

仍然存在极大的困难。我们必须考虑到细胞器和细胞体积的差别、抗体渗透性的差别、以及固定的有效率等。不过,在同样实验系统的比较研究中,它还是可以提供相对定量的信息。

植物激素的免疫定位将能加深我们对植物激素作用的理解,它的普及和应用必将给植物激素作用机理的研究带来突破性的进展。目前,它是惟一可以用来研究植物激素原位分布的研究方法,而且它对定量分析的结果提供了有力的补充。应该注意的是,这种方法所能提供的仅仅是这些化合物时空分布的信息,其本身并不能得出作用位点和源库关系的结论。在突变体或转基因研究中采用这种方法,可以勾划出更加完整的植物激素作用机理的模型,以及鉴别特定的靶细胞或细胞器。而且,结合免疫细胞化学技术和利用报告基因或融合蛋白的研究手段,将是植物激素信号转导机制研究的强有力工具。

摘 要

本文介绍了植物激素免疫细胞化学定位研究中常用抗原与抗体的制备、激素小分子在植物细胞内的原位固定及其在激素生理研究中的应用。

参 考 文 献

- [1] Iten. M. et al. , 1999, *J. Recept Signal Transduct Res.* , 19:41 - 58.
- [2] Coons. A. H. et al. , 1942, *J. Immunol.* , 45:159 - 170.
- [3] 李宗霆、周燮, 1996, 植物激素及其免疫检测技术, 江苏科学技术出版社, 250 - 298.
- [4] Pastor, A. et al. , 1995, *Plant Growth Regul.* , 16: 287 - 292.
- [5] Fuchs, S. and Fuchs, Y. , 1969, *Biochem. Biophys. Acta* , 192:528 - 530.
- [6] Weiler. E. W. , 1981, *Planta* , 153:319 - 325.
- [7] 张能刚等, 1990, 南京农业大学学报, 13:116 - 119.
- [8] Marcussen, J. et al. , 1989, *Plant Physiol.* , 89: 1071 - 1078.
- [9] Mertens. R. et al. , 1985, *Planta* , 100:389 - 393.
- [10] Atzorn, R. and Weiler, E. W. , 1983, *Planta* , 159:1 - 6.
- [11] Knox, J. P. et al. , 1987, *Planta* , 170:86 - 91.
- [12] 郑志富等, 1995, 生物工程学报, 11:310 - 314.
- [13] Weiler, E. W. , 1980, *Planta* , 148:262 - 272.
- [14] Sossountzov, L. , et al. , 1988, *Planta* , 175:291 - 304.
- [15] Strnad, M. , et al. , 1992b In: Kaminek, M. , et al. , (eds) *Physiology and Biochemistry of Cytokinins in Plants*. The Hague:SPB Academic Publishing, 437 - 447.
- [16] Brandon, D. L. , et al. , 1992, In: Kaminek, M. , et al. , (eds) *Physiology and Biochemistry of Cytokinins in Plants*. The Hague:SPB Academic Publishing, 447 - 453.
- [17] Dewitte, W. , et al. , 1999, *Plant Physiol.* , 119: 111 - 121.
- [18] Galway, M. E. , et al. , 1993, *J. Cell Sci.* , 106: 847 - 858.
- [19] Zhang, G. F. , et al. , 1993, *J. Cell Sci.* , 104:819 - 831.
- [20] Kaneko, Y. and Walther, P. , 1995, *J. Electron Microscopy* , 44:104 - 109.
- [21] Pool. C. W. , et al. , 1983, In: Cuello, A. C. , (ed) *Immunocytochemistry*. Chicester: Wiley. 1 - 46.
- [22] Chriqui, D. , et al. , 1999, In: Altman, A. , et al. , (eds) *Plant Biotechnology and in vitro Biology in the 21st Century*, Dordrecht:Kluwer Academic Publishers, 33 - 37.
- [23] Ohmiya, A. , et al. , 1990, *Plant Cell Physiol.* , 31: 711 - 715.
- [24] 陈以峰等, 1999a, 武汉植物学研究, 17:297 - 302.
- [25] 陈以峰等, 1999b, 植物学报, 41:1145 - 1149.
- [26] Sotta, B. , et al. , 1985, *J. Histochem. Cytochem.* , 33: 201 - 208.
- [27] Sossountzov, L. , et al. , 1986, *Planta* , 168:471 - 481.
- [28] Welbaum, G. E. , et al. , 1997, *J. Exp. Bot.* , 48: 643 - 654.
- [29] Pastor, A. , et al. , 1999, *Physiol. Plant.* , 105:272 - 279.
- [30] 贾文锁等, 1994, 植物生理学报, 20:380 - 328.
- [31] 王学臣、贾文锁, 1995, 植物生理学报, 21:324 - 328.
- [32] Yamaguchi, I. and Weiler, E. W. , 1991 In: Takahashi, N. , et al. , (eds) *Gibberellins*. Springer-Verlag, 146 - 165.
- [33] 陈以峰等, 1998, 植物学报, 40:478 - 480.
- [34] Zavala, M. E. and Brandon. D. L. , 1983, *J. Cell Biol.* , 97:1235.
- [35] Berner. G. , et al. , 1993, *Plant Cell* , 5:1147 - 1155.
- [36] Ivanova, M. I. , 1994, *J. Exp. Bot.* , 45:1009 - 1017.
- [37] 何宇炯等, 1998, 植物生理学报, 24:24 - 30.
- [38] Taylor, P. E. , et al. , 1993, *Planta* , 189:91 - 100.
- [39] Fischer-Iglesias, C. , et al. , 2001, *Plant J.* , 26: 115 - 129.
- [40] Senger. S. , et al. , 2001, *Plant Cell Rep.* , 20:112 - 120

单核苷酸多态性(SNPs)原理及其在植物功能基因组学中的应用前景

刘健毅 潘建伟 朱睦元*

(浙江大学生命科学学院 杭州 310012)

顾青

(杭州商学院生物工程系 杭州 310035)

2001年11月诞生了人类基因组SNPs(Single Nucleotide Polymorphisms, 即单核苷酸多态性)图谱,这是第一张精确的SNPs图谱^[1]。同时在植物基因组学研究中,各国科学家在拟南芥(*Arabidopsis*

thaliana)基因组测序完成后,也正积极绘制其

基金项目:国家自然科学基金(39770420, 30100115)和浙江省自然科学基金(300255)资助项目。

* 通讯作者。

SNPs图谱。我国超级杂交水稻基因组计划也把完成SNPs图谱作为既定目标之一(<http://south.genomics.org.cn/study/rice.htm>)。可以说,SNPs已经成为当今基因组学的一个研究热点。

SNPs是生物基因组多态性的一种。随着基因组学兴起,特别是人类基因组计划发展,人们越来越相信基因组中的这种多态性将有助于解释个体的表型差异,不同群体或个体对疾病,尤其对复杂疾病的易感性和对环境因子的不同反应^[2]。SNPs不仅可使基因图谱分辨率更高,基因定位更精确,而且将会在基因功能分析,尤其在复杂性状的关联分析、基因定位和种群遗传学研究中发挥重要的作用,大大推动植物功能基因组学的发展。

本文就SNPs及其在植物功能基因组学中的应用前景作一下探讨。

一、SNPs原理

SNPs主要是指近缘种群或同种个体基因组水平上由于单核苷酸变异所引起的DNA序列多态性,其多态性频率大于1%。但在实际研究中,也常把位于cDNA、频率低于1%的单核苷酸变异称为SNPs^[3]。

单核苷酸的多态性常表现为核苷酸的转换、颠换、插失和缺失,但通常所指的SNPs不包括后两种情况。因此,SNPs主要有四种类型,一种为转换: $C \leftrightarrow T(G \leftrightarrow A)$,三种为颠换: $C \leftrightarrow A(G \leftrightarrow T)$ 、 $C \leftrightarrow G(G \leftrightarrow C)$ 、 $T \leftrightarrow A(A \leftrightarrow T)$ ^[3]。由于核苷酸的5-甲基胞嘧啶脱氨基反应相对比较频繁,使得四种SNPs在基因组中出现的频率是不同的,在生物体中约2/3是 $C \leftrightarrow T(G \leftrightarrow A)$,而在拟南芥中 $C \leftrightarrow T(G \leftrightarrow A)$ 约占52%^[4]。不仅如此,SNPs在单个基因或整个基因组中的分布也是不均匀的,多存在于非转录序列中。SNPs在基因组中的分布比RFLP、STRs更为广泛,而且绝大多数DNA多态性为SNPs,譬如人类基因组中90%的多态性属于SNPs^[3]。

目前,鉴定SNPs的方法主要有四种^[5]:1)单链构象多态性(SSCPs, single strand conformation polymorphisms)分析,把待测DNA片段PCR扩增产物变性为单链,然后以非变性聚丙烯酰胺凝胶为介质进行电泳,利用含SNPs的单链二级结构构象变化导致异常迁移来检测SNPs,检出率可达70% - 95%。然而,这种方法效率低,劳动强度大。2)异源双链分析(heteroduplex analysis),PCR扩增后,利用高效液相层析(high performance liquid chro-

matographic column, HPLC)中错配位点的杂合双链区域比无错配的同源配对区更易洗脱来检测SNPs,检出率可达95% - 100%。这种方法成本低,效率高。3)DNA直接测序(direct DNA sequencing),这是最彻底的方法。由于基因序列数据库,尤其是基因表达序列标记序列数据库的建立,可对不同来源的同一序列进行对比来高效地鉴定SNPs。4)DNA芯片技术(DNA chip),利用DNA荧光杂交和PCR技术鉴定SNPs。这种方法简单、快速,但目前该技术还不是很成熟,成本较高。

总的来说,SNPs有别于RFLP、STRs等DNA标记,不再以“长度”的差异为检测手段,而直接以序列的差异作为标记。因此,在技术上,它完全可以摆脱电泳分型的瓶颈,而采用最新的非电泳分型技术,同时也就意味着研究者通过SNPs可以“直接”掌握基因序列差别的信息。

二、SNPs在植物功能基因组学中的应用

功能基因组学作为植物基因组学的一个分支,它利用结构基因组学研究获得的大量数据与信息来评价基因的功能,包括生化功能、细胞功能、发育功能、适应功能等,其主要研究手段结合了高通量的大规模的实验方法、统计和计算机分析技术。SNPs将会以其优良的特性在植物功能基因组学中得到应用,促进植物基因组学性状表达的基因网络研究,了解环境与表型的内在联系,加速突变体的筛选,掌握突变体的分子遗传信息,促进种群遗传学与植物功能基因组学的结合。

SNPs在人类遗传学已得到广泛的应用,而其在植物学中应用刚刚起步,国内外这方面报道仍很少,但这并不能抹杀其在植物研究中的广阔前景。植物与人体一样,其基因组中也存在大量SNPs。Kanazin V等^[6]对五个大麦品系的54个基因分别进行测序时发现38个基因存在SNPs,共有112种多态。M. I. Tenailon等^[7]测算出玉米一号染色体上平均每104bp中就有一个SNPs,远高于人体平均1,500bp一个SNPs。华大基因研究中心根据其完成的92%籼稻(*Oryza sativa* L. ssp. *indica*)基因组草图,估算SNPs丰度为0.43%,即170万个左右^[8];与此相比,瑞士Syngenta公司测序粳稻(*Oryza sativa* L. ssp. *japonica*)基因组时仅发现48351个STRs^[9]。可见,植物中SNPs数量明显高于STRs等分子标记。

1. SNPs在多基因性状研究中的作用

早在20世纪80年代,分子生物学家研究限制性内切酶的酶切位点时就发现单核苷酸变异的存在,并进一步用于遗传标记,即RFLP标记(第一代分子标记)。进入90年代,出现第二代分子标记STRs(simple tandem repeats,简单串联序列,即微卫星)。到了90年代后期,由于人类基因组计划的实施和基因组学的发展,分子生物学家又重新对SNPs产生了浓厚的兴趣。

因为许多生物复杂性状是多基因控制的,每一个基因或基因的改变都对性状起着微效作用,而且有些等位基因的外显率较低,环境因子起着较大的作用,所以用经典的连锁分析来鉴定复杂性状所牵扯到的多基因及其相互关系似乎是不可能的^[5]。研究多基因性状比较好的方法是基于一定群体数量的连锁不平衡(linkage disequilibrium, LD,指同一个基因组上两个或两个以上的分子标记的关联程度):选择相对隔离种群,对性状相关调节通路的候选基因或通过基因组定位所锁定的候选基因,进行分子标记的关联研究。

由此,研究者开始探求适合此类研究的分子标记。这种分子标记必须适合大规模基因实验分析,具有较高的丰度和稳定性,最好能够满足自动化操作要求。研究表明STRs只是基于连锁分析的单基因性状研究的理想分子标记,杂合率(heterozygosity)较高,并且当STRs与基因所在区域距离小于1cM时,实验误差较大^[5]。另外,STRs还存在同型异源性(homoplasy)^[10],即种群起源不同,STRs大小却相同。所以,STRs不太适合作为连锁不平衡的分子标记。

为此,分子生物学家开始重新审视SNPs。SNPs不仅丰度比STRs高,更重要的是杂合率低、稳定性高,是稳定遗传的早期突变,比STRs更适合自动化操作。每个SNPs位点通常仅含2种等位基因(这虽然反映其信息量少,但其高密度和稳定遗传的特性弥补了信息量的不足),是一种二态标记,故在基因组中鉴定SNPs往往只需十/一的分析,而不用分析片段的长度,这就为发展自动筛选鉴定SNPs的技术提供了基础,同时也有利于基因分型。因此,SNPs适合作为连锁不平衡研究的分子标记。

植物中有许多重要性状都是由多基因控制的复杂性状,譬如穗型、成熟期等。要了解这些复杂性状的分子机理,不但要对单个基因,而且也要对多基因,甚至整个基因组在同一时间内表达的所有基因进行分析研究。转基因研究表明只通过孤立地改变单个基因的方式很难有效地改善农作物的产量和品质,我

们只有掌握多基因作用知识才能创造新的种质资源和重新设计品种^[11]。因此,研究者有必要在植物功能基因组学开展SNPs的连锁不平衡研究。

值得指出的是植物中有相当数量物种为自花传粉,可以减小由于重组所造成的连锁不平衡降低甚至消失^[12],这使得植物种群具有更高、更稳定的连锁不平衡,基因的纯合率较高。这对自花传粉的拟南芥、水稻等模式植物功能基因组研究是非常有利的。同时Nordborg^[13]认为在拟南芥基因组中连锁不平衡的失效距离为100至200kb,也就是说2000个左右的SNPs就可以较充分地研究拟南芥基因组,这对目前的技术和基因组信息研究来说是令人鼓舞的。

如今,“基因网络”研究正在兴起,研究者已认识到基因表达的大量数据反映了基因网络这个复杂信息处理系统在分子层次上的运转^[14]。基因网络的重要目的就是对某一物种或组织中全部基因的表达关系进行整体性研究,它的研究处于生物学、数学和信息学三大学科的交叉点上。显然,丰富的二态标记SNPs很适合基因网络的数学和信息学处理。

2. SNPs在单基因性状研究中的作用

SNPs同样可作为单基因连锁分析的标记。STRs本身的多态性对生物性状没有什么贡献,但许多SNPs存在于基因的编码序列和调节序列中,因而这些SNPs本身就是功能序列的一部分。据统计,在外显子中的SNPs有一半会造成非同义密码子突变^[3]。用SNPs作为分子标记,可以直接把生物性状与基因突变联系起来,也就是所谓的cSNPs(coding SNPs)和启动子区域SNPs(promoter region SNPs)。

Von Buren M等^[15]克隆小麦 γ -麦醇溶蛋白(γ -gliadin)基因时注意到不同品种间有SNPs的存在。Perry Creganpedeng等研究大豆SCN(soybean cyst nematode)抗性基因时发现一个位点(Satt309)的核苷酸突变是造成各品系抗性差异的原因(<http://bldg6.arsusda.gov/pberkum/Public/sarl/cregan/snps.htm>)。水稻的WAXY基因(编码结合在淀粉粒上的淀粉合成酶,Granule-bound starch synthase, GBSS)其5'端前导内含子倘若从AGGTATA突变成了AGTTATA,将直接影响GBSS的产生,导致水稻颗粒中直链淀粉含量的明显下降^[16]。

同时,在任何已知或未知的基因周围都可能找到众多的SNPs,研究者就可以把它用于单倍型(haplotype)诊断和基因定位。Jose M等^[17]在利用

SSLP 进行基因定位时,结合 SNPs 标记在 15kb 间距范围内定位了拟南芥的 *untracurvata* 突变体基因 *UCU1* 和 *UCU2*。Cho RJ 等^[18] 则利用基于 SNPs 的图谱定位了拟南芥的 *Eds16* 基因,此基因涉及拟南芥对 *Erysiphe orontii* 病原体的防御反应。

此外,在植物突变体检测中 SNPs 也有巨大意义。虽然研究者已经建立了拟南芥的受精、细胞分化及抗病等各种模型,但是缺乏高密度的基因图谱还是制约了突变体的筛选。常用的插入突变方法并不能鉴定内部发生自然或化学诱导突变的基因。因此,各国科学家和研究机构正积极地建立其 SNPs 库来促进拟南芥研究的深入。在 <http://godot.ncgr.org> 的 Cereon 库中已经收集了拟南芥 37 344 个 SNPs(2002/7),并可免费查询。

在实际研究中,Barton A 等^[19] 认为对照分析(case-control)比较适合 SNPs 的应用,而家系分析(family-based)有利于功能性 SNPs 的定位,但 Kanazin V^[7] 等认为两种方法没有多大区别。

3. 加强与种群遗传学的联系

RFLP、RAPD、STRs 等分子标记技术已经广泛地应用于植物物种进化和物种间亲缘关系的鉴定上。SNPs 同样将会在植物遗传多样性、分类和进化分析上发挥作用。研究者把 SNPs 作为参数可以简化种群遗传学最大相似性数学模型的复杂度^[20],还可以利用 SNPs 来研究植物适应性进化的历史^[13]。此外,通过 SNPs 对比来寻找染色体上不正常的低变化区域,可以了解自然选择的基因组信号(genomic signature)^[21]。

更重要的是,SNPs 在种群遗传学中的应用也常基于连锁不平衡分析,这使得种群遗传学、分子遗传学、DNA 技术更紧密地结合到一起,将会促进功能进化遗传学(functional evolutionary genetics)的兴起^[3,13]。种群遗传学所取得的成果将促进植物基因组学利用 SNPs 进行基因型-表型相互关系的分析。如果研究者把进化方面的知识和信息用于基因功能的预测,来研究基因在进化过程中如何变得相似而不是研究序列相似性本身,还将有助于提高预测基因功能的能力^[22]。

三、展望

SNPs 正成为基因组作图继 RFLP、STRs 后的第三代分子标记,使得生物遗传图谱日趋精确,将推动植物遗传学分子标记的发展。目前在植物基因组研究中,由于缺乏有用的、大规模准确的、有普遍意

义的植物 SNPs 数据库,不能提供有用的 SNPs 位点信息供研究者使用,使得 SNPs 的应用仍处于探索阶段。在具体应用 SNPs 时尚有一些理论上的困难需要解决,譬如在高等生物精卵形成时,双拷贝的 DNA 重组为单拷贝,这种重组时 SNPs 的界定(mapping 与 association test)十分棘手^[23];同时还可能需要开发相关的软件和数据包,譬如如何用软件方便地辨别 PCR 扩增产物 SNPs 的真实性。这些问题的存在都制约了 SNPs 在植物功能基因组学中的应用。但相信,随着水稻测序的完成、DNA 直接测序技术的不断简便及几种重要植物 SNPs 库的建立和理论的成熟,SNPs 将会在植物学研究领域中得到广泛的应用。

摘要

近缘种群或同种个体之间的单核苷酸多态性(SNPs)是存在于生物基因组上由单个核苷酸的变异所引起的一种 DNA 序列多态性。它具有高丰度、高信息量和良好稳定性的特点,而且比较适合自动化操作。SNPs 将会在植物功能基因组学研究中得到广泛应用,并加强功能基因组学与种群遗传学的联系。本文就 SNPs 及其在植物功能基因组学中的应用前景作一下探讨。

参考文献

- [1] The international SNP map working group, *Nature*, 2001, **409**:928-933.
- [2] 张思仲, 1999, *中华医学遗传学杂志*, **16**(2):119-122.
- [3] Anthony J. Brookes, 1999, *Gene*, **234**(2):177-186.
- [4] The Arabidopsis Initiative, *Nature*, 2000, **408**:796-815.
- [5] Ian C. Gray, et al., 2000, *Human Molecular Genetics*, **9**(13):2403-2408.
- [6] Kanazin V, et al., 2002, *Plant Molecular Biology*, **48**(5):529-537.
- [7] M. I. Tenaillon, et al., 2001, *Proc. Natl. Acad. Sci. U. S. A.* **98**(16):9161-9166.
- [8] Jun Yu, et al., 2002, *Science*, **296**(13):79-92.
- [9] S. A. Goff, et al., 2002, *Science*, **296**(13):92-100.
- [10] Estoup A., et al., *Mol Biol Evol*, 1995, **12**(6):1074-1084.
- [11] 薛勇彪, 2001, *植物学报*, **43**(7):769-770.
- [12] Nancy A. Eckardt, et al., 2001, *Plant Cell*, **13**(6):1249-1254.
- [13] Magnus Nordborg, 2000, *Genetics*, **154**(2):923-929.
- [14] 彭华正等, 2001, *生物化学与生物物理进展*, **28**(6):815-818.
- [15] Von Buren M, et al., 2000, *Theoretical & Applied Genetics*, **100**(2):271-279.
- [16] Larkin Patrick D, et al., 1999, *Plant Molecular Biology*, **40**(4):719-727.