

# 2012 值得关注的技术：单倍体因果突变

高通量测序技术的快速发展将研究焦点从数据获取转移到了数据解读。人类基因组领域的研究人员在初级序列分析方面获得了很大的进步——不仅能针对短片段，单核苷酸多态性，插入缺失进行图谱绘制，甚至可以解析更复杂的结构突变。

《Nature Methods》盘点 2011 年度技术，选出了最受关注的技术成果：人工核酸酶介导的基因组编辑（genome editing with engineered nucleases）技术。

除了基因组编辑以外，《Nature Methods》也整理出了 2011 年最值得关注的几项技术，分别为：单细胞技术（Single-cell methods）、功能基因组资源（Functional genomic resources）、糖蛋白组学（Glycoproteomics）、单倍体因果突变（Causal mutations in a haploid landscape）、单层光生物成像（Imaging life with thin sheets of light）、非模式生物（Non - model organisms）、光基础电生理学（Light-based electrophysiology）和 RNA 结构（RNA structures）。

高通量测序技术的快速发展将研究焦点从数据获取转移到了数据解读。人类基因组领域的研究人员在初级序列分析方面获得了很大的进步——不仅能针对短片段，单核苷酸多态性，插入缺失进行图谱绘制，甚至可以解析更复杂的结构突变。

但是这方面仍然存在两个挑战，第一个是确定单体型——个体基因组中每个染色体中各自的序列，尤其是没有相关基因组信息的情况下。其次是分析所有已发现的突变的功能作用，并从中找到与疾病相关的突变。

2011 年取得了个体单体型基因组分析的一些成果，比如来自马普研究院的 fosmid 克隆高通量测序，以及由斯坦福大学等处的研究人员完成的首个个人单体型分析成果。这篇发表于 Nature Biotechnology 杂志上的成果，与其它文章一道，介绍了几种测定出个人单倍型的新方法，这将可用于揭示某些疾病的遗传原因，并应用于医学诊断。

其中来自华盛顿大学的研究人员在 40 万个碱基增量中对单体型分辨基因组进行了测序。来自斯坦福大学的研究人员对来自单细胞的单倍染色体进行了基因分型。

在第一项研究中，研究人员利用了一个带有 4 万个插入碱基的印度人基因组 DNA fosmid 文库，将其裂解生成了 115 个包含超过 5000 个克隆的混合池。这些混合池 99% 衍生于同源染色体的其中一条染色体序列，每个混合池均覆盖大约 3% 的基因组。在对这些混合池进行测序后研究人员确定了 40 万个碱基的单体型，证实其与目前单体型图计划（HapMap）取得的 99% 的结果相关联。

另外一个研究组取得了一些基础的种群遗传学研究结果，找到了过去在 HapMap 或千人基因组计划中未发现的单体型，并证实其富集于一些新型变异中。此外研究人员还改良了从头基因组组装技术。近期针对个体基因组的鸟枪测序数据揭示了一些从前在参考基因组中未发现的序列。研究人员将

从他们的克隆中获得的读数定位到这些基因序列群中，并针对其来源的 fosmid 库进行了追踪，进而找到了它们的共享位点。此外，研究人员还将一些读数锚定到了从前未曾定位的基因序列群中。

这两种方法的主要差别之一在于前者是利用细菌进行 fosmid 克隆扩增，而后者则主要是在体外进行多重置换扩增（MDA）。

这些研究成果表明在许多基因中存在新突变，分析这些突变将有助于评估它们对于个体的影响。

了解单倍体突变分布还只是第一步，我们还需要更深入的分析这些突变对于人体有何益处——

这正是 Gregory Cooper 和 Jay Shendure 在 *Nat. Rev. Genet.* 上发表的文章中提到的：“一旦杂草被清理干净了，在一堆针中找到自己需要的针就简单了”

在分析蛋白编码序列，和基因组非编码区域潜在突变的时候，也许可以借助于基于序列进化保守性的计算机分析方法，蛋白序列生物化学特性，以及结构信息分析。但是这些方法也仅仅是在候选名单中，我们仍然需要能分析这些突变，以及分子表型的大规模实验方法。

（生物通：万纹）