

# 果树全基因组测序研究进展

乔鑫, 李梦, 殷豪, 李雷廷, 吴俊, 张绍铃\*

(南京农业大学梨工程技术研究中心, 南京 210095)

**摘要:** 回顾了葡萄、番木瓜、苹果、香蕉、梨、甜橙等 10 种果树的全基因组测序的发展历程, 基于全基因组测序结果探讨了不同果树的起源进化, 并对果树全基因组测序的后续研究进行了概述, 进一步讨论了果树全基因组测序面临的挑战以及今后的研究重点。

**关键词:** 果树; 全基因组测序; 基因组学; 进化

**中图分类号:** S 66

**文献标志码:** A

**文章编号:** 0513-353X (2014) 01-0165-13

## Advances on Whole Genome Sequencing in Fruit Trees

QIAO Xin, LI Meng, YIN Hao, LI Lei-ting, WU Jun, and ZHANG Shao-ling\*

(Centre of Pear Engineering Technology Research, Nanjing Agricultural University, Nanjing 210095, China)

**Abstract:** At present, 10 complete genome sequences are available for fruit trees such as grape, papaya, apple, banana, pear and sweet orange. In this review, we first surveyed the historical development of genomic sequence of fruit trees. Then we investigated the evolution history of fruit trees, and summarized the progress of fruit trees post-genomic research. Challenges in genomics for fruit trees are raised and the future of post-genomics research interests in fruit trees was discussed in the end.

**Key words:** fruit trees; whole genome sequencing; genomics; evolution

自 2007 年葡萄 (*Vitis vinifera*) 基因组测序完成以来, 已有 10 种果树的全基因组测序工作相继完成。果树全基因组测序有望从基因组层面上突破传统分子生物学的研究瓶颈, 可以根据全基因组序列, 利用分子生物学及生物工程技术手段研究基因的功能, 为进一步解析重要农艺性状提供良好的基因组数据平台, 同时也将为培育高产、优质和抗病的果树新品种奠定坚实基础, 推动果树基因组学研究进入一个新时代。本文中回顾了果树全基因组测序的发展历程, 并结合果树全基因组测序结果探讨了不同果树的起源与进化, 对果树全基因组测序的后续研究进行了概述, 讨论了果树全基因组测序面临的挑战以及今后的研究重点。

## 1 果树全基因组测序及树种起源研究

### 1.1 测序技术的发展

20 世纪 90 年代早期, 随着自动化测序技术的发展, 以双脱氧链终止法为核心的 Sanger 测序技

**收稿日期:** 2013-07-15; **修回日期:** 2013-12-11

**基金项目:** 南京农业大学青年科技创新基金项目 (KJ2013014); 国家自然科学基金青年科学基金项目 (31301748); 南京农业大学科研启动费项目; 江苏省农业科技自主创新资金项目 [CX (11) 1013]

\* 通信作者 Author for correspondence (E-mail: slzhang@njau.edu.cn; Tel: 025-84396485)

术诞生, 该技术的应用使得大量基因被发现, 这标志着基因组学的开端。2000 年, 国际拟南芥基因组机构 (*Arabidopsis* Genome Initiative) 基于 Sanger 测序技术, 利用 BAC-by-BAC (bacterial artificial chromosome) 策略完成拟南芥全基因组测序 (The *Arabidopsis* Genome Initiative, 2000), 这是该策略在植物基因组测序中的首次应用。完成测序的拟南芥基因组成为开展其它植物基因组研究的模式。随着毛细管测序技术的开发以及组装算法和计算能力的改进, Sanger 测序平台进一步升级, 产出量大幅提升, 测序时间缩短, 成本降低, 这就使得对更大的真核生物基因组 (Adams et al., 2000) 进行测序成为可能。2007 年, Jaillon 等 (2007) 基于这种升级的 Sanger 测序技术, 利用 WGS 策略完成了葡萄全基因组测序, 这是首个完成测序的果树基因组。随后番木瓜 (*Carica papaya*) 基因组的成功测序 (Ming et al., 2008) 成为 Sanger WGS 测序方法在果树全基因组测序中应用的又一实例。Sanger 测序平台有其自身的优势, 测序读取片段长, 从而准确性较高; 然而由于技术、费用以及时间上的限制, 这些测序完成的基因组都是不完整的, 组装的序列中仍存在一些缺口和错误 (Bennett et al., 2003)。

在完成拟南芥测序之后的十多年间, 全基因组测序技术有了很大发展。自 2005 年以来, 以罗氏公司 (Roche) 的 454 系统 (<http://www.roche.com>)、Applied Biosystems 公司的 SOLiD™ 系统以及 Illumina 公司的 Solexa 系统为代表的第二代高通量测序技术 (表 1) 渐渐成为大规模全基因组测序的技术主导。

Roche 454 测序平台基于一种高通量的焦磷酸合成测序法, 利用乳液聚合酶链反应 (PCR), 在一个皮升 (pL) 大小的流动池内进行测序。2012 年, 454 系统升级为 FLX + 测序平台, 该平台一次最大能产生 100 万的读取量, 最大读取长度达到 1 000 bp (<http://my454.com/products/gs-flx-system/index.asp>), 可以代替 Sanger 测序平台读取更长的序列, 这对于破译重复序列具有重要价值。Roche 454 测序技术已广泛应用于植物全基因组测序中 (<http://my454.com/applications/whole-genome-sequencing/index.asp>)。2010 年 Nature genetics 公布了基于 Roche 454 和 Sanger 两种测序平台利用 WGS 策略完成的苹果 (*Malus × domestica*) 全基因组测序结果 (Velasco et al., 2010), 之后 Roche 454 技术在果树全基因组测序中得到广泛应用 (表 2)。

SOLiD™ 是另外一种高通量的测序平台, 利用乳液 PCR 扩增模板, 再通过荧光标记探针连接模板, 在流动池内进行测序反应。2012 年初 SOLiD™ 测序平台每天的产出量已达到 7 ~ 20 Gb, 最大读取长度为 75 bp。

Illumina 测序平台基于高通量的合成测序法, 在测序过程中使用可逆染料终止剂, 该测序平台的特别之处在于使用桥式 PCR (bridge PCR) 技术。2011 年底, 在 HiSeq2000 测序平台采用双末端测序模式, 一个测序通道能够产出全长序列中大于 2.5 亿的读取量 ([http://res.illumina.com/documents/products/datasheets/datasheet\\_hiseq\\_systems.pdf](http://res.illumina.com/documents/products/datasheets/datasheet_hiseq_systems.pdf))。目前 Illumina 测序平台作为一种主流的测序平台, 在果树全基因组测序中得到广泛使用。与第一代 Sanger 测序技术相比, 第二代测序技术降低了测序成本, 提高了测序速率, 且测序覆盖度更高, 因此在果树全基因组研究中具有广阔的应用前景。

近来, 以单分子测序为特点的第三代测序技术出现。Helicos 公司的 Heliscope 单分子测序仪、Pacific Biosciences 公司的 SMRT 技术和 Oxford Nanopore Technologies 公司的纳米孔单分子技术都属于第三代测序技术 (Munroe & Harris, 2010)。与前两代技术相比, 第三代测序技术省略了克隆步骤, 极大地提高了测序速度, 降低了测序成本 (Delseny et al., 2010)。但是第三代测序技术以及配套的软件、算法尚未发展成熟, 规模化应用尚需时日。

表 1 不同测序平台的比较\*

Table 1 Comparison of different sequencing platforms

测序平台 Sequencing platform	DNA 扩增方式 DNA amplification method	产出量/(Mb · d <sup>-1</sup> ) Throughput	最大读长/ nt Max read length	平均错误率/% Average error rate	测序成本/(\$ · Mb <sup>-1</sup> ) Costs
Sanger	大肠杆菌内的质粒扩增 Plasmid replication in <i>Escherichia coli</i>	6	800	10 <sup>-4</sup> ~ 10 <sup>-5</sup>	~ 500
Roche 454	乳液 PCR Emulsion PCR	750	400	10 <sup>-3</sup> ~ 10 <sup>-4</sup>	~ 20
SOLiD™	乳液 PCR Emulsion PCR	5000	50	10 <sup>-2</sup> ~ 10 <sup>-3</sup>	~ 0.50
Illumina	桥式 PCR Bridge PCR	5000	100	10 <sup>-2</sup> ~ 10 <sup>-3</sup>	~ 0.50
Helicos	单分子测序 Single-molecule sequencing	5000	32	10 <sup>-2</sup>	< 0.50

\*: Kircher &amp; Kelso, 2010.

表 2 10 种果树全基因组数据比较

Table 2 The comparison of 10 fruit trees whole-genome data

物种 Species	测序材料 Sequencing material	测序方法 Sequencing methods	基因组/ Mb Genome size	组装全长/ Mb Assembly	基因数量 Numbers of gene	重复序列 比率/% Repeat	总序列 骨架数 Scaffold total	文献 Published agazine
葡萄 <i>Vitis vinifera</i>	PN 40024	Sanger, WGS	475	487.1	30 434	41	3 514	Jaillon et al., 2007
番木瓜 <i>Carica papaya</i>	Sun Up	Sanger, WGS	372	370	28 629	43	17 764	Ming et al., 2008
苹果 <i>Malus × domestica</i>	Golden Delicious	Sanger, 454, WGS	742.3	603.9	57 386	67	1 629	Velasco et al., 2010
草莓 <i>Fragaria vesca</i>	Hawaii 4	454, SOLiD, Illumina WGS	240	209.8	34 809	23	3 263	Shulaev et al., 2010
海枣 <i>Phoenix dactylifer</i>	Khalas variety	Illumina, WGS	658	381	25 059	-	57 277	Al-Dous et al., 2011
可可 <i>Theobroma cacao</i>	Belizean Criollo	Sanger, 454, Illumina WGS	430	326.9	28 798	-	4 792	Argout et al., 2011
香蕉 <i>Musa acuminata</i>	DH-Pahang	Sanger, 454, Illumina WGS	523	472.2	36 542	44	7 513	D'Hont et al., 2012
梨 <i>Pyrus bretschneideri</i>	Dangshansuli	Illumina HiSeq2000 BAC-by-BAC	527	512	42 812	53.1	2 103	Wu et al., 2012
甜橙 <i>Citrus sinensis</i>	Valencia sweet orange	Illumina, GAII WGS	367	320.5	29 445	20	4 811	Xu et al., 2012
桃 <i>Prunus persica</i>	Lovell	Sanger, WGS	265	224.6	27 852	37.14	202	Verde et al., 2013

## 1.2 果树基因组特点及果树物种起源的研究

继葡萄 (Jaillon et al., 2007)、番木瓜 (Ming et al., 2008)、苹果 (Velasco et al., 2010)、草莓 (Shulaev et al., 2010)、海枣 (Al-Dous et al., 2011)、可可 (Argout et al., 2011)、香蕉 (D'Hont et al., 2012)、梨 (Wu et al., 2012)、甜橙 (Xu et al., 2012) 基因组草图完成后, 第 10 种果树桃 (The International Peach Genome Initiative et al., 2013) 的基因组草图也于近期发表, 10 种果树全基因组测序研究为进行其他果树作物全基因组测序研究提供了大量的参考信息 (表 2)。此外, 梅 (*Prunus mume*) 全基因组测序结果已经公布 (Zhang et al., 2012), 这将为果梅 (*Prunus mume*) 研究提供重要参考。

在测序方法上, 10 种果树作物以第一代和第二代测序方法为主, 其中葡萄、番木瓜和桃采用了第一代测序方法; 草莓、海枣、梨、甜橙采用了第二代测序方法; 而苹果、可可和香蕉则采用了一、二代测序方法相结合的策略。梨全基因组测序采用了 BAC-by-BAC 策略, 这是该策略在果树全基因

组测序中的首次应用, 解决了高度杂合重复基因组组装难的问题 (Wu et al., 2012)。该组装技术对于其他果树或高度杂合的高等植物基因组研究具有很好的借鉴价值。在组装长度上, 苹果基因组的组装全长最长, 为 603.9 Mb, 这是目前已完成全基因组测序果树中最大的基因组, 而草莓基因组的组装全长最短, 为 209.8 Mb。在基因数量上, 苹果基因组中注释的基因数量最多, 共 57 386 个, 其次是梨, 注释出 42 812 个基因, 海枣基因组中注释的基因数量最少, 共 25 059 个基因。此外, 重复序列的大小直接影响了基因组的大小, 例如苹果和梨全基因组大小的差异主要是由重复序列造成的 (Wu et al., 2012), 重复序列在两个基因组中所占比例分别为 67%和 53.1%。

利用 NCBI 分类数据库 (<http://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi>) 构建已测序的 10 种果树与梅、拟南芥、杨树 (*Populus trichocarpa*) 之间的进化树, 并结合果树全基因组测序分析结果标出了部分物种的分化时间, 以及发生全基因组复制事件 (whole-genome duplication, WGD) 的次数, 结果如图 1 (Jaillon et al., 2007; Argout et al., 2011; D'Hont et al., 2012; Wu et al., 2012; Xu et al., 2012)。

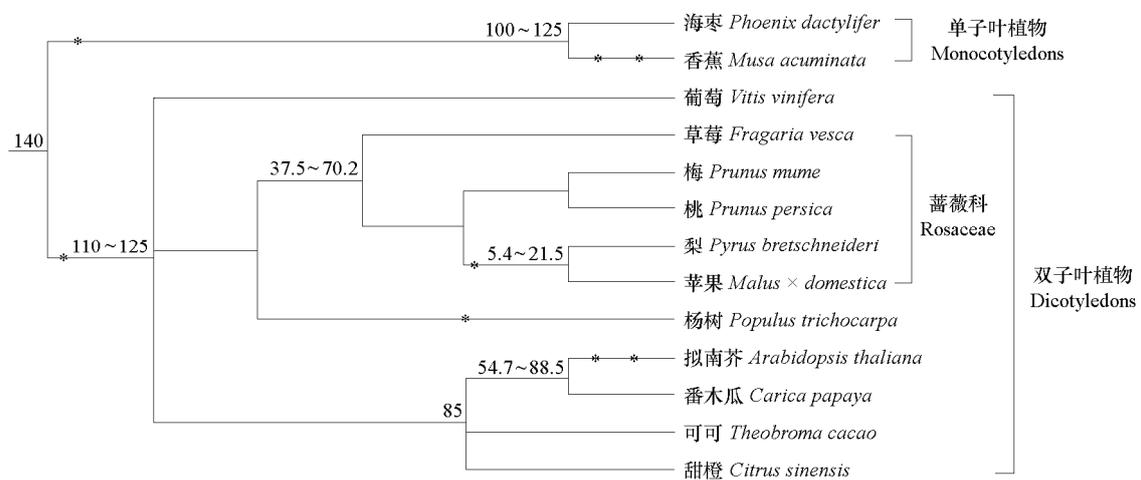


图 1 不同物种之间的进化关系及全基因组复制事件

\* 个数表示 WGD 发生的次数; 数字表示物种分化时间, 单位为百万年以前。

Fig. 1 Evolutionary relationships among different species and whole genome duplication events

The number of \* indicates the occurrences number of WGD;

Numbers in the figures indicate species divergence time, Unit: MYA (million years ago).

大多数植物在进化过程中都经历了全基因组复制事件, 全基因组复制存在古老的全基因组复制 (Old WGD) 和近代的全基因组复制 (Recent WGD) 两种方式, 它是植物重要的进化资源 (Jiao et al., 2011)。几乎所有的被子植物基因组都有全基因组复制的证据, 这表明多倍体化是被子植物基因组进化的重要推动因素 (Soltis et al., 2008)。在已经完成测序的 10 种果树基因组中都发现有全基因组复制的痕迹, 而且香蕉经历了 3 次全基因组复制 (D'Hont et al., 2012)。

Jaillon 等 (2007) 根据葡萄全基因组测序结果进行了系统进化分析, 结果表明葡萄在进化早期发生了古六倍体化复制事件, 葡萄基因组在近代未发生全基因组复制事件, 葡萄、杨树、拟南芥都属于蔷薇类植物 (Rosids) (Bremer et al., 2009), 通过比较葡萄、杨树、拟南芥基因组发现, 杨树、拟南芥也发生过类似的古六倍体化事件, 这一结果表明, 蔷薇类 (Rosids) 植物都起源于一个共同的古六倍体化祖先。这一结论在后续的梨、苹果、草莓以及桃基因组测序中得到验证。

梨与其他蔷薇科植物的比较基因组及进化分析表明, 梨与苹果、草莓都经历了 14 000 万年前双

子叶植物所共有的古六倍体化事件，之后大约在 3 000 到 4 500 万年前，梨和苹果又经历了一次全基因组复制事件，梨与苹果的分化大约发生在 540 万到 2 150 万年前左右 (Velasco et al., 2010; Wu et al., 2012)。此外，Wu 等 (2012) 利用不同物种之间的共线性关系，模拟了双子叶植物到蔷薇科的演化过程，结果表明 7 条双子叶植物的祖先染色体可能经历了复制和重组，产生了 9 条蔷薇科祖先染色体。苹果全基因组测序发现，一个近期的全基因组复制事件使得苹果亚科从 9 条古老染色体变为 17 条染色体，现存的苹果染色体可能来源于 9 个古老染色体的复制、易位以及缺失 (Velasco et al., 2010)。这些研究也为大多数双子叶植物具有共同的六倍体祖先提供了进一步证据。此外，全基因组测序研究还发现草莓和桃没有发生近代全基因组复制事件，现存的染色体是由蔷薇科 9 条古老染色体进化而来，但与葡萄类似，发生过古六倍体化事件 (Shulaev et al., 2010; The International Peach Genome Initiative et al., 2013)。

对于非蔷薇科物种甜橙、番木瓜、可可，通过全基因组测序也发现它们具有共同的古六倍体祖先。甜橙与所有核心真双子叶植物一样发生过古老的基因组复制事件 ( $\gamma$  基因组复制事件)，但没有发生过近代基因组复制事件 (Xu et al., 2012)。通过比较番木瓜和拟南芥的基因组，发现番木瓜大约在 7 200 万年前从拟南芥中分化出来，推测在双子叶植物出现早期发生的  $\beta$  基因组复制事件是拟南芥和番木瓜分化的真正原因，并且番木瓜没有发生近代全基因组复制事件 (Ming et al., 2008)。基因组分析结果表明可可是由双子叶植物的 7 条古老祖先染色体进化而来，发生过  $\gamma$  古六倍体化事件，再一次确认了双子叶植物具有共同古六倍体祖先 (Argout et al., 2011)。

值得一提的是，香蕉属于单子叶植物，发生过 3 次全基因组复制事件 ( $\alpha$ 、 $\beta$ 、 $\gamma$  基因组复制事件)，其基因组测序进一步阐明了单子叶植物的系统发育关系，揭示了禾本科植物特异的进化历程，并且发现了在单子叶植物和双子叶植物分化之前已经存在的高度保守的非编码序列，这些序列成为鉴定具有基因调控作用的新功能域的宝贵资源 (D'Hont et al., 2012)。

### 1.3 果树基因组重要基因挖掘

#### 1.3.1 糖

糖的组成和含量对于果实品质和风味有重要影响。与非蔷薇科物种不同的是，蔷薇科果树以山梨(糖)醇而非蔗糖作为主要的光合产物以及韧皮部运输的主要成分。把梨中山梨醇代谢相关基因同其它物种进行比较，结果发现 6-磷酸山梨醇脱氢酶(sorbitol-6-phosphate dehydrogenase, *S6PDH*)，山梨醇脱氢酶(sorbitol dehydrogenase, *SDH*)，山梨醇转运酶(sorbitol transport, *SOT*) 3 个基因家族在梨中的表达量比非蔷薇科物种更高，*S6PDH*、*SDH* 和 *SOT* 基因家族已经在梨和苹果基因组中扩张，并且这 3 个基因家族都属于苹果亚科特异分支(Maloideae-specific clade)，此外，在梨中发现 4 个 *S6PDH* 基因，转录组数据表明所有这 4 个 *S6PDH* 基因在果实中确实是表达的，这表明山梨醇能由单糖重新合成，尤其是在果实发育后期 (Wu et al., 2012)。

#### 1.3.2 挥发性物质

果实香气属于挥发性物质，是代表果实品质的重要性状。挥发性化合物主要来自于脂肪酸代谢、氨基酸代谢和碳水化合物代谢，果实的香气物质主要包括酯类、醇类、酮类、醛类、萜类和挥发性酚类物质等。与其它物种相比，梨中与  $\alpha$ -亚麻酸( $\alpha$ -linolenic acid)代谢途径相关的脂氧合酶(Lipoxygenase, *LOX*) (25) 和乙醇脱氢酶(Alcohol Dehydrogenase, *ADH*) (10) 基因有更多的数量，在果实发育过程中 1/3 的 *LOX* 同源基因高度表达，并且在果实发育中期表达量达到顶峰，与此同时，在果实发育阶段 *ADH* 的表达量随着乙醇的形成而增加 (Wu et al., 2012)。因此  $\alpha$ -亚麻酸代谢途径对于梨果实香气的形成可能是非常重要的。

### 1.3.3 维生素 C

维生素 C 又名抗坏血酸 (Ascorbic acid, AsA), 是生物体内重要的抗氧化剂和许多酶的辅因子。植物体内的 AsA 不仅对自身生长发育和抗逆性发挥重要作用, 而且与人类的健康密切相关。甜橙果实富含维生素 C, 甜橙基因组测序对维生素 C 代谢通路进行了深入研究, 结果表明控制半乳糖代谢途径分支的 3 个基因 (*PG*, *PME*, *GalUR*) 是显著上调的, 在这 3 个基因中编码 D-半乳糖醛酸还原酶的基因是该途径的限速步骤, 研究数据还表明在甜橙果实中半乳糖代谢途径是高度上调的, *GalUR* 对于甜橙 AsA 积累可能发挥尤为重要的作用 (Xu et al., 2012)。

### 1.3.4 木质素

木质素可以间接影响某些果实 (如梨, 番木瓜) 的品质。在梨果实中, 石细胞存在于果肉中, 是评价梨果实品质的重要指标, 在其它果实中很少存在。木质素是梨果实石细胞的主要组成成分, 其生物合成直接影响石细胞的含量 (Tao et al., 2009), 从而影响果实品质。研究表明, *HCT* 基因能够促进木质素的合成 (Hoffmann et al., 2004)。对梨果实发育不同阶段的 RNAseq 数据分析表明 (Wu et al., 2012), 在果实发育早期羟肉桂酰基转移酶 (hydroxycinn amoyl transferases, *HCT*) 基因的表达水平很高, 并与高度表达的 *p*-香豆酰-莽草酸/奎尼酸 3'-羟化酶 (*p*-coumaroyl-shikimate/quinate 3'-hydroxylases, C3'H) 基因和咖啡酰-辅酶 A O-甲基转移酶 (caffeoyl-CoA O-methyltransferase, *CCOMT*) 基因一起促使 *p*-香豆酰-辅酶 A (*p*-coumaroyl-CoA, *PCC*) 向咖啡酰-辅酶 A (caffeoyl-CoA, *CFC*) 和阿魏酰-辅酶 A (feruloyl-CoA, *FC*) 高水平的转化, 并最终导致 G 型木质素和 S 型木质素 (G-lignin 和 S-lignin) 的积累。基于以上发现, Wu 等 (2012) 认为梨果实石细胞中含有高水平的 G 型木质素和 S 型木质素, 而不含有 P 型木质素, 此外还发现咖啡酰-辅酶 A 向阿魏酰-辅酶 A 转变的这一反应过程是梨中木质素生物合成的限速步骤。

### 1.3.5 果实发育

果实的发育和成熟是一个重要的生物学过程并且有着巨大的经济价值。*MADS-box* 基因与果实发育密切相关, 可能调节梨果实发育, 因为在目前所分析的所有植物种类中 *MADS-box* 基因决定花组织最终的发育程度 (Shan et al., 2009)。大量的苹果 *type II MADS-box* 基因在系统发生上属于 *StMADSII* 亚分支, 该分支可能与梨果发育相关 (Velasco et al., 2010)。甜橙全基因组研究表明呼吸跃变型果实 (番茄) 和非呼吸跃变型果实 (甜橙) 存在两种不同的成熟机制, *MADS-RIN* 可能作为一种关键调节因子调节果实的成熟 (Xu et al., 2012)。

### 1.3.6 抗性

果树全基因组测序在抗病性研究方面的成果可以作为抗性育种的宝贵参考资源。植物基因组中包含大量的抗病基因 (*R gene*), 这类基因在植物对病虫害的防御过程中发挥重要作用 (Heath, 2000), 大多数 *R* 基因编码的蛋白质都包含 1 个核苷酸结合位点 (NBS) 和 1 个的富含亮氨酸重复 (LRR) 区域 (Hulbert et al., 2001), 在梨中, 总共鉴定出 396 个 *R* 基因 (Wu et al., 2012)。在其它果树基因组也发现了大量的 *R* 基因, 如葡萄基因组中包含 341 个 NBS-LRR 基因 (Jaillon et al., 2007)。

## 2 全基因组测序的后续研究

### 2.1 重测序研究

通过基因组重测序, 可以对栽培种和野生种基因组之间的差异进行比较, 从而揭示物种起源以及驯化过程, 为鉴定有价值的遗传资源以及果树育种提供重要参考。

果树全基因组测序以及重测序, 对果树的起源和驯化过程进行了广泛而深入的研究, 其中对苹

果、甜橙、桃驯化栽培过程的研究取得了重大进展。Velasco 等 (2010) 对苹果属 (*Malus*) 以及梨属 (*Pyrus*) 共 74 个种进行重测序分析表明, 现代栽培苹果与新疆野苹果 (*M. sieveesii*) 的亲缘关系比与欧洲苹果 (*M. sylvestris*)、山荆子 (*M. baccata*)、西府海棠 (*M. micromalus*) 和楸子 (*M. prunifolia*) 更近。因此, Velasco 等 (2010) 认为, 现代栽培苹果是由新疆野苹果进化而来, 而不是早先的认为的起源于欧洲苹果。为了探究甜橙的杂种起源, Xu 等 (2012) 首先分析了甜橙中的 307 个 SSR 标记的遗传来源, 初步确定柚 (*Citrus grandis*) 和柑橘 (*C. reticulata*) 与甜橙之间的特殊遗传关系, 然后进一步对 3 个柚品种和 3 个柑橘品种进行重测序 (覆盖深度 > 30×), 确定了甜橙基因组中有 39.7 Mb 的序列来源于柚, 118.2 Mb 的序列来源于柑橘, 符合 1:3 的遗传分离规律, 进而推测甜橙来源于柚子 (母本) 和柑橘 (父本) 的回交杂种 [sweet orange = (pummelo × mandarin) × mandarin]。为了阐明桃的驯化历程, 国际桃基因组机构对 11 个普通桃 (*Prunus persica*) 品种以及分别来自新疆桃 (*P. ferganensis*), 甘肃桃 (*P. kansuensis*), 山桃 (*P. davidiana*) 和西藏桃 (*P. mira*) 的一个品种进行了重测序, 基于全基因组 SNPs 的系统进化分析表明新疆桃与 11 个普通桃品种同属一个类群, 这表明新疆桃与桃栽培种可能属于同一个种, 新疆桃可能只是一种未经驯化的野生桃, 也可能代表桃驯化过程中的一个中间态基因组 (The International Peach Genome Initiative et al., 2013)。

## 2.2 SNP 芯片的开发

SNP 是基因组多态性最丰富的 DNA 分子标记, 其数量多, 分布均匀, 具有密度高、分布范围广和分型简便的特点。Clark 等 (2007) 利用高通量 SNP 芯片技术, 对模式植物拟南芥不同品系进行了多态性研究。随着不同果树基因组测序的完成, 在果树方面也陆续开展了相关 SNP 芯片研发和利用。Chagne 等 (2012) 对 27 个苹果品种进行低深度重测序检测并确认全基因组范围的 SNP, 开发了苹果 8K 的 SNP 芯片, 可用于苹果幼苗大规模检测, 这将会促进标记一位点—性状之间关联性的发现, 进一步阐明质量性状的遗传结构特性, 推动遗传变异研究。Cabezas 等 (2011) 利用覆盖全基因组的 48 个高分辨率的 SNP 标记开发了基因芯片用于葡萄品种的鉴定; Lijavetzky 等 (2007) 结合重测序以及 SNPlex 技术挖掘高通量 SNP, 并进行葡萄基因分型。这些 SNP 芯片的开发为果树遗传多样性研究、品种鉴定、连锁图谱构建、标记与性状之间的关联分析、数量性状定位及分子标记辅助选择等提供了必要的技术平台。

## 2.3 高精度图谱构建及 QTL 定位

果树全基因组序列对于遗传图谱以及物理图谱的构建具有重要参考价值, 利用果树基因组中的 SNP 标记或多态性 SSR 标记构建高精度遗传图谱及物理图谱已成为果树全基因组测序后续研究的重要内容, 已相继构建出了苹果、葡萄、番木瓜等多个不同类型的遗传图谱和物理图谱。Antanaviciute 等 (2012) 利用苹果全基因组 SNP 基因芯片构建了一个苹果砧木 ‘M432’ 后代的高密度遗传图谱, 该图谱将用于具有经济效益的 QTL 分析, 并且改善基因组序列的组装精度; Vezzulli 等 (2008) 利用来自于 3 个葡萄栽培种的 283 个 SSR 标记以及 501 个 SNP 标记构建了综合遗传图谱; Scalabrin 等 (2010) 构建了一个高度杂合基因组的物理图谱: ‘Pinot Noir’ 葡萄物理重叠群图。

结合全基因组测序信息进行果树重要质量性状位点 (QTL) 定位, 对于果树育种和果实品质改良具有重要意义。在测序完成的果树物种 (如苹果、葡萄、草莓等) 中已经开展了大量的 QTL 定位研究。其中在挥发性物质相关的 QTL 定位方面, Khan 等 (2012) 通过对苹果代谢物的遗传分析, 揭示酚类物质 QTL 热点位于 16 号连锁群; Battilana 等 (2009) 利用 QTL 分析与候选基因方法, 结合代谢表达谱, 对控制葡萄香气形成的关键遗传因子进行了鉴定, 最终两个候选基因定位到葡萄 10

号连锁群上控制芳樟醇含量 QTL 区域, 还有 1 个候选基因 1-脱氧-d-木酮糖 5-磷酸盐 (1-deoxy-d-xylulose 5-phosphate) 合酶基因与 5 号连锁群上控制单萜含量的主效 QTL 被共定位; 在果实品质、大小形状相关的 QTL 定位方面, Longhi 等 (2012, 2013) 利用 QTL 定位解析复杂的苹果质地生物学, 并进一步鉴定出影响果实质地的主效 QTL 为 Md-PG1。Zorrilla-Fontanesi 等 (2011) 对控制八倍体草莓 (*Fragaria × ananassa*) 农艺学和果实品质性状的质量性状位点和潜在的候选基因进行了鉴定; Lerceteau-Kohler 等 (2012) 解析了八倍体草莓栽培种果实品质性状的遗传特性, 发现 homoeo-QTL 对于控制品质性状具有突出作用。Blas 等 (2012) 对番木瓜中控制果实大小和形状的质量性状位点 (QTL) 进行遗传作图; 此外, 还有对抗病相关 QTL 的定位研究, Moreira 等 (2011) 利用葡萄两个种间杂交种的遗传连锁图谱来定位抗霜霉病的质量性状位点。可见基因组测序数据为解析数量性状的遗传机制提供了更多的信息和更高的平台。

## 2.4 转录组学研究

早期的转录组学研究利用表达序列标签 (ESTs) 进行转录组分析, 随着基因芯片的发明 (microarrays), 研究者可以对成千上万个基因同时进行表达分析。最近出现的转录组深度测序 (RNA-Seq), 可以用更低的成本对不同植物或同一植物不同组织、不同发育时期、不同条件下的转录组进行大规模深度测序, 大大提高了转录组研究的准确度以及表达谱覆盖率 (Martínez-Gómez et al., 2011)。目前 RNA-Seq 技术已经被广泛应用到果树转录组研究中, 在葡萄 (Zenoni et al., 2010; Venturini et al., 2013)、甜橙 (Yu et al., 2012)、香蕉 (Wang et al., 2012) 等研究中已经取得了较大进展。Zenoni 等 (2010) 利用 RNA-Seq 技术对葡萄栽培种 ‘Corvina’ 浆果发育 3 个阶段的转录组变化进行了研究, 检测到 17 324 个基因在浆果发育过程中表达, 其中 6 695 个在某个特定阶段特异性表达, 这表明了众多功能基因的表达特异性以及转录组的高度复杂性。Yu 等 (2012) 对甜橙品种 ‘Anliu’ (WT) 以及它的红色果肉突变体 ‘Hong Anliu’ (MT) 果实 4 个发育阶段的转录组进行了比较, 对这两个品种果实发育和成熟过程中基因表达的动态变化以及差异表达基因进行探究, 在 MT 和 WT 中分别检测到 19 440 和 18 829 个表达的基因, 在 WT 中超过 89% 的基因在果实发育和成熟过程中是差异表达的, 进一步对这些差异表达基因进行功能分类表明, 这些基因主要参与细胞壁生物合成、碳水化合物以及柠檬酸代谢、类胡萝卜素代谢、胁迫响应调控过程。

## 2.5 数据库构建

随着测序技术的不断发展, 全基因组测序的趋势已在向更多物种更多个体上进行, 通过全基因组测序获得了包括全基因组序列, 转录组及表达数据, 蛋白质组和代谢组数据, 基因型和表型数据, 遗传图谱与物理图谱在内的大量生物学信息。因此, 利用生物信息学工具和数据库分析、贮存、整合以及传播这些数据, 为基础研究和应用研究提供实用的数据是非常必要的。利用全基因组测序信息已经构建了许多重要果树基因组学和生物信息学数据库, 综合性数据库有 GDR (Genome Database for Rosaceae, <http://www.rosaceae.org/node/1>), 其中包括苹果、梨、桃、草莓、杏、樱桃、玫瑰 7 个物种, 包含遗传图谱, EST 数据, 基因组序列以及一些相应的数据分析工具等。其他果树也有相应的网站, 如梨 (<http://peargenome.njau.edu.cn/>)、甜橙 (<http://citrus.hzau.edu.cn/orange/>)、葡萄 (<http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/>) 等, 在这些网站也可以进行全基因组序列、代谢数据等搜索及下载。这些数据库将为深入开展果树基因组学研究提供丰富的资源 and 数据。

### 3 果树全基因组测序面临的挑战及今后研究展望

植物基因组在进化过程中发生全基因组复制、多倍体化的频率比哺乳动物更高，因此多基因家族以及重复序列大量存在 (Lockton & Gaut, 2005)，在大量同源基因和重复序列存在的情况下，高通量测序平台产生的短序列读取片段不能够准确地定位到参考基因组上，这就给全基因组测序及组装造成巨大困难。许多果树基因组的高度重复性是由于转座子的高拷贝数和扩张特性造成的，例如，利用 BAC-by-BAC 策略进行测序的梨基因组，转座子占到基因组序列 37.2% (Wu et al., 2012)。此外对组装过程中基因组缺失的部分进行质量评价，结果表明大部分未组装到基因组上的序列都是由重复序列构成的 (Huang et al., 2009; Paterson et al., 2009)。因此，即使使用现有最好的组装算法来组装这些高度重复的序列也同样具有很高的挑战性。

果树基因组的复杂性还在于大多数果树并不是纯合的自交系二倍体，而是高度杂合的二倍体。利用 WGS 策略对一些高度杂合的果树基因组进行组装仍然面临很大的困难。为了克服这个问题，研究者们选用了基因型纯合度高的材料进行测序，例如葡萄在进行全基因组测序时利用了高度纯合的自交系 ‘PN40024’ (Jaillon et al., 2007)，在番木瓜全基因组测序时也选用了高度纯合的近交系品种 (Ming et al., 2008)，从而降低了杂合度对测序及组装的干扰；类似地在进行甜橙 (Xu et al., 2012) 和桃 (The International Peach Genome Initiative et al., 2013) 全基因组测序时，选用了基因型纯合的双单倍体品系，大大降低了基因组测序的复杂度。对于果树基因组从头测序和组装来说，染色体的倍性及来源是最大的挑战，测序质量直接取决于被测物种是同源多倍体还是异源多倍体，以及多倍体化事件发生的时间。到目前为止，所有对多倍体生物基因组测序进行的尝试都是通过降低倍性或者染色体的物理分离进行的，例如栽培种草莓 (*Fragaria × ananassa*) 是来源 4 个不同祖先种的八倍体，为了避开对多倍体基因组进行测序的困难，在进行草莓全基因组测序时选择了二倍体野生草莓 (*Fragaria vesca*) 作为测序材料，最终成功地完成了测序 (Shul'yaev et al., 2010)。因此，在能够阐明科学问题以及满足科学研究价值的前提下，应尽量选择基因型纯合度高、基因组相对较小的品系作为测序材料，尤其是要优先选用单倍体材料或二倍体材料，而不宜采用多倍体材料。

通过全基因组测序虽然得到了大量数据，但是如何破解并利用这些数据成为了摆在研究者面前的巨大挑战。因此，在今后很长的一段时间内研究重点将从表现型依赖型研究转入基因型依赖型研究，从对单一基因的研究转入对整个基因组的研究。在全基因组测序结果的基础上，选择不同品种、品系或个体进行全基因组重测序，探索物种的起源与驯化历程，填补果树参考基因组中的缺口，提高组装精度；对同种果树作物的不同时期、不同组织材料，野生型与突变体材料，胁迫处理与各种胁迫处理（低温、干旱、高盐、ABA）材料之间的转录组学 (Transcriptomics)、代谢组学 (Metabolomics)、蛋白质组学 (Proteomics) 和降解组学 (Degradomics) 进行比较研究 (施季森等, 2012)；对重要基因家族进行比较基因组学分析，并进行后续的功能验证；探索全基因组关联分析 (Genome-wide association studies, GWAS)，评价决定个体基因型的成千上万单核苷酸多态性 (SNPs)，解析控制复杂的农艺性状的遗传位点 (Huang et al., 2010; Brachi et al., 2011)，这些将是果树全基因组测序后研究的重点和方向。随着测序技术的不断发展，相信在不久的将来对各种果树进行全基因组测序将成为现实，果树基因组学研究也必将走向深入。

### References

- Adams M D, Celniker S E, Holt R A, Evans C A, Gocayne J D, Amanatides P G, Scherer S E, Li P W, Hoskins R A, Galle R F. 2000. The genome sequence of *Drosophila melanogaster*. *Science*, 287: 2185 - 2195.
- Al-Dous E K, George B, Al-Mahmoud M E, Al-Jaber M Y, Wang H, Salameh Y M, Al-Azwani E K, Chaluvadi S, Pontaroli A C, DeBarry J,

- Aronel V, Ohlrogge J, Saie I J, Suliman-Elmeer K M, Bennetzen J L, Kruegger R R, Malek J A. 2011. De novo genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). *Nat Biotechnol*, 29: 521 - 527.
- Antanaviciute L, Fernandez-Fernandez F, Jansen J, Banchi E, Evans K M, Viola R, Velasco R, Dunwell J M, Troggio M, Sargent D J. 2012. Development of a dense SNP-based linkage map of an apple rootstock progeny using the *Malus* infinium whole genome genotyping array. *Bmc Genomics*, 13: 203.
- Argout X, Salse J, Aury J M, Guiltinan M J, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova S N, Abrouk M, Murat F, Fouet O, Poulain J, Ruiz M, Roguet Y, Rodier-Goud M, Barbosa-Neto J F, Sabot F, Kudrna D, Ammiraju J S, Schuster S C, Carlson J E, Sallet E, Schiex T, Dievart A, Kramer M, Gelley L, Shi Z, Berard A, Viot C, Boccara M, Risterucci A M, Guignon V, Sabau X, Axtell M J, Ma Z, Zhang Y, Brown S, Bourge M, Golser W, Song X, Clement D, Rivallan R, Tahiri M, Akaza J M, Pitollat B, Gramacho K, D'Hont A, Brunel D, Infante D, Kebe I, Costet P, Wing R, McCombie W R, Guiderdoni E, Quetier F, Panaud O, Wincker P, Boes S, Lanaud C. 2011. The genome of *Theobroma cacao*. *Nat Genet*, 43: 101 - 108.
- Battilana J, Costantini L, Emanuelli F, Sevini F, Segala C, Moser S, Velasco R, Versini G, Grando M S. 2009. The 1-deoxy-d-xylulose 5-phosphate synthase gene co-localizes with a major QTL affecting monoterpene content in grapevine. *Theoretical and Applied Genetics*, 118: 653 - 669.
- Bennett M, Leitch I, Price H J, Johnston J S. 2003. Comparisons with *Caenorhabditis* (approximately 100 Mb) and *Drosophila* (approximately 175 Mb) using flow cytometry show genome size in *Arabidopsis* to be approximately 157 Mb and thus approximately 25% larger than The *Arabidopsis* Genome Initiative Estimate of approximately 125 Mb. *Annals of Botany*, 91: 547 - 557.
- Blas A L, Yu Q Y, Veatch O J, Paull R E, Moore P H, Ming R. 2012. Genetic mapping of quantitative trait loci controlling fruit size and shape in papaya. *Molecular Breeding*, 29: 457 - 466.
- Brachi B, Morris G P, Borevitz J O. 2011. Genome-wide association studies in plants: The missing heritability is in the field. *Genome Biology*, 12: 232.
- Bremer B, Bremer K, Chase M, Fay M, Reveal J, Soltis D, Soltis P, Stevens P. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society*, 161: 105 - 121.
- Cabezas J A, Ibanez J, Lijavetzky D, Velez D, Bravo G, Rodriguez V, Carreno I, Jermakow A M, Carreno J, Ruiz-Garcia L, Thomas M R, Martinez-Zapater J M. 2011. A 48 SNP set for grapevine cultivar identification. *Bmc Plant Biology*, 11: 153.
- Chagne D, Crowhurst R N, Troggio M, Davey M W, Gilmore B, Lawley C, Vanderzande S, Hellens R P, Kumar S, Cestaro A, Velasco R, Main D, Rees J D, Iezzoni A, Mockler T, Wilhelm L, Van de Weg E, Gardiner S E, Bassil N, Peace C. 2012. Genome-Wide SNP detection, validation, and development of an 8K SNP array for apple. *Plos One*, 7: e31745.
- Clark R M, Schweikert G, Toomajian C, Ossowski S, Zeller G, Shinn P, Warthmann N, Hu T T, Fu G, Hinds D A. 2007. Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. *Science*, 317: 338 - 342.
- Delseny M, Han B, Hsing Y I. 2010. High throughput DNA sequencing: The new sequencing revolution. *Plant Science*, 179: 407 - 422.
- D'Hont A, Denoeud F, Aury J M, Baurens F C, Carreel F, Garsmeur O, Noel B, Boes S, Droc G, Rouard M, Da Silva C, Jabbari K, Cardi C, Poulain J, Souquet M, Labadie K, Jourda C, Lengelle J, Rodier-Goud M, Alberti A, Bernard M, Correa M, Ayyampalayam S, McKain M R, Leebens-Mack J, Burgess D, Freeling M, Mbeguie A M D, Chabannes M, Wicker T, Panaud O, Barbosa J, Hribova E, Heslop-Harrison P, Habas R, Rivallan R, Francois P, Poirion C, Kilian A, Burthia D, Jenny C, Bakry F, Brown S, Guignon V, Kema G, Dita M, Waalwijk C, Joseph S, Dievart A, Jaillon O, Leclercq J, Argout X, Lyons E, Almeida A, Jeridi M, Dolezel J, Roux N, Risterucci A M, Weissenbach J, Ruiz M, Glaszmann J C, Quetier F, Yahiaoui N, Wincker P. 2012. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature*, 488: 213 - 217.
- Heath M C. 2000. Hypersensitive response-related death. *Plant Molecular Biology*, 44: 321 - 334.
- Hoffmann L, Besseau S, Geoffroy P, Ritzenthaler C, Meyer D, Lapierre C, Pollet B, Legrand M. 2004. Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. *The Plant Cell Online*, 16: 1446 - 1465.
- Huang S, Li R, Zhang Z, Li L, Gu X, Fan W, Lucas W J, Wang X, Xie B, Ni P. 2009. The genome of the cucumber, *Cucumis sativus* L. *Nature Genetics*, 41: 1275 - 1281.
- Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z. 2010. Genome-wide association studies of 14 agronomic traits

- in rice landraces. *Nature Genetics*, 42: 961 - 967.
- Hulbert S H, Webb C A, Smith S M, Sun Q. 2001. Resistance gene complexes: Evolution and utilization. *Annual review of phytopathology*, 39: 285 - 312.
- Jaillon O, Aury J M, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Huguency P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyere C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro F, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G, Laucou V, Chatelet P, Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pe M E, Valle G, Morgante M, Caboche M, Adam-Blondon A F, Weissenbach J, Quetier F, Wincker P, French-Italian Public Consortium for Grapevine Genome C. 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, 449: 463 - 467.
- Jiao Y, Wickett N J, Ayyampalayam S, Chanderali A S, Landherr L, Ralph P E, Tomsho L P, Hu Y, Liang H, Soltis P S. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature*, 473: 97 - 100.
- Khan S A, Chibon P Y, de Vos R C H, Schipper B A, Walraven E, Beekwilder J, van Dijk T, Finkers R, Visser R G, van de Weg E W, Bovy A, Cestaro A, Velasco R, Jacobsen E, Schouten H J. 2012. Genetic analysis of metabolites in apple fruits indicates an mQTL hotspot for phenolic compounds on linkage group 16. *Journal of Experimental Botany*, 63: 2895 - 2908.
- Kircher M, Kelso J. 2010. High-throughput DNA sequencing - concepts and limitations. *Bioessays*, 32: 524 - 536.
- Lerceteau-Kohler E, Moing A, Guerin G, Renaud C, Petit A, Rothan C, Denoyes B. 2012. Genetic dissection of fruit quality traits in the octoploid cultivated strawberry highlights the role of homoeo-QTL in their control. *Theoretical and Applied Genetics*, 124: 1059 - 1077.
- Lijavetzky D, Cabezas J A, Ibanez A, Rodriguez V, Martinez-Zapater J M. 2007. High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. *Bmc Genomics*, 8: 424.
- Lockton S, Gaut B S. 2005. Plant conserved non-coding sequences and paralogue evolution. *Trends in Genetics*, 21: 60 - 65.
- Longhi S, Hamblin M T, Trainotti L, Peace C P, Velasco R, Costa F. 2013. A candidate gene based approach validates Md-PG1 as the main responsible for a QTL impacting fruit texture in apple (*Malus × domestica* Borkh.). *Bmc Plant Biology*, 13: 37.
- Longhi S, Moretto M, Viola R, Velasco R, Costa F. 2012. Comprehensive QTL mapping survey dissects the complex fruit texture physiology in apple (*Malus × domestica* Borkh.). *Journal of Experimental Botany*, 63: 1107 - 1121.
- Martínez-Gómez P, Crisosto C H, Bonghi C, Rubio M. 2011. New approaches to *Prunus* transcriptome analysis. *Genetica*, 139: 755 - 769.
- Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw J H, Senin P, Wang W, Ly B V, Lewis K L, Salzberg S L, Feng L, Jones M R, Skelton R L, Murray J E, Chen C, Qian W, Shen J, Du P, Eustice M, Tong E, Tang H, Lyons E, Paull R E, Michael T P, Wall K, Rice D W, Albert H, Wang M L, Zhu Y J, Schatz M, Nagarajan N, Acob R A, Guan P, Blas A, Wai C M, Ackerman C M, Ren Y, Liu C, Wang J, Wang J, Na J K, Shakirov E V, Haas B, Thimmapuram J, Nelson D, Wang X, Bowers J E, Gschwend A R, Delcher A L, Singh R, Suzuki J Y, Tripathi S, Neupane K, Wei H, Irikura B, Paidi M, Jiang N, Zhang W, Presting G, Windsor A, Navajas-Perez R, Torres M J, Feltus F A, Porter B, Li Y, Burroughs A M, Luo M C, Liu L, Christopher D A, Mount S M, Moore P H, Sugimura T, Jiang J, Schuler M A, Friedman V, Mitchell-Olds T, Shippen D E, dePamphilis C W, Palmer J D, Freeling M, Paterson A H, Gonsalves D, Wang L, Alam M. 2008. The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature*, 452: 991 - 996.
- Moreira F M, Madini A, Marino R, Zulini L, Stefanini M, Velasco R, Kozma P, Grando M S. 2011. Genetic linkage maps of two interspecific grape crosses (*Vitis* spp.) used to localize quantitative trait loci for downy mildew resistance. *Tree Genetics & Genomes*, 7: 153 - 167.
- Munroe D J, Harris T J. 2010. Third-generation sequencing fireworks at Marco Island. *Nature Biotechnology*, 28: 426 - 428.
- Paterson A H, Bowers J E, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberger G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti A K, Chapman J, Feltus F A, Gowik U, Grigoriev I V, Lyons E, Maher C A, Martis M, Narechania A, Otillar R P, Penning B W, Salamov A A, Wang Y, Zhang L, Carpita N C, Freeling M, Gingle A R, Hash C T, Keller B, Klein P, Kresovich S, McCann M C, Ming R, Peterson D G, Mehboob-ur-R, Ware D, Westhoff P, Mayer K F, Messing J, Rokhsar D S. 2009. The *Sorghum bicolor* genome and the diversification of grasses. *Nature*, 457: 551 - 556.

- Scalabrin S, Troglio M, Moroldo M, Pindo M, Felice N, Coppola G, Prete G, Malacarne G, Marconi R, Faes G, Jurman I, Grando S, Jesse T, Segala C, Valle G, Policriti A, Fontana P, Morgante M, Velasco R. 2010. Physical mapping in highly heterozygous genomes: A physical contig map of the 'Pinot Noir' grapevine cultivar. *Bmc Genomics*, 11: 204.
- Shan H, Zahn L, Guindon S, Wall P K, Kong H, Ma H, DePamphilis C W, Leebens-Mack J. 2009. Evolution of plant MADS box transcription factors: Evidence for shifts in selection associated with early angiosperm diversification and concerted gene duplications. *Molecular Biology and Evolution*, 26: 2229 - 2244.
- Shi Ji-sen, Wang Zhan-jun, Chen Jin-hui. 2012. Progress on whole genome sequencing in woody plants. *Hereditas*, 34 (2): 145 - 156. (in Chinese)  
施季森, 王占军, 陈金慧. 2012. 木本植物全基因组测序研究进展. *遗传*, 34 (2): 145 - 156.
- Shulaev V, Sargent D J, Crowhurst R N, Mockler T C, Folkerts O, Delcher A L, Jaiswal P, Mockaitis K, Liston A, Mane S P, Burns P, Davis T M, Slovin J P, Bassil N, Hellens R P, Evans C, Harkins T, Kodira C, Desany B, Crasta O R, Jensen R V, Allan A C, Michael T P, Setubal J C, Celton J-M, Rees D J G, Williams K P, Holt S H, Rojas J J R, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin S A, Troglio M, Viola R, Ashman T-L, Wang H, Dharmawardhana P, Elser J, Raja R, Priest H D, Bryant D W, Fox S E, Givan S A, Wilhelm L J, Naithani S, Christoffels A, Salama D Y, Carter J, Girona E L, Zdepski A, Wang W, Kerstetter R A, Schwab W, Korban S S, Davik J, Monfort A, Denoyes-Rothan B, Arus P, Mittler R, Flinn B, Aharoni A, Bennetzen J L, Salzberg S L, Dickerman A W, Velasco R, Borodovsky M, Veilleux R E, Folta K M. 2010. The genome of woodland strawberry (*Fragaria vesca*). *Nature Genetics*, 43: 109 - 116.
- Soltis D E, Bell C D, Kim S, Soltis P S. 2008. Origin and early evolution of angiosperms. *Annals of the New York Academy of Sciences*, 1133: 3 - 25.
- Tao S, Khanizadeh S, Zhang H, Zhang S. 2009. Anatomy, ultrastructure and lignin distribution of stone cells in two *Pyrus* species. *Plant Science*, 176: 413 - 419.
- The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, 408: 796 - 815.
- The International Peach Genome Initiative, Verde I, Abbott A G, Scalabrin S, Jung S, Shu S, Marroni F, Zhebentyayeva T, Dettori M T, Grimwood J, Cattonaro F, Zuccolo A, Rossini L, Jenkins J, Vendramin E, Meisel L A, Decroocq V, Sosinski B, Prochnik S, Mitros T, Policriti A, Cipriani G, Dondini L, Ficklin S, Goodstein D M, Xuan P, Fabbro C D, Aramini V, Copetti D, Gonzalez S, Horner D S, Falchi R, Lucas S, Mica E, Maldonado J, Lazzari B, Bielenberg D, Pirona R, Miculan M, Barakat A, Testolin R, Stella A, Tartarini S, Tonutti P, Arus P, Orellana A, Wells C, Main D, Vizzotto G, Silva H, Salamini F, Schmutz J, Morgante M, Rokhsar D S. 2013. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat Genet*, 45: 487 - 494.
- Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar S K, Troglio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald L M, Gutin N, Lanchbury J, Macalma T, Mitchell J T, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu V T, King S T, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater M M, Masiero S, Lasserre P, Lespinasse Y, Allan A C, Bus V, Chagne D, Crowhurst R N, Gleave A P, Lavezzo E, Fawcett J A, Proost S, Rouze P, Sterck L, Toppo S, Lazzari B, Hellens R P, Durel C E, Gutin A, Bumgarner R E, Gardiner S E, Skolnick M, Egholm M, Van de Peer Y, Salamini F, Viola R. 2010. The genome of the domesticated apple (*Malus × domestica* Borkh.). *Nat Genet*, 42: 833 - 839.
- Venturini L, Ferrarini A, Zenoni S, Tornielli G B, Fasoli M, Dal Santo S, Minio A, Buson G, Tononi P, Zago E D, Zamperin G, Bellin D, Pezzotti M, Delledonne M. 2013. De novo transcriptome characterization of *Vitis vinifera* cv. Corvina unveils varietal diversity. *Bmc Genomics*, 14: 41.
- Vezzulli S, Troglio M, Coppola G, Jermakow A, Cartwright D, Zharkikh A, Stefanini M, Grando M S, Viola R, Adam-Blondon A F, Thomas M, This P, Velasco R. 2008. A reference integrated map for cultivated grapevine (*Vitis vinifera* L.) from three crosses, based on 283 SSR and 501 SNP-based markers. *Theoretical and Applied Genetics*, 117: 499 - 511.
- Wang Z, Zhang J B, Jia C H, Liu J H, Li Y Q, Yin X M, Xu B Y, Jin Z Q. 2012. De Novo characterization of the banana root transcriptome and

- analysis of gene expression under *Fusarium oxysporum* f. sp. *ubense* tropical race 4 infection. *Bmc Genomics*, 13: 650.
- Wu J, Wang Z, Shi Z, Zhang S, Ming R, Zhu S, Khan M A, Tao S, Korban S S, Wang H, Chen N J, Nishio T, Xu X, Cong L, Qi K, Huang X, Wang Y, Zhao X, Wu J, Deng C, Gou C, Zhou W, Yin H, Qin G, Sha Y, Tao Y, Chen H, Yang Y, Song Y, Zhan D, Wang J, Li L, Dai M, Gu C, Wang Y, Shi D, Wang X, Zhang H, Zeng L, Zheng D, Wang C, Chen M, Wang G, Xie L, Sovero V, Sha S, Huang W, Zhang S, Zhang M, Sun J, Xu L, Li Y, Liu X, Li Q, Shen J, Wang J, Paull R E, Bennetzen J L, Wang J, Zhang S. 2012. The genome of the pear (*Pyrus bretschneideri* Rehd.). *Genome Res*, 23: 396 - 408.
- Xu Q, Chen L L, Ruan X, Chen D, Zhu A, Chen C, Bertrand D, Jiao W B, Hao B H, Lyon M P, Chen J, Gao S, Xing F, Lan H, Chang J W, Ge X, Lei Y, Hu Q, Miao Y, Wang L, Xiao S, Biswas M K, Zeng W, Guo F, Cao H, Yang X, Xu X W, Cheng Y J, Xu J, Liu J H, Luo O J, Tang Z, Guo W W, Kuang H, Zhang H Y, Roose M L, Nagarajan N, Deng X X, Ruan Y. 2012. The draft genome of sweet orange (*Citrus sinensis*). *Nat Genet*, 45: 59 - 66.
- Yu K, Xu Q, Da X, Guo F, Ding Y, Deng X. 2012. Transcriptome changes during fruit development and ripening of sweet orange (*Citrus sinensis*). *BMC Genomics*, 13: 10.
- Zenoni S, Ferrarini A, Giacomelli E, Xumerle L, Fasoli M, Malerba G, Bellin D, Pezzotti M, Delledonne M. 2010. Characterization of transcriptional complexity during berry development in *Vitis vinifera* using RNA-Seq. *Plant Physiology*, 152: 1787 - 1795.
- Zhang Q, Chen W, Sun L, Zhao F, Huang B, Yang W, Tao Y, Wang J, Yuan Z, Fan G, Xing Z, Han C, Pan H, Zhong X, Shi W, Liang X, Du D, Sun F, Xu Z, Hao R, Lv T, Lv Y, Zheng Z, Sun M, Luo L, Cai M, Gao Y, Wang J, Yin Y, Xu X, Cheng T, Wang J. 2012. The genome of *Prunus mume*. *Nat Commun*, 3: 1318.
- Zorrilla-Fontanesi Y, Cabeza A, Dominguez P, Medina J J, Valpuesta V, Denoyes-Rothan B, Sanchez-Sevilla J F, Amaya I. 2011. Quantitative trait loci and underlying candidate genes controlling agronomical and fruit quality traits in octoploid strawberry (*Fragaria × ananassa*). *Theoretical and Applied Genetics*, 123: 755 - 778.

## 征 订

### 欢迎订阅 2014 年《保鲜与加工》

《保鲜与加工》是由国家农产品保鲜工程技术研究中心(天津)主办的农产品采后技术研究领域的科技期刊,为中国科技核心期刊和中国北方优秀期刊,本刊已被美国《化学文摘》(CA)、中国学术期刊(光盘版)杂志收录。主要报道农产品保鲜与加工相关领域基础理论、新技术、新工艺、新设备、新材料的研究成果及国内外相关行业的动态与信息。主要设置专家论坛、保鲜研究、加工研究、专题论述、技术指南、行业资讯、科普沙龙、政策法规等栏目。适于科技人员、农业技术推广人员、相关企业管理和技术人员、大专院校师生及广大从事保鲜与加工技术研发领域的人士参阅。

本刊为双月刊,大 16 开本,64 页,逢单月 10 日出版,国内外公开发行,每期定价 10 元,全年 60 元,国内统一连续出版物号:CN12-1330/S,国际标准连续出版物号:ISSN1009-6221,邮发代号:6-146。

欢迎在全国各地邮局(所)或本编辑部订阅,欢迎广大读者踊跃投稿,并诚邀刊登各类相关广告。

通讯地址:天津市西青区津静公路 17 公里处,国家农产品保鲜工程技术研究中心(天津)《保鲜与加工》编辑部,邮编:300384

电话:022-27948711,联系邮箱:bxyjg@163.com,投稿平台:www.bxyjg.com。