张　岚,袁　敏,王振怡,等. 无油樟与葡萄、拟南芥、水稻基因组的多倍化及共线性分析[J]. 江苏农业科学,2018,46(8):45 - 49.
doi:10.15889/j.issn.1002 - 1302.2018.08.011

无油樟与葡萄、拟南芥、水稻基因组的多倍化及共线性分析

张　岚, 袁　敏, 王振怡, 潘玉欣
(华北理工大学生命科学学院/华北理工大学基因组学与计算生物学研究中心,河北唐山 063210)

摘要:以无油樟、葡萄、拟南芥和水稻为研究对象,基于比较基因组学的研究方法,对物种基因组内和基因组间进行同源结构和共线性分析。结果表明,无油樟与葡萄基因组之间的同源共线性片段最长,葡萄基因组保留的共线性基因数最多(50.9%),无油樟与葡萄、水稻基因组之间保留的无油樟基因数最多(30.6%)。同时构建了多物种基因组联合比对图谱,明确基因组经历的多倍化事件对单、双子叶植物同源共线性基因的影响,为探讨单、双子叶植物的共同祖先以及被子植物的进化过程提供了重要的参考依据。

关键词:无油樟;葡萄;多倍化;同源性;共线性
中图分类号: Q941^{+}.2;S188　　**文献标志码:** A　　**文章编号:**1002 - 1302(2018)08 - 0045 - 05

全基因组加倍,又称伴随基因丢失的多倍化,长期以来已经成为动物、真菌和其他生物,尤其是植物基因组进化的一个重要驱动力[1]。随着全基因组加倍的进行,基因组结构常常会发生明显改变,如染色体重组、基因倒位、基因丢失等,而多倍化产生的大量重复基因成为遗传创新的重要材料来源[2]。被子植物是迄今为止陆地植物的最大族群,有超过 300 000 种生物,大多数开花植物的进化过程均受到多倍化影响[3]。研究表明,在被子植物的早期进化过程中甚至存在古多倍化事件。无油樟是纵观被子植物历史的一个重要的参照,是已测序的开花植物中最古老的物种之一,基因组结构和种系基因组学分析表明,被子植物的祖先是多倍体,祖先基因在被子植物生物学中起着关键作用[4-6]。双子叶植物的共同祖先有 7 条染色体,在 13 000 万年前经历了双子叶植物共有的三倍乘(γ)事件后,又经历染色体的融合等过程,葡萄基因组形成现有的 19 条染色体[7-8]。作为首先被测序完成的开花植物,拟南芥在和葡萄一样经历了双子叶植物共同祖先的三倍乘(γ)事件后,又发生了最近的 2 次二倍乘(α,β)事件[9-12]。水稻是单子叶禾本科植物中重要的模式植物,基因组拼接较为完整,其结构也最接近禾本科植物基因组。水稻基因组在约 1 亿年前经历了禾本科植物共有的全基因组加倍事件,而后又先后发生了 2 次自己独立的全基因组二倍乘事件[13-14]。这些多倍化产生的重复基因为大范围分析基因进化和功能提供了有价值的参考。

本研究对无油樟、葡萄、拟南芥、水稻基因组进行比较基因组学分析,通过 4 个物种基因组内和基因组间的同源共线性分析,明确多倍化过程对重复基因造成的影响,解析基因组在进化过程中发生的变化,对于探寻单、双子叶植物的共同起源以及研究被子植物的进化过程具有重要的意义。

1　材料与方法

1.1　全基因组序列数据

葡萄(*Vitis vinifera*,简称 *Vv*)的全基因组序列来自 Phytozome(https://phytozome.jgi.doe.gov),无油樟(*Amborella trichopoda*,简称 *Ar*)、拟南芥(*Arabidopsis thaliana*,简称 *At*)和水稻(*Oryza sativa*,简称 *Os*)的全基因组数据来自 PGDD(http://chibba.agtec.uga.edu/duplication)。其中,无油樟基因组涉及 1 210 个 scaffolds,26 846 个基因;葡萄有 19 条染色体,26 346 个基因;拟南芥有 5 条染色体,27 416 个基因;水稻有 12 条染色体,39 049 个基因。

1.2　全基因组同源结构分析

为了更准确地寻找各物种基因组内以及基因组间的同源结构片段,本研究对获取的原始基因组数据进行了数据预处理。对于无油樟,选取前 100 个长度 2 Mb 以上的 scaffolds,其他物种均去除没有锚定到染色体上的基因。利用基本局部序列比对工具 BLASTP($E \leqslant 10^{-5}$)[15],对单一物种基因组内以及 2 个物种基因组间作双序列比对,获取基因组内以及基因组间的同源基因对。

结合基因组内及基因组间双序列比对结果与各个物种染色体的长度信息、染色体上各个基因的位置信息,编写 Perl 程序绘制同源结构点阵图,并且标记出基因对间匹配最好、次好以及其余的 3 个基因对。根据同源结构点阵图上基因对组成的成线性的片段,验证各个物种在进化历程中发生的全基因组加倍事件,同时为共线性数据分析提供图示化依据。

1.3　全基因组共线性分析

利用共线性分析工具 MCScan[16] 和 Colinearscan($P < 0.05$)[14],获取基因组内以及基因组间同源共线区域。为了描述同源共线片段的相似性,在基因组内和基因组间,考虑到

收稿日期:2017 - 07 - 19
基金项目:华北理工大学青年科学研究基金(编号:Z201528)。
作者简介:张　岚(1983—),女,山西长治人,硕士,讲师,主要从事植物比较基因组学研究。E - mail:zhanglan1374@sohu.com。
通信作者:潘玉欣,博士,副教授,主要从事植物学研究。E - mail:panyu - xin@163.com。

大的基因家族(重复基因对数≤30 个)[17]以及串联重复基因(基因对间相对位置 $<2\times10^6$ bp)[18]的影响,按照同源共线区域内包含的基因对数量,统计在不同长度范围的同源共线区域的数量和区域内包含的同源共线基因对数量、无油樟与其他物种基因组间共线性基因的保留情况,以及无油樟与拟南芥、葡萄、水稻基因组间共有的无油樟基因占共线性基因的百分比。

1.4　多物种基因组联合比对图谱的构建

本研究以葡萄基因组三倍乘事件的共线性基因作为参考列,结合基因组内和基因组间同源共线性分析结果,建立多物种基因组联合比对列表,并编写 python 程序,利用 Matplotlib 绘制多物种基因组联合比对图谱,以圈图的形式反映基因组间的相似性。

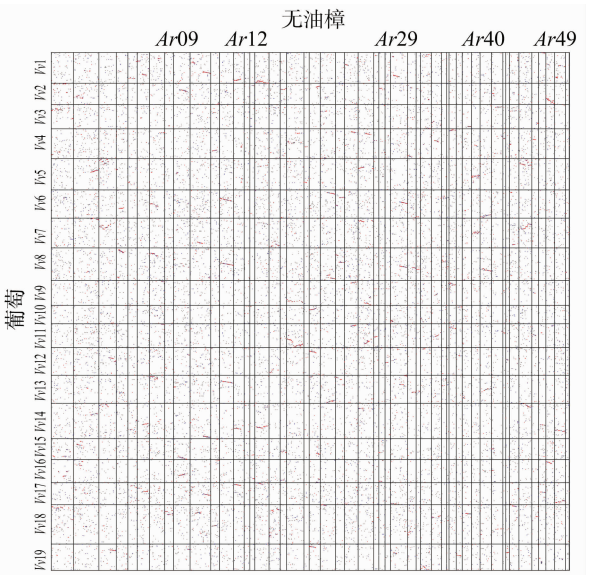
2　结果与分析

2.1　基因组内及基因组间同源结构分析

以 1 个已知的较为古老的物种或者全基因组发生加倍次数较多的物种基因组内的同源结构点阵图为依据,结合 2 个物种基因组间的同源结构点阵图,可以基本推断后一物种在进化过程中在全基因组层面发生的主要加倍事件,而且能够统计出在不同加倍事件中参与的基因数量。

在同源结构点阵图中,红色的点表示在物种基因组间进行的蛋白质序列比对中,处于最佳匹配的同源基因对,表示物种间的直系同源片段;蓝色的点表示的是次好匹配的同源基因对,而灰色的点表示的是其余的同源基因对,均表示种间旁系同源片断。

无油樟是被子植物中已测序的最古老的物种,其基因组没有被锚定到某一条染色体上,从无油樟基因组内的同源结构点阵图中,看到的是一些非常短小的线段,很难分析出其自身的加倍过程。而在无油樟和葡萄 2 个物种基因组间的同源结构分析中,发现无油樟的 9 号 scaffold 与葡萄的 2、5、16 号染色体之间存在 1∶3 同源共线关系,无油樟的 29 号 scaffold 与葡萄的 6、8、13 号染色体之间存在的 1∶3 同源共线关系。在无油樟的第 12、40 号 scaffold 与葡萄的 6、8、13 号染色体之间,以及第 49 号 scaffold 与葡萄的第 1、14、17 号染色体间也存在明显的 1∶3 的同源共线关系。同样,在其他 scaffold 与葡萄染色体之间也存在类似情况(图 1)。这与 *Amborella* Genome Project 报道的结果[4]相符。



横、纵坐标分别表示无油樟(*Ar*1~*Ar*50)和葡萄(*Vv*1~*Vv*19)的染色体,按照染色体号从左到右、从上到下依次排列(为显示方便,无油樟涉及前 50 个 scaffolds)。在二维坐标系中,每个点代表的是1对同源基因对,由基因在染色体上的位置所决定。其中,红色线表示1个无油樟的染色体片段对应 3 个葡萄的染色体片段,蓝色线表示其他匹配情况

图1　无油樟与葡萄基因组间同源结构点阵

2.2　同源共线片段推断

本研究使用共线性分析工具推测无油樟与葡萄、拟南芥、水稻基因组内和基因组间的同源共线片段。结果表明,同源性往往依赖于长的同源共线性片段,而不同基因组间的高相似性对解析基因组结构、功能与基因家族的分析提供了帮助[18]。

在无油樟基因组内,存在 102 个长度≥4 个基因对的同源共线区域,包含 473 个基因对,这说明其同源共线性较差。同样,在同源结构点阵图中也看不到大段的连续区域。当同源共线区域≥20 个基因对长度时,葡萄基因组内部包含 21 个同源共线区域,595 个同源共线基因对,最长的片段位于葡萄的 5 号和 7 号染色体之间,包含 52 个基因对;拟南芥基因组内部包含52个同源共线区域,有2 472个同源共线基因对,最长的片段位于拟南芥的 3 号和 2 号染色体之间,包含 181 个基因对;水稻基因组内包含 34 个同源共线区域,共 2 774 个同源共线基因对,最长的片段位于水稻的 1 号和 5 号染色体之间,包含 250 个基因对。而在水稻基因组内,在不同的基因对长度范围,得到的同源共线性区域以及同源共线性基因对的数量相对较多,超过 200 个基因对长度的区域有 2 个,包含 499 个同源基因对。大量的长片段存在于水稻和拟南芥基因组内(表 1),这与同源结构点阵图得到的结论相符。研究表明,无油樟与其他几个物种相比,拥有最短的共线性同源区域,它是最古老的物种,在漫长的历史进程中,基因组的结构遭受了严重的破坏,而其他几种模式植物则相对保存了较好的同源共线性。

在无油樟基因组与其他物种的共线性分析中,发现在不同长度同源共线区域限定时,无油樟与葡萄的基因组间保存相对较好的同源共线区域及同源共线基因对数量。当同源共线区域≥10 个基因对长度时,在无油樟与葡萄基因组间有 139 个同源共线区块,包含 2 311 个基因对;无油樟与拟南芥基因组间包含 14 个同源共线区块,包含 164 个基因对;在无油樟与水稻基因组间包含 49 个同源共线区块,包含 622 个基因对。当同源共线区域≥20 个基因对长度时,在无油樟与葡萄基因组间有 33 个共线性区块,包含 928 个基因对,最长的片段位于无油樟的第 2 号 scaffold 和葡萄的第 5 号染色体之间,包含 52 个基因对;在无油樟和水稻基因组间包含 2 个共线性区块,包含 42 个基因对,最长的片段位于无油樟的第 25 号 scaffold 和水稻的第 2 号染色体之间,包含 22 个基因对;无油樟和拟南芥基因组间没有发现超过 20 个同源共线基因对的同源共线区域。在无油樟与葡萄间甚至存在 2 个超过 50 个基因对长度的区域,共有 102 个同源共线基因对(表 1)。研究发现,以只经历了古老的被子植物全基因组加倍事件的

表 1　无油樟(Sca1 ~ Sca100)与其他物种基因组内与基因组间同源共线性区域以及基因对数量统计结果

物种	同源共线区块长度≥4 个		同源共线区块长度≥10 个		同源共线区块长度≥15 个		同源共线区块长度≥20 个	
	同源共线区域数(个)	同源共线基因对数(个)	同源共线区域数(个)	同源共线基因对数(个)	同源共线区域数(个)	同源共线基因对数(个)	同源共线区域数(个)	同源共线基因对数(个)
无油樟(*Ar*)	102	473	—	—	—	—	—	—
拟南芥(*At*)	725	6 570	111	3 234	65	2 689	52	2 472
水稻(*Os*)	812	7 392	88	3 442	46	2 982	34	2 774
葡萄(*Vv*)	352	2 868	80	1 431	48	1 040	21	595
无油樟和拟南芥(*Ar vs At*)	819	4 279	14	164	2	33	—	—
无油樟和水稻(*Ar vs Os*)	874	4 959	49	622	13	228	2	42
无油樟和葡萄(*Ar vs Vv*)	812	6 034	139	2 311	63	1 422	33	928
水稻和拟南芥(*Os vs At*)	1 806	10 098	70	817	7	110	—	—
拟南芥和葡萄(*At vs Vv*)	1 183	10 956	309	5 697	149	3 753	90	2 856
水稻和葡萄(*Os vs Vv*)	1 276	8 803	164	2 289	60	1 104	18	410

物种	同源共线区块长度≥50		同源共线区块长度≥100		同源共线区块长度≥200		最大同源共线区块含基因对数(个)	最大同源共线区块所在位置
	同源共线区域数(个)	同源共线基因对数(个)	同源共线区域数(个)	同源共线基因对数(个)	同源共线区域数(个)	同源共线基因对数(个)		
无油樟(*Ar*)	—	—	—	—	—	—	7	*Ar*16 *vs* *Ar*57;*Ar*24 *vs* *Ar*48;*Ar*61 *vs* *Ar*49
拟南芥(*At*)	22	1 591	2	360	—	—	181	*At*3 *vs* *At*2
水稻(*Os*)	19	2 378	12	1 930	2	499	250	*Os*1 *vs* *Os*5
葡萄(*Vv*)	—	—	—	—	—	—	52	*Vv*5 *vs* *Vv*7
无油樟和拟南芥(*Ar vs At*)	—	—	—	—	—	—	17	*Ar*2 *vs* *At*4
无油樟和水稻(*Ar vs Os*)	—	—	—	—	—	—	22	*Ar*25 *vs* *Os*2
无油樟和葡萄(*Ar vs Vv*)	2	102	—	—	—	—	52	*Ar*2 *vs* *Vv*5
水稻和拟南芥(*Os vs At*)	—	—	—	—	—	—	18	*Os*1 *vs* *At*5
拟南芥和葡萄(*At vs Vv*)	8	515	1	101	—	—	101	*At*3 *vs* *Vv*14
水稻和葡萄(*Os vs Vv*)	—	—	—	—	—	—	31	*Os*1 *vs* *Vv*8

注:“—”表示不存在同源共线区域或基因对。

无油樟基因组为参考,水稻和拟南芥基因组在进化历程中经历的加倍事件相对葡萄基因组复杂,二者所保留的同源共线区域和同源共线基因对数比葡萄基因组少。

2.3　同源共线性基因的保留情况分析

本研究涉及的 4 个物种在经历了各自的加倍事件后,其基因组不可避免地发生了分离、融合现象,从而导致各物种的基因存在不同程度的丢失,其中葡萄基因组丢失得最少(49.1%),水稻的次之(62.1%),拟南芥丢失得最为严重(65.2%)。这可能是由于葡萄有 19 条染色体,其基因组保留了大量的祖先基因,基因组比较稳定。在葡萄、拟南芥、水稻 3 个物种之间,拟南芥和葡萄间保留的同源共线性基因比例达到了 40.8%,而水稻和葡萄、水稻和拟南芥之间保留的同源共线性基因则相对较少,显然这与被子植物单、双子叶间的亲缘关系相关,亲缘关系越近,同源共线性基因丢失得越少(表 2)。

表 2　无油樟、葡萄、拟南芥和水稻基因组间同源共线性基因保留情况统计结果

物种 无油樟	葡萄			拟南芥			水稻		
	共线性基因保留数(个)	共线性基因保留比例(%)	2 个物种基因组间共线性基因数(个)	共线性基因保留数(个)	共线性基因保留比例(%)	2 个物种基因组间共线性基因数(个)	共线性基因保留数(个)	共线性基因保留比例(%)	2 个物种基因组间共线性基因数(个)
无油樟 —	3 073	50.9	6 034	1 490	34.8	4 297	1 877	37.9	4 959
葡萄 —	—	—	—	—	—	—	—	—	—
拟南芥 —	4 475	40.8	10 956	—	—	—	—	—	—
水稻 —	2 546	28.9	8 803	2 213	21.9	10 098	—	—	—

注:“—”表示不考察其共线基因的保留情况。下表同。

本研究统计了无油樟和拟南芥与无油樟和水稻、无油樟和拟南芥与无油樟和葡萄、无油樟和葡萄与无油樟和水稻保留的无油樟的基因数。发现无油樟和葡萄与无油樟和水稻间保留的无油樟基因数最多(30.6%),无油樟和拟南芥与无油樟和水稻间保留的无油樟的基因数最少(26.7%),详见表 3,这说明被子植物的祖先基因在进化过程中发生了大规模丢失,可能是由于物种各自的加倍事件造成的。

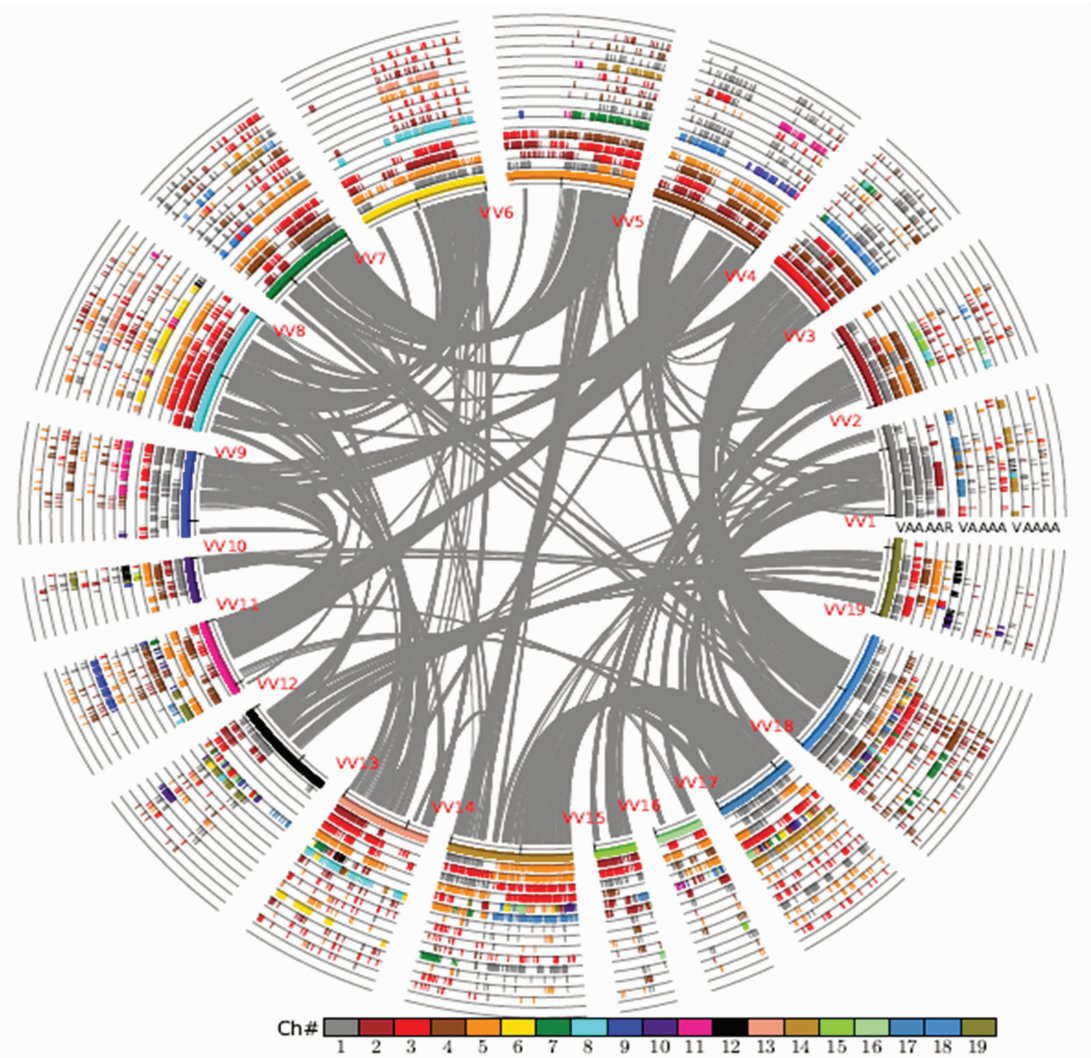
2.4　多物种基因组联合比对图谱的构建

作为双子叶植物中重要的模式植物,葡萄基因组保存了双子叶植物最完整的 γ 事件,在 13 000 万年前经历了双子叶植物共同祖先的全基因组三倍乘(γ)事件,而后并未发生单独的多倍化事件,因此以葡萄为参考,基于物种基因组内和基因组间的同源共线性分析结果,构建多物种基因组联合比对图谱。

表 3　无油樟与葡萄、拟南芥、水稻基因组间保留的 *Amborella* 同源共线基因保留情况统计结果

物种	无油樟和拟南芥	无油樟和水稻			无油樟和葡萄		
		保留的无油樟同源共线基因数(个)	保留的无油樟同源共线基因所占比例(%)	无油樟前 100 个 scaffolds 基因总数(个)	保留的无油樟同源共线基因数(个)	保留的无油樟同源共线基因所占比例(%)	无油樟前 100 个 scaffolds 基因总数(个)
无油樟和拟南芥	—	5 526	26.7	20 660	6 045	29.3	20 660
无油樟和水稻	—	—	—	—	6 328	30.6	20 660
无油樟和葡萄	—	—	—	—	—	—	—

由于葡萄和拟南芥共同经历了双子叶植物共同祖先的三倍乘事件，而拟南芥又先后发生了自己的 2 次二倍乘事件，因此，图 2 中的 3 圈葡萄分别对应 3 组拟南芥，而每组拟南芥又分别对应自己的 2 次二倍乘事件。由图 2 还可以看出，由内往外圈颜色分布越来越淡，这说明物种间的直系同源关系在进化过程中保持了相对完整，而物种内的旁系同源基因对则发生了大规模的丢失。由于无油樟与葡萄间存在 1∶3 的共线性关系，因此对应有 1 圈无油樟，可以看出图 2 在此圈有大量空白，说明存在大量基因片段的缺失。在多物种基因组联合比对图谱中，可以根据不同颜色来判定同源基因在物种间的分布情况，还可以用于分析各物种染色体的重组现象，其中每条染色体上不同的颜色区域都预示可能的染色体重组现象。图谱直观地阐释了各个物种基因组在进化过程中发生的基因丢失情况，同时也反映各物种染色体的结构变化。物种基因组的加倍规模和次数对基因组的影响在图谱中清晰可见，重复基因越多，可能造成基因组的不稳定以及基因丢失的增多。



最内圈逆时针排列的是葡萄的 19 条染色体，分别用图例所示 19 种不同的颜色表示，根据同源共线性分析结果，最内圈的灰线表示葡萄基因组内的旁系同源关系，其余各圈染色体基因分别对应最内圈葡萄染色体颜色。由内而外第 1、7、12 圈分别表示葡萄(V)的三倍乘事件，第 2~5、8~11、13~16 圈分别对应3组拟南芥(A)，其中，第 2~5 圈的拟南芥与葡萄是直系同源的关系，而第 2、3组拟南芥分别表示与葡萄间的种外旁系关系。为显示方便，第 6 圈的无油樟(R)只选取前 20 个 scaffolds，表示与葡萄基因组间的同源共线关系

图2　多物种基因组联合比对图谱

3　讨论与结论

多倍化是被子植物进化史上的重要特征，全基因组加倍对基因组结构及基因家族的进化有重要的影响[19]。本研究对无油樟、葡萄、拟南芥和水稻基因组进行了比较基因组学分析，在同源基因点阵图中发现无油樟基因组内的同源性较差，而在无油樟与其他几个物种的点阵图中也得到类似结果。而由无油樟与葡萄的点阵图得到与已有报道相同的结果，结构分析表明，无油樟和葡萄基因组之间存在很明显的 1∶3 的关系，这与古多倍体的 γ 事件相契合，表明无油樟发生的全基因组加倍不是特殊的谱系事件，可能是发生在 2 个物种的共同祖先中的，由此推断无油樟的分化要早于古 γ 事件。这可能是受无油樟基因组测序质量的影响，并与无油樟是已测序的最古老的开花植物有关[4]。

本研究在推测同源共线片段时，采用了严格的阈值限制，包括合适的 *E* 值、*P* 值、删除可能的大的重复基因家族和串联重复基因，同时区分同源共线片段的大小，从而保证了研究结果的可靠性[17-18,20]。本研究结果表明，在与无油樟的比较分析中，拟南芥和水稻同源共线性区域以及基因对数都比葡萄少，葡萄基因组所拥有的大片段数也是最多的，这说明物种基因组的加倍次数越多，加倍情况越复杂，所拥有的同源共线区域以及基因对数就越少，拥有的祖先基因也就越少。同时，在 2 个物种基因组间双向 BLASTp 得到的最佳匹配基因在共线基因的保留情况统计结果中[19]，发现葡萄基因组所保留的基因数最多(50.9%)，无油樟与拟南芥、水稻间丢失的无油樟的基因数最多(73.3%)，无油樟与葡萄、水稻间丢失的无油樟的基因数最少(69.4%)，整体来说，保留的无油樟基因数在 1/3 左右。本研究结果表明，在经历了物种基因组各自共有和单独加倍事件后，基因组的稳定性明显降低，并伴随着基因的丢失，而丢失仍将继续，该过程大多存在随机性。实际上，从多倍化的推断、非正常重组的分析、基因家族的进化分析等，在很大程度上都依赖于共线性分析，可以说，物种基因组内和基因组间共线性分析是比较基因组研究最重要的内容之一，是解析基因组结构和功能、研究基因、基因家族，以及调控路径等进化规律的重要契机，对基因组学许多研究工作都发挥着至关重要的作用。

参考文献：

[1] Edger P P, Pires J C. Gene and genome duplications: the impact of dosagesensitivity on the fate of nuclear genes [J]. Chromosome Research, 2009, 17(5): 699-717.

[2] Otto S P. The evolutionary consequences of polyploidy [J]. Cell, 2007, 131(3): 452-462.

[3] Lyons E, Pedersen B, Kane J, et al. Finding and comparing syntenic regions among *Arabidopsis* and the outgroups papaya, poplar, and grape: CoGe with rosids [J]. Plant Physiology, 2008, 148(4): 1772-1781.

[4] Amborella Genome Project. The *Amborella* genome and the evolution of flowering plants [J]. Science, 2013, 342(6165): 1241089.

[5] Chamala S, Chanderbali A S, Der J P, et al. Assembly and validation of the genome of the nonmodel basal angiosperm *Amborella* [J]. Science, 2013, 342(6165): 1516-1517.

[6] Rice D W, Alverson A J, Richardson A O, et al. Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella* [J]. Science, 2013, 342(6165): 1468-1473.

[7] Tang H B, Bowers J E, Wang X Y, et al. Synteny and collinearity in plant genomes [J]. Science, 2008, 320(5875): 486-488.

[8] Jaillon O, Aury J M, Noel B, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla [J]. Nature, 2007, 449(7161): 463-467.

[9] Vision T J, Brown D G, Tanksley S D. The origins of genomic duplications in *Arabidopsis* [J]. Science, 2000, 290(5499): 2114-2117.

[10] Barker M S, Vogel H, Schranz M E. Paleopolyploidy in the Brassicales: analyses of the cleome transcriptome elucidate the history of genome duplications in *Arabidopsis* and other Brassicales [J]. Genome Biology and Evolution, 2009, 1(1): 391-399.

[11] Jiao Y N, Leebens-Mack J, Ayyampalayam S, et al. A genome triplication associated with early diversification of the core eudicots [J]. Genome Biol, 2012, 13(1): R3.

[12] Kaul S, Koo H L, Jenkins J, Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana* [J]. Nature, 2000, 408(6814): 796-815.

[13] Wang X Y, Shi X L, Hao B L, et al. Duplication and DNA segmental loss in the rice genome: implications for diploidization [J]. New Phytologist, 2005, 165(3): 937-946.

[14] Wang X Y, Shi X L, Li Z, et al. Statistical inference of chromosomal homology based on gene colinearity and applications to *Arabidopsis* and rice [J]. BMC Bioinformatics, 2006, 7(1): 447-459.

[15] Altschul S F, Gish W, Miller W, et al. Basic local alignment search tool [J]. Journal of Molecular Biology, 1990, 215(3): 403-410.

[16] Tang H B, Wang X Y, Bowers J E, et al. Unraveling ancient hexaploidy through multiply-aligned angioserm gene maps [J]. Genome Research, 2008, 18(12): 1944-1954.

[17] Wang X Y, Wang J P, Jin D C, et al. Genome alignment spanning major Poaceae lineages reveals heterogeneous evolutionary rates and alters inferred dates for key evolutionary events [J]. Molecular Plant, 2015, 8(6): 885-898.

[18] Wang J P, Yu J X, Sun P C, et al. Comparative genomics analysis of rice and pineapple contributes to understand the chromosome number reduction and genomic changes in grasses [J]. Frontiers in Genetics, 2016, 7: 174.

[19] Jiao Y N, Wickett N J, Ayyampalayam S, et al. Ancestral polyploidy in seed plants and angiosperms [J]. Nature, 2011, 473(7345): 97-100.

[20] Wang X Y, Tang H B, Paterson A H. Seventy million years of concerted evolution of a homoeologous chromosome pair, in parallel, in major Poaceae lineages [J]. Plant Cell, 2011, 23(1): 27-37.