

•综述•

InDel标记的研究和应用进展

杨洁 赫佳 王丹碧 施恩 杨文字 耿其芳 王中生*

(南京大学生命科学院, 南京 210023)

摘要: InDel是指在近缘种或同一物种不同个体之间基因组同一位点的序列发生不同大小核苷酸片段的插入或缺失(insertion-deletion), 是同源序列比对产生空位(gap)的现象。InDel在基因组中分布广泛、密度大、数目众多。InDel多态性分子标记是基于插入/缺失位点两侧的序列设计特异引物进行PCR扩增的标记, 其本质仍属于长度多态性标记, 可利用便捷的电泳平台进行分型。InDel标记准确性高、稳定性好, 避免了由于特异性和复杂性导致的后续分析模糊。此外, InDel标记能扩增混合DNA样品和高度降解的微量DNA样品, 并进行有效分型。InDel标记目前已开始应用于动植物群体遗传分析、分子辅助育种以及人类法医遗传学、医学诊断等领域。随着位于功能基因上InDel标记的开发, 结合染色体步移和基因精细定位, 可将这些标记应用于相关物种经济性状的功能基因的筛选, 有利于优良基因的进一步开发和利用。

关键词: 分子标记; InDel; SNP; SSR

Progress in research and application of InDel markers

Jie Yang, Jia He, Danbi Wang, En Shi, Wenyu Yang, Qifang Geng, Zhongsheng Wang*

School of Life Sciences, Nanjing University, Nanjing 210023

Abstract: InDel indicates insertions or deletions (insertion-deletion) of nucleotide fragments of different sizes at the same site in the genome sequence between the same or closely related species and is a gap in sequence derived from alignment of the homologous sequence. InDel is widely distributed across the genome and occurs in a high density and large numbers in a genome. The InDel polymorphic molecular marker is a PCR-amplified marker that is based on specific primers designed from both sides of the site of sequence of insertion / deletion. It is essentially a length polymorphic marker still, and one can use the convenient electrophoresis platform for genotyping. InDel molecular markers have the advantage of high accuracy and good stability, which help to avoid confusion in subsequent analysis due to marker specificity and complexity, as is often seen in other length polymorphic markers. Furthermore, mixed or highly degraded DNA samples can be successfully amplified with InDel markers, and effectively typed. Because of its abundance, convenient typing platform and other advantages, InDel molecular markers have been applied to genetic analyses of animal and plant populations, molecular assisted crops and farmed animal breeding, human forensic genetics, medical diagnostics and other research areas. The development of the InDel molecular marker located on functional genes, combined with chromosome walking and fine gene mapping, has enabled the application of these molecular markers in the screening of genes related to important economic traits, which is conducive to the further development and utilization of these valuable genes. In this review, on the basis of an overview of the InDel marker development and applications, we discuss some of the technical limitations of the development and limited efficiency of genetic analysis, as well as potential future applications in the fine mapping and genetic structure of large numbers of individuals.

Key words: molecular marker; InDel; SNP; SSR

插入/缺失(insertion-deletion, InDel)是指在近缘种或同一物种不同个体之间基因组同一位点的序

收稿日期: 2015-07-15; 接受日期: 2015-12-17

基金项目: 国家自然科学基金(31100270)

* 通讯作者 Author for correspondence. E-mail: wangzs@nju.edu.cn

列发生了不同大小核苷酸片段的插入或缺失,即一个序列上某一位点相比同源的另一个序列插入或缺失了一个或多个碱基(Weber et al, 2002)。InDel是同源序列比对产生空位(gap)的现象,但大多数情况下无法获知祖先序列(Fan et al, 2007),很难判断空位位点是哪个序列发生了插入突变,或哪个序列发生了缺失突变,所以一般统称它们为插入/缺失突变。InDel标记基于PCR扩增技术,本质上属于长度多态性标记(Hyten et al, 2010)。InDel标记因其稳定性好、多态性高、分型系统简单(Jander et al, 2002),开始应用于动植物群体遗传分析、分子辅助育种以及人类法医遗传学、医学诊断等领域。

1 InDel的产生机制及特征

1.1 InDel的产生机制

InDel产生主要与基因组的特征和DNA复制错误有关。Kondrashov和Rogozin (2004)根据对人类编码区外显子序列的研究,发现InDel的产生频率与所在序列的碱基类型有一定关系,当突变热点含GTAAGT序列或DNA两条链上的嘌呤嘧啶含量不平衡时会增加缺失发生的几率,序列包含AT(A/C)(AC)GCC和TACCRC时会增加插入发生的几率,而TATCGC和GCGG序列分别是缺失和插入的冷点。有学者在家族性高胆固醇血症研究中发现低密度脂蛋白(LDL)受体编码基因外显子发生5 kb的*Alu*重组缺失,当缺失片段一侧或两侧为*Alu*序列的反向重复时,两个反向重复序列可能会形成双茎环结构(double stem-loop),从而导致序列缺失(Lehrman et al, 1986)。另有研究发现*Alu*重复序列有很高的碱基突变率(Britten et al, 2003)。

除基因组序列本身特征以外,InDel的产生与DNA复制错误也有关联。人类男性基因组中,InDel数目在X染色体上较常染色体上少,这可能是因为X染色体复制次数比常染色体少,产生复制错误的次数也相应较少,InDel的产生几率降低(Kvikstad et al, 2007)。基因序列中PolyA结构热稳定性较差,在DNA复制过程中解链时很容易发生复制滑移;高GC含量的序列会增加复制滑移的几率,而复制滑移会导致序列发生插入/缺失突变(Sjödin et al, 2010)。InDel的长度与其形成机制也有一定关系,一般较长的InDel是由于转座元件复制或异常重组造成的,较短的InDel则是由于复制滑移所导致。

除上述产生机制以外,还有转座子复制和插入、移动元件插入、序列异常重组和同类重复拷贝不等交换等各种因素(Britten et al, 2003)。然而大部分插入/缺失突变的产生机制现在仍旧未知。

1.2 InDel的特性

Mills等于2006年创建了第一个人类基因组的插入/缺失图谱(Mills et al, 2006)。他们通过对InDel位点序列分析将InDel分成5大类: (1)单碱基对插入/缺失; (2)单碱基插入/缺失; (3) 2–15 bp重复单元的多碱基对插入/缺失; (4)转座子插入; (5)随机DNA序列插入/缺失。

InDel在基因组中分布广泛、密度大、数目众多。就分布密度而言,InDel仅次于SNP,但远高于SSR。在人类基因组中,平均密度为每7.2 kb包含一个InDel (Mills et al, 2006); 在鸡和狗中,平均密度为每5 kb包含一个InDel (Brandström & Ellegren, 2007); 在水稻中平均密度为每953 bp包含一个InDel (Shen et al, 2004)。

虽然InDel在基因组中分布频率高,但其分布并不均匀,在不同染色体上分布的密度有所差异(Sjödin et al, 2010)。人类性染色体上InDel的密度远低于常染色体上的密度,如X染色体上InDel位点约为常染色体的一半,Y染色体则仅为常染色体的5%,且常染色体之间InDel分布密度也有差异(Mills et al, 2006)。此外InDel在同一染色体上分布密度也有偏倚性(Bhangale et al, 2005),因其编码区序列相对保守,InDel大部分分布于非编码区。在Brandstrom和Ellegren (2007)关于原鸡(*Gallus gallus*)的研究中发现,在基因间的InDel数量比在上游序列、5'-UTR、首个外显子中分布得多,在串联重复序列中缺失发生频率也比其他序列中高。

InDel长度变化很大,平均长度为36 bp,最长可达10,000 bp,但99%以上的InDel长度小于50 bp (Britten et al, 2003)。因测序技术的发展与结构变异检测方法的进步,被定义为InDel的插入和缺失的长度在不断缩小,一些长度较长的插入和缺失被归入结构变异的概念(Tuzun et al, 2005),将InDel定义为50 bp以下插入和缺失的总称(Alkan et al, 2011)。尽管如此,InDel依旧是人类基因组中含量最丰富的变异类型之一。

1.3 InDel标记的优点

高通量测序技术的发展使得InDel标记开发成

本降低(Varshney et al, 2009), 作为含量最丰富的多态性突变之一, InDel适用于全基因组分子标记的开发(Gao et al, 2012)。InDel作为高通量分子标记, 兼具各类遗传标记的优点。传统分子标记开发一般只基于一份序列, 而InDel标记开发完全基于序列差异, 所以开发过程中无多型位点较少。与其他分子标记相比, 基因组同一位点发生相同长度大小的InDel突变的几率极小, 可看作同一起来源 (identify-by-descent, IBD)(Shedlock & Okada, 2000), 所以InDel标记准确性高、变异稳定, 避免了由于特异性和复杂性导致的后续分析模糊。Bastos-Rodrigues等(2006)对世界人群HGD-CEPH多样性进行研究显示, 40个InDel位点的遗传标记效率相当于65个*Alu*或60个SNP标记。此外, InDel标记在种内和种间都有多态性, 其通用性较强(Väli et al, 2008)。

InDel标记本质上仍属于长度多态性遗传标记, 可基于PCR扩增技术对InDel进行分型。与分型系统复杂的SNP标记相比, InDel检测简单便捷, 对仪器设备和技术要求较低, 在电泳技术平台上即可进行。作为序列中分布密度最大的两种分子标记, InDel和SNP分布并不均匀, 它们的位置具有一定互补性, 可根据目的基因所在区域设计不同的分子标记(Li et al, 2014)。此外, Santos等(2015)通过SNP和InDel标记技术对混合DNA样品进行法医血统分析, 发现与基于SNaPshot分型平台的SNP标记相比, InDel标记能更准确地鉴定样品间的亲缘关系。

与SSR标记相比, 冯芳君等(2005)实验发现InDel标记的扩增产物带型清晰简单, 其稳定性和产物分离效果均明显优于SSR标记。SSR标记扩增产物片段长度一般为100–450 bp, 扩增结果直接受DNA质量的影响, 针对高度降解的DNA样品分型有一定难度(Da et al, 2013)。而一些扩增产物较短的InDel标记对DNA质量要求较低, 且对样品量要求较少, 能扩增高度降解的DNA (Manta et al, 2012)。

2 InDel标记的开发及分型方法

2.1 InDel标记的开发

InDel标记开发方法一般是先获取相关DNA序列, 通过BLAST比对分析筛选InDel位点, 进而根据InDel位点两端的序列设计引物来扩增这些插入/缺失位点。InDel位点序列主要可从公共数据库或通过重测序获得, 也可从已发表的相关文献中查找。

从公共基因序列数据库中查找InDel位点, 是最经济快捷的一种方法。主要有三大国际一级生物数据库: The National Center for Biotechnology Information (NCBI)、Europe Molecular Biology Laboratory (EMBL)和DNA Data Bank of Japan (DDBJ)。此外还有许多实验室针对不同物种分别建立的数据库, 例如蔷薇科、松科松属(*Pinus*)植物、拟南芥(*Arabidopsis thaliana*)、水稻(*Oryza sativa*)和玉米(*Zea mays*)等。从公共数据库中查找InDel位点虽然省时省力, 但只限于已有序列数据发布的物种, 且这些物种多为模式生物或经济物种, 对于大多数生物来说, 已知序列相对较少, 这就限制了利用InDel标记对于这些生物的研究。而通过测序获得相关序列的方法则不受物种限制。

开发InDel标记可通过同源序列比对分析获得。目前InDel标记开发对象主要包括不同亚种之间、同种双亲之间、不同地理宗之间以及大量同种个体之间的同源序列比对分析。利用BLAST分析同源序列差异, 寻找InDel位点, 并根据此位点设计相应的InDel标记引物。在比对过程中有学者发现, 有些序列虽然在SSR位点不存在差异, 但是在其侧翼序列中存在着InDel差异, 他们利用这些序列开发出了相应的InDel标记(仪泽会等, 2012)。除了通过BLAST序列比对筛选InDel位点之外, 还有部分学者利用软件查找序列中InDel位点, 如修正版的autoSNP软件等(Barker et al, 2003; Savage et al, 2005)。

2.2 InDel标记的分型方法

InDel标记是基于基因组中插入/缺失位点两侧的序列设计特异引物进行PCR扩增的标记, 目前大多是利用电泳平台进行分型, 该分型平台快捷经济, 不需要复杂的实验设备, 可操作性强。电泳分型平台有琼脂糖凝胶电泳、变性或非变性聚丙烯酰胺凝胶电泳以及毛细管电泳。当InDel位于某种限制性内切酶的酶切位点上时, 可转化为扩增片段酶切长度多态性(CAPS), 先用该限制性内切酶酶切后再电泳分型(Konieczny & Ausubel, 1993)。InDel长度变化较大, 选择InDel位点时, 应优先选择多态性差异5–20 bp的InDel, 这类InDel设计的成功率高, 而且便于检测。为了节约高通量分型检测成本, Salathia等(2007)利用InDel阵列对拟南芥*Landsberg erecta*与*Columbia*两种生态型进行InDel分型, 通过比较基

基因组杂交精确分析InDel多态性。也有研究利用测序技术进行InDel分型, 错误率远低于1% (Bhangale et al, 2005), 但成本偏高。

3 InDel标记的应用

3.1 种质资源分析与分子辅助遗传育种

在水稻遗传育种中, 籼稻与粳稻种质资源的鉴定具有特殊的价值, RAPD、RFLP等传统分子标记难以对水稻籼粳属性进行准确鉴定(龙雯虹和许明辉, 2002; Lu et al, 2002), 而InDel标记技术在水稻籼粳属性鉴定中准确率高, 且快捷简便。Lu等(2009)基于籼稻(93-11)和粳稻(日本晴)两个水稻亚种全基因组序列比对获得的45个InDel位点, 分别对21个典型的籼稻品种与21个典型的粳稻品种进行PCR扩增和聚丙烯酰胺凝胶电泳, 对各InDel标记的基因型数据矩阵进行中性检测, 发现其中的34个InDel位点与籼、粳属性分化关系紧密。再进一步分别对来自亚洲11个国家随机42个籼、粳属性未知的栽培品种和12种野生稻进行标记鉴定, 确定了各个样品的籼、粳属性。结果表明, 通过对不同品种的InDel位点上籼型或粳型等位基因平均频率的统计能够准确鉴别水稻的籼、粳属性, 解决水稻育种中准确选用籼稻或粳稻种质资源的问题。Steele等(2008)利用42个InDel位点从21种香米中将印度香米准确鉴别出来。Hayashi等(2006)将抗稻瘟品种与不抗稻瘟品种作为双亲, 将双亲的9个常见水稻抗稻瘟基因序列Piz、Piz-t、Pit、Pik、Pik-m、Pik-p、Pita、Pita-2及Pib进行比对分析开发InDel标记, 利用针对不同水稻抗稻瘟基因开发的InDel标记对大量F₂代个体进行分析, 成功地开发了水稻抗稻瘟个体的InDel分子标记。此外, InDel标记在其他作物的分子辅助遗传育种领域中同样应用广泛, 如结球甘蓝(*Brassica oleracea* var. *capitata*) (Lv et al, 2013)和黄瓜(*Cucumis sativus*) (张圣平等, 2011)等。

3.2 群体遗传分析

目前关于群体遗传分析的研究大多是利用SSR分子标记技术, 但SSR位点易发生突变使得后续结果分析模棱两可(Ellegren, 2004)。已有学者开始应用InDel分子标记进行群体遗传分析。Vali等(2008)基于两个品种狗的全基因组和9个品种狗的鸟枪法测得短序列作为开发InDel标记的序列资源, 通过比对筛选得到200,000个InDel位点, 并利用其中100

个长度为4 bp的非重复碱基序列的InDel位点分析7个互不关联品种的狗和分布于世界各地的18匹狼。分型结果表明, 100个InDel标记中分别有94个和76个InDel位点在狗和狼中有多型, 总平均观测杂合度(*H_o*)在狗和狼中分别为26.8%和19.4%, 期望杂合度(*H_e*)分别为35.5%和26.1%。此外, Vali等(2008)利用在18匹狼中有多型的76个InDel位点来分析濒危的瑞典狼的种群, 其中51个InDel位点有多型, 且总平均观测杂合度和期望杂合度分别为25.3%和17.0%。该研究表明利用狗的序列开发的InDel位点可用于同一属狼的群体遗传分析, 也可用于遗传多样性受限的濒危群体的遗传多样性研究。可见InDel标记的通用性较强, 且在群体遗传分析中应用范围广泛。

3.3 图位克隆、基因定位及遗传图谱的构建

在物种功能基因的图位克隆中, 通常需要分布密度高的分子标记, 而目前已有的各种分子标记并没有达到令人满意的分布密度。InDel标记分布密度大, 可满足目的基因的精细定位。潘存红等(2007)通过日本晴和93-11两个水稻品种的序列比对分析, 开发了25个InDel位点, 利用InDel位点对不完全隐性的卷叶基因成功地完成了精细定位和克隆。

利用InDel位点与某些功能基因或QTL间的连锁关系, 可以将功能基因或QTL定位在染色体上。Vasemägi等(2010)根据大西洋鲑鱼EST序列开发了76个InDel可用位点, 对来自两个相关联生境的大西洋鲑鱼(*Salmo salar*)处于生活史早期的个体进行分析, 成功构建了该时期的大西洋鲑鱼的QTL定位。此外, Lv等(2013)将抗枯萎病洋白菜(*Brassica oleracea* var. *capitata*)和枯萎病敏感型洋白菜作为双亲, 通过全基因组序列比对分析开发InDel标记, 再对F₁和F₂代个体进行分析, 分析得到抗枯萎病基因(FOC)位于C06染色体上, 且在M10和A1两个InDel位点之间0.6–1.2 cM。

遗传学图谱包括遗传图谱、物理图谱和转录图谱, 只有构建出遗传图谱后才能进一步构建物理图谱和转录图谱。遗传图谱是指某一物种的染色体图谱, 也就是连锁图谱, 它显示所知的基因或遗传标记的相对位置。InDel多态性较高, 可以作为分子标记用来构建遗传图谱。Bhatramakki等(2002)对8株玉米重测序, 根据EST序列开发的InDel位点, 对B73品种和Mo17品种个体及其杂交后代进行了遗

传分析。利用InDel标记得到的遗传图谱显示, 共有22个多态位点, 其中5号染色体上有5个位点, 在1号染色体上有4个位点, 2、3号染色体上有3个位点, 7、8、9号染色体上有2个位点, 10号染色体上有1个位点。

3.4 法医遗传学及医学诊断

InDel作为兼具SNP和STR优点的分子标记技术, 已应用于法医遗传学领域。Pereira等(2009)利用38个位于常染色体非编码区长度为2–5 bp的二等位基因InDel位点对来自非洲、欧洲及亚洲的人类群体进行遗传分析, 结果显示所有的InDel位点多态性均较好, 随机匹配概率在 10^{-14} 到 10^{-15} 之间, 表明InDel标记能有效地进行人类个体识别。38个InDel位点的扩增产物最长不超过160 bp, 能成功分析仅为0.3 ng的DNA样品, 而且能利用高度降解的DNA样品进行准确分型。此外, InDel标记在人类群体遗传学和人群结构学研究中已有所应用。Santos等(2010)利用InDel标记来分析人群混合程度并确定人群结构, 使用48个InDel位点作为祖先多态位点(ancestry informative marker, AMI), 分析593份来自欧洲、非洲和美洲互无关联的个体, 并对380个来源已知的巴西混合人群进行验证分析, 结果显示与预期高度一致; 这48个InDel位点构成的AMI体系可用于区别不同地区的人种, 能够有效识别混合人群中的遗传亚结构。

InDel标记不仅应用于个体识别和亲缘关系鉴定, 而且在医学方面也有所应用。Lee等(2006)利用InDel标记技术增加微嵌合体诊断的准确性。微嵌合体是指来自于遗传上与宿主不同的细胞在宿主的身体内残留, 具体的产生机制还不太清楚。以往微嵌合体的诊断是利用HLA或Y染色体作为目标基因, 然而HLA与Y染色体由于自身局限等原因往往无法全面检测, 而InDel标记信息量更高、准确性更强, 利用InDel标记能检测出HLA与Y染色体无法检测的微嵌合体。

4 InDel标记的应用局限及前景展望

InDel作为一种高通量遗传标记, 目前已开始应用于动植物群体遗传分析、分子辅助育种以及人类法医遗传学、医学诊断等领域, 但其大规模的开发应用仍存在较大的阻碍。InDel位点是通过同源序列比对分析获得, 但序列已知的生物种类有限, 大

多为模式物种或经济作物, 对于非模式生物而言, 大量基因组序列未知, 导致InDel引物设计与开发存在一定的难度。此外, 与其他多等位基因遗传标记相比, InDel标记作为二等位基因遗传标记, 携带的遗传信息量仍然有限。目前已有研究将连锁在同一染色体上的2个或3个InDel位点作为multi-InDel标记, 使得InDel标记的遗传分析效率进一步提高。Huang等(2014)将位于外显子上间隔不超过100 bp的2个或3个InDel同时作为标记位点应用于法医分析, 结果显示累积排除率和累计个体鉴别率分别为0.9989和0.99999999999994。此外也有学者将多个InDel位点集中于同一扩增体系以达到足够高的分辨效能。如Santos等(2010)将来自于不同染色体或同一染色体不同区域的16个InDel位点集中于同一扩增体系来分析人群混合程度并识别混合人群中的遗传亚结构。这些方法不仅可以缩短实验周期、降低实验成本, 而且使InDel标记具有更为丰富的遗传信息量, 但在技术操作上也增加了一定难度。

目前, InDel标记开始初步应用于染色体步移和基因精细定位(fine mapping) (Ma et al, 2005), 有研究结合SSR标记和InDel标记对奶牛控制产奶量(Schnabel et al, 2005)以及水稻控制每穗颖花数(Liu et al, 2009)、铝耐受性(Xue et al, 2007)等基因成功进行了精细定位。随着位于功能基因上InDel标记的开发, 可将这些标记应用于相关物种经济性状的功能基因筛选, 有利于优良基因的进一步开发利用。

随着InDel标记应用范围的拓展, InDel标记能更全面地表征多杂交和大谱系群体的遗传结构, 解决遗传进化研究中的基本问题, 如同一QTL在群体中的离散程度, 以及在自然种群中有多少QTLs共同作用于相关性状等(Vasemägi et al, 2010)。

总之, 随着测序技术的迅速发展, 测序成本进一步降低, 国际公共数据库序列信息将越来越丰富, InDel标记开发的可利用资源逐步增加, InDel标记将在动植物分子生物学和人类生物医学等研究领域发挥更大的作用。

参考文献

- Alkan C, Coe BP, Eichler EE (2011) Genome structural variation discovery and genotyping. *Nature Reviews Genetics*, 12, 363–376.
- Barker G, Batley J, O'Sullivan H, Edwards KJ, Edwards D (2003) Redundancy based detection of sequence polymorphisms in expressed sequence tag data using autoSNP. *Bio-*

- informatics, 19, 421–422.
- Bastos-Rodrigues L, Pimenta JR, Pena SD (2006) The genetic structure of human populations studied through short insertion-deletion polymorphisms. *Annals of Human Genetics*, 70, 658–665.
- Bhangale TR, Rieder MJ, Livingston RJ, Nickerson DA (2005) Comprehensive identification and characterization of diallelic insertion-deletion polymorphisms in 330 human candidate genes. *Human Molecular Genetics*, 14, 59–69.
- Bhattaramakki D, Dolan M, Hanafey M, Wineland R, Vaske D, Register JC III, Tingey SV, Rafalski A (2002) Insertion-deletion polymorphisms in 3' regions of maize genes occur frequently and can be used as highly informative genetic markers. *Plant Molecular Biology*, 48, 539–547.
- Brandström M, Ellegren H (2007) The genomic landscape of short insertion and deletion polymorphisms in the chicken (*Gallus Gallus*) genome: a high frequency of deletions in tandem duplicates. *Genetics*, 176, 1691–1701.
- Britten RJ, Rowen L, Williams J, Cameron RA (2003) Majority of divergence between closely related DNA samples is due to indels. *Proceedings of the National Academy of Sciences, USA*, 100, 4661–4665.
- Da SC, Matos S, Costa HA, Morais P, Santos RM, Espinheira R (2013) Genetic portrait of south Portugal population with InDel markers. *Forensic Science International: Genetics*, 7, e101–e103.
- Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics*, 5, 435–445.
- Fan YH, Wang WJ, Ma GJ, Liang LJ, Shi Q, Tao SH (2007) Patterns of insertion and deletion in mammalian genomes. *Current Genomics*, 8, 370–378.
- Feng FJ, Luo LJ, Li Y, Zhou LG, Xu XY, Wu JH, Chen HW, Chen L, Mei HW (2005) Comparative analysis of polymorphism of InDel and SSR markers in rice. *Molecular Plant Breeding*, 3, 725–730. (in Chinese with English abstract) [冯芳君, 罗利军, 李荧, 周立国, 徐小艳, 吴金红, 陈宏伟, 陈亮, 梅捍卫 (2005) 水稻InDel和SSR标记多态性的比较分析. *分子植物育种*, 3, 725–730.]
- Gao Q, Yue G, Li W, Wang J, Xu J, Yin Y (2012) Recent progress using high-throughput sequencing technologies in plant molecular breeding. *Journal of Integrative Plant Biology*, 54, 215–227.
- Hayashi K, Yoshida H, Ashikawa I (2006) Development of PCR-based allele-specific and InDel marker sets for nine rice blast resistance genes. *Theoretical and Applied Genetics*, 113, 251–260.
- Huang J, Luo H, Wei W, Hou Y (2014) A novel method for the analysis of 20 multi-Indel polymorphisms and its forensic application. *Electrophoresis*, 35, 487–493.
- Hytien DL, Cannon SB, Song Q, Weeks N, Fickus EW, Shoemaker RC, Specht JE, Farmer AD, May GD, Cregan PB (2010) High-throughput SNP discovery through deep resequencing of a reduced representation library to anchor and orient scaffolds in the soybean whole genome sequence. *BMC Genomics*, 11, 1248–1251.
- Jander G, Norris SR, Rounsley SD, Bush DF, Levin IM, Last RL (2002) *Arabidopsis* map-based cloning in the post-genome era. *Plant Physiology*, 129, 440–450.
- Kondrashov AS, Rogozin IB (2004) Context of deletions and insertions in human coding sequences. *Human Mutation*, 23, 177–185.
- Konieczny A, Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *The Plant Journal*, 4, 403–410.
- Kvikstad EM, Tyekucheva S, Chiaromonte F, Makova DK (2007) A macaque's-eye view of human insertions and deletions: differences in mechanisms. *PLoS Computational Biology*, 3, 1772–1782.
- Lee TH, Chafets DM, Reed W, Wen L, Yang Y, Chen J, Utter GH, Owings JT, Busch MP (2006) Enhanced ascertainment of microchimerism with real-time quantitative polymerase chain reaction amplification of insertion-deletion polymorphisms. *Transfusion*, 46, 1870–1878.
- Lehrman MA, Russell DW, Goldstein JL, Brown MS (1986) Exon-*Alu* recombination deletes 5 kilobases from the low density lipoprotein receptor gene, producing a null phenotype in familial hypercholesterolemia. *Proceedings of the National Academy of Sciences, USA*, 83, 3679–3683.
- Li X, Gao W, Guo H, Zhang X, Fang DD, Lin Z (2014) Development of EST-based SNP and InDel markers and their utilization in tetraploid cotton genetic mapping. *BMC Genomics*, 15, 1046–1056.
- Liu T, Mao D, Zhang S, Xu C, Xing Y (2009) Fine mapping *SPPI*, a QTL controlling the number of spikelets per panicle, to a BAC clone in rice (*Oryza sativa*). *Theoretical and Applied Genetics*, 118, 1509–1517.
- Long WH, Xu MH (2002) The genetic difference between Indica rice and Japonica rice on RAPD. *Journal of Yunnan Agricultural University*, 17, 245–247. (in Chinese) [龙雯虹, 许明辉 (2002) 籼稻和粳稻品种在RAPD上的遗传差异. *云南农业大学学报*, 17, 245–247.]
- Lu BR, Cai X, Jin X (2009) Efficient Indica and Japonica rice identification based on the InDel molecular method: its implication in rice breeding and evolutionary research. *Progress in Natural Science*, 19, 1241–1252.
- Lu BR, Zheng KL, Qian HR, Zhuang JY (2002) Genetic differentiation of wild relatives of rice as assessed by RFLP analysis. *Theoretical and Applied Genetics*, 106, 101–106.
- Lv H, Yang L, Kang J, Wang Q, Wang X, Fang Z, Liu Y, Zhuang M, Zhang Y, Lin Y, Yang Y, Xie B, Liu B, Liu J (2013) Development of InDel markers linked to Fusarium wilt resistance in cabbage. *Molecular Breeding*, 32, 961–967.
- Ma JF, Nagao S, Huang CF, Nishimura M (2005) Isolation and characterization of a rice mutant hypersensitive to Al. *Plant and Cell Physiology*, 46, 1054–1061.
- Manta F, Caiafa A, Pereira R, Silva D, Amorim A, Carvalho EF, Gusmão L (2012) Indel markers: genetic diversity of 38 polymorphisms in Brazilian populations and application in a paternity investigation with post mortem material. *Forensic*

- Science International Genetics, 6, 658–661.
- Mills RE, Luttig CT, Larkins CE, Beauchamp A, Tsui C, Pittard WS, Devine SE (2006) An initial map of insertion and deletion (INDEL) variation in the human genome. *Genome Research*, 16, 1182–1190.
- Pan CH, Wang ZB, Ma YY, Yin YJ, Zhang YF, Zuo SM, Chen ZX, Pan XB (2007) InDel and SNP markers and their application in map-based cloning of rice genes. *Chinese Journal of Rice Science*, 21, 447–453. (in Chinese with English abstract) [潘存红, 王子斌, 马玉银, 殷跃军, 张亚芳, 左示敏, 陈宗祥, 潘学彪 (2007) InDel和SNP标记在水稻图位克隆中的应用. *中国水稻科学*, 21, 447–453.]
- Pereira R, Phillips C, Alves C, Amorim A, Carracedo A, Gusmão L (2009) A new multiplex for human identification using insertion/deletion polymorphisms. *Electrophoresis*, 30, 3682–3690.
- Salathia N, Lee HN, Sangster TA, Morneau K, Landry CR, Schellenberg K, Behere AS, Gunderson KL, Cavalieri D, Jander G, Queitsch C (2007) Indel arrays: an affordable alternative for genotyping. *The Plant Journal*, 51, 727–737.
- Santos C, Fondevila M, Ballard D, Banemann R, Bento AM, Børsting C, Branicki W, Brisighelli F, Burrington M, Capal T, Chaitanya L, Daniel R, Decroyer V, England R, Gettings KB, Grosso TE, Haas C, Harteveld J, Hoff-Olsen P, Hoffmann A, Kayser M, Kohler P, Linacre A, Mayr-Eduardoff M, McGovern C, Morling N, O'Donnell G, Parson W, Pascalis VL, Porto MJ, Roseth A, Schneider PM, Sijen T, Stenzl V, Syndercombe Court D, Templeton JE, Turanska M, Vallone PM, van Oorchot RAH, Zatkalikova L, The RUROFORGEN-NoE Consortium, Carracedo A, Phillips C (2015) Forensic ancestry analysis with two capillary electrophoresis ancestry informative marker (AIM) panels: results of a collaborative EDNAP exercise. *Forensic Science International: Genetics*, 19, 56–67.
- Santos NPC, Ribeiro-Rodrigues EM, Ribeiro-Dos-Santos AKC, Pereira R, Gusmão L, Amorim A, Guerreiro JF, Zago MA, Matte C, Hutz MH, Santos SEB (2010) Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. *Human Mutation*, 31, 184–190.
- Savage D, Batley J, Erwin T, Logan E, Love CG, Lim GAC, Mongin E, Barker G, Spangenberg GC, Edwards D (2005) SNPServer: a real-time SNP discovery tool. *Nucleic Acids Research*, 33, 493–495.
- Schnabel RD, Kim JJ, Ashwell MS, Sonstegard TS, Van Tas-sell CP, Connor EE, Taylor JF (2005) Fine-mapping milk production quantitative trait loci on BTA6: analysis of the bovine osteopontin gene. *Proceedings of the National Academy of Sciences, USA*, 102, 6896–6901.
- Shedlock AM, Okada N (2000) SINE insertions: powerful tools for molecular systematics. *Bioessays*, 22, 148–160.
- Shen YJ, Jiang H, Jin JP, Zhang ZB, Xi B, He YY, Wang G, Wang C, Qian L, Li X, Yu QB, Liu HJ, Chen DH, Gao JH, Huang H, Shi TL, Yang ZN (2004) Development of genome-wide DNA polymorphism database for map-based cloning of rice genes. *Plant Physiology*, 135, 1198–1205.
- Sjödin P, Bataillon T, Schierup MH (2010) Insertion and deletion processes in recent human history. *PloS ONE*, 5, e8650.
- Steele KA, Ogden R, McEwing R, Briggs H, Gorham J (2008) InDel markers distinguish Basmati from other fragrant rice varieties. *Field Crops Research*, 105, 81–87.
- Tuzun E, Sharp AJ, Bailey JA, Kaul R, Morrison VA, Pertz LM, Haugen E, Hayden H, Albertson D, Pinkel D, Olson MV, Eichler EE (2005) Fine-scale structural variation of the human genome. *Nature Genetics*, 37, 727–732.
- Väli U, Brandström M, Johansson M, Ellegren H (2008) Insertion-deletion polymorphisms (Indels) as genetic markers in natural populations. *BMC Genetics*, 9, 715–720.
- Varshney RK, Nayak SN, May GD, Jackson SA (2009) Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends in Biotechnology*, 27, 522–530.
- Vasemägi A, Gross R, Palm D, Paaver T, Primmer CR (2010) Discovery and application of insertion-deletion (INDEL) polymorphisms for QTL mapping of early life-history traits in Atlantic salmon. *BMC Genomics*, 11, 156–166.
- Weber JL, David D, Heil J, Fan Y, Zhao CF, Marth G (2002) Human diallelic insertion/deletion polymorphisms. *American Journal of Human Genetics*, 71, 854–862.
- Xue Y, Jiang L, Su N, Wang JK, Deng P, Ma JF, Zhai HQ, Wan JM (2007) The genetic basic and fine-mapping of a stable quantitative-trait loci for aluminium tolerance in rice. *Planta*, 227, 255–262.
- Yi ZH, Lu YF, Guo XQ, Hui MX, Zhang LG, Zhang MK (2012) Development of simple sequence repeat (SSR) and insertion/deletion (InDel) markers in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) and analysis of their transferability. *Journal of Agricultural Biotechnology*, 20, 1398–1406. (in Chinese with English abstract) [仪泽会, 卢有飞, 郭晓芹, 惠麦侠, 张鲁刚, 张明科 (2012) 大白菜简单序列重复(SSR)和插入/缺失(InDel)标记的开发及通用性分析. *农业生物技术学报*, 20, 1398–1406.]
- Zhang SP, Miao H, Cheng ZC, Zhang ZH, Wu J, Sun RF, Gu XF (2011) The insertion-deletion (Indel) marker linked to the fruit bitterness gene (Bt) in cucumber. *Journal of Agricultural Biotechnology*, 19, 649–653. (in Chinese with English abstract) [张圣平, 苗晗, 程周超, 张忠华, 武剑, 孙日飞, 顾兴芳 (2011) 黄瓜果实苦味(Bt)基因的插入缺失(Indel)标记. *农业生物技术学报*, 19, 649–653.]

(责任编辑: 卢宝荣 责任编辑: 时意专)