

苹果果实可溶性固形物含量相关基因的全基因组筛选与分析

苑克俊*, 程来亮, 牛庆霖, 王江勇

山东省果树研究所, 山东泰安271000

摘要: 可溶性固形物含量是苹果果实品质的重要指标。在果实成熟期对苹果杂交后代的果实可溶性固形物含量进行测定, 从中选择6个可溶性固形物含量低和4个含量高的植株分成2组, 在转录组水平上对其果实的基因表达进行比较分析。结果筛选出14个差异表达基因, 其中包括蛋白激酶基因和甲基转移酶基因, 表明蛋白质磷酸化和甲基化可能和苹果果实可溶性固形物含量相关。另外, 差异表达基因还包括转录因子、细胞色素 b_6-f 复合体亚单元7、形成素类蛋白和1-脱氧-*D*-木酮糖-5-磷酸合成酶等基因。这些结果为通过深入研究找出影响苹果果实可溶性固形物含量的关键基因提供了线索, 苹果果实可溶性固形物含量和上述这些基因的关系值得进一步研究。本研究还初步证实了基于转录组测序技术的数量性状集群分离分析法是可行的。

关键词: 苹果; 果实; 可溶性固形物; 转录组; 基因

苹果果实可溶性固形物的主要成分是糖分, 其含量是一个重要的果实品质和商品价值指标, 研究果实可溶性固形物含量相关基因具有重要意义。首先, 通过在全基因组范围内研究相关基因, 找出影响果实可溶性固形物积累的关键基因, 有助于在分子辅助育种中应用或者采用基因工程手段控制果实可溶性固形物含量。过去由于实验方法和技术的限制, 在全基因组范围内进行这方面的研究很困难。近年来转录组测序技术的发展使得这方面的研究成为可能。本研究采用转录组测序技术筛选苹果果实可溶性固形物含量相关基因, 可以在转录组水平上了解影响苹果果实可溶性固形物累积的相关基因, 也有助于今后针对苹果相关基因开展深入研究, 找出关键基因, 解决目前进行苹果分子育种和遗传改良时目标基因不明确的问题。

以往对苹果果实可溶性固形物的研究主要集中在其积累规律和糖分转运机理上。研究表明, 山梨醇和蔗糖是苹果树光合作用的主要产物, 在苹果生长过程中, 由叶子制造的光合产物主要以山梨醇的形式运输到果实中, 果实内的糖以积累为主, 少部分用于生长与生命活动的维持, 大部分转变为淀粉或其他糖类物质储藏果实内(秦静远和张佐省1997; 邓丽莉和生吉萍2012)。当苹果接近成熟期或者进入成熟期时, 由于淀粉酶、转化酶及蔗糖合成酶活性的提高, 淀粉转化为葡萄糖、果糖和蔗糖等可溶性糖, 前期积累的淀粉含量急剧下降, 这些可溶性糖含量增加(丁平海和王中英1997; Berüter 2004; 邓丽莉和生吉萍2012)。

研究表明, 糖的转运载体存在于苹果果肉细胞的质膜上, 糖跨苹果果肉细胞的质膜和液泡膜运输, 是一个需要能量的主动过程(吕英民等1999)。目前, 山梨醇和蔗糖转运子均已被发现。龚晏(2008)从‘皇家嘎拉’苹果叶片中分离得到一种以跨膜蛋白结构形式存在的山梨醇转运子。Fan等(2009)在苹果中发现了一种定位于质膜上的蔗糖转运子(MdSUT1)和一种山梨醇转运子(MdSOT6), 这两种不同的转运子能与锚定于内质网的细胞色素 b_5 (MdCYB5)结合, 形成的复合物能够起到上调转运子亲和力的功能。这些结果说明, 果实可溶性固形物的形成和积累涉及一些酶和转运子, 理论上这些酶和转运子影响果实可溶性固形物含量, 也可能有未知因子影响果实可溶性固形物含量。但上述这些酶和转运子是否是关键基因有待进一步探讨。有研究者试图通过QTL定位分析找出果实可溶性固形物含量相关基因, 但结果在不同品种上并不一致(Kunihisa等2014)。

总之, 目前影响果实可溶性固形物含量的相关基因研究报道甚少。本研究首次在全基因组范围内筛选苹果果实可溶性固形物含量的直接相关基因。在分析比较可溶性固形物含量低和含量高的2组植株果实基因表达差异的基础上, 试图找出有显著性差异的基因, 探讨这些基因在苹果果实可溶性固形物累积过程中的作用, 为通过深入研

收稿 2015-02-25 修定 2016-04-28

资助 山东省泰山学者建设工程专项经费。

* 通讯作者(E-mail: yuankj@vip.sina.com)。

究找出关键基因、阐明苹果果实可溶性固形物累积的机制提供有价值的信息。

材料与方法

1 试验材料准备、可溶性固形物测定和RNA提取

试验材料采自山东省果树研究所天平湖试验果园,为9株‘富士’和‘嘎拉’苹果(*Malus domestica* Borkh.)杂交后代果实,以及1株‘富士’和未知红色苹果杂交后代(W21)果实,是从81株苹果杂交后代中筛选的。2015年8月和9月在果实成熟期采摘果实,用WZ-103手持式折光仪测定81株苹果杂交后代果实可溶性固形物含量,然后根据测定结果并参考往年测定结果,选择6个可溶性固形物含量低和4个含量高的植株,于2015年9月17日采摘果实,在其果肉部位取样,在液氮中冷却, -76°C保存,以干冰包装运输到北京在百迈客生物科技有限公司进行转录组测序实验。总RNA利用天根公司植物总RNA试剂盒提取,提取后分别采用Nanodrop、Qubit 2.0、Agilent 2100方法检测RNA样品的纯度、浓度和完整性等,以保证使用合格的样品进行转录组测序。

2 RNA文库构建和测序数据分析

样品检测合格后,构建RNA文库,建库过程中采用了RNA样品准备试剂盒[E7490, NEBNext® Poly(A) mRNA Magnetic Isolation Module]。建库主要流程如下:(1)用带有Oligo(dT)的磁珠富集真核生物mRNA。(2)将Fragmentation缓冲液和mRNA溶液置于一个PCR管中随机打断mRNA。(3)以mRNA为模板,用六碱基随机引物合成第一条cDNA链。然后加入缓冲液、dNTPs、RNase H和DNA polymerase I合成第二条cDNA链,利用AMPure XP beads纯化cDNA。(4)纯化的双链cDNA再进行末端修复、加A尾并连接测序接头,用AMPure XP beads进行片段大小选择。(5)最后通过PCR富集得到cDNA文库。

文库构建完成后,分别使用Qubit2.0和Agilent 2100对文库的浓度和插入片段大小进行检测:(1)使用Qubit2.0进行初步定量,使用Agilent 2100对文库的插入片段大小进行检测,插入片段大小符合预期后才可进行下一步实验。(2)用Q-PCR方法对文库的有效浓度进行准确定量,完成库检。库检合格

后,用IlluminaHiSeq测序平台进行高通量测序。

对原始数据进行过滤,去除其中的接头序列及低质量读数获得高质量的合格数据(clean data)。利用TopHat软件(Trapnell等2009),将合格数据与指定的参考基因组进行序列比对,获得比对到参考基因组转录本上的基因组匹配数据(mapped data)。本研究主要是利用MalDomGD1.0版参考基因组(http://www.ncbi.nlm.nih.gov/assembly?LinkName=genome_assembly&from_uid=358)进行分析。基于基因组匹配数据,进行插入片段长度检验、随机性检验等测序文库质量评估,进行表达量分析、新基因发掘等。根据基因在不同样品或不同样品组中的表达量进行差异表达分析、差异表达基因功能注释等高级分析。为了解某些信息也利用GDR数据库(<http://www.rosaceae.org/>)中的苹果参考基因组V1.0p进行了分析,基因功能注释是2个参考基因组的综合分析结果。

3 差异表达基因的鉴定

为了让片段数目能真实地反映转录本表达水平,需要对样品中的基因组匹配读数(mapped reads)的数目和转录本长度进行归一化。本研究采用Cufflinks软件的Cuffquant和Cuffnorm组件进行分析,采用FPKM (fragments per kilobase of transcript per million fragments mapped)作为衡量转录本或基因表达水平的指标(Trapnell等2010),FPKM计算公式如下:

$$\text{FPKM} = \frac{\text{比对到某一转录本上的片段数目}}{[\text{比对到转录本上的片段总数}(\text{百万}) \times \text{转录本长度}(\text{kb})]} \times 10^6$$

公式中,比对到某一转录本上的片段数目指双端Reads数目;比对到转录本上的片段总数,以 10^6 为单位;转录本长度,以 10^3 个碱基为单位。

差异表达分析使用DESeq软件进行(Anders和Huber 2010),获得两个样品之间的差异表达基因集。在差异表达基因检测过程中,将差异倍数(fold change) ≥ 2 且错误发现率(false discovery rate, FDR) < 0.01 作为筛选标准。这里,差异倍数表示两样品(组)间表达量的比值,错误发现率是通过差异显著性P值(P-value)进行校正得到的。由于转录组测序的差异表达分析是对大量的基因表达值进行独立的统计假设检验,会存在假阳性问题,因此在进行差异表达分析过程中,采用了公认的Benjami-

ni-Hochberg校正方法对原有假设检验得到的显著性 P 值进行校正。

4 基因注释

采用Blast软件中的blastx程序, 以获得的差异表达基因核酸序列作为query序列, 与对应的蛋白序列数据库进行比对, 从比对结果中选取分值最高的一条序列作为比对结果, 并附上对应数据库的注释信息作为获得基因的注释信息。对应的蛋白序列数据库包括NR和Swiss-Prot数据库, 其全称分别为非冗余蛋白质序列数据库(non-redundant protein sequences)和UniProtKB/Swiss-Prot数据库(Altschul等1997; 邓泱泱等2006; Apweiler等2004)。

表1 10株杂交后代植株的果实可溶性固形物含量(%)

Table 1 The soluble solids content (%) in the fruits of 10 apple hybrid plants

采收日期	B17	B7	B25	71	W21	118	B5	108	129	142
8月31日	11.15	9.82	9.57	10.78		11.03			18.60	15.63
9月15日	11.57	11.13	10.67	11.90	11.72	11.03	15.73	15.03	22.37	16.22
9月22/23日	13.23	11.53	12.40	11.10	12.20	12.57	16.80			17.23

11.34%, 含量较高一组植株果实的可溶性固形物含量为15.03%~22.37%, 平均值17.34%, 不包括129号的高含量组其他3个植株果实的可溶性固形物含量平均值为15.66%; 高含量组中最低的可溶性固形物含量与低含量组中最高可溶性固形物含量差异达到3.13%以上(15.03%~11.90%=3.13%)。

从图1可看出, 可溶性固形物含量基本上呈正态分布; 含量较低一组植株果实在全部样品中处于可溶性固形物含量的低端, 含量较高一组植株果实在全部样品中处于可溶性固形物含量的高端。 t 测验结果表明, 2组果实的可溶性固形物含量差异达到极显著水平, 故选择用来进行转录组分析筛选差异基因。需要说明的是, 无论包含还是不包含129号样品数据, t 测验时2组果实的可溶性固形物含量差异均达到极显著水平。

2 转录组测序结果

10个杂交后代植株果实样品的主要测序结果见表2。每个样品获得945~1 161万个Q30大于90.6%的质量合格读数(clean reads), 72.18%~79.74%的读数被比对到基因数据库中的参考基因组Mal-DomGD1.0版。

实验结果

1 果实可溶性固形物含量

从表1可看出, B17、B7、B25、71、W21和118这6个杂交后代植株的果实可溶性固形物含量较低, B5、108、129和142这4个杂交后代植株的果实可溶性固形物含量较高。相同采收日期结果比较, 2组果实的可溶性固形物含量差异达到3%以上。需要说明的是, 129号杂交后代植株果实可溶性固形物含量明显高于其他杂交后代植株果实, 这是由于其成熟期较早, 较早开始糖分积累。

9月15日结果比较, 含量较低一组植株果实的可溶性固形物含量为10.67%~11.90%, 平均值

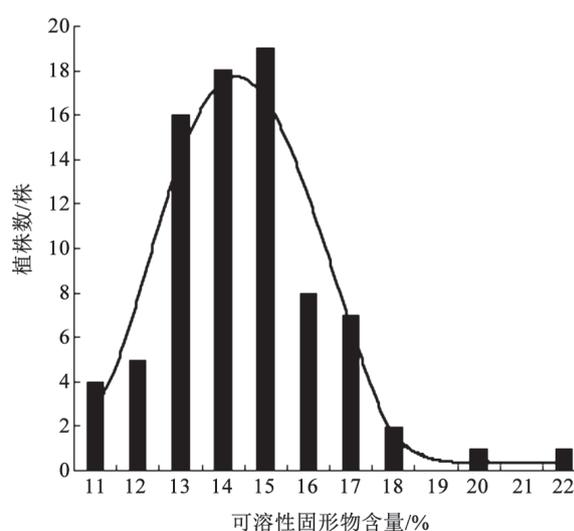


图1 81个苹果杂交植株果实的可溶性固形物含量
Fig.1 The soluble solids content in the fruits of 81 apple hybrid plants

3 基因表达分析

从表3可看出, 以差异倍数 ≥ 2 且错误发现率 < 0.01 作为筛选标准, 共找出14个差异表达基因, 其中10个基因上调表达, 4个基因下调表达(图4)。

表2 10个杂交后代植株果实样品测序数据统计

Table 2 Summary of transcriptome sequencing results of the fruits from 10 hybrid plants

杂交植株	原始读数	质量合格读数	Q30/%	基因组匹配读数/%	质量合格碱基数	GC含量/%
B17	18 416 418	11 609 847	90.61	79.74	3 504 752 877	47.64
B7	15 640 357	9 969 382	90.98	77.55	3 008 560 355	47.44
B25	14 817 632	9 452 079	90.97	77.21	2 852 259 318	48.00
71	11 096 246	10 907 821	91.23	72.75	3 271 141 874	48.26
W21	15 483 934	9 893 208	91.05	76.95	2 984 732 497	47.97
118	10 902 854	10 727 123	90.66	72.18	3 217 264 092	48.37
B5	15 257 841	9 729 228	90.98	77.68	2 936 563 708	47.61
108	9 983 893	9 756 447	91.26	72.85	2 925 717 106	48.72
129	16 540 988	10 712 015	91.13	76.73	3 231 062 807	47.45
142	16 800 806	10 834 683	90.92	77.78	3 268 818 230	47.75

表3 两组果实的差异表达基因及其表达量(FPKM)

Table 3 The differentially expressed genes of the fruits from two group of hybrid plants and their expressions (FPKM)

基因编号	B17	B7	B25	71	W21	118	B5	108	129	142
gene56697	71	57	106	65	63	44	18	12	20	18
gene28440	17	18	26	16	27	29	6	3	10	6
gene16639	3	1	2	1	2	1	0	0	0	0
gene39722	33	43	56	34	30	39	12	13	4	18
gene5817	1	1	1	1	2	0	4	9	5	8
gene22267	1	2	0	3	0	0	13	6	10	9
gene26341	0	8	0	0	0	7	20	13	35	23
apple_new gene1574	4	7	3	5	5	10	20	24	27	17
gene18590	8	1	13	15	22	6	30	44	44	35
gene21069	35	13	13	65	56	44	82	119	123	102
gene11018	69	44	8	41	13	47	164	93	276	219
gene7517	0	0	0	0	0	0	1	1	1	1
gene51961	1	0	0	0	0	0	4	5	10	4
gene6513	0	0	0	0	0	0	3	5	9	6

从表4中差异倍数的自然对数值可看出, 2组果实间有2个蛋白激酶的基因表达存在明显差异, 富含半胱氨酸的受体类蛋白激酶基因下调表达, 富含亮氨酸的重复受体类丝氨酸/苏氨酸蛋白激酶基因上调表达, 表明蛋白质磷酸化可能和苹果果实可溶性固形物含量相关。另外还可看出, 2组果实间有2个甲基转移酶(茉莉酸O-甲基转移酶和S-腺苷蛋氨酸依赖型甲基转移酶)的基因表达存在明显差异, 且皆上调表达, 表明甲基化可能和苹果果实可溶性固形物含量相关。

两组果实间基因表达存在明显差异的基因还有一个形成素类蛋白5和一个转录因子, 其中形成素类蛋白5基因上调表达, 转录因子基因下调表达。

此外, 2组果实间基因表达存在明显差异的基

因还有一个丝氨酸/精氨酸重复基质蛋白1 (serine/arginine repetitive matrix protein 1)基因、一个节-1类同源框蛋白3 (homeobox protein knotted-1-like 3)基因、一个细胞色素b₆-f复合体亚单元7基因、一个蛋白质KRI1的同源蛋白基因、一个1-脱氧-D-木酮糖-5-磷酸合成酶2基因、一个茎特异表达蛋白基因、一个胁迫相关蛋白基因和一个未知基因gene6513, 表明这些基因也可能和苹果果实可溶性固形物含量相关。

讨 论

果实可溶性固形物含量是一个重要的苹果品质指标。过去通过QTL分析进行果实可溶性固形物含量相关基因的基因座定位, 然后再通过染色

表4 两组果实的差异表达基因注释

Table 4 The annotation of differentially expressed genes of the fruits from two groups of hybrid plants

基因编号	差异倍数的自然对数值	蛋白质数据库Swissprot中的基因注释	基因数据库NR中的基因注释
gene56697	-2.02304	细胞色素b ₆ -f复合体亚单元7	细胞色素b ₆ -f复合体亚单元7 [苹果]
gene28440	-1.82244	-	丝氨酸/精氨酸重复基质蛋白质1 [苹果]
gene16639	-5.59714	富含半胱氨酸的受体类蛋白激酶3	富含半胱氨酸的受体类蛋白激酶2 [苹果]
gene39722	-1.73181	转录因子	转录因子 [苹果]
gene5817	3.068956	富含亮氨酸的重复受体类丝氨酸/苏氨酸蛋白激酶	富含亮氨酸的重复受体类丝氨酸/苏氨酸蛋白激酶 [苹果]
gene22267	3.245541	茉莉酸O-甲基转移酶	茉莉酸O-甲基转移酶类 [苹果]
gene26341	3.132912	可能的1-脱氧-D-木酮糖-5-磷酸合成酶2	可能的1-脱氧-D-木酮糖-5-磷酸合成酶2 [苹果]
apple_new	1.894068	形成素类蛋白5	形成素类蛋白5 [梨]
gene1574			
gene18590	1.828385	茎特异表达蛋白TSJT1 [烟草]	茎特异表达蛋白TSJT1 [苹果]
gene21069	1.592769	胁迫相关蛋白	胁迫相关蛋白 [苹果]
gene11018	2.326549	可能的S-腺苷蛋氨酸依赖型甲基转移酶	可能的S-腺苷蛋氨酸依赖型甲基转移酶 [苹果]
gene7517	4.725276	蛋白质KRI1的同源蛋白	蛋白质KRI1的同源蛋白 [苹果]
gene51961	4.425881	节-1类同源框蛋白	节-1类同源框蛋白3 [苹果]
gene6513	6.724206	-	未鉴定蛋白质

-: 无注释。

体步移找出关键基因, 费时费力。近年来转录组测序技术的发展使得在全基因组范围内直接筛选相关基因不仅成为可能, 而且获得的信息量大。为此, 本研究采用转录组测序技术筛选苹果果实可溶性固形物含量相关基因。本方法是基于如下假设: 在大量样品中筛选出的2组样品, 一组样品果实可溶性固形物含量较低, 另一组样品果实可溶性固形物含量较高, 2组样品的相关基因表达量差异大, 包括主效基因在内的相关基因造成2组样品可溶性固形物含量的差异。本方法的关键是不采用可溶性固形物含量在中间范围的大量样品, 而是在全部样品中筛选出可溶性固形物含量在低端和高端的2组样品进行实验, 以保证存在影响2组样品可溶性固形物含量差异的主要基因。本方法实际上是集群分离分析法(Bulk Segregation Analysis)在数量性状上的应用, 是我们在转录组测序技术出现后发展出的数量性状集群分离分析法。本研究结果已初步证明了方法的可行性。

本研究在转录组水平上对可溶性固形物含量低和含量高的2组果实基因表达进行比较分析, 结果筛选出14个差异表达基因, 包括蛋白激酶基因、甲基转移酶基因、转录因子、细胞色素b₆-f复合体亚单元7、蛋白质KRI1的同源蛋白、形成素类蛋白和1-脱氧-D-木酮糖-5-磷酸合成酶2等基

因。这些结果为深入研究苹果果实可溶性固形物含量相关基因提供了线索, 苹果果实可溶性固形物含量和上述这些基因的关系值得进一步研究。目前的问题是, 在这些基因中哪些基因影响了苹果果实可溶性固形物的积累, 哪些基因表达的差异是由于苹果果实可溶性固形物含量的差异造成的。

有报道糖可以诱导质膜联系的Ca²⁺依赖型蛋白激酶(CDPKs)的活性, 活化后的CDPKs通过催化蛋白质磷酸化而激活糖的载体或H⁺-ATPase (Ohto和Nakamura 1995; 吕英民等1999)。本研究未发现差异表达的质膜联系的Ca²⁺依赖型蛋白激酶, 但发现了一个富含半胱氨酸的受体类蛋白激酶基因和一个富含亮氨酸的重复受体类丝氨酸/苏氨酸蛋白激酶基因。既然差异表达基因包括两种蛋白激酶基因以及两种甲基转移酶基因, 就表明蛋白质磷酸化和甲基化可能与苹果果实可溶性固形物含量相关。目前, 在苹果上有蛋白质磷酸化对淀粉酶激活的研究报道(张凌云等2002), 未见有关甲基化的研究报道。今后可溶性固形物高含量和低含量果实中的磷酸化蛋白质、甲基化基因和蛋白质应该是研究的一个重点。

有研究表明, VvSUC11和VvSUC12两种糖转录因子的转录水平在果实成熟期会出现上调表达现象(Davies等1999; 王西成等2015)。但若于葡萄果

实着色度达50%时对其进行1-MCP处理, 不仅会导致果实中蔗糖水平的下降, 同时还会抑制*VvSUC11*和*VvSUC12*基因的表达(Chervin等2006; 王西成等2015)。本研究发现的一个转录因子基因则下调表达。

众所周知, 苹果果实糖分主要通过蔗糖和山梨醇进行转运。Fan等(2009)在苹果中发现了一种定位于质膜上的蔗糖转运子(MdSUT1)和一种山梨醇转运子(MdSOT6), 这两种不同的转运子能与锚定于内质网的细胞色素b5 (MdCYB5)结合, 形成的复合物能够起到上调转运子亲和力的功能。本研究未发现细胞色素b5, 但发现了细胞色素b_{6-f}。苹果果实中的细胞色素b_{6-f}和苹果果实可溶性固形物积累的关系值得进一步探讨。

1-脱氧-D-木酮糖-5-磷酸合成酶2 (DXS)是MEP途径中的第一个关键酶, 在萜类化合物合成过程中发挥重要作用(魏麟等2014)。本研究在可溶性固形物含量低和含量高的2组果实间发现该DXS基因的表达有显著性差异, 其作用非常值得进一步探讨。

综合前述, 我们通过研究获得了可溶性固形物含量高和含量低的2组果实中的14个差异表达基因, 在前人研究苹果果实糖分积累和转运机理的基础上, 进一步探讨了影响苹果果实可溶性固形物积累的相关基因, 这其中一些基因与苹果果实含糖量或可溶性固形物含量的关系未见报道, 例如茉莉酸O-甲基转移酶和S-腺苷蛋氨酸依赖型甲基转移酶基因。这些结果为深入研究苹果果实可溶性固形物积累机制提供了有价值的信息。例如, 根据这些结果非常有必要进一步探讨蛋白质磷酸化和甲基化在苹果果实可溶性固形物积累过程中的作用以及影响苹果果实可溶性固形物积累的关键基因。今后应通过深入研究, 进一步探明已找出的众多基因与苹果果实可溶性固形物积累的关系。最后需要指出的是, 本研究获得的只是转录组初步筛选结果, 有些基因需要在后续研究中进行进一步的验证。

参考文献

- Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997). Gapped BLAST and PSI BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, 25 (17): 3389–3402

- Anders S, Huber W (2010). Differential expression analysis for sequence count data. *Genome Biol*, 11 (10): R106
- Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, et al (2004). UniProt: the universal protein knowledgebase. *Nucleic Acids Res*, 32 (Database issue): D115–D119
- Berüter J (2004). Carbohydrate metabolism in two apple genotypes that differ in malate accumulation. *J Plant Physiol*, 161 (9): 1011–1029
- Chervin C, Terrier N, Ageorges A, Ribes F, Kuapunyakoon T (2006). Influence of ethylene on sucrose accumulation in grape berry. *Am J Enol Viticult*, 57 (4): 511–513
- Davies C, Wolf T, Robinson SP (1999). Three putative sucrose transporters are differentially expressed in grapevine tissues. *Plant Sci*, 147 (2): 93–100
- Deng LL, Sheng JP (2012). Research progress in sugar metabolism and its regulation in apple. *Stor Proc*, 12 (1): 1–5, 11 (in Chinese with English abstract) [邓丽莉, 生吉萍(2012). 苹果果实糖代谢过程及其调控研究进展. 保鲜与加工, 12 (1): 1–5, 11]
- Deng YY, Li JQ, Wu SF, Zhu YP, Chen YW, He FC (2006). Integrated nr database in protein annotation system and its localization. *Comput Engin*, 32 (5): 71–73, 76 (in Chinese with English abstract) [邓泱泱, 荔建琦, 吴松峰, 朱云平, 陈耀文, 贺福初(2006). nr数据库分析及其本地化. 计算机工程, 32 (5): 71–73, 76]
- Ding PH, Wang ZY (1997). A dynamic study of carbohydrates in the leaf and fruit of Starkrimson. *Hebei J Forest Orchard Res*, 12 (1): 48–52 (in Chinese with English abstract) [丁平海, 王中英(1997). 新红星苹果叶片和果实中碳水化合物动态变化. 河北林果研究, 12 (1): 48–52]
- Fan RC, Peng CC, Xu YH, Wang XF, Li Y, Shang Y, Du SY, Zhao R, Zhang XY, Zhang LY, Zhang DP (2009). Apple sucrose transporter SUT1 and sorbitol transporter SOT6 interact with cytochrome b5 to regulate their affinity for substrate sugars. *Plant Physiol*, 150 (4): 1880–1901
- Gong Y (2008). Cloning and sequence analysis of cDNA fragments of sorbitol transporters and sobitol dehydrogenase from apple leaf (Master's thesis). Yangling: Northwest A&F University (in Chinese with English abstract) [龚晏(2008). 山梨醇转运子和SDH基因cDNA的克隆及其序列分析(硕士论文). 杨凌: 西北农林科技大学]
- Kunihisa M, Moriya S, Abe K, Okada K, Haji T, Hayashi T, Kim H, Nishitani C, Terakami S, Yamamoto T (2014). Identification of QTLs for fruit quality traits in Japanese apples: QTLs for early ripening are tightly related to preharvest fruit drop. *Breed Sci*, 64 (3): 240–251
- Lü YM, Zhang DP, Yan HY (1999). Sugar unloading mechanisms in the developing apple fruit. *Acta Horti Sin*, 26 (3): 3–8 (in Chinese with English abstract) [吕英民, 张大鹏, 严海燕(1999). 糖在苹果果实中卸载机制的研究. 园艺学报, 26 (3): 3–8]
- Ohto M, Nakamura K (1995). Sugar-induced increase of calcium-dependent protein kinases associated with the plasma membrane in leaf tissues of tobacco. *Plant Physiol*, 109: 973–998

- Qin JY, Zhang ZS (1997). The transmit of the material of apple in mature period. *Acta Bot Boreal-Occident Sin*, 17 (6): 173–177 (in Chinese) [秦静远, 张佐省(1997). 苹果成熟过程中的物质转化. *西北植物学报*, 17 (6): 173–177]
- Trapnell C, Pachter L, Salzberg SL (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, 25 (9): 1105–1111
- Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*, 28 (5): 511–515
- Wang XC, Wu WM, Zhao MZ, Qian YM, Wang ZW (2015). Progress in the research of physiological and molecular mechanism in development and ripening of grape fruit regulated by hormones. *Acta Agric Univ Jiangxiensis*, 37 (4): 604–611 (in Chinese with English abstract) [王西成, 吴伟民, 赵密珍, 钱亚明, 王壮伟 (2015). 激素调控葡萄果实发育与成熟的生理与分子机制研究进展. *江西农业大学学报*, 37 (4): 604–611]
- Wei L, Wu XJ, Li SH, Liu SG, Tang YL, He AN, Wang D, Ning PF, Li ZJ (2014). Cloning and expression analysis of 1-deoxy-*D*-xylulose-5-phosphate synthase 1 gene in *Houttuynia cordata*. *Chin Tradit Herba Drugs*, 45 (11): 1607–1612 (in Chinese with English abstract) [魏麟, 伍贤进, 李胜华, 刘胜贵, 唐玉莲, 贺安娜, 王丹, 宁鹏飞, 李昭君(2014). 鱼腥草1-脱氧-*D*-木酮糖-5-磷酸合成酶1基因克隆与表达分析. *中草药*, 45 (11): 1607–1612]
- Zhang LY, Zhang N, Pan QH, Duan CQ, Zhang DP (2002). Amylase activities are stimulated by protein phosphorylation in apple fruit. *Acta Horti Sin*, 30 (4): 455–458 (in Chinese with English abstract) [张凌云, 张娜, 潘秋红, 段长青, 张大鹏(2002). 苹果果实中蛋白质磷酸化对淀粉酶的激活. *园艺学报*, 30 (4): 455–458]

Genome-wide screen and analysis of the genes related to soluble solids content in apple fruits

YUAN Ke-Jun*, CHENG Lai-Liang, NIU Qing-Lin, WANG Jiang-Yong
Shandong Institute of Pomology, Tai'an, Shandong 271000, China

Abstract: Soluble solids content in fruits is an important apple quality factor, it is very important to have a better understanding of the genes related to soluble solids content in apple fruits. In order to make clear whether there are the genes related to soluble solids content in apple fruits, one group of 6 hybrid plants with low and another group of 4 hybrid plants with high content of soluble solids were selected, a genome-wide transcriptome analysis was applied to analyze the gene expression of the fruits from these 10 hybrid plants. Comparative analysis of the gene expression profile revealed 14 significant differentially expressed genes between two groups of hybrid plants. Among them, some genes were recognised as protein kinase and methyltransferase, indicating that there are protein phosphorylation and methylation in apple fruits. In addition, transcription factor, cytochrome b_6 - f complex subunit 7, formin-like protein, 1-deoxy-*D*-xylulose-5-phosphate synthase and some other genes were also recognised. These results may provide some help for further studies and some clues for screening key genes related to soluble solids content in apple fruits. Further research is needed to test the role of these 14 genes. Our results also confirmed the feasibility of bulk segregation analysis in analysing quantitative traits based on transcriptome sequencing.

Key words: apple; fruit; soluble solids; transcriptome; gene

Received 2016-02-25 Accepted 2016-04-28

This work was supported by Construction Project for Taishan Scholar of Shandong Province, China.

*Corresponding author (E-mail: yuankj@vip.sina.com).