

# 在 Opentrons Flex® 自动化移液工作站上运行自动化 PacBio® HiFi 测序的文库制备



## 作者

D'Andra Howell,<sup>1</sup> Matthew Akana,<sup>1</sup> Christine Lambert,<sup>2</sup>

Gregory Young,<sup>2</sup> and Kinnari Watson, PhD<sup>1</sup>

<sup>1</sup>Opentrons Labworks, Inc., <sup>2</sup>Pacific Biosciences

## 摘要

长读长测序技术对于组装全基因组的科学家来说是一项颠覆性技术。PacBio 的 HiFi 测序是一种长读长测序方法，它能提供类似短读长的准确性，但读长可达数千碱基。在 Opentrons Flex 移液工作站上为 HiFi 测序制备样本涉及两个环节：(1) 通过其短读长去除 (SRE) 试剂盒消除短 DNA 并对剩余的 DNA 进行机械片段化，(2) 构建用于 HiFi 测序的 DNA 文库。在本应用说明中，我们演示了如何在 Opentrons Flex 上自动化这一流程，从而以更少的用户人为干预和手动操作时间生产高质量的文库。

## 主要亮点：

- 在八小时内自动化完成 SRE、DNA 剪切和文库制备，手动操作时间仅需一小时。
- 自动化文库制备可提供与手动方法相似的高质量结果，满足 PacBio 在产量、条形码 (barcode) 和碱基质量方面的基准。

## 引言

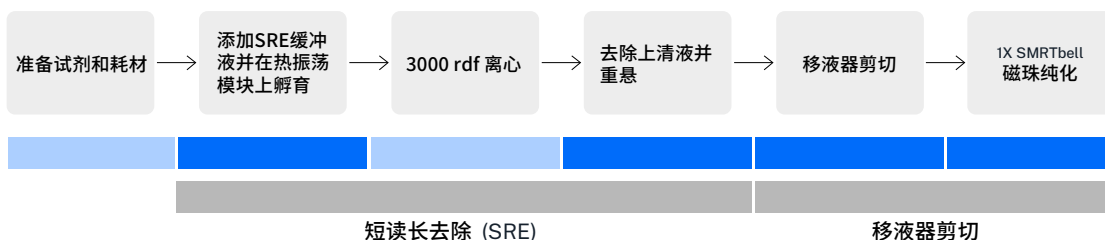
PacBio 长读长 HiFi 测序技术能够对长度范围从 0.5 kb 到超过 25 kb 的 DNA 文库进行准确可靠的测序。此外，该技术还能同时检测碱基修饰，从而能够表征基因组和表观基因组。这使得研究人员能够通过一次实验即可实现从头基因组组装、解析单倍型、检测复杂变异，并基于 Fiber-seq 实验获得的甲基化信息和染色质可及性，对表观遗传状态进行分类。

HiFi 文库制备工作流程首先通过使用 SRE 去除样本中降解的短 DNA 片段来提高样本质量。如果起始 DNA 质量高或 DNA 量有限，此步骤是可选的。接下来对剩余的高分子量 DNA 进行片段化步骤（即剪切），以优化用于 HiFi 测序的片段大小分布。然后，工作流程继续进行文库制备，涉及一系列酶促反应，类似于传统的 NGS 工作流程：末端修复、A 尾添加和接头连接。连接后，采用核酸酶反应去除未连接的片段。最终产物是一种称为 SMRTbell® 文库的环状分子。为了简化此过程，我们使用两个连续的运行脚本在 Opentrons Flex 上调整了此工作流程：(1) 去除短读长并将 DNA 剪切至适当大小，以及 (2) 对剪切后 DNA 进行文库制备。

使用如 Opentrons Flex 自动化移液工作站这样的液体处理系统自动化此工作流程具有许多优势，包括：(1) 减少科学家的手动操作时间，(2) 减少用户错误，以及 (3) 通过使用自动化移液器进行剪切，无需专门的 DNA 剪切设备。在此，我们演示了如何利用自动化在 8 小时内生成 HiFi 测序文库，且手动操作时间仅需 1 小时（图 1）。

## 实验方法

### 短读长去除和移液器剪切工作流程

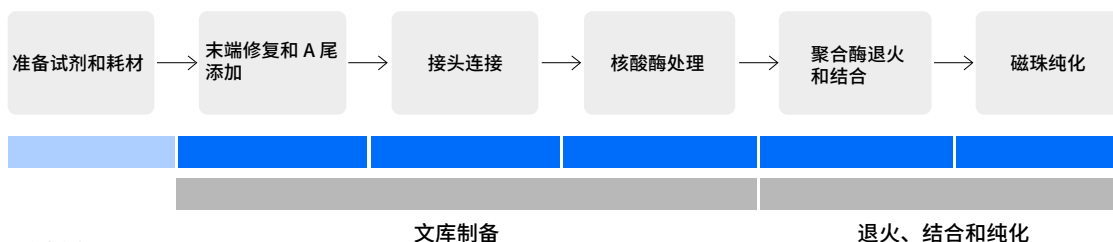


8 个样本：总时间 3.5 小时

48 个样本：总时间 6.5 小时

约 20 分钟手动操作时间

### HIFI 文库制备工作流程



8 个样本：总时间 4.5 小时

48 个样本：总时间 8.5 小时

约 30 分钟手动操作时间

图 1. 短读长去除和 DNA 片段化方案（上图）以及 HIFI 文库制备方案（下图）的工作流程步骤。

工作流程每个阶段下方显示了手动步骤（浅蓝色）和 Opentrons Flex 上的自动化步骤（深蓝色）。

**样本制备：**高分子量（HMW）DNA 来源于 Coriel 医学研究所（目录号：HH01891），以及使用 PacBio Nanobind PanDNA kit（货号：103-260-000）从人类唾液和全血中提取。

**短读长去除（SRE）和 DNA 片段化：**我们使用 PacBio SRE HT kit（货号：103-124-500）去除短读长，并在 Opentrons Flex 自动化移液工作站上使用 PacBio 短读长去除和移液器剪切方案（[library.opentrons.com/p/pacbio-sre-shearing](https://library.opentrons.com/p/pacbio-sre-shearing)），通过移液器剪切将高分子量 DNA 片段化至 15-20 kb。

**片段化 DNA 分析：**使用 Agilent Femto Pulse 系统（货号：M5330AA）和 Genomic DNA 165 kb kit（货号：FP-1002-0275）对剪切前和剪切后的 DNA 进行分析。

**全基因组测序 HIFI 文库制备：**使用片段化 DNA（ $\leq 3 \mu\text{g}$ ），通过 PacBio HIFI prep kit 96（货号：103-381-200）和 Revio SPRQ polymerase kit 96（货号：103-497-000）生成可用于测序的 HIFI 文库。这些试剂可以一起购买，即 Revio® SPRQ™ HIFI prep kit 96（货号：103-522-600）。该方案在 Opentrons Flex NGS 工作站（货号：991-00354）上使用 PacBio 全基因组测序 HIFI 文库制备方案实现了自动化。（[library.opentrons.com/p/pacbio-hifi-prep](https://library.opentrons.com/p/pacbio-hifi-prep)）

**HIFI 测序：**使用 PacBio Revio 系统对 HIFI 文库进行测序，每个 SMRT® Cell 上样一个文库复合物。使用 Revio SPRQ sequencing plate and polymerase kit，测序运行约 24 小时。文库以 250 pM 的浓度采用适应性上样方式进行上样。

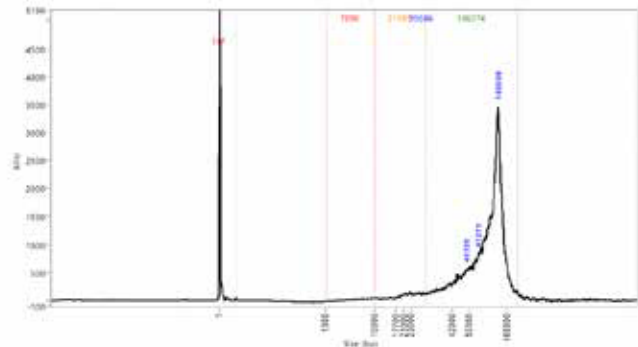
**HiFi 测序文库分析:** 使用 Qubit™ IX dsDNA High Sensitivity Assay Kit (货号: Q33231) 和 Qubit 4 荧光计对 DNA 进行定量, 以评估输入 DNA 质量、HiFi 文库质量以及退火、结合和纯化步骤后的文库复合物质量。使用 SMRT® Link 分析软件处理原始测序数据, 以生成质量控制报告并确定检测到的条形码读长百分比、条形码/样本交叉污染以及碱基质量。

## 结果与讨论

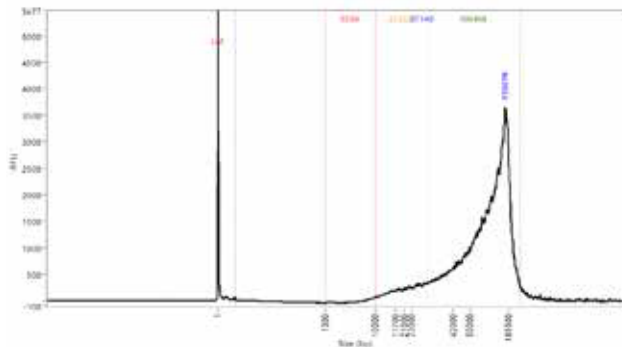
### 短读长去除和 DNA 剪切

为了制备用于 PacBio HiFi 文库制备和测序的高分子量 DNA, 我们使用 PacBio 短读长去除和移液器剪切工作流程 ([library.opentrons.com/p/pacbio-sre-shearing](https://library.opentrons.com/p/pacbio-sre-shearing)) 在 Opentrons Flex 上自动化运行。该方案首先逐步去除高达 25 kb 的片段, 这提高了样本质量, 特别是在样本部分降解的情况下。短片段去除后, 高分子量 DNA 随后通过移液器剪切至 15-20 kb, 并使用 SMRTbell 磁珠进行纯化。我们使用 Agilent Femto Pulse 系统分析了这些片段。图 2A 显示了输入 DNA, 图 2B 显示了短读长去除后的输出。图 2C 显示了移液器剪切后的大小分布。如图所示, 经过此工作流程后的 DNA 大小约为 15 kb (平均 15,472 bp, 峰值在 14,739 bp), 这是 PacBio HiFi 文库制备所需的理想输入大小。

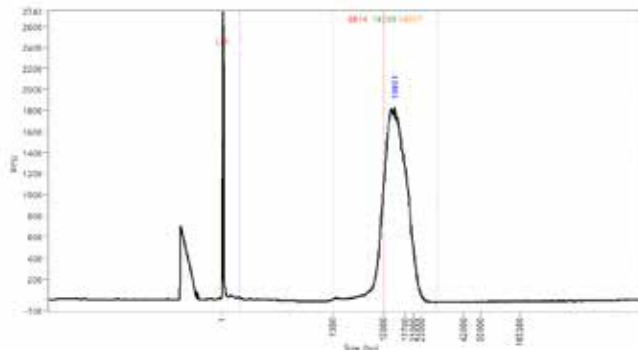
(A) 输入高分子量 DNA



(B) 短读长去除后的高分子量 DNA



(C) 剪切和 1X SMRTbell 磁珠纯化后的高分子量 DNA



**图 2. 短读长去除和移液器剪切过程中高分子量片段大小的分析。** (A) 工作流程开始前的输入高分子量 DNA。 (B) 短读长去除后的高分子量 DNA。 (C) 剪切和 1X SMRTbell 磁珠纯化后的高分子量 DNA。

在 Opentrons Flex 上自动化此过程的一个优点是无需购买专门的 DNA 剪切设备。并且, 我们证明了能够使用移液器剪切生成片段化 DNA, 从而简化流程并降低成本。总的来说, 我们能够在 3.5 小时内完成五个样本的 SRE 和剪切。

### 全基因组测序 HiFi 文库制备

工作流程的下一步是使用 PacBio 全基因组测序 HiFi 文库制备工作流程 ([library.opentrons.com/p/pacbio-hifi-prep](https://library.opentrons.com/p/pacbio-hifi-prep)) 制备 HiFi DNA 文库。分析文库制备成功与否的关键指标需要比较制备文库后回收的 DNA 量 (文库质量)、加载聚合酶后 (ABC 复合质量) 与输入 DNA 的量。我们发现总回收率 (即加载聚合酶后剩余的输入 DNA 百分比) 符合预期的 10-25% 回收率范围 (表 1)。

**表 1. 全基因组测序 HiFi 文库制备方案后的 DNA 质量和回收百分比。**输入质量代表片段化后的 DNA 质量。文库质量是制备文库后、与聚合酶复合前回收的 DNA 量。ABC（退火、结合、纯化）质量是最终纯化后与聚合酶复合的 DNA 量。总回收率 % = (ABC 复合质量) / (输入质量) × 100。

样本 ID	输入质量 (ng)	文库质量 (ng)	ABC 复合质量 (ng)	总回收率 (%)
Coriel_1	1070	178.5	141	13.2
Coriel_2	1120	180.7	146	13
WB_1	1049.6	224.25	207.25	19.7
WB_2	1090.6	286.5	280	25.7

制备文库后，下一步是对文库进行测序并评估质量控制指标。为此，我们首先查看了条形码。每个样本都被分配一个唯一的条形码，以便识别样本。我们发现，在我们的五个重复样本中，每个样本都只包含一种条形码，表明样本之间没有交叉污染（表2）。平均而言，93.4% 的读长包含条形码。

**表 2. HiFi 测序指标**

样本 ID	HiFi 产量 (Gb)	P1 Loading	HiFi 读长 N50 (bp)	碱基质量 ≥Q30 (%)	覆盖度
Coriel_1	98.33	59.17	17.18	93.88	30.7
Coriel_2	113.71	64.49	15.14	95.14	35.5
WB_1	116.25	72.51	15.76	93.79	36.3
WB_2	116.29	72.14	15.52	93.7	36.3

接下来，我们转向测序指标，如 HiFi 产量、Q30 质量分数和读长，以评估文库的测序效果。产量是指测序的碱基数量，理想情况下应在 100 Gb 到 120 Gb 之间。我们发现我们的数据基本符合这种情况，因为我们的重复样本产量范围在 98.33 Gb 到 115.29 Gb 之间。我们的数据显示读长 N50 为 15.9 kb，在推荐的最佳全基因组测序结果的 15-20 kb 大小范围内。对于我们的重复样本，我们发现 93-95% 的碱基高于 Q30，符合 PacBio 的 ≥90% Q30 的成功标准。

## 结论

我们发现，自动化 PacBio HiFi 测序的全基因组测序文库制备可产生高质量的文库和测序结果。这些工作流程的 Opentrons 方案可在 8 小时内完成，且手动操作时间仅需 1 小时。

## 致谢

作者感谢 Psomagen 公司的 Cristian Bravo、Alice Jung 和 Elly Lee 对方法开发提供的宝贵反馈。

[www.opentrons.com.cn](http://www.opentrons.com.cn)

☎ 18098952246

✉ [Marketing.china@opentrons.com](mailto:Marketing.china@opentrons.com)

