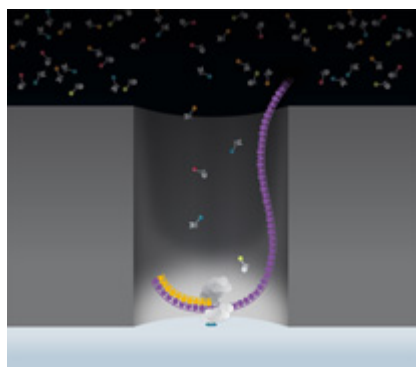


纵观第三代测序之 Pacific Biosciences

众所周知，天然的 DNA 复制本身就是一个非常高效且准确的过程。真核生物 DNA 复制的速度达几十个核苷酸/秒。若能以这样的神速测序，那可比 Sanger 测序快了几万倍。然而，知易行难。如何在不影响聚合酶活性的前提下实时观察 DNA 合成，这可是个棘手的问题。要知道，DNA 聚合酶的直径只有 15 nm。而且，这是一个涉及酶学、表面化学和检测光学的多学科难题。

早在 10 多年前，美国康奈尔大学的研究生 Steven Turner 和 Jonas Korlach 就开始为解决这一难题而做出不断努力。如今，Korlach 和 Turner 创办了一家名为 Pacific Biosciences（以下简称 PacBio）的公司，并且终于将他们的想法转变成现实产品。让我们看看他们主要克服了哪些挑战。

实时观察 DNA 聚合酶的主要挑战之一是如何能在 DNA 合成期间检测到单个核苷酸的掺入。因为在显微镜实时记录 DNA 链上荧光的时候，周围众多的荧光标记的核苷酸形成了非常强大的荧光背景。这种强大的荧光背景使单分子的荧光探测成为不可能。正在他们一筹莫展之际，突然，微波炉门给了他们启发。微波炉门上的金属筛布满了小洞，它们比微波的波长要小得多。因此，这些洞能阻止微波通过并穿透玻璃。然而，比波长更小的可见光就能够通过，让我们看到食物。他们的 SMRT 技术应用了相同的原理，不过规格就缩小至纳米级，称之为 ZMW（zero-mode waveguide，零模式波导）。



ZMW 是一个直径为几十纳米的小孔，它阻止可见的激光(波长大约为 600 nm)完全透过 ZMW。激光在进入 ZMW 后迅速衰减，因此，只有底部 30 nm 被照亮(下图)。在每个 ZMW 中，单个 DNA 聚合酶分子利用专利技术锚定在底部玻璃的表面。随后核苷酸涌入 ZMW 中，并在阵列表面扩散。当聚合酶检测到正确的核苷酸时，便将其掺入新生链中，这个过程需要几毫秒，而单纯的扩散只需要几微秒。这种时间差使掺入的核苷酸产生了很高的信号强度，类似于脉冲信号。因此，ZMW 有能力在荧光标记核苷酸的背景下检测单个掺入事件。

测序是在专利的 SMRT Cell 中开展的，每个 Cell 中有一个阵列，上面大约有 75,000 个 ZMW。每个 ZMW 都能够包含一个 DNA 聚合酶及一条不同的 DNA 样品链。这样，SMRT Cell 能够平行检测大约 75,000 个单分子测序反应。

Korlach 与 Turner 还有一项重大发明。传统的核苷酸标记方法是将荧光标记连接到核苷酸的碱基上，也就是掺入 DNA 链中。这对于实时观察 DNA 合成也是有问题的，因为大分子的染料会干扰 DNA 聚合酶的活性，或造成聚合反应提前终止。因此，Korlach 和 Turner 就开发出一种新型的核苷酸标记方法，即在核苷酸的磷酸链上进行标记，而不是碱基，这样，一旦核苷酸掺入到新生 DNA 链中，标记基团就会自动脱落。

当然，要将 ZMW 中的反应记录下来，还需要一台高精尖的仪器。PacBio RS 于是诞生了，它能够实时开展、监控并分析单分子生化反应。PacBio RS 使用一个大孔径物镜和四个单光子照相机来收集荧光所发射的光脉冲，观察整个过程；并使用一套优化的算法，将光学系统所捕获的信息翻译成 ACGT 碱基。一旦测序开始，实时的数据就传送到初步分析流水线，生成碱基身份和质量值。

以上三种新技术结合，就能让研究人员在短短的几分钟内对长片段 DNA 进行测序。它的速度、读长、灵活性和费用让其从第三代测序中脱颖而出。从样本制备到测序，所需的时间还不到一天。典型的测序运行时间低至 30 分钟。序列数据在几分钟之内就能产生，而不是几天，这对于传染病监控和分子病理学尤为重要。PacBio RS 目前的读长超过 1kb，比第二代测序要长得多。此外，试剂消耗和样本制备极少，不需要常规的 PCR 扩增，失误也大大减少，聚合酶动力学的直接观察赋予了测序之外的更多应用。凭借这些优势，PacBio 今年被美国麻省理工学院（MIT）的《技术评论》（Technology Review）杂志评为全球 50 家最具创新力的企业之一。

一系列的文章也证明了 SMRT 技术。Korlach 与 Turner 于 2009 年 2 月在《科学》杂志上发表了一篇介绍 PacBio 单分子 DNA 测序技术的文章。这篇文章代表了首个第三代测序技术的“原理验证”，因此引用率非常高。而后，他们又利用 SMRT 技术，直接测定了 DNA 的甲基化，这相对目前流行的第二代测序技术又前进了一大步。文章发表在 今年 5 月的《Nature Methods》上。

荧光脉冲的到达时间和持续时间反映了有关聚合酶动力学的信息，从而允许直接检测 DNA 模板链中的修饰核苷酸，包括 N6-甲基腺嘌呤、5-甲基胞嘧啶和 5-羟甲基胞嘧啶。各种修饰对聚合酶动力学的影响不一，从而能够将它们区分开。研究人员使用这些动力学特征，鉴定出基因组样品中的腺

嘌呤甲基化，并发现再结合 circular consensus sequencing，他们能够在单碱基分辨率上鉴定出表观遗传学修饰（mA、mC 和 hmC）。研究人员还预计，其它的表现遗传学修饰，以及各种形式的 DNA 伤害，也能用这种方法检测。

看到这里，大家可能迫不及待地想知道 PacBio 的测序仪何时上市，价格多少。其实，PacBio RS 已成型，目前正在早期试用阶段。今年 2 月份，PacBio 宣布了 10 位北美的早期试用客户，包括贝勒医学院、冷泉港实验室、美国能源部联合基因组研究所、华盛顿大学基因组中心和斯坦福大学等等，并与第二季度开始发货。8 月份，PacBio 又迎来了它的首位欧洲试用客户，大名鼎鼎的 Wellcome Trust Sanger 研究院。

整个测序系统包括 PacBio RS 仪器以及专利的耗材。PacBio RS 在美国的目录价为 695,000 美元。耗材部分则包括 SMRT Cell 和试剂盒。PacBio RS 上的每个反应消耗一个 SMRT Cell。8 个 SMRT Cell 包装成 8Pac 的形式。单个 SMRT Cell 在美国的售价大约为 99 美元，8Pac 的包装则低于 800 美元。

另外，他们还提供了三种试剂盒，用途各不同。Template Preparation Kit 将 DNA 转化成 SMRTbell 文库形式。Binding Kit 则将这个文库与聚合酶结合，为测序做准备。Sequencing Kit 包含了实时测序所需的试剂。每个样本可在单个 SMRT Cell 上测序，也可根据项目需求在多个 SMRT Cell 中测序。因此，每个反应的价格取决于实验设计。

PacBio 预计，到 2013 年，个人基因组的测序能在 15 分钟内完成，费用低于 1000 美元，人人都可以消费得起。说实话，2008 年我第一次听到这个消息时，心中饱存疑惑，看来我实在是低估了 PacBio 的实力，也低估了测序技术发展的脚步。

（生物通 余亮）