染色体数目减少及B染色体产生的进化基因组学模型

王振怡¹, 王希胤^{1,2*}

1. 华北理工大学生命科学学院, 基因组学与计算生物学研究中心, 唐山 063210;

2. 成都中医药大学基因组学与生物大数据学院, 成都 610075

* 联系人, E-mail: wangxiyin@vip.sina.com

收稿日期: 2020-02-17; 接受日期: 2020-04-01; 网络版发表日期: 2020-04-29

国家重点研发计划(批准号: 2016YFD0101001)和河北省高等学校科学技术研究项目(批准号: QN2017123)资助

摘要 在长期进化过程中, 染色体数目展示出动态性变化. 关于真核生物染色体进化仍有许多谜团, 一个是如何维持染色体数目, 另外一个B染色体是如何产生的. 全基因组测序工作的开展为认识染色体数目变化提供了新的机会, 特别是反复的多倍化事件后染色体的改变有助于人们理解相关生物学规律和机制. 通过比较基因组学分析, 本文提出了染色体数目变化与B染色体产生的模型, 认为染色体数目的减少主要是由于染色体融合, 并且端粒的丢失及B染色体的产生在基因组重新整合和染色体数目减少中起着重要作用.

关键词 进化基因组学, B染色体, 染色体数目减少, 多倍化, 全基因组加倍, 核型进化

测序技术的不断发展, 为人们从染色体DNA水平认识生命现象提供了机会, 而分子和计算分析工具也为探测基因组历史提供了新的方法. 关于染色体的进化仍有许多让人们困惑的地方. 全基因组分析发现, 很多真核生物特别是植物, 在进化史上都经历了一次或多次多倍化事件, 而每次多倍化造成染色体加倍之后, 常常会发生复杂的染色体重构成, 最终形成现有物种的细胞核型和染色体构成. 那么, 真核生物是如何维持染色体数目在一个小的范围内? 另外, 在真核生物中, 除一般的“正常”染色体之外, 还有一种很小的染色体, 被称为B染色体, B染色体的产生被认为是生理之谜. 染色体数减少的现象可能只发生在很短的时间内, 或者在很长的时间里仅发生一次, 使得基

于显微镜的生物学观察很难发现和解释相关的规律和机制.

1 真核生物染色体及染色体数目

1.1 真核生物的染色体

真核生物一般拥有线性的核染色体, 这些线性染色体的两端是由重复序列构成的端粒, 保护染色体的完整性. 线性染色体比较靠中间的位置是着丝粒, 着丝粒是一个独特的染色体结构域, 与有丝分裂和减数分裂过程中染色体的分离有关. 染色体可能发生断裂、融合和重排, 引发基因组不稳定, 改变生物的细胞核型^[1]. 即使亲缘关系相近的不同生物, 其染色体的

引用格式: 王振怡, 王希胤. 染色体数目减少及B染色体产生的进化基因组学模型. 中国科学: 生命科学, 2020, 50: 524–537
Wang Z Y, Wang X Y. Evolutionary genomics model of chromosome number reduction and B chromosome production (in Chinese). Sci Sin Vitae, 2020, 50: 524–537, doi: 10.1360/SSV-2020-0042

数目也可能不同. 很显然, 细胞核型是不断变化的, 细胞核型的变化造成了物种分化.

1.2 一个生理之谜——B染色体

除了“正常”的染色体外(称为A染色体), 有些物种中还有一些多余的染色体, 称为B染色体(图1). B染色体的起源一直是一个生理谜团^[2]. 1907年, 威尔逊在研究半翅目昆虫的染色体时发现了B染色体^[3]. B染色体对于一个物种的生命活动来说不是必需的, 并且在很多物种中没有这类染色体^[4]. 大多数B染色体含有丰富的异染色质(因此大部分是非编码的), 主要由重复序列组成^[5], 但有些比较特殊, 如玉米的B染色体含有大量的常染色体片段^[6]. B染色体存在于许多植物、动物和真菌中^[7,8](图1), 不与任何A染色体配对, 不遵循孟德尔遗传定律^[9,10]. B染色体是核型多余的可有可无的部分, 通常被认为是没有功能基因的“垃圾DNA”或基因组寄生虫. 但有研究认为, B染色体可携带转录活性基因序列, 可能影响其宿主基因组的转录组谱^[11]. 各种研究表明, B染色体对不同的宿主具有中性、有害或有益的影响^[7,9,12,13]. 在自然群体中, 个体可能携带1~34个B染色体^[7,14]. 约15%的真核生物有B染色体, 但在分子和功能水平上尚不清楚^[15].

B染色体与A染色体形态不同, 且B染色体数目和机制在不同物种和个体之间变化很大, 研究起来较为复杂. 很多学者认为B染色体的起源可能是多种途径的, 并提出了多种假说. 通常认为B染色体来源于A染色体^[15~20], 其起源与A染色体的基因组不稳定有关, 涉及到B染色体的最初形成和后来重复序列的扩展^[11,17,18]. 其中, 重复序列包括转座子^[21~24], 代表大多数真核基因组的大部分, 也是B染色体构成的主要成分^[7,22]. B染色体结构的进化历来在细胞遗传学水平上引起关注^[25], 最近关注更多的是分子和基因组水平的研究^[26,27], 以及重复DNA对B染色体进化的影响^[4].

2 全基因组加倍现象

2.1 有花植物全基因组加倍

染色体数目是真核生物进化的一个显著动态特征. 在染色体数目变化的机制中, 多倍化(polyploidization)由于影响了动物^[28]、真菌^[29]和原生动物^[30]在内的多种真核生物而备受关注. 多倍化或全基因组加倍

图1 植物和动物中具有B染色体的例子^[7]. A: 黑麦(*Secale cereale*); B: 雏菊(*Brachycome dichromosomatica*); C: 毛叶还阳参(*Crepis capillaris*); D: 大林姬鼠(*Apodemus peninsulae*). 箭头指的是B染色体

Figure 1 Examples of plant and animal species possessing B chromosomes^[7]. A: The B chromosomes in rye (*Secale cereale*); B: the B chromosomes in *Brachycome dichromosomatica*; C: the B chromosomes in *Crepis capillaris*; D: the B chromosomes in *Apodemus peninsulae*. Arrows indicate B chromosomes

(whole-genome duplication, WGD)被认为是高等植物物种形成和进化的主要原材料. 大约70%的被子植物在演化过程中经历过一次或多次多倍化^[31]. 许多重要作物都是多倍体或被认为是多倍体起源, 如小麦、棉花、玉米、大豆、香蕉、咖啡、甘蔗等.

反复的多倍化是植物基因组进化和物种分化的重要推动力, 模式植物水稻和拟南芥基因组序列都揭示了古多倍化现象. 研究推测, 所有的现代植物都有一个多倍化的祖先^[32]. 开花植物和种子植物的产生、快速分化及在地球上的优势生长也归因于多倍化事件^[33]. 这在很大程度上解答了当年达尔文对于被子植物快速分歧产生大量类群的迷惑^[34]. 利用最新的化石证据和共线性基因对禾本科主要进化事件发生的时间进行重新估算, 表明禾本科大概起源于1亿年前左右, 即一次古多倍化发生之时^[35].

多倍体植物含有两套以上的基因组, 其染色体重组过程复杂. 多倍体通常在形态、生理或生活史特征上与其祖先有显著性差异^[36,37], 这些差异可能有助于多倍体物种在新环境中的生存和分化^[38]. 多倍体植物拥有较强的可塑性, 其基因组具有较强的缓冲能力, 极大提高了其抗逆性与环境适应性, 例如小麦通过两

次多倍化形成普通小麦(六倍体小麦), 在世界各地广泛种植。

2.2 脊椎动物全基因组加倍

在脊椎动物谱系中, 多倍体最常见于两栖动物和硬骨鱼^[39-41]。自20世纪30年代以来, 基因和基因组加倍一直被认为在进化过程中起着重要作用^[42-44]。1970年, Ohno^[44]提出脊椎动物的复杂性和基因组大小是由早期脊椎动物进化中的两次全基因组加倍所导致的, 为脊椎动物进化的多样化提供了原材料。硬骨鱼类经历了另一轮基因组加倍, 即所谓的鱼类特异基因组加倍^[45-47]。这次多倍化事件发生的时间和硬骨鱼类发生辐射扩增的时间几乎是一致的(2.53亿~4.04亿年前), 硬骨鱼类巨大的形态多样性可能与这一事件有因果关系^[45,46,48-51]。这一事件之后重复基因的丢失广泛发生, 重复基因的差异保留和丢失是基因组复制后物种形成过程中的常见现象, 最终可能导致物种分化。此外, 一些硬骨鱼类还受到进一步多倍化影响, 如鲤形目^[52-55]和鲢形目等^[56]。

虽然鱼类多倍化的发生没有植物进化中那样频繁, 但其基因组也受到多倍化的影响^[57]。硬骨鱼的祖先在3亿多年前发生了多倍化^[58], 鲑鱼祖先在5000万~8000万年前发生了多倍化^[59], 比较近的加倍事件(约800万~1200万年前)使鲤鱼和草鱼有一个多倍化的祖先^[57,60,61]。同源基因分析可以明确地把其基因组分成两个亚基因组, 表明这一加倍事件是一次异源的基因组加倍^[61,62]。

哺乳动物和鱼类在大约4.5亿年前分歧, 没有证据表明哺乳动物发生多倍化^[63]。哺乳动物基因组的大小几乎没有种间变异^[64,65], 哺乳动物基因组进化中可能也没有发生全基因组加倍事件^[66]。在哺乳动物中, 只有有限数量的组织是由多倍体细胞组成的。

在两栖类和爬行类的研究中, 目前仅有关于非洲爪蟾四倍体的起源和基因组进化的报道^[67]。非洲爪蟾是拥有从二倍体到十二倍体的多倍体物种之一, 是研究基因组加倍影响的理想选择^[68]。四倍体爪蟾在1700万~1800万年前经历了异源多倍化事件, 两个亚基因组的进化是不对称的, 其中一个染色体组较好地保持了原始状态, 另一个染色体组则经历了更多的基因丢失、缺失、重排和基因表达减少^[67]。此研究为解析脊椎动物古多倍化事件和多倍体组学研究提供了参考。

2.3 染色体数量的多变性

多倍化在基因组结构、基因含量和进化过程中都起着重要作用^[69], 全基因组加倍后染色体数目加倍, 是真核生物进化的一个重要特征。真核生物的多样性与基因组加倍密切相关。

(1) 植物多倍化与染色体数目。大多数开花植物被证明是古老的多倍体植物的后代, 或者说祖先基因组在进化过程中经历了一轮或多轮的多倍化事件。植物基因组的一个重要特点是反复的全基因组加倍, 造成染色体数目的倍增。基因组加倍之后, 如果是异源多倍化, 新产生的所谓多倍体植物本质上还是二倍体, 也就是基因组执行二倍化遗传, 如小麦、油菜等; 而同源多倍体在一定的时期可能执行多倍化遗传, 之后很多植物会恢复二倍化遗传。研究表明, 二倍化遗传的恢复过程, 可能与多倍化产生的重复或者部分同源染色体间的非正常DNA重组有关。这种非正常的遗传重组引起基因组的不稳定, 可能造成大规模的基因丢失和重排。二倍化遗传恢复的同时, 染色体的数目比基因组加倍初期可能有所减少, 甚至恢复到基因组加倍之前的数目, 如玉米。反复发生的全基因组加倍使得植物基因组比大多数其他真核生物的基因组更加不稳定, 染色体重排的规模大、数目减少的幅度也大, 因此为理解染色体结构的演化提供了更多的信息。

比如, 禾本科物种的一个重要特征是, 共同的祖先物种在约1亿年前发生了一次全基因组加倍^[35], 生成一个四倍体祖先。祖先基因组的多倍化引发了以DNA序列急剧变化为特征的“基因组风暴”, 发生了大规模的染色体重排、基因倒位、基因丢失等^[32,70]。这为禾本科物种的进化提供了极为丰富的变异材料, 为自然选择出成功的物种提供了机会。一个明显的事实是在多倍化之后, 禾本科物种在地球上迅速扩增, 成为进化上最成功的类群之一。在多倍化之前, 禾本科植物祖先的基本染色体数目为7条。水稻的全基因组序列可以清晰地展示禾本科的全基因组多倍化事件(p事件), 所以经常用水稻作为参考基因组来分析其他禾本科物种。玉米的基因组比较复杂, 除了禾本科祖先发生的那次全基因组加倍事件外, 玉米又单独进行了一次全基因组加倍, 如果玉米染色体在两次全基因组加倍后都保留下来, 那么基本染色体数目将达到28条($7 \times 2 \times 2 = 28$), 而不是现在的10条基本染色体(表1)。二

表 1 植物基因组中染色体数目的减少^{a)}
Table 1 The reduction of chromosome number in plant genomes^{a)}

物种分类	物种(常用名字)	现存基本染色体数	全基因组二倍化/ 三倍化次数	古代基本染色体数目	染色体如果没有减少的数目
单子叶植物	二穗短柄草 <i>Brachypodium distachyon</i> (purple false brome)	5	1(ρ)	7 ρ	14
	水稻 <i>Oryza sativa</i> (rice)	12	1(ρ)	7 ρ	14
	高粱 <i>Sorghum bicolor</i> (sorghum)	10	1(ρ)	7 ρ	14
	谷子 <i>Setaria italica</i> (foxtail millet)	9	1(ρ)	7 ρ	14
	玉米 <i>Zea mays</i> (maize)	10	2(ρ, 1 WGD)	7 ρ	28
双子叶植物	番木瓜 <i>Carica papaya</i> (papaya)	12	1(γ)	7 γ	21
	黄瓜 <i>Cucumis sativus</i> (cucumber)	7	1(γ)	7 γ	21
	葡萄 <i>Vitis vinifera</i> (grape)	19	1(γ)	7 γ	21
	杨树 <i>Populus trichocarpa</i> (poplar)	19	2(γ, 1 WGD)	7 γ	42
	苹果 <i>Malus x domestica</i> (apple)	17	2(γ, 1 WGD)	7 γ	42
	琴叶拟南芥 <i>Arabidopsis lyrata</i> (lyrate rockcress)	8	3(γ, 2 WGD)	7 γ	84
	拟南芥 <i>Arabidopsis thaliana</i> (mouse-ear cress)	5	3(γ, 2 WGD)	7 γ	84
	大豆 <i>Glycine max</i> (soybean)	20	3(γ, 2 WGD)	7 γ	84
	白菜 <i>Brassica rapa</i> (Chinese cabbage)	10	4(γ, 2WGD, 1WGT)	7 γ	252
	甘蓝 <i>Brassica oleracea</i> (cabbage, etc.)	9	4(γ, 2WGD, 1WGT)	7 γ	252

a) ρ: 推测的禾本科祖先染色体二倍化事件; γ: 推测的双子叶祖先染色体三倍化事件; WGD: 全基因组二倍化事件; WGT: 全基因组三倍化事件

穗短柄草、高粱、谷子等其他禾本科物种都只经历了禾本科祖先的那次全基因组加倍，后来他们各自发生了不同的染色体融合事件，最终形成了现在的染色体数目(表1)。

葡萄的全基因组序列清晰地展示了一次六倍体形成事件(γ事件)，说明很多双子叶植物共有一个六倍体的祖先^[71]，这些双子叶涵盖了拟南芥、木瓜和杨树等主要双子叶植物^[72~74](表1)。值得注意的是，拟南芥是第一个被测得了全基因组序列的开花植物，其基因组较小，只有约125 Mbp。在拟南芥全基因组测序完成之前，几乎没有人预期到其祖先基因组会经历多次加倍。

在所有这些多倍化事件之前，这些被子植物的祖先有7条染色体^[75]，如果在经历了1次三倍化(γ)和2次二倍化(WGD)事件后，所有染色体都保留至今而没有发生融合，那么拟南芥基本染色体数目应达到84条(7×3×2=84)，而不是5条基本染色体。

(2) 动物染色体数量的多变性。鱼类的染色体大多为20~100条，60%的海水鱼类有48条染色体(2n)，21%的淡水鱼类有54条染色体^[76]。哺乳动物的染色体数目大多介于30~60条之间，有袋类二倍体染色体数目明显偏低，有蹄类和食肉类染色体数较高，灵长类具有较低的染色体数。人类、大猩猩和黑猩猩有着共同的

灵长类祖先, 在他们的祖先中, 染色体的基因结构和人类染色体构成相似^[77]. 人类2号染色体是染色体融合事件的产物, 其稳定性是通过抑制一个着丝粒实现的, 这导致人类(46条染色体)及其最密切相关的类人猿(48条染色体)之间存在染色体数目的差异^[77~79]. 染色体融合后, 2p着丝粒保持活跃, 而2q着丝粒失活, 但有着丝粒DNA的残留^[80~82]. 大多数哺乳动物都是以特定的染色体数目为特征, 但有时一个物种内二倍体数目的变化是由于涉及到近端着丝粒染色体的融合(罗伯逊易位)而形成多态性^[65]. 罗伯逊易位引成染色体融合在许多物种中都有发现, 包括小家鼠, 所有的二倍体染色体数目都在22~40条之间^[83]. 在奔原鼠中发现了24种不同的核型^[84], 在尼日利亚沙鼠中有20种核型^[85].

3 染色体数目变化与B染色体产生的模型

3.1 基因组测序为理解染色体重构提供了新机会

染色体数目减少主要是由于染色体发生融合造成的. 两条染色体之间可能发生端部与端部的合并, 在这一过程中, 一条染色体变成了端部或近端部着丝粒染色体, 另外一条染色体在其着丝粒附近发生断裂^[86], 此外, 一条染色体可以合并到另一条染色体的着丝粒附近^[86]. 然而这些推论所提出的机制仍然是不完善的, 不能很好解释这个复杂的动态过程.

多倍体在植物中很常见, 染色体重排多, 染色体数在加倍后减少, 为人们理解真核生物染色体数目的变化机制提供了思路. 植物基因组内常有很多共线性基因, 是多倍化的产物, 这有助于认识染色体重构过程. 基因组测序技术的发展以及海量基因组数据的产生, 为理解染色体重构提供了新机会.

3.2 一个以端粒为中心的模型

基于对多种植物的比较基因组学分析, 人们提出了植物和其他真核生物基因组重构的新模型, 以助于进一步阐明核型进化的分子动力学. 其机理是, 从染色体上去除两个端粒, 形成一条过渡型的双自由端染色体, 最终插入到另一条染色体上; 或者通过两条各自去除一个端粒的染色体的融合来产生更大的染色体(图2). 在这个过程中, 由两个端粒形成的小染色体丢失或被忽视, 指明了染色体数目减少的本质. 以端粒为中心的基因组重组决定了在真核生物进化过程中,

图2 染色体核型进化的模型. 红色圆圈代表端粒; 绿色圆圈代表着丝粒. A: 染色体花束期形成环状染色体, 一条染色体的两个端粒附近交叉, 产生具有两个自由端的主要染色体和一条可能在以后丢失的小染色体. 粉色叉代表一条染色体两个端粒附近发生交叉. B: 两条染色体在端粒区附近交叉, 产生一条大的染色体, 以及一条可能会丢失的小染色体. CEJ(chromosome end-end joining)代表染色体端端连接; def.(defunctionized)指示着丝粒失活, 失去功能. C: 两条染色体在末端交叉, 产生两条新的染色体. cross代表相互交叉易位.

Figure 2 Model of karyotype evolution. Red circles, telomeres; green circles, centromeres. A: A ring chromosome was formed during the chromosome bouquet stage, and two telomeres of a chromosome crossed, resulting in a main chromosome with two free ends and a small chromosome that may be lost in the future. Pink crosses show cross-overs. B: Two chromosomes crossed over near the telomeric regions to produce a large chromosome, and a small chromosome, which may be lost. "CEJ", showing chromosome end-end joining; "def.", showing defunctionized centromeres. C: Two chromosomes crossed over near one of two ends of each chromosome to produce two novel chromosomes. "cross", showing reciprocal translocation

特别是多倍化之后, 染色体数目的减少^[75], 强调了端粒在基因组整合中的重要作用.

上述模型所描述的现象应发生于细胞减数分裂过程中的同源染色体配对时期. 这一时期, 染色体花束结构^[87,88]的形成使端粒接近, 引起同源染色体配对和基因组重组. 染色体末端的物理邻近性^[85]使染色体能够识别同源染色体或部分同源染色体^[87], 最终引发同源染色体的配对和融合^[89,90]. 端粒的重复序列是有效正

确识别染色体的主要因素^[88]。在染色体聚集过程中, 染色体的两端可以彼此靠近, 促进环状结构染色体^[91,92]的形成。同时, 不同染色体的端部靠近^[93], 可能使它们发生交叉(尽管机率可能很小)。

在“染色体花束”期, 通过染色体内交叉产生的环状染色体^[91](ring chromosome, RC)可能形成一条过渡型的含有两个黏性末端(自由端)的染色体, 包含原染色体的主体部分, 以及一条包含两个端粒及基因很少的小染色体(图2A)。小染色体可能会丢失, 这对植物的适应性几乎没有影响, 但可能会导致人类出现严重的症状^[91]。而黏性端染色体会引起基因组不稳定, 其一端可能插入到另一条染色体的着丝粒区域(着丝粒区域的重复序列可能有助于插入的发生)。这种插入实际上是一种“侵入”过程, 会导致“被侵入”的染色体断裂。一条染色体的一个断裂端连接到另一条染色体的断裂端, 可能在双链断裂后通过一种经典的重组机制完成^[94]。染色体在空间上的接近有利于两个染色体另外的黏性自由端之间的连接, 从而造成入侵的染色体嵌套式插入“被侵入”的染色体中, 导致嵌套染色体融合(nested chromosome fusion, NCF), 或者更精确地说, 嵌套染色体融合是由产生的环形染色体介导形成的。“侵入”和“被侵入”的染色体往往是同源的, 也就是起源于一个共同祖先的染色体, 具有比无同源关系的染色体更大的DNA相似性。融合后的染色体有两个着丝粒, “侵入”染色体的着丝粒保留下来, 而“被侵入”染色体的着丝粒则失活, 这可能是为了保护插在“被侵入”染色体的着丝粒区的“侵入”染色体的基因密集区, 而自然选择使“被侵入”染色体的着丝粒的DNA不断丢失。

此外, 在“染色体花束”期, 两条不同染色体之间的交叉可以导致染色体臂的相互易位(reciprocal translocation of chromosome arms), 这是基因组重组产生新染色体的另一种机制。在染色体花束期, 靠近的两条染色体如在其末端发生交叉, 可能使染色体端端连接(chromosome end-end joining, CEJ)形成一个包含大部分DNA的主要染色体(这条双着丝粒染色体其中的一个着丝粒会失活), 和一个包含端粒的小染色体。这条小染色体可能会丢失(图2B), 导致染色体数目的减少。另外, 如果两条染色体在与端粒有一定距离的位置交叉, 可能形成两条新的染色体(图2C)。

在一些受到全基因组加倍影响的被子植物中, 这

个模型很好地解释了在全基因组加倍后染色体数目是如何减少的。嵌套染色体融合利用“侵入”染色体的端粒和“被侵入”染色体的着丝粒, 形成新的染色体。嵌套染色体融合和染色体端端连接融合, 都可能产生两个端粒构成的小染色体, 而且一般会丢失。这一端粒为中心的染色体重构模型充分考虑了端粒的作用。端粒特殊的帽式结构有助于保护染色体免受分裂或其他染色体融合^[95-97]。如果端粒未被移除, 嵌套染色体融合或染色体端端连接融合式染色体融合就不会发生^[86,98]。

以端粒为中心的染色体重构模型解释了单子叶植物小麦、二穗短柄草、谷子、水稻、玉米、高粱, 以及双子叶植物拟南芥、豆科、葫芦科等植物类群的核型进化。这一模型也可以解释脊椎动物^[99]和酵母^[100]的核型进化, 以及大多数植物是如何保存少量染色体数目的, 比如上文所涉及的人类2号染色体就是类人猿祖先两条染色体经CEJ式端端融合的产物。在此过程中可能产生了小染色体, 其在进化过程中的丢失导致人类染色体的数目减少了1对。

以前的研究虽然讨论了禾本科基因组重组, 提及了嵌套染色体融合, 但忽略了染色体端端连接融合, 没有明确端粒在染色体重构中的重要意义, 特别是没有注意到产生“黏性末端”或“自由端”染色体的重要性, 也没有考虑到小染色体的产生。本文提出的模型阐明, 正是由于端粒的去除, “自由端”染色体产生并侵入了其他染色体, 或者两条不同染色体端端融合, 产生小染色体并丢失, 导致了染色体数目的减少。

本文提出的模型在一定程度上获得了Hi-C数据上的支持。2017年, Mascher等人组装了高质量大麦的参考基因组, 利用Hi-C数据获得大麦细胞核染色体三维结构, 发现了细胞核间期的Rabl结构^[101]。在这种结构中, 染色体折叠形成环状染色体, 所有染色体的着丝粒和端粒聚集在细胞核相反的两极^[102], 即所有大麦染色体的端粒之间或者着丝粒之间在空间上相互接近, 端粒和端粒聚集在一起, 同时着丝粒和着丝粒聚集在一起, 染色体的端粒存在物理邻近性。

3.3 染色体数目的减少与B染色体产生的模型

上文提出的染色体数目变化的端粒中心模型, 初步解决了染色体数目减少和B染色体产生这两个问题, 该模型中产生的小染色体可能就是长期无法解释起源

原因的B染色体. 如果将该模型中产生的小染色体计算入内, 染色体的总数目并没有减少. 但由于小染色体的丢失或被忽视, 染色体数目就减少了, 即染色体数目减少必然伴随着B染色体的产生和丢失.

3.4 染色体数目的减少与B染色体产生的模型应用

(1) 禾本科植物染色体的演化路径. 通过水稻和高粱基因组比较分析发现, 有7次染色体断裂、倒位, 2次发生在水稻和高粱共同的多倍体祖先基因组, 5次发生在它们分化后的高粱中; 而水稻基因组自从与其他禾本科植物分开后, 其基因组中没有显示出较大的染色体结构变化, 表明其染色体核型与禾本科共同祖先非常相似. 在水稻12条染色体(Os1~12)中, 按基因组分析推断出五组重复染色体对应关系, 分别是: Os1:Os5; Os2:Os4, Os6; Os3:Os7, Os10; Os8:Os9; Os11:Os12^[98,103].

那么, 禾本科祖先在基因组加倍之前到底有几条染色体呢? 一个关键的问题是Os2是由加倍后的一组Os4(部分)与Os6的同源物融合形成, 还是Os2的同源物断裂产生现在的Os4(部分)和Os6(图3A). 同样地, Os3是由祖先的Os7与Os10的(部分)同源物经融合形成, 还是Os3同源物的裂变产生了染色体Os7和Os10(部分). Os2和Os3与其具有部分同源关系的染色体有交替性的同源关系, 分别对应着Os6-Os4-Os6和Os10-Os7-Os12-Os7.

香蕉基因组^[104]的序列提供了禾本科染色体核型演变的证据. 以水稻Os2, Os4和Os6为例, 如果Os4和Os6起源于香蕉和禾本科植物的共同祖先的不同染色体, 它们的同源片段应该更可能位于不同香蕉染色体中. 相比之下, 如果Os2同源物断裂为Os4和Os6, 那么Os4和Os6对应的香蕉同源片段更可能出现在相同染色体上. 然而虽然经历多次多倍化, 在香蕉基因组中, Os4和Os6的同源区几乎总是互相独立的, 而不是具有相关联的同源区^[75]. 比如, 香蕉7号染色体(Mu7)与Os6是同源染色体, 这两条染色体之间的共线性同源关系达到了接近整条染色体长度, 但是Os7却没有与Os4的同源性. 同样地, 香蕉Mu6, Mu8和Mu11都与Os4同源, 共线性同源关系也接近达到整条染色体, 但是与Os6只有有限的同源信息. 因此, 最简单的解释是: Os4与Os6起源于不同的祖先染色体, 而它们在禾本科祖先

多倍化中产生的部分同源染色体融合形成了Os2.

与香蕉染色体比较基因组学分析表明, Os3是由于Os7和Os10的部分同源物融合形成, 并且又融合了一个来自Os12的小片段. 实际上, Os7与Mu6, Mu7和Mu9同源, Os10与Mu1和Mu11具有显著同源性, 表明Os7与Os10分别来源于禾本科祖先中两条不同的染色体.

简言之, 比较水稻和香蕉的染色体很大程度上支持本文的推断: Os2和Os3分别由不同祖先的染色体融合产生(图3A). 而且它们的形成是经嵌套染色体融合, 随后Os3染色体中DNA发生倒位, 解释了其长臂末端显著的重叠片段富集是由于部分着丝粒区重置造成的.

本文在上述推断中分析了Os2与Os3分别由两条祖先染色体融合形成, 结合上文中水稻染色体同源性的分组, 可以推断禾本科植物共同祖先单倍型基因组有7条染色体(GAK1~7), 而基因组加倍之后, 单倍型有14条染色体(这里基于一定的证据^[105], 假设祖先多倍化事件为异源多倍化). 禾本科植物祖先染色体的核型与现存水稻染色体的对应关系可表示为: GAK1: Os1, Os5; GAK2: Os6; GAK3: Os4; GAK4: Os7; GAK5: Os10; GAK6: Os8, Os9; GAK7: Os11, Os12. 两次嵌套染色体融合使水稻基本染色体数目变成12条(p1~12), 这也代表了禾本科多倍化又经染色体合并后的共同祖先的植物核型.

嵌套染色体融合完全解释了在禾本科植物祖先加倍后, 14条染色体减少到10条的高粱和5条的二穗短柄草的过程. 例如, 由于两次嵌套染色体融合事件, GAK3, GAK6和GAK7合并产生二穗短柄草的1号染色体. 高粱染色体S3与S8的祖先禾本科植物的染色体(P3和P8)合并产生了玉米3号染色体^[75]. 同时也可以解释玉米特有多倍化后大部分的染色体融合(80%), 而染色体端端连接融合占了20%. 玉米和高粱同源基因点图中, 共线性断裂的发生暗示祖先黍亚科(Panicoidae)染色体P1和P10经换臂式DNA易位, 分别形成了玉米染色体M9和M5的一部分. 换臂式DNA易位也解释了玉米染色体M1和M10的产生^[75].

总的来说, 染色体的融合与相互易位解释了观察到的染色体生成的过程和数目的减少. 染色体融合过程中产生的小染色体可能就是禾本科的B染色体, 染色体数目的减少伴随着禾本科B染色体的产生和丢失.

图 3 染色体的核型进化。A: 水稻2号染色体(Os2)的核型进化。水稻染色体Os2, Os4和Os6是同源染色体。祖先染色体GAK2和GAK3发生NCF形成染色体p(Os2)。B: 麦类和二穗短柄草部分染色体的核型进化。麦类染色体T4和T5形成涉及一次NCF、一次CEJ和一次相互易位, 大麦(Hv)作为代表, 并展示了其过程。二穗短柄草染色体Bd4形成涉及2次NCF事件, 并展示了其进化过程。C: 拟南芥部分染色体的核型进化。拟南芥At2和At3以及At4和At5形成过程类似, 都发生了一次相互易位和一次CEJ, 并展示了其进化过程。注: 一条染色体上的双箭头代表发生交叉形成环状染色体。NCF(nested chromosome fusion)代表嵌套染色体融合; CEJ(chromosome end-end joining)代表染色体端端连接; def.(defunctionized)代表失活, 失去功能; cross代表相互易位; IV(inversion)代表倒位; Os代表水稻(*Oryza sativa*); Hv代表大麦(*Hordeum vulgare*); Bd代表二穗短柄草(*Brachypodium distachyon*); At代表拟南芥(*Arabidopsis thaliana*); Al代表琴叶拟南芥(*Arabidopsis lyrata*)

Figure 3 Karyotype evolution of chromosome. A: Karyotype evolution of rice chromosome 2. OS2, OS4 and OS6 are homologous. A fusion of ancestral chromosomes GAK2 and GAK3 formed p(Os2) by nested chromosome fusion (NCF). B: Karyotype evolution of some chromosomes in Triticeae and *Brachypodium distachyon*. The formation of Triticeae chromosomes T4 and T5 involved one NCF, one CEJ and one reciprocal translocation. Barley (Hv) is the representative of Triticeae, which shows the process. The formation of chromosome Bd4 involved two NCF events and showed its evolutionary process. C: Karyotype evolution of *Arabidopsis thaliana* chromosomes. In *Arabidopsis thaliana*, At2 and At3, as well as At4 and At5, are formed in a similar process, with one reciprocal translocation and one CEJ, showing the evolutionary process. Note: The double arrows on one chromosome indicate that a cross-over has occurred to form a ring chromosome. NCF, nested chromosome fusion; CEJ, chromosome end-end joining; def., defunctionized; cross, reciprocal translocation; IV, inversion; Os, *Oryza sativa*; Hv, *Hordeum vulgare*; Bd, *Brachypodium distachyon*; At, *Arabidopsis thaliana*; Al, *Arabidopsis lyrata*

(2) 麦类染色体的演化路径。相对于水稻来说, 同属于早熟禾亚科的麦类和二穗短柄草, 系统进化关系比较接近。麦类的基本染色体数目为7条, 二穗短柄草的基本染色体数目为5条, 那么两者的核型进化过程是不是比较一致呢?

按照水稻和早熟禾亚科(麦类和二穗短柄草)基因共线性, 分析直系同源染色体(或直系同源染色体片段)的对应关系发现, 麦类染色体T3和T6分别对应水稻Os1和Os2(p1和p2, 以下类似), 保持了祖先染色体的结构。而其他染色体都发生了融合性重构, 生成了

新的染色体: T1是Os10和Os5经嵌套染色体融合形成, T2是由Os4和Os7经嵌套染色体融合形成, T7是由Os6和Os8经嵌套染色体融合形成。T4和T5的进化过程比较复杂。首先Os11和Os3经嵌套染色体融合, 形成中间产物Os11/3和一条小染色体。另外, Os12和Os9经端端连接的染色体融合, 形成中间产物Os12/9和一条小染色体。两个中间产物Os11/3和Os12/9经相互易位形成T4和T5(图3B)。麦类染色体在进化过程中, 涉及4次嵌套染色体融合事件, 同时可能形成了4条小染色体; 还涉及到2条染色体端端融合形成1条新的染色体和1条

小染色体. 在麦类进化过程形成的5条小染色体有可能在进化过程中丢失, 最终导致由12条染色体变成现存麦类的7条基本染色体^[106].

二穗短柄草2号染色体Bd2是水稻染色体Os5和Os1经嵌套染色体融合形成的. 而二穗短柄草1号染色体Bd1的形成涉及到2次嵌套染色体融合事件, 水稻Os3和Os7发生嵌套染色体融合形成中间产物Os3/7, Os6染色体两个端粒交叉形成的主要染色体侵入到中间产物Os3/7的着丝粒区域, 发生嵌套染色体融合, 最终形成Bd1. 染色体Bd3和Bd4的形成过程和Bd1类似, 发生了2次嵌套染色体融合事件, Os8, Os2和Os10这3条染色体合并产生Bd3, 而Bd4则是由Os9, Os12和Os11合并产生(图3B). Bd5的核型大尺度上变化不大, 与Os4基本相同. 在二穗短柄草染色体形成过程中, 发生了7次嵌套染色体融合, 产生了7条小染色体, 损失的这些小染色体导致染色体数目从12条减少到5条^[106].

通过共线性和点阵图分析发现, 虽然二穗短柄草是模式生物, 并且和麦类亲缘关系很近, 但是麦类和二穗短柄草的核型进化是完全不同的、独立的途径. 这意味着这两类物种的进化不是单一的事件, 比如共享交叉或融合形成的中间世系或现存的染色体, 而是两者经历了不同的进化途径, 并且进化过程没有任何的相同过程.

此外, 在二穗短柄草中, 入侵染色体和被侵入的染色体是同源的, 起源于同一个共同祖先. 也就是说, 相较于随机染色体, 其具有更广泛的DNA水平上的对应性. 而麦类入侵染色体和被侵入的染色体是非同源的. 这表明两类物种经过了不同的进化途径, 在进化过程中有共同的祖先, 即以水稻作为两者的共同祖先.

一个物种的1号或排号靠前的染色体大体上都是多次染色体融合产生的, 而排号靠后的染色体更可能是完整的古老染色体. 二穗短柄草1~4号染色体都是融合产生, 5号染色体却保留了完整的古老染色体的核型.

(3) 拟南芥染色体的演化路径. 拟南芥(*Arabidopsis thaliana*)基本染色体数目是5条, 它的祖先物种琴叶拟南芥(*Arabidopsis lyrata*)基本染色体数目为8条. 拟南芥在染色体重构过程中发生了3次染色体融合事件.

拟南芥祖先3号染色体(A13)和5号染色体(A15)发生相互易位^[107], 产生了现存的拟南芥3号染色体(At3)和中间染色体A13/5. 中间染色体A13/5和拟南芥祖先4

号染色体(A14)通过端端连接融合, 产生了拟南芥的2号染色体(At2)(图3C). A16与A17发生相互易位产生了拟南芥4号染色体(At4)和一个中间染色体, A18和中间染色体通过端端连接融合, 产生了拟南芥的5号染色体(At5)(图3C). A11与A12通过端端连接融合产生了拟南芥的1号染色体(At1). 拟南芥发生的这3次端端连接融合事件是通过两条染色体之间的相互易位实现的, 且每次都会产生一条大的染色体和一条小染色体, 由此造成了拟南芥的染色体数目从8条减到现在的5条^[75].

(4) 豆科植物染色体的演化路径. 豆科作物是人类粮、油等食品的重要来源之一, 具有重要的经济价值, 包括大豆、花生等. 多倍体基因组为花生和其他豆科植物染色体的进化提供了依据.

在花生基因组中保留着约5900万年前的豆科植物共同四倍体祖先^[108](legume-common tetraploidy, LCT)和约1.3亿年前双子叶植物共同六倍体祖先(core-eudicot-common hexaploidy, ECH)的痕迹. 花生核型很大程度上独立于其他豆科植物.

把葡萄和豆科植物基因组进行共线性比较分析, 确定了5个独立的染色体融合事件, 包括3次嵌套染色体融合和2次端端连接融合, 在LCT之前产生了16个基本的豆科祖先染色体. 通过自下而上的方法发现, 5条常见的豆类染色体和11个染色体块, 在很大程度上保存在不同的豆科植物中, 确定了16条LCT后的祖先染色体. 这意味着在LCT加倍后, 原始染色体数目在基因组重新定位后恢复, 这与玉米相类似^[109]. 与使用普通菜豆基因重建的16条LCT后染色体(称为Lu)的比较显示, 花生祖先染色体A1, A3, A4, A5, A6和A7由来自Lu染色体的片段组成, 且6次融合事件导致染色体数量减少. 染色体A2, A8, A9和A10通过Lu染色体的两个交叉产生. 在从花生A基因组分裂后, 在花生B基因组中杂交产生其特异性染色体7和8^[110].

4 展望

近年来, 基因组分析的最新进展加深了人们对染色体来源和组成的理解. 以上内容表明, 植物基因组测序为人们提供了一个关于染色体进化的基本答案. 预计新的测序技术将解决经典染色体生物学中的一些困惑, 但要解答所有的问题, 需要更多的细胞学和生理学上的证据. 本团队将致力于更多的多倍体物种在某

些属(类)中存在的多倍体变化的研究, 期望获得足够的数据进一步解析染色体的结构变化. 同时将构建有关植物、动物和真菌等核型进化的在线数据库, 动态地重构地球上生物染色体在过去数亿年来的演化过程. 这一工作已在多种植物类群中展开, 并取得了一定

的成果, 在动物和真菌中的相关研究也已初步开展. 当然, 这一数据库是一个长期的项目, 需要不断地更新和完善. 本团队希望进一步研究染色体数目的变化和B染色体产生的生物学机制, 并在未来的工作中对B染色体在细胞学和生理学上开展进一步研究.

参考文献

- Eichler E E, Sankoff D. Structural dynamics of eukaryotic chromosome evolution. *Science*, 2003, 301: 793–797
- Bougourd S M, Jones R N. B chromosomes: A physiological enigma. *New Phytol*, 2010, 137: 43–54
- Ueshima N. Supernumerary chromosomes in the human bed bug, *Cimex lectularius* linn. (*Cimicidae:Hemiptera*). *Chromosoma*, 1967, 20: 311–331
- Marques A, Klemme S, Houben A. Evolution of plant B chromosome enriched sequences. *Genes*, 2018, 9: 515
- Beukeboom L W. Bewildering Bs: An impression of the 1st b-chromosome conference. *Heredity*, 1994, 73: 328–336
- Fu S, Lv Z, Gao Z, et al. *De novo* centromere formation on a chromosome fragment in maize. *Proc Natl Acad Sci USA*, 2013, 110: 6033–6036
- Houben A, Banaei-Moghaddam A M, Klemme S, et al. Evolution and biology of supernumerary B chromosomes. *Cell Mol Life Sci*, 2014, 71: 467–478
- D'Ambrosio U, Alonso-Lifante M P, Barros K, et al. B-chrom: A database on b-chromosomes of plants, animals and fungi. *New Phytol*, 2017, 216: 635–642
- Camacho J P M, Sharbel T F, Beukeboom L W. B-chromosome evolution. *Phil Trans R Soc Lond B*, 2000, 355: 163–178
- Jones R N. B-chromosome drive. *Am Natist*, 1991, 137: 430–442
- Banaei-Moghaddam A M, Martis M M, Macas J, et al. Genes on B chromosomes: Old questions revisited with new tools. *Biochim Biophys Acta*, 2015, 1849: 64–70
- Yoshida K, Terai Y, Mizoiri S, et al. B chromosomes have a functional effect on female sex determination in lake victoria cichlid fishes. *PLoS Genet*, 2011, 7: e1002203
- Ploskaya-Chaibi M, M. Voitovich A, V. Novitsky R, et al. B-chromosome and V-shaped spot asymmetry in the common frog (*Rana temporaria* L.) populations. *Comptes Rendus Biol*, 2015, 338: 161–168
- Randolph L F. Genetic characteristics of the B chromosomes in maize. *Genetics*, 1941, 26: 608–631
- Ruiz-Ruano F J, Cabrero J, López-León M D, et al. Satellite DNA content illuminates the ancestry of a supernumerary (B) chromosome. *Chromosoma*, 2017, 126: 487–500
- Martis M M, Klemme S, Banaei-Moghaddam A M, et al. Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proc Natl Acad Sci USA*, 2012, 109: 13343–13346
- Silva D M Z A, Pansonato-Alves J C, Utsunomia R, et al. Delimiting the origin of a B chromosome by fish mapping, chromosome painting and DNA sequence analysis in *astyanax paranae* (teleostei, characiformes). *PLoS ONE*, 2014, 9: e94896
- Valente G T, Conte M A, Fantinatti B E A, et al. Origin and evolution of B chromosomes in the cichlid fish *astatotilapia latifasciata* based on integrated genomic analyses. *Mol Biol Evol*, 2014, 31: 2061–2072
- Croll D, McDonald B A. The accessory genome as a cradle for adaptive evolution in pathogens. *PLoS Pathog*, 2012, 8: e1002608
- Navarro-Domínguez B, Ruiz-Ruano F J, Cabrero J, et al. Protein-coding genes in B chromosomes of the grasshopper *cyreprocnemis plorans*. *Sci Rep*, 2017, 7: 45200
- Charlesworth B, Sniegowski P, Stephan W. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature*, 1994, 371: 215–220
- Shapiro J A, von Sternberg R. Why repetitive DNA is essential to genome function. *Biol Rev*, 2005, 80: 227–250
- Wicker T, Sabot F, Hua-Van A, et al. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet*, 2007, 8: 973–982
- Rebollo R, Romanish M T, Mager D L. Transposable elements: An abundant and natural source of regulatory sequences for host genes. *Annu Rev Genet*, 2012, 46: 21–42
- Valente G T, Nakajima R T, Fantinatti B E A, et al. B chromosomes: From cytogenetics to systems biology. *Chromosoma*, 2017, 126: 73–81

- 26 Ruban A, Schmutzer T, Scholz U, et al. How next-generation sequencing has aided our understanding of the sequence composition and origin of B chromosomes. [Genes](#), 2017, 8: 294
- 27 Makunin A I, Dementyeva P V, Graphodatsky A S, et al. Genes on B chromosomes of vertebrates. [Mol Cytogenet](#), 2014, 7: 99
- 28 Furlong R F, Holland P W H. Polyploidy in vertebrate ancestry: Ohno and beyond. [Biol J Linnean Soc](#), 2004, 82: 425–430
- 29 Kellis M, Birren B W, Lander E S. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. [Nature](#), 2004, 428: 617–624
- 30 Aury J M, Jaillon O, Duret L, et al. Global trends of whole-genome duplications revealed by the ciliate paramecium tetraurelia. [Nature](#), 2006, 444: 171–178
- 31 Masterson J. Stomatal size in fossil plants: Evidence for polyploidy in majority of angiosperms. [Science](#), 1994, 264: 421–424
- 32 Paterson A H, Bowers J E, Chapman B A. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. [Proc Natl Acad Sci USA](#), 2004, 101: 9903–9908
- 33 Jiao Y, Wickett N J, Ayyampalayam S, et al. Ancestral polyploidy in seed plants and angiosperms. [Nature](#), 2011, 473: 97–100
- 34 Soltis D E, Bell C D, Kim S, et al. Origin and early evolution of angiosperms. [Ann New York Acad Sci](#), 2008, 1133: 3–25
- 35 Wang X, Wang J, Jin D, et al. Genome alignment spanning major poaceae lineages reveals heterogeneous evolutionary rates and alters inferred dates for key evolutionary events. [Mol Plant](#), 2015, 8: 885–898
- 36 Ramsey J, Schemske D W. Neopolyploidy in flowering plants. [Annu Rev Ecol Syst](#), 2002, 33: 589–639
- 37 Levin D A. Polyploidy and novelty in flowering plants. [Am Natist](#), 1983, 122: 1–25
- 38 Mayrose I, Barker M S, Otto S P. Probabilistic models of chromosome number evolution and the inference of polyploidy. [Systat Biol](#), 2010, 59: 132–144
- 39 Mable B K, Alexandrou M A, Taylor M I. Genome duplication in amphibians and fish: An extended synthesis. [J Zool](#), 2011, 284: 151–182
- 40 Otto S P. The evolutionary consequences of polyploidy. [Cell](#), 2007, 131: 452–462
- 41 Comai L. The advantages and disadvantages of being polyploid. [Nat Rev Genet](#), 2005, 6: 836–846
- 42 Bridges C B. The bar “gene” a duplication. [Science](#), 1936, 83: 210–211
- 43 Stephens S G. Possible significance of duplication in evolution. [Adv Genet](#), 1951, 4: 247–265
- 44 Ohno S. Evolution by gene duplication. [Am J Hum Genet](#), 1971, 23: 541
- 45 Amores A, Force A, Yan Y L, et al. Zebrafish hox clusters and vertebrate genome evolution. [Science](#), 1998, 282: 1711–1714
- 46 Taylor J S, Braasch I, Frickey T, et al. Genome duplication, a trait shared by 22,000 species of ray-finned fish. [Genome Res](#), 2003, 13: 382–390
- 47 Meyer A, Van de Peer Y. From 2R to 3R: Evidence for a fish-specific genome duplication (FSGD). [BioEssays](#), 2005, 27: 937–945
- 48 Taylor J S, Van de Peer Y, Braasch I, et al. Comparative genomics provides evidence for an ancient genome duplication event in fish. [Phil Trans R Soc Lond B](#), 2001, 356: 1661–1679
- 49 Christoffels A, Koh E G L, Chia J M, et al. Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. [Mol Biol Evol](#), 2004, 21: 1146–1151
- 50 Thomas-Chollier M, Ledent V. Comparative phylogenomic analyses of teleost fish hox gene clusters: Lessons from the cichlid fish *Astatotilapia burtoni*: Comment. [BMC Genomics](#), 2008, 9: 35
- 51 Vandepoele K, De Vos W, Taylor J S, et al. Major events in the genome evolution of vertebrates: Paraneome age and size differ considerably between ray-finned fishes and land vertebrates. [Proc Natl Acad Sci USA](#), 2004, 101: 1638–1643
- 52 Tsigenopoulos C S, Ráb P, Naran D, et al. Multiple origins of polyploidy in the phylogeny of southern African barbs (Cyprinidae) as inferred from mtDNA markers. [Heredity](#), 2002, 88: 466–473
- 53 Luo J, Gao Y, Ma W, et al. Tempo and mode of recurrent polyploidization in the *Carassius auratus* species complex (Cypriniformes, Cyprinidae). [Heredity](#), 2014, 112: 415–427
- 54 Yang L, Sado T, Vincent Hirt M, et al. Phylogeny and polyploidy: Resolving the classification of cyprinine fishes (Teleostei: Cypriniformes). [Mol Phylogenets Evol](#), 2015, 85: 97–116
- 55 Wang X, Gan X, Li J, et al. Cyprininae phylogeny revealed independent origins of the tibetan plateau endemic polyploid cyprinids and their diversifications related to the neogene uplift of the plateau. [Sci China Life Sci](#), 2016, 59: 1149–1165
- 56 Macqueen D J, Johnston I A. A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. [Proc R Soc B](#), 2014, 281: 20132881

- 57 Xu P, Xu J, Liu G, et al. The allotetraploid origin and asymmetrical genome evolution of the common carp *Cyprinus carpio*. *Nat Commun*, 2019, 10: 4625
- 58 Nakatani Y, Takeda H, Kohara Y, et al. Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Res*, 2007, 17: 1254–1265
- 59 Guyomard R, Boussaha M, Krieg F, et al. A synthetic rainbow trout linkage map provides new insights into the salmonid whole genome duplication and the conservation of synteny among teleosts. *BMC Genet*, 2012, 13: 15
- 60 Wang Y, Lu Y, Zhang Y, et al. The draft genome of the grass carp (*Ctenopharyngodon idellus*) provides insights into its evolution and vegetarian adaptation. *Nat Genet*, 2015, 47: 625–631
- 61 Xu P, Zhang X, Wang X, et al. Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. *Nat Genet*, 2014, 46: 1212–1219
- 62 Chen Z, Omori Y, Koren S, et al. *De novo* assembly of the goldfish (*Carassius auratus*) genome and the evolution of genes after whole-genome duplication. *Sci Adv*, 2019, 5: eaav0547
- 63 Chalopin D, Volff J N. Analysis of the spotted gar genome suggests absence of causative link between ancestral genome duplication and transposable element diversification in teleost fish. *J Exp Zool (Mol Dev Evol)*, 2017, 328: 629–637
- 64 Kapusta A, Suh A, Feschotte C. Dynamics of genome size evolution in birds and mammals. *Proc Natl Acad Sci USA*, 2017, 114: E1460–E1469
- 65 Graphodatsky A S, Trifonov V A, Stanyon R. The genome diversity and karyotype evolution of mammals. *Mol Cytogenet*, 2011, 4: 22
- 66 Moriyama Y, Koshiba-Takeuchi K. Significance of whole-genome duplications on the emergence of evolutionary novelties. *Briefings Funct Genomics*, 2018, 17: 329–338
- 67 Session A M, Uno Y, Kwon T, et al. Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature*, 2016, 538: 336–343
- 68 Kobel H R, Du Pasquier L. Genetics of polyploid *Xenopus*. *Trends Genets*, 1986, 2: 310–315
- 69 Wendel J F. Genome evolution in polyploids. *Plant Mol Biol*, 2000, 42: 225–249
- 70 Soltis P S, Soltis D E. Polyploidy and Genome Evolution. Heidelberg: Springer, 2012
- 71 Jaillon O, Aury J M, Noel B, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, 2007, 449: 463–467
- 72 Arabidopsis Genome I. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, 2000, 408: 796–815
- 73 Ming R, Hou S, Feng Y, et al. The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature*, 2008, 452: 991–996
- 74 Tuskan G A, Difazio S, Jansson S, et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. AAA Gray). *Science*, 2006, 313: 1596–1604
- 75 Wang X, Jin D, Wang Z, et al. Telomere-centric genome repatterning determines recurring chromosome number reductions during the evolution of eukaryotes. *New Phytol*, 2015, 205: 378–389
- 76 Nirchio M, Rossi A R, Foresti F, et al. Chromosome evolution in fishes: A new challenging proposal from neotropical species. *Neotrop Ichthyol*, 2014, 12: 761–770
- 77 Yunis J J, Prakash O. The origin of man: A chromosomal pictorial legacy. *Science*, 1982, 215: 1525–1530
- 78 Lejeune J, Dutrillaux B, Rethoré M O, et al. Comparaison de la structure fine des chromatides d’Homo sapiens et de Pan troglodytes. *Chromosoma*, 1973, 43: 423–444
- 79 Ijdo J W, Baldini A, Ward D C, et al. Origin of human chromosome 2: An ancestral telomere-telomere fusion. *Proc Natl Acad Sci USA*, 1991, 88: 9051–9055
- 80 Avarello R, Pedicini A, Caiulo A, et al. Evidence for an ancestral alphoid domain on the long arm of human chromosome 2. *Hum Genet*, 1992, 89: 247–249
- 81 Baldini A, Ried T, Shridhar V, et al. An alphoid DNA sequence conserved in all human and great ape chromosomes: Evidence for ancient centromeric sequences at human chromosomal regions 2q21 and 9q13. *Hum Genet*, 1993, 90: 577–583
- 82 Chiatante G, Giannuzzi G, Calabrese F M, et al. Centromere destiny in dicentric chromosomes: New insights from the evolution of human chromosome 2 ancestral centromeric region. *Mol Biol Evol*, 2017, 34: 1669–1681
- 83 Capanna E. Robertsonian numerical variation in animal speciation: *Mus musculus*, an emblematic model. *Prog Clin Biol Res*, 1982, 96: 155–177
- 84 Fagundes V, Vianna-Morgante A M, Yonenaga-Yassuda Y. Telomeric sequences localization and G-banding patterns in the identification of a polymorphic chromosomal rearrangement in the rodent *Akodon cursor* (2n=14, 15 and 16). *Chromosome Res*, 1997, 5: 228–232

- 85 Volobouev V, Vogt N, Viegas-Péguignot E, et al. Characterization and chromosomal location of two repeated DNAs in three *Gerbillus* species. [Chromosoma](#), 1995, 104: 252–259
- 86 Schubert I, Lysak M A. Interpretation of karyotype evolution should consider chromosome structural constraints. [Trends Genets](#), 2011, 27: 207–216
- 87 Ding D Q, Yamamoto A, Haraguchi T, et al. Dynamics of homologous chromosome pairing during meiotic prophase in fission yeast. [Dev Cell](#), 2004, 6: 329–341
- 88 Bozza C G, Pawlowski W P. The cytogenetics of homologous chromosome pairing in meiosis in plants. [Cytogenet Genome Res](#), 2008, 120: 313–319
- 89 Niwa O, Shimanuki M, Miki F. Telomere-led bouquet formation facilitates homologous chromosome pairing and restricts ectopic interaction in fission yeast meiosis. [EMBO J](#), 2014, 19: 3831–3840
- 90 Scherthan H. A bouquet makes ends meet. [Nat Rev Mol Cell Biol](#), 2001, 2: 621–627
- 91 Sodré C P, Guilherme R S, Meloni V F A, et al. Ring chromosome instability evaluation in six patients with autosomal rings. [Genet Mol Res](#), 2010, 9: 134–143
- 92 McClintock B. The behavior in successive nuclear divisions of a chromosome broken at meiosis. [Proc Natl Acad Sci USA](#), 1939, 25: 405–416
- 93 Barzel A, Kupiec M. Finding a match: How do homologous sequences get together for recombination? [Nat Rev Genet](#), 2008, 9: 27–37
- 94 Barton N H. Mutation and the evolution of recombination. [Phil Trans R Soc B](#), 2010, 365: 1281–1294
- 95 Ferreira M G, Miller K M, Cooper J P. Indecent exposure: When telomeres become uncapped. [Mol Cell](#), 2004, 13: 7–18
- 96 Lundblad V. Molecular biology: Telomeres keep on rappin. [Science](#), 2000, 288: 2141–2142
- 97 Lundblad V. Telomeres: A tale of ends. [Nature](#), 2000, 403: 149–151
- 98 Murat F, Xu J H, Tannier E, et al. Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. [Genome Res](#), 2010, 20: 1545–1557
- 99 Ijdo J W, Wells R A, Baldini A, et al. Improved telomere detection using a telomere repeat probe (TTAGGG)_n generated by PCR. [Nucl Acids Res](#), 1991, 17: 4780
- 100 Gordon J L, Byrne K P, Wolfe K H. Mechanisms of chromosome number evolution in yeast. [PLoS Genet](#), 2011, 7: e1002190
- 101 Tiang C L, He Y, Pawlowski W P. Chromosome organization and dynamics during interphase, mitosis, and meiosis in plants. [Plant Physiol](#), 2012, 158: 26–34
- 102 Mascher M, Gundlach H, Himmelbach A, et al. A chromosome conformation capture ordered sequence of the barley genome. [Nature](#), 2017, 544: 427–433
- 103 Salse J, Bolot S, Throude M, et al. Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution. [Plant Cell](#), 2008, 20: 11–24
- 104 D'Hont A, Denoeud F, Aury J M, et al. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. [Nature](#), 2012, 488: 213–217
- 105 Murat F, Zhang R, Guizard S, et al. Shared subgenome dominance following polyploidization explains grass genome evolutionary plasticity from a seven protochromosome ancestor with 16k protogenes. [Genome Biol Evol](#), 2014, 6: 12–33
- 106 Wang Z, Wang J, Pan Y, et al. Reconstruction of evolutionary trajectories of chromosomes unraveled independent genomic repatterning between triticeae and brachypodium. [BMC Genomics](#), 2019, 20: 180
- 107 Lysak M A, Berr A, Pecinka A, et al. Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related brassicaceae species. [Proc Natl Acad Sci USA](#), 2006, 103: 5224–5229
- 108 Bowers J E, Chapman B A, Rong J, et al. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. [Nature](#), 2003, 422: 433–438
- 109 Schnable J C, Springer N M, Freeling M. Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. [Proc Natl Acad Sci USA](#), 2011, 108: 4069–4074
- 110 Zhuang W, Chen H, Yang M, et al. The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. [Nat Genet](#), 2019, 51: 865–876

Evolutionary genomics model of chromosome number reduction and B chromosome production

WANG ZhenYi¹ & WANG XiYin^{1,2}

1 Center for Genomics and Computational Biology, School of Life Sciences, North China University of Science and Technology, Tangshan 063210, China;

2 College of Genomics and Big Data Biology, Chengdu University of TCM, Chengdu 610075, China

In the long-term evolution, the number of chromosomes shows dynamic changes. There are still many mysteries about eukaryote chromosome evolution, one is how to maintain a small chromosome number, the other is how to produce B chromosome. The development of whole genome sequencing provides a new opportunity to understand chromosome number changes, especially after repeated polyploidy events. Through comparative genomics analysis, we propose an evolutionary model of chromosome number change and B chromosome production, and we propose that the reduction of chromosome number is mainly due to chromosome fusion, which shows that the loss of telomere and the formation of B chromosome play an important role in genome repatterning and chromosome number reduction.

evolutionary genomics, B chromosomes, chromosome number reduction, polyploidy, whole-genome duplication, karyotype evolution

doi: [10.1360/SSV-2020-0042](https://doi.org/10.1360/SSV-2020-0042)