

DNA 杂交测序法

米志勇 柯 岫 王丽霞

(中国科学院发育生物学研究所植物发育分子生物学实验室, 北京 100080)

自从 Sanger 双脱氧末端终止法和 Maxam-Gilber 化学法问世以来, 迄今已测定的 DNA 核苷酸序列累计达 1×10^8 bp 以上。但随着各种基因组计划的实施, 这两种传统的测序方法, 无论在效率或成本上, 都不足以满足将来的测序需要。因而客观上急需发展更加快捷经济的新的 DNA 序列测定技术。

在八十年代末九十年代初, 先后就有若干实验小组报道了一种新的 DNA 序列测定技术, 即杂交测序法 (SBH)。这一技术的基本原理是通过单链标记的 DNA 样品, 分别与许多种特定长度的随机寡核苷酸片段进行充分杂交, 形成完全互补的双链, 从而排列出样品的碱基序列。

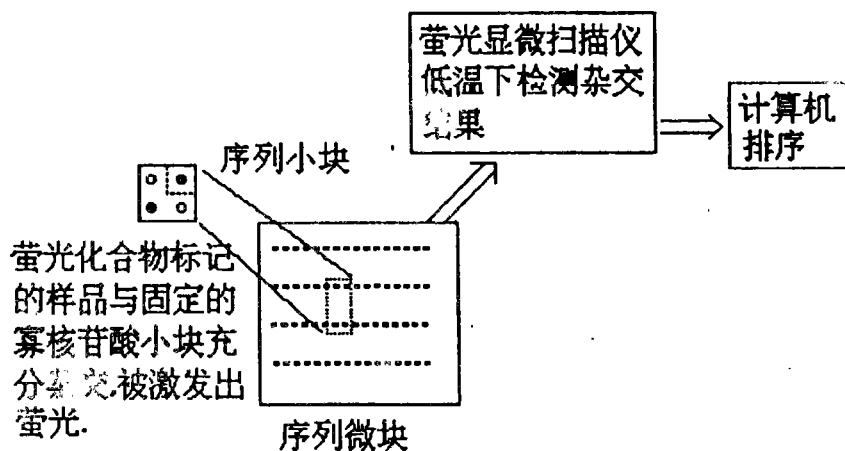


图 1

一、固定的随机寡核苷酸矩阵

早期的实验发现, 过短的寡核苷酸片段不能充分紧密地结合在硝酸纤维素膜或尼龙膜上。现在通行的做法是利用共价结合在寡核苷酸 3' 或 5' 末端的衔接物与玻璃板或凝胶结合,

从而形成固定住的寡核苷酸小块, 不同碱基组合的寡核苷酸小块排列在一起形成序列微块。寡核苷酸小块的数目取决于寡核苷酸片段的长度。若寡核苷酸长为 8bp 则需要 4^8 (65536) 种寡核苷酸小块的排列。现在的技术已可将 4^8 种寡

核苷酸小块排列成 3×3cm 大小的序列微块。这一序列微块可用来杂交测序。就目前来说分别合成成千上万种不同的寡核苷酸片段,价格非常昂贵,不易推广。但部分实验室已展开了新的合成技术的研究,采用光驱动的化学反应在玻璃板上直接平行地合成寡核苷酸片段,为工业化生产序列微块开创了前景。

二、杂交的检测

单链 DNA 样品可以用放射性同位素或荧光化合物标记,然后用放射性检测仪或荧光显微扫描仪在几分钟之内得出杂交结果。现在所面临的问题是如何快速地检测或消除寡核苷酸片段与样品间不完全的配对(错配)。已有实验证明在寡核苷酸内部一个碱基的错配即可造成杂交信号的明显差异。因而双链内部的错配很

容易同完全互补的双链区分开。尽管双链的稳定性随寡核苷酸长度的减少而降低,然而越短也就越易区分出内部错配的双链。但 3'端或 5'端的错配却很难用杂交的条件来消除。同时由于不同寡核苷酸的 G-C 含量有时差异极大,所以寻找到一种最佳的杂交条件非常困难。现在通常用增加 AT 含量多的寡核苷酸片段长度或减少 G-C 含量多的寡核苷酸片段长度来调节,使众多的杂交反应有一个均一的 T_m 值。同时可添加 TMAC (tetramethylammonium) 增加 AT 的稳定性,或利用甜菜碱(甘氨酸三甲内盐)降低由碱基组而造成 T_m 值的不同的效应。最近 C. R. Cantor 提出了端位 SBH(termed positional SBH)来消除 5'端或 3'端的错配(图 2)。

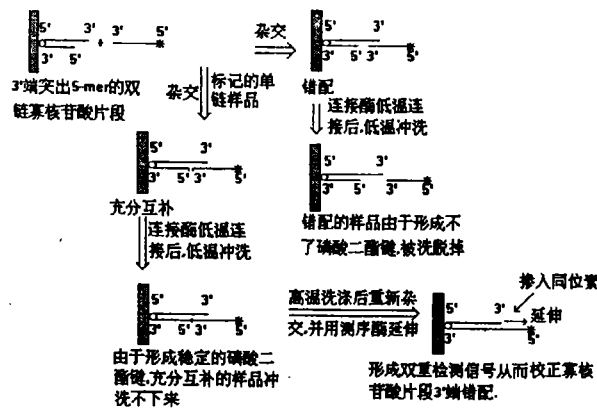


图 2

端位 SBH 是采用双链突出的寡核苷酸片段替代单链寡核苷酸。同时用连接酶来消除 5'端的错配和用序列级的 DNA 聚合酶来消除 3'的错配。因为连接酶对连接位点的错配极其敏感。而测序酶对延伸点的稳定也极为敏感。

三、SBH 的应用范围

SBH 还存在许多局限性,目前对单链 DNA 样品自身的二级、三级结构还无法突破,同时也无法应用于重复序列及卫星 DNA 等的序列测定。但假若重复区段短于寡核苷酸片段

长度,同时采用精密的定量杂交,则对重复单元长度及重复次数还是可以测定的。

现在 SBH 主要用于同源序列的比较测定。对于同源序列可通过选取特定的寡核苷酸组成从而缩小寡核苷酸小块的排列数目。例如,对三种同源性,编码原始人 T 细胞 β 受体 116bp 长的序列仅用挑选的 156 种 8-mer 寡核苷酸就可取得清晰的结果。

SBH 测定的精度取决于样品序列的长度。

在容许的杂交错误率中,6-mer 寡核苷酸对应的样品长为 500 bp。7-mer 对应 2000bp,8-mer 对应 8000bp。最近发展的连续堆积杂交(Continuous-Stacking hybridizations)通过利用 8-mer 的固定寡核苷酸和 5-mer 的非固定寡核苷酸与样品进行多次杂交可使 8+5 测定长度达

到 4000bp,9+6 测定长度为 16000bp。

就目前来说,SBH 进行同源序列的比较测定,点突变的检测,已日趋成熟。SBH 作为一种快速、简捷、廉价、自动化的测序方法,相信在将来几年中会有迅速的发展,为 DNA 序列的快速分析提供了一条可行之路。

DNA Sequencing by hybridization

Mi Zhiyong Ke Shen Wang Lixia

Abstract

Up to now, The DNA sequence data has accumulated over 1×10^8 bp since the introduction of the chain-termination DNA sequencing and Maxam-Gilbert chemical methods. However, these two kinds of traditional methods will not widely meet with the need of the future DNA sequencing with the every kinds of genomic programmes, no matter in efficiency or in cost. During 1990's, Sequencing by hybridization--SBH was reported by several labs. Through hybridization between marked single strand DNA template and specific length random oligo-nucleotides, We can obtain DNA sequence data quickly.