

从水稻基因组序列中挖掘生物信息

樊龙江^{1,2}, 郭兴益²

(1. 浙江大学 作物科学研究所, 浙江 杭州 310029; 2. 浙江大学 生物信息学研究所, 浙江 杭州 310029)

摘要: 重点介绍了在水稻基因组序列分析方面取得的几项最新研究结果. 内容主要包括: 水稻基因组被证实曾发生过全基因组倍增; 水稻基因组大小变化的趋势; 水稻籼粳亚种的分化时间比以前估计的 100 万年前要迟得多; 对水稻基因产生 GC 含量负梯度现象的解释.

关键词: 水稻基因组; 全基因组倍增; 籼粳分化时间; 基因组大小; GC 含量梯度; 综述
中图分类号: Q78 **文献标识码:** A

FAN Long-jiang^{1,2}, GUO Xing-yi² (1. Institute of Crop Science, Zhejiang University, Hangzhou 310029, China; 2. Institute of Bioinformatics, Zhejiang University, Hangzhou 310029, China)

Data mining of rice genome. Journal of Zhejiang University (Agric. & Life Sci.), 2005, 31(4): 355-361
Abstract: In this review, we focus on four recent advances in rice genome analysis: an whole genome duplication which is believed to occur in rice genome; evolution of rice genome size; the *indica-japonica* divergence that was likely to occur later than the previous estimate of 1 million years ago; hypothesis for GC compositional gradients in rice genes.

Key words: rice genome; whole genome duplication; *indica-japonica* divergent time; genome size; GC compositional gradient; review

水稻是第一个被全基因组测序的作物. 共有 2 个水稻亚种被测序, 一个是粳稻品种日本晴分别通过全基因组鸟枪法^[1]和逐步克隆方法^[2~4]测序, 另一个是籼稻品种“9311”通过全基因组鸟枪法测序^[5]. 除了核基因组外, 水稻的叶绿体基因组序列早在 15 年前就已测序完成^[6], 同时, 其线粒体基因组最近也被测序完成.

在获得基因组序列后, 一项艰巨的研究任务是如何从巨量的水稻基因组序列中挖掘出潜藏的遗传事件、进化机制等重要生物信息. 为此

本文结合我们自身的一些研究工作, 重点介绍了近年来在水稻基因组序列分析中获得的几项最新的研究结果.

1 现代的二倍体, 古老的多倍体

2004 年水稻基因组研究的一个重要进展, 是获得清晰的证据表明水稻基因组曾发生过全基因组倍增. Paterson 等^[7]、Guyot 等^[8]和我们的^[9,10]的研究结果也一致表明, 在禾本科作物分化前发生过一次全基因组倍增 (whole-

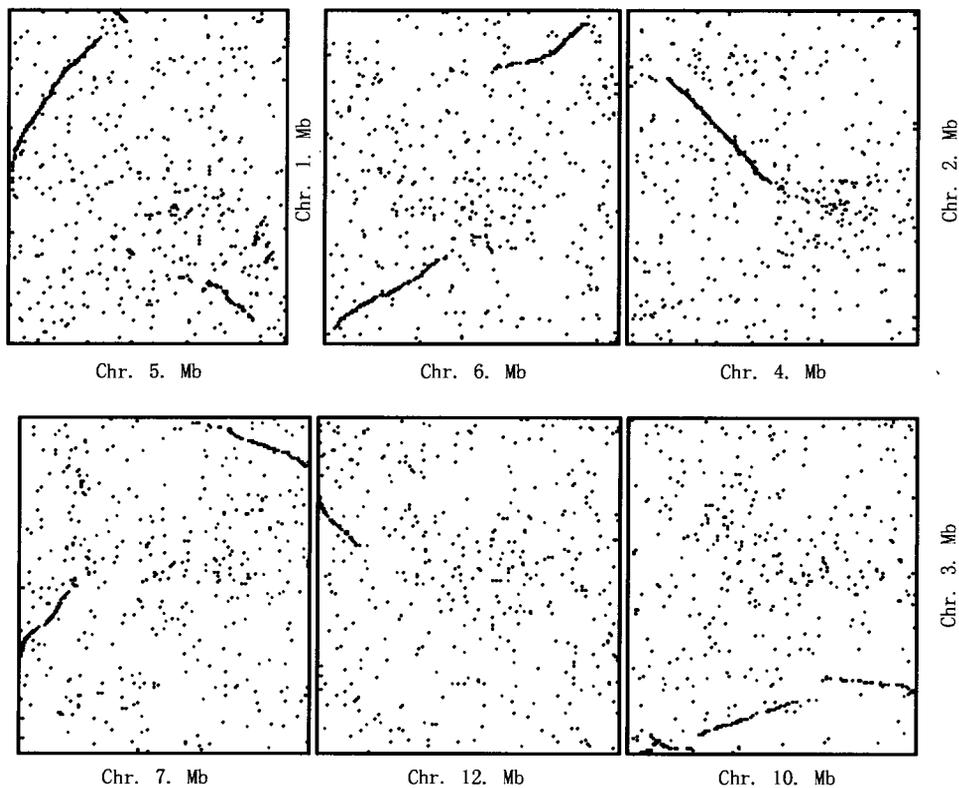
收稿日期: 2005-03-30

基金项目: 国家自然科学基金资助项目(30270810/30471067).

作者简介: 樊龙江(1965—), 男, 教授, 博士生导师, 主要从事禾本科作物基因组学及其生物信息学方法、遗传育种方面的研究. Tel: 0571-86971730; E-mail: fanlj@zju.edu.cn.

genome duplication). 早在 2002 年, 根据最初的水稻基因组草图序列, Goff 等^[1] 利用同义替换率分布方法(*K_s-based age distribution*)提出水稻基因组可能发生过一次全基因组倍增. 而在此之前, 利用分子标记、DNA 重复元件等方法对水稻部分染色体区段的研究, 也提出水稻基因组的一些染色体间可能发生过片段倍增(block or segmental duplication). 2003 年两篇重要文章相继发表, 对水稻基因组起源和倍增事件做出了初步分析和有益探索^[11,12]. 随着水稻基因组序列数据的增加, 特别是美国基因组研究院(TIGR)利用逐步克隆(clone by clone)测序的数据首次拼成 12 条水稻染色体序列, 利用 TIGR 的数据和基因相似性矩阵方法(GHM, gene homology matrix), 检测到大量染色体间的倍增片段, 这些倍增片段几乎覆盖了水稻全基因组(图 1、图中包括水稻第 2 号染色体与第 4 和 6 号染色体、第 3 号染色体与第 7、10 和 12 号染色体和第 1 与 5 号染色体间的

倍增片段. 另外第 8 与 9 号染色体、第 11 与 12 号染色体间的倍增片段未列出). 这是全基因组倍增的有力证据. 根据倍增片段上同源基因的分子进化分析, 全基因组倍增大致发生在 7000 万年前, 在禾本科作物分化前^[7]. 我们在 2004 年初利用 TIGR 的第一版水稻基因组数据(osal, version 1)和 GHM 方法就已发现了这一水稻基因组倍增的证据并投稿(论文摘要已递交上海-合肥举行的系统与进化研讨会,^[9]). 但就在 6 月底-7 月初, Paterson 等^[7]和 Guyot 等^[8]的文章相继发表. 后我们利用 TIGR 更新的数据(osal, version 2)对水稻染色体间倍增片段进行了更新, 并以此为基础, 利用同义替换率分布方法检测到另一次更古老的(单双子叶植物分化前)基因组倍增事件^[10]. 该研究的最新进展是中科院北京基因组研究所(华大)刚刚发表的水稻基因组精细图分析结果也同样证实了水稻基因组的倍增^[13], 同时, 另外一个独立的课题组最近也获得了同样的结论^[14].



引自文献^[10].

图 1 部分水稻基因组倍增片段

Fig. 1 Duplicate segments of rice genome

全基因组倍增或整倍体化过程被认为是植物尤其是禾本科作物物种形成和进化过程中非常普遍和重要的事件, 50%~70% 的开花植物在进化过程中均经历了一次或多次染色体加倍过程^[15]. 基因组加倍后, 再经历所谓的二倍体化过程(diploidization), 进化成当代的二倍体物种. 大量的复制基因将在二倍体化过程中丢失. 整倍体化过程一般可通过同源加倍(autohexyploid)和异源加倍(allohexyploid)两种方式发生. 已测序完成的模式植物拟南芥, 经全基因组序列分析发现, 至少发生过 3 次全基因组自身复制^[16]; 玉米被认为在其与高粱分化后发生一次异源加倍过程, 即起源于异源四倍体(allotetrapolyloid)^[17]. 利用同义替换率分布方法检测和最新序列数据库数据, Blank and Wolfe^[18] 在很多重要作物中均发现了全基因组倍增的证据.

水稻全基因组倍增片段是迄今为止发现的在动植物基因中最为清晰、完整的基因组倍增的遗迹. 拟南芥基因组在更近代的时候也发生过全基因组倍增, 但它的倍增片段都比较短且凌乱^[16, 19]. 水稻之所以保存得这么完整可能与水稻基因组相对比较稳定有关^[20].

2 最小的核基因组: 基因组在扩增还是在缩小?

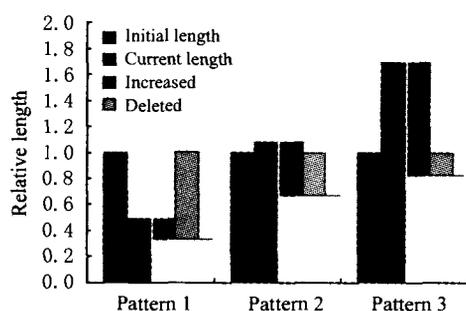
植物界基因组中 DNA 含量差异很大, 它们的差异性与生物的复杂性程度并不完全相关, 这种现象称为 C 值悖理. 如大麦(*Hordeum vulgare*)、水稻和拟南芥的生物复杂性比较相似, 但大麦基因组分别为水稻和拟南芥基因组的 11 倍和 35 倍. 众多因素(机制)决定了基因组的膨胀和缩小^[21], 早在 19 世纪 30 年代, 基因复制就被认为是增长遗传物质的首要机制^[22]. 在植物界中, 基因数目的增加通常归因于基因复制、DNA 片断或基因组复制. 基因组膨胀的最主要因素为基因组的倍增^[15, 23]. 而转座因子的扩增则是另一个推动基因组增加的关键因素. 在禾本科内, 已报道在最近的 1 千万年内大多数基因组的膨胀由 LTR 逆转座因子的扩增所导致^[24, 25]. 很明显的, 这一机制只

能导致基因组膨胀^[26], 而基因组只是这样一味地膨胀进化吗? 并非如此. 后来发现了抵制这一膨胀的机制: 异常重组(illegitimate recombination)和非同源性重组(unequal homologous recombination)可以减少 LTR 逆转座序列从而抵制基因组的膨胀^[27~29]. 最近已发现水稻和拟南芥基因组中的 LTR 逆转座序列的大量丢失^[29, 30]. 在最近的 800 万年, 水稻基因组中至少有 190 Mb 的 LTR 逆转座序列被删除^[29]. 利用非洲栽培稻进行的比较基因组研究表明, 亚洲栽培水稻的籼粳稻基因组大小均增加了 2% 和 6%^[25]. 但该研究的结论仅是根据约 1 Mb 长度的基因组片段(水稻 430 Mb 基因组的 0.2%)得出. 根据 non-LTR 逆转座研究, Petrov 和他的同事得出非平衡性的少量删除和插入导致昆虫类的基因组缩小^[31]. 然而, 在植物基因组中是否存在同样相似的机制作用于转座因子, 或者其它机制导致非重复序列的丢失仍然没有明确的答案.

为了探索基因组大小改变的潜在进化机制, 一种较理想的途径是比较基因组间大小差异很大的相近物种. 通过比较果蝇(165 Mb)和其它两个基因组极大的相近物种 *Laupala crickets* (1910 Mb) 和 *Podisma grasshoppers* (18150 Mb), 发现果蝇 DNA 的大量丢失^[31]. 最近, 通过比较异源多倍体物种棉花(*Gossypium hirsutum*)不同基因组序列片断, 探索了该物种基因组大小变化的进化机制^[23].

在有花植物中, 全基因组倍增是普遍发生的现象, 并且被认为在物种进化和分化中起着重要作用^[15]. 一旦染色体倍增过后, 古老多倍体的基因组进化速率加快, 在“二倍化”过程中伴随着大量的 DNA 序列的消失以及染色体重排现象^[2]. 水稻基因组测序工作的完成^[2, 4], 为研究水稻基因组的进化史提供了一个前所未有的机会. 水稻基因组多倍体的起源已被证实^[7, 10, 11]. 多倍化事件估计发生在 7000 万年前, 在禾本科分化之前^[7]. 这一结论是基于许多非重叠的倍增块几乎覆盖了整个基因组这一事实而得出. 该研究结果为研究水稻基因和基因组倍增后的二倍体化的进化机制提供了非常好的素材.

当一次复制事件发生,两对应的复制片断或染色体在初始阶段通常应具有同样的大小.但经过长期的进化,其同源的复制片断的大小有可能存在差异.由基因组复制产生的复制块(同源复制块)将经历一次“二倍体化”的剧烈进化过程,伴随着大量的 DNA 序列的丢失.同源复制片断间存在的巨大长度差异为分析基因组膨胀或缩小进化机制提供了有效的途径.在我们的研究中,从水稻全基因组倍增产生的同源复制片断(分别来自第 2, 3, 6, 7 和 10 号染色体),由于它们存在着巨大的差异性而被选择为研究对象,用于探索水稻经历多倍化后基因组大小的进化机制.我们的研究表明,在最近 70 百万年里,水稻染色体以不均衡的模式(即染色体长度存在膨胀、平衡和减小 3 种情况)进化着,影响复制片断大小的差异主要由非重复序列的 DNA 丢失引起的,且 LTR 因子的扩增也起着重要作用^[32](图 2).



引自文献^[32].

图 2 水稻基因组染色体长度变化的三种进化模式

Fig. 2 Three patterns of chromosomal size evolution in rice genome

3 籼粳稻分化时间比原来估计的要迟得多

水稻 (*Oryza sativa* L.) 属于禾本科 (Gramineae 或 Poaceae), 也是 3 大谷类植物之一, 即水稻, 小麦 (*Triticum aestivum*) 和玉米 (*Zea mays*). 为人类提供了主要食源. 大约在 77 百万年前禾本科从同一祖先分化而来, 其两个亚科 Ehrhartoideae (水稻) 和 Panicoideae (玉米和高粱) 大约在 5000 万年分开^[39]. 水稻化石

的研究可追溯到约 4000 万年前. 22 个水稻物种中已发现 9 个物种为 2 倍体类型 ($2n=24$) 以及由不同重组形成的异源 4 倍体 ($2n=48$) 等. *O. rufipogon* 是栽培稻 (*Oryza sativa* L., AA 基因组) 的野生祖先, 后被驯化为 *O. sativa*, 其驯化时间可能起源于 9000 年前. 栽培稻有 2 个主要亚种籼稻和粳稻, 基于来自 2 个亚种的 29 kb 的同源片断, Bennetzen^[26] 认为它们约在 1 百万年前分开, 但是他未给出这一时间估计的详细信息. 这一分化时间估计后来在水稻基因组的研究中被广泛引用^[33, 34].

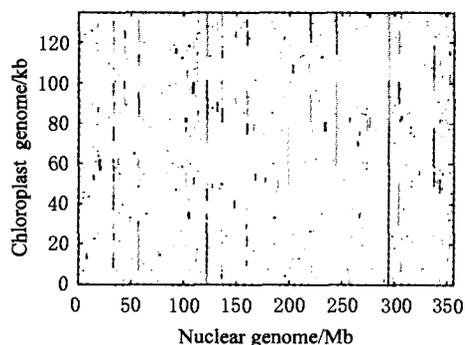
水稻基因组测序的工作已基本完成. 栽培稻粳稻日本晴通过全基因组鸟枪法^[1] 和利用遗传图和物理图的逐步克隆方法被测定^[2~4]. 栽培稻籼稻“9311”通过全基因组鸟枪法被测定^[5]. 除了核基因组, 水稻叶绿体基因组早在 15 年前就被测序完成^[6]. 同样地, 玉米和小麦地叶绿体测序工作最近也已完成.

细胞核、叶绿体和线粒体间 DNA 序列的插入很早就被发现^[35]. 粳稻第 10 号染色体上的 2 个长的叶绿体基因组序列插入已被检测到^[4]. 同时, 籼稻基因组序列中也同样发现大量的叶绿体序列的插入^[36].

植物细胞核和细胞器基因的同义替换率 (ds) 通常被用于进化事件的时间估计^[37]. 考虑到叶绿体的一些有利因素, 如母系遗传、很少或没有重组等^[38], 叶绿体 DNA 已被广泛地应用于植物分化时间的估计^[37~39]. 同时, 核基因序列也被用于分化时间的研究中, 如 Bennetzen 等人的研究^[26]. 非同义替换率 (氨基酸改变, dn) 与同义替换率的 (氨基酸不改变, ds) 的比值 (dn/ds) 也经常用于分化分析. dn/ds 的比值为 1 表示所研究的基因在中性选择 (neutral selection) 下进化, 小于 0.25 意味着纯化选择 (purifying selection) 下进化, 当比值大于 1 时则被认为进行正向选择 (positive selection) 下的进化^[40, 41].

我们利用大片断叶绿体 DNA 的插入 (图 3). 图中为水稻核基因组序列——12 条染色体依次拼接在一起——与叶绿体基因组联配结果. 蓝线表示叶绿体片断顺式插入核基因组, 绿线表示叶绿体片断反式插入核基因组) 来估计

水稻 2 个亚种籼稻和粳稻的分化时间. 通过 PCR 扩增和籼稻基因组层次上对叶绿体大片断的搜索, 确立了籼稻—粳稻分化之前叶绿体的最近一次插入并根据同义替换率推断出 2 个亚种分化时间至少在近 70 万年内, 很有可能发生在距今 50 万年前^[42]. 该结果与最新利用核基因做出的推断相一致^[25].



引自文献^[42].

图 3 水稻核基因组中叶绿体 DNA 的插入情况

Fig. 3 Chloroplast DNA insertions in nuclear genome of rice

4 高 GC 含量基因与 GC 负梯度现象

禾本科基因沿转录方向上 GC(鸟嘌呤+胞嘧啶)组成上存在负梯度现象最近被发现, 而在双子叶植物基因却无此现象^[5, 43]. 这是一个明显和有趣的现象, 但其产生的机制尚无合理的解释. GC 含量作为基因组的一个重要识别标志, 已被用于基因组的基本组成的分析, 编码序列的进化以及密码子使用偏好性上^[44]. 例如, CpG 岛(GC 富含区)被用于真核生物的基因一个路标信息^[45]. 物种间基因平均 GC 含量的变化幅度较大, 甚至在同一类物种(如细菌)中也是如此. 物种中的这种 GC 含量差异产生的原因尚不清楚.

禾本科内包括了将近 10000 物种, 可被分成 700 个属^[39], 表现为独立的遗传体系^[46]. 最近的比较基因组学研究表明, 所有禾本科植物都追溯到一个共同的“禾本等位基因”(Grass alleles)群体^[47]. 有报道指出, 禾本科有一次 GC 含量提高过程并且在玉米和水稻中可分成

两类基因(高 GC 和低 GC)^[48]. 通过考察来自 4 个禾本科物种(水稻, 玉米, 小麦和大麦)的 25 个基因家族, 每个家族成员的基因 GC 含量存在着巨大差异^[49]. 同时, 最近也有报道指出, 微卫星分布的一个新特点也沿着基因转录方向呈现梯度变化. 对于水稻基因, 通常在基因 5' 端能探测到富含 GC 的微卫星, 如(CCG)_n 等^[50]. 通过水稻基因组内 CpG 岛的分析, 同样也大量出现在表达基因的 5' 端^[45].

基于水稻 28000 个全长 cDNA(来源于实验)和基因组序列以及其它物种的类似数据, 我们详细研究了禾本科以及其它物种的基因 GC 含量和梯度. 根据水稻转录组 GC 含量的分布, 我们得出了水稻 GC 梯度变化模式和明显的两组基因类型(图 4). 由此我们提出一个假设解释水稻等禾本科基因 GC 含量梯度产生的机制: 禾本科可能存在着一个异源起源过程, GC 含量存在明显差异的两个古老物种参与了禾本科的物种形成. 水稻基因组中插入的外源基因(如叶绿体和细菌)GC 含量的变化支持上述假设.

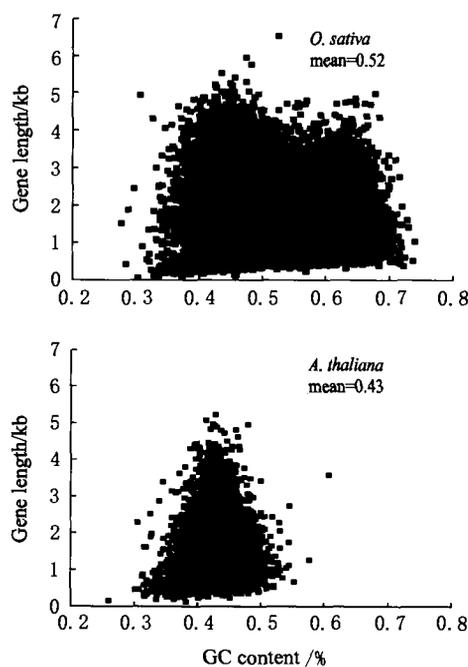


图 4 水稻和拟南芥基因组基因的 GC 含量分布

Fig. 4 GC contents distributions of transcriptomes of rice and *Arabidopsis*

References:

- [1] Goff S A, Rick D, Lan T H, *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*) [J]. **Science**, 2002, 296, 92-100.
- [2] Sasaki T, Matsumoto T, Yamamoto K, *et al.* The genome sequence and structure of rice chromosome 1 [J]. **Nature**, 2002, 420: 312-316.
- [3] Feng Q, Zhang Y, Hao P, *et al.* Sequence and analysis of rice chromosome 4 [J]. **Nature**, 2002, 420: 316-320.
- [4] The Rice Chromosome 10 Sequencing Consortium. In-depth view of structure, activity, and evolution of rice chromosome 10 [J]. **Science**, 2003, 300: 1566-169.
- [5] Yu J, Hu S, Wang J, *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*) [J]. **Science**, 2002, 296: 79-92.
- [6] Hiratsuka J, Shimada H, Whittier R, *et al.* The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals [J]. **Mol. Gen. Genet.**, 1989, 217: 185-194.
- [7] Paterson A H, Bowers J E, Chapman B A. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics [J]. **Proc. Natl. Acad. Sci. USA**, 2004, 101: 9903-9908.
- [8] Guyot R, Keller B. Ancestral genome duplication in rice [J]. **Genome**, 2004, 47(3): 610-4.
- [9] Fan L, Xu G, Zhang Y, *et al.* Duplication of rice (*Oryza sativa*) genome [A]. **Proceedings of 8th National Systematic and Evolutionary Botany Youth Seminar and Systematic & Evolutionary Biology Conference**[C], 2004. 7.
- [10] Zhang Y, Xu G H, Guo X Y, *et al.* Two ancient rounds of polyploidy in rice genome [J]. **Journal of Zhejiang University SCIENCE**, 2005a, 6(2): 87-90.
- [11] Paterson A H, Bowers J E, Peterson J, *et al.* Structure and evolution of cereal genomes [J]. **Curr. Opin. Genet. Dev.**, 2003, 13: 644-650.
- [12] Vandepoele K, Simillion C, Van de Peer Y. Evidence that rice and other cereals are ancient aneuploids [J]. **Plant Cell**, 2003, 15: 2192-2202.
- [13] Yu J, Wang J, Lin W, *et al.* The genome of *Oryza sativa*: A history of duplication [J]. **PLoS. Biol.**, 2005, 3(2): e38.
- [14] Wang X, Shi X, Hao B, *et al.* Duplication and DNA segmental loss in the rice genome: implications for diploidization [J]. **New Phytologist**, 2005, 165: 937-946.
- [15] Wendel J F. Genome evolution in polyploids [J]. **Plant Mol. Biol.**, 2000, 42: 225-249.
- [16] Bowers J E, Chapman B A, Rong J, *et al.* Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events [J]. **Nature**, 2003, 422: 433-438.
- [17] Gaut B S, Doebley J F. DNA sequence evidence for the segmental allotetraploid origin of maize [J]. **Proc. Natl. Acad. Sci. USA**, 1997, 94: 6809-6814.
- [18] Blanc G, Wolfe K H. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes [J]. **Plant Cell**, 2004, 16: 1667-1678.
- [19] Simillion C, Vandepoele K, Van Montagu M C E, *et al.* The hidden duplication past of *Arabidopsis thaliana* [J]. **Proc. Natl. Acad. Sci. USA**, 2002, 99: 13627-13632.
- [20] Ilic K, SanMiguel P J, Bennetzen J L. A complex history of rearrangement in an orthologous region of the maize, sorghum, and rice genomes [J]. **Proc. Natl. Acad. Sci. USA**, 2003, 100: 12265-12270.
- [21] Bennetzen J F. Mechanisms and rates of genome expansion and contraction in flowering plants [J]. **Genetica**, 2002, 115: 29-36.
- [22] Betran E, Long M. Expansion of genome coding region by acquisition of new genes [J]. **Genetica**, 2002, 115: 65-80.
- [23] Grover C E, Kim H, Paterson A H, *et al.* Incongruent patterns of local and global genome size evolution in cotton [J]. **Genome Res.**, 2004, 14: 1474-1482.
- [24] San Miguel P, Tikhonov A, Jin Y-K, *et al.* Nested retrotransposons in the intergenic regions of the maize genome [J]. **Science**, 1996, 274: 765-768.
- [25] Ma J, Bennetzen J L. Rapid recent growth and divergence of rice nuclear genomes [J]. **Proc. Natl. Acad. Sci. USA**, 2004, 101: 12404-12410.
- [26] Bennetzen J L. Comparative sequence analysis of plant nuclear genomes: microcolinearity and its many exceptions [J]. **Plant Cell**, 2000, 12: 1021-1029.
- [27] Vicient C M, Suoniemi A, Anamthawas-Jonsson K, *et al.* Retrotransposon BARE-1 and its role in genome evolution in the genus *Hordeum* [J]. **Plant Cell**, 1999, 11: 1769-1784.
- [28] Shirasu K, Schulman A H, Lahaye T, *et al.* A contiguous 66 kb barley DNA sequence provides evidence for reversible genome expansion [J]. **Genome Res.**, 2000, 10: 908-915.
- [29] Ma J, Devos K, Bennetzen J L. Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice [J]. **Genome Res.**, 2004, 14: 860-869.

- [30] Devos K M, Brown J K M, Bennetzen J F. Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis* [J]. **Genome Res.**, 2002, 12: 1075-1079.
- [31] Petrov D A. DNA loss and evolution of genome size in *Drosophila* [J]. **Genetica**, 2002, 115: 81-91.
- [32] Fan L, Guo X, Xu G, *et al.* Incongruent chromosomal size evolution in rice. 2005a [J]. **Genomics**, Submitted.
- [33] Song R, Llaca V, Messing J. Mosaic organization of orthologous sequences in grass genomes [J]. **Genome Res.**, 2003, 12: 1549-1555.
- [34] Han B, Xue Y. Genome-wide intraspecific DNA-sequence variations in rice [J]. **Curr. Opin. Genet. Dev.**, 2003, 13: 134-138.
- [35] Notsu Y, Masood S, Nishikawa T, *et al.* The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants [J]. **Mol. Genet. Genomics**, 2002, 268: 434-445.
- [36] Shahmuradov I A, Akbarova Y Y, Solovyev V V, *et al.* Abundance of plastid DNA insertions in nuclear genomes of rice and *Arabidopsis* [J]. **Plant Mol. Biol.**, 2003, 52: 923-934.
- [37] Wolfe K H, Gouy M, Yang Y-W, *et al.* Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data [J]. **Proc. Natl. Acad. Sci. USA**, 1989, 86: 6201-6205.
- [38] Sall T, Jakobsson M, Lind-hallden C, *et al.* Chloroplast DNA indicates a single origin of the allotetraploid *Arabidopsis suecica* [J]. **J. Evol. Biol.**, 2003, 16: 1019-1029.
- [39] Gaut B S. Evolutionary dynamics of grass genomes [J]. **New Phytologist**, 2002, 154: 15-28.
- [40] Hurst L D. The Ka/Ks ratio: diagnosing the form of sequence evolution [J]. **Trends Genet.**, 2002, 18: 486-487.
- [41] Swanson W J. Adaptive evolution of genes and gene families [J]. **Curr. Opin. Genet. Dev.**, 2003, 13: 617-622.
- [42] Fan L, Guo X, Zhang X, *et al.* Timing of *indica-japonica* subspecies divergence of rice (*Oryza sativa* L.) based on average synonymous substitution rates of chloroplast DNA insertions in nuclear genome [J]. **Mol. Genet. Genomics**, 2005b, Submitted.
- [43] Wong G K, Wang J, Tao L, *et al.* Compositional Gradients in Gramineae Genes [J]. **Genome Res.**, 2002, 12: 851-856.
- [44] Bernardi G. Isochores and the evolutionary genomics of vertebrates [J]. **Gene**, 2000, 241: 3-17.
- [45] Ashikawa I. Gene-associated CpG islands in plants as revealed by analyses of genomic sequences [J]. **Plant J**, 2001, 26: 617-625.
- [46] Bennetzen J L, Freeling M. Grasses as a single genetic system-genome composition, colinearity and compatibility [J]. **Trends Genet.**, 1993, 9: 259-261.
- [47] Freeling M. Grasses as a single genetic system [J]. **Plant Physiol.**, 2001, 125: 1191-1197.
- [48] Carels N, Bernardi G. Two classes of genes in plants [J]. **Genetics**, 2000, 154: 1819-1825.
- [49] Zhang L, Pond K S, Gaut B S. A survey of the molecular evolutionary dynamics of twenty-five multigene families from four grass taxa [J]. **J. Mol. Evol.**, 2001, 52: 144-156.
- [50] Fujimori S, Washio T, Higo K, *et al.* A novel feature of microsatellites in plants: a distribution gradient along the direction of transcription [J]. **FEBS Lett.**, 2003, 554: 17-22.