 无核酸酶水洗脱（涡旋混合，孵育 2 min）

分析流程

在实验研究中，根据性别（雄性/雌性）和处理方式（生理盐水/VPA）对 4 组小鼠（每组 8 只）进行了分类。根据上述方案共处理了 64 个样品（32 个血液和 32 个大脑）。在 NextSeq550 (Illumina) 的中通量流动槽上进行测序，每次运行对 16 个样品进行多重分析，参数如下：配对末端，读长 151，双标签，每个样品产生 1000 万个 read。

以两种方式进行分析：

- 命令行脚本
- 商业软件

为分析创建的命令行管道主要使用 Bismark (<https://github.com/FelixKrueger/Bismark>) 进行清理、比对（使用 Bowtie2 (<https://github.com/BenLangmead/bowtie2>) 比对程序）和用于识别阶段。然后使用 BiSeq R 软件包 (<https://www.bioconductor.org/packages/release/bioc/html/BiSeq.html>) 计算覆盖率，检查测序和比对的质量，并通过分析确定差异性甲基化区域 (DMRs)。

比对统计数据

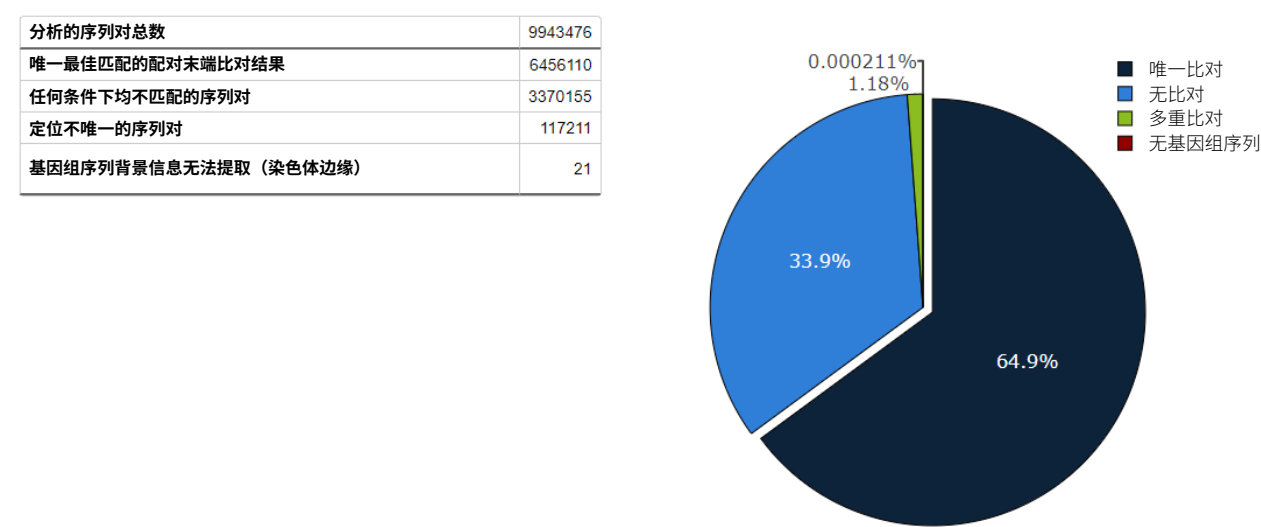


图 1. 重亚硫酸盐处理样品 read 比对后的 Bismark 报告。它根据使用的基因组构建对基本比对指标进行分组

在捕获区域上，read 深度通常超过 100X（图 2）。但是，根据定制设计策略的预期，每个样品的 read 总数有限（每个样品仅 1000 万个 read）（详见定制设计章节）。所有样品的在靶率 > 75%，显示出了 Agilent SureSelect 3 Mb 甲基化测序定制设计的高效捕获性能。

基于 Bismark 的详细结果，可以根据研究的不同参数表示数据，然后突出显示胞嘧啶、基因或目标区域（数据未显示）。

对于所有与单个胞嘧啶或整个基因相关的分析，开发了特定的 Perl 脚本。

已创建了一个 Github 专区来共享用于此项目的所有脚本。请参阅：

[Github-qRp/Methylation_scripts](#)。

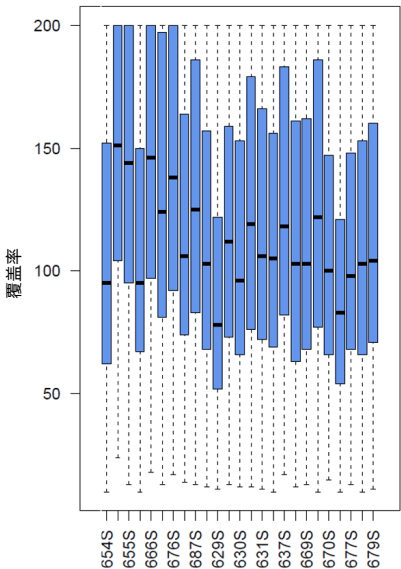


图 2. 每个样品在捕获区域的平均覆盖率。平均获得 100X 以上的测序深度

另外还使用商业软件StrandNGS (<https://www.strand-ngs.com/>) 分析了结果。对于 StrandNGS 中的分析,使用了之前步骤中生成的 BAM 文件。StrandNGS 收集 Bismark 在测序 read 上生成的“标签”,并将其整合到内置分析流程中。通过将定制设计靶向 BED 文件合并到系统中,可以将分析限制在目标区域。其余的分析遵循 StrandNGS 向导驱动的甲基化测序分析工作流程。对于此处列出的研究,使用默认设置。

这种方法的优势体现在非生物信息学专家可轻松使用。分析的连续步骤包括检测每个样品中的甲基化,然后通过差异甲基化位点 (DMCs) 评价甲基化差异,并将这些 DMC 分析转化为差异甲基化区域 (DMRs)。然后对结果进行注释(通过相关基因、GO 分析或通路分析)。StrandNGS 软件还包括各种图形数据展示和数据库集成,非统计专家也可轻松使用。

使用 StrandNGS 集成基因组浏览器(图 3)、直方图(图 4)和棒棒糖状图(图 5)将本研究结果进行了可视化。组间比较可以生成 DMCs 和 DMRs 列表,可使用 StrandNGS 软件轻松编译为文氏图。

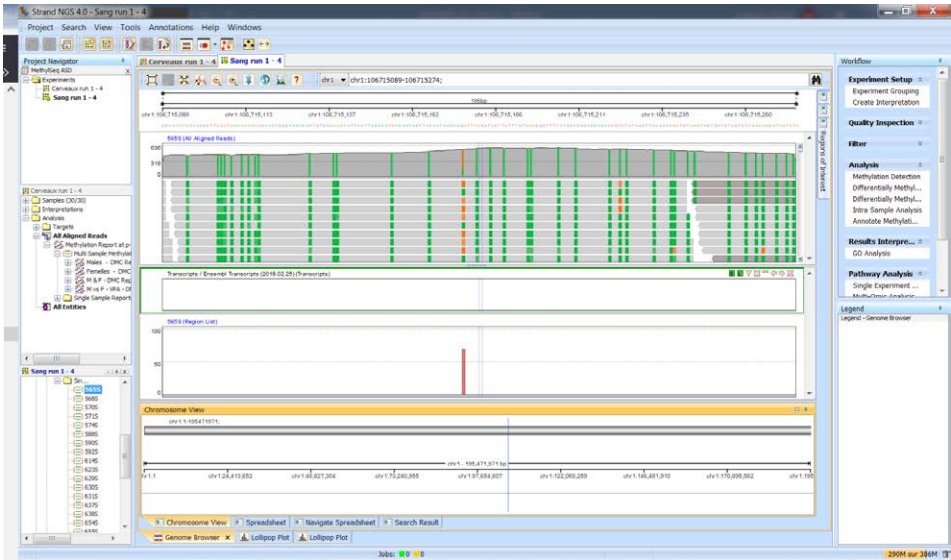


图 3. 1 号染色体上捕获的一个区域在 StrandNGS 集成基因组浏览器中的可视化结果,展示了甲基化率约为 75% 的胞嘧啶

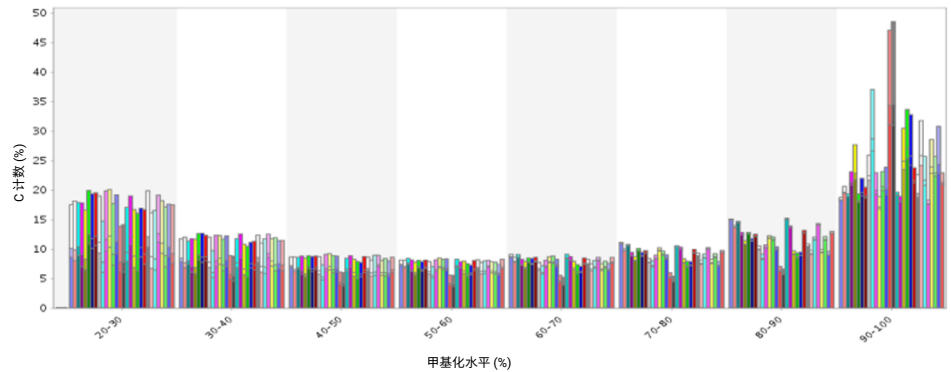


图 4. 分析样品中甲基化整体分布以直方图形式的可视化结果 (n = 30)。分析的大多数胞嘧啶甲基化率 > 90% 或 < 30%

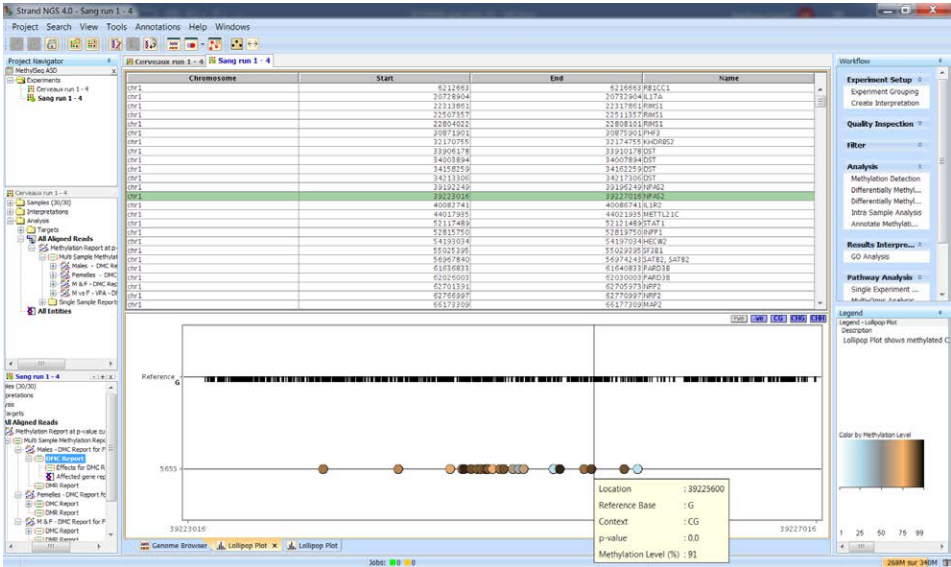


图 5. 棒棒糖状图,显示一个血样中 1 号染色体对应于 NPAS2 基因区域的 DMC 分布

这两种互补生物信息学方法的结合可在至少有一个 DMC 的所有组之间鉴定出 67 个常见基因（图 6）。这种联合分析证实，大脑中甲基化的差异很大程度上发生在“分离”的胞嘧啶上（并非位于 CpG 岛），而在血液中未观察到甲基化分布的 CpG 或 CHG/CHH 背景信息的差异（图 7）。此外，利用靶向序列捕获甲基化测序通过聚焦于特定的目标区域，有助于显示非常小的甲基化变异。

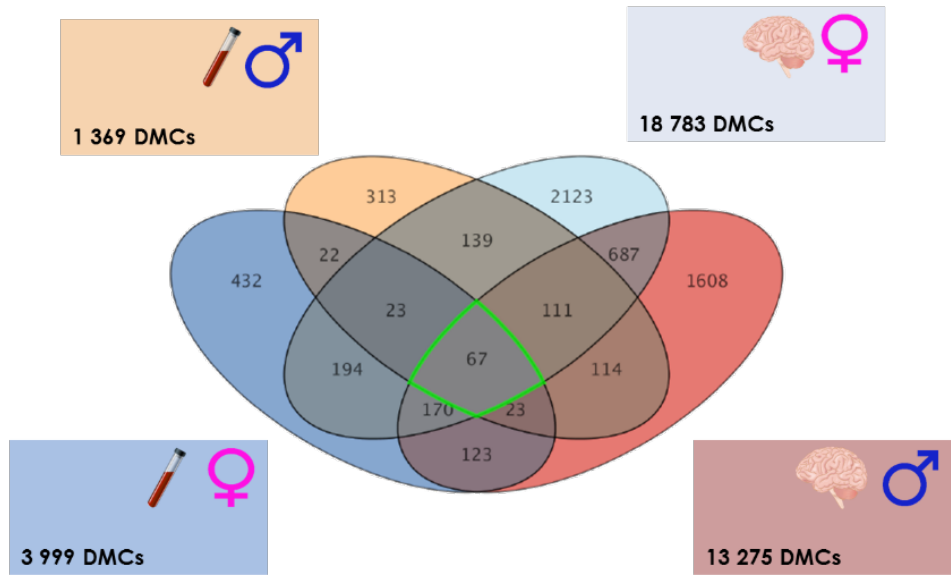


图 6. 文氏图，显示组间受 DMC 影响的基因转录本

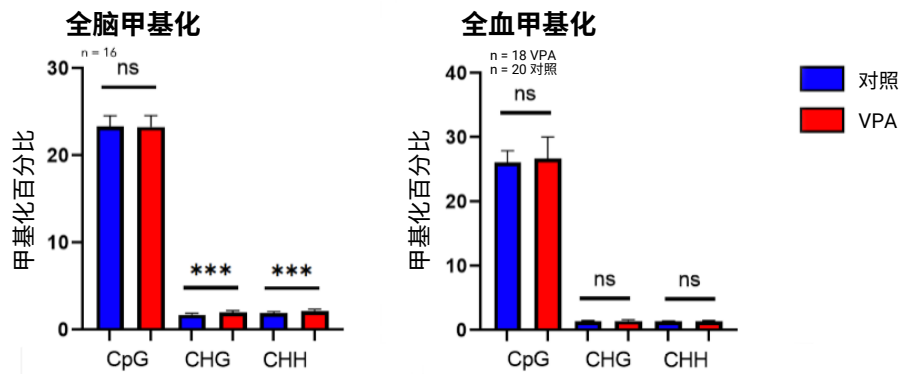


图 7. 根据胞嘧啶定位的背景信息（CpG 岛或 CpG 外：CHG 或 CHH），对照组小鼠与 VPA 小鼠大脑（左）和血液（右）中整体甲基化水平的比较。Mann-Whitney 统计检验

文库混合和测序

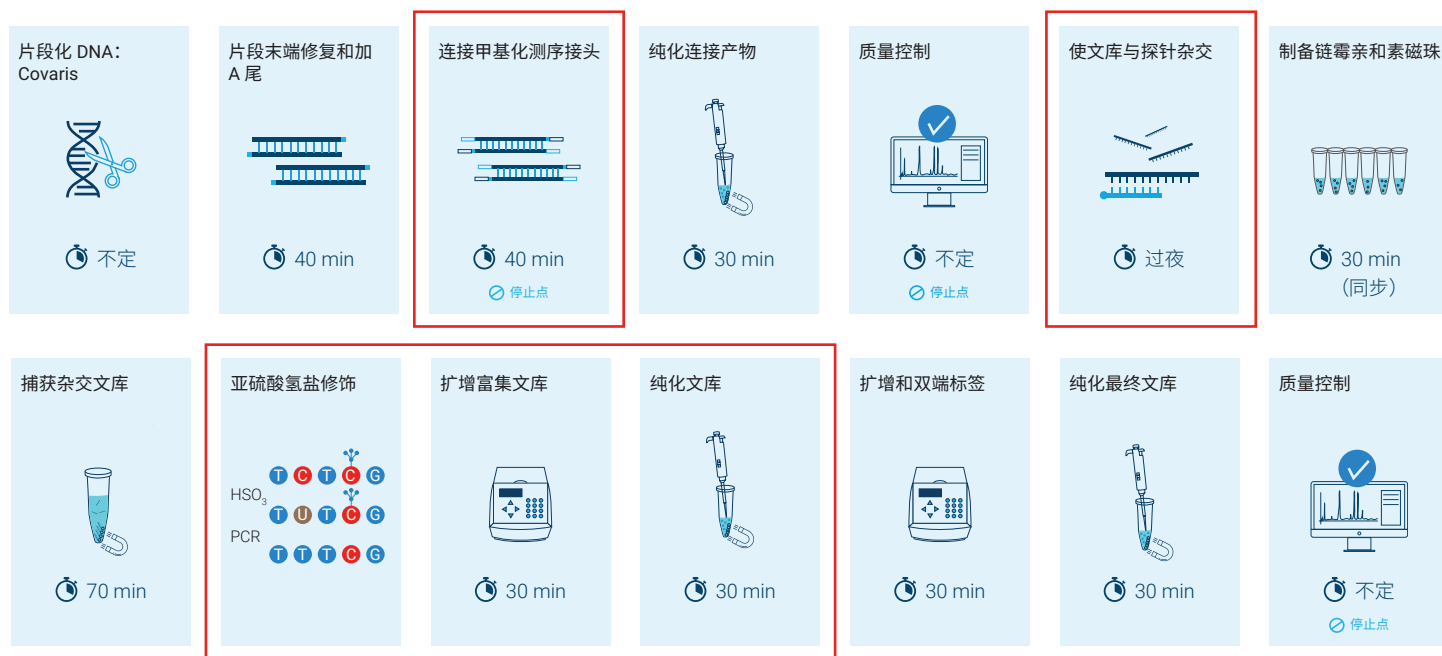


图 8. 改良的甲基化测序 SureSelect 序列靶向捕获方案概览

结论

本项目使用定制捕获 Panel 有三大优点。首先，它让分析预选的候选基因和基因组区域列表成为了可能，因此允许对更多样品进行多重分析，并最终节约了整体测序成本。其次，捕获能够分析 CpG 岛以外的胞嘧啶，这在研究胚胎干细胞或研究在 CpG 背景信息之外呈现高水平甲基化的神经元时特别有用^[3]。最后，通过使用靶向序列捕获，获得的测序深度有所增加（平均 > 100X，在某些目标区域 > 400X），能够可靠地检测仅有几个百分点的甲基化差异。

本研究的完整结果将在即将举行的研讨会上讨论，并将发表。

参考文献

1. Moore, L. D.; Le, T.; Fan, G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology* **2012**, 38 (1), 23–38. <https://doi.org/10.1038/npp.2012.112>.
2. Jang, H. S.; Shin, W.; Lee, J. E.; Tae, J. CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function. *Genes* **2017**, 8 (6), 148. <https://doi.org/10.3390/genes8060148>.
3. Lister, R.; Mukamel, E. A.; Nery, J. R.; Urich, M. A.; Puddifoot, C. A.; Johnson, N.; Lucero, J.; Huang, Y.; Dwork, A. J.; Schultz, M. D.; et al. R. Global Epigenomic Reconfiguration during Mammalian Brain Development. *Science* **2013**, 341 (6146). <https://doi.org/10.1126/science.1237905>.

www.agilent.com

仅供科研使用。不用于临床诊断用途。
PR7001-1658

本文中的信息、说明和指标如有变更，恕不另行通知。

© 安捷伦科技（中国）有限公司，2023
2023 年 12 月 5 日，中国出版
5994-6841ZHCN

查找当地的安捷伦客户中心：

www.agilent.com/chem/contactus-cn

免费专线：

800-820-3278，400-820-3278（手机用户）

联系我们：

LSCA-China_800@agilent.com

在线询价：

www.agilent.com/chem/erfq-cn

 **Agilent**
Trusted Answers