Biosci. Biotechnol. Biochem., **77** (11), 2169–2174, 2013

Transcriptome Analysis to Identify Putative Floral-Specific Genes and Flowering Regulatory-Related Genes of Sweet Potato

Xiang TAO,[1,2] Ying-Hong GU,[1] Yu-Song JIANG,[1] Yi-Zheng ZHANG,[1] and Hai-Yan WANG[1,†]

[1]*College of Life Sciences, Key Laboratory of Bio-resources and Eco-environment, Ministry of Education, Sichuan Key Laboratory of Molecular Biology and Biotechnology, Sichuan University, 29 Wangjiang Road, Chengdu, Sichuan, 610064, China*
[2]*Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, Sichuan, 610041, China*

Received March 18, 2013; Accepted July 30, 2013; Online Publication, November 7, 2013
[doi:10.1271/bbb.130218]

Sweet potato flowers were collected for a transcriptome analysis to identify the putative floral-specific and flowering regulatory-related genes by using the RNA-sequencing technique. Pair-end short reads were *de novo* assembled by an integrated strategy, and then the floral transcriptome was carefully compared with several published vegetative transcriptomes. A total of 2595 putative floral-specific and 2928 putative vegetative-specific transcripts were detected. We also identified a large number of transcripts similar to the key genes in the flowering regulation network of *Arabidopsis thaliana*.

Key words: sweet potato; transcriptome; floral-specific gene; flowering regulation

Sweet potato [*Ipomoea batatas* L. (Lam.)] is one of the seven most important crops in the word for annual hectareage and total production. It is widely grown around the world due to its strong adaptability, easy management, rich nutrient content and multiple usage.[1)] The annual tuberous root production of this crop is more than 105 million metric tons, 95% of which is produced in developing countries. Sweet potato is a hexaploid and generally performs asexual reproduction;[2)] the process of sexual reproduction, especially the flowering regulation, is therefore less studied.

Sweet potato is a thermophilic, short day root crop. Most sweet potato cultivars can bloom under natural growth conditions in tropical and subtropical areas, although, in the temperate zone, *i.e.*, north of 23 degrees North latitude, sweet potato does not generally bloom because of the short growing season. However, it can bloom in the temperate zone when exposed to biotic or abiotic stress, including the stress generated by grafting to a closely allied species, being subjected short days or to a growth regulator, drought or salt stress. *Ipomoea carnea* has been found as the best stock for grafting[3)] and an 8 h light/16 h dark day, and GA3/GA7 are good choices for short-day and growth regulator treatments, respectively.[4)] However, all of these treatments more or less change the normal living conditions for sweet potato, and the flowering induction efficiency is usually extremely low.

Why do these artificial treatments induce sweet potato blooming, and what kinds of genes are involved in this process?[5)] To get some putative floral-specific genes and flowering regulatory genes that would provide guidance for studies on flowering regulation and further provide assistance for hybrid breeding, we induced sweet potato flowering by drought stress, and collected flowers for a transcriptomic analysis by the next-generation sequencing (NGS) technique. We compared two published vegetative transcriptomes[6,7)] and the digital gene expression (DGE) profiles of seven vegetative tissues[6)] and identified 2595 putative floral-specific expressed transcripts (F-SETs). We also mined such flowering regulator encoding genes as *LEAFY* (*LFY*), *AGAMOUS* (*AG*), *TERMINAL FLOWER* (*TFL*), *FLOWERING LOCUS C* (*FLC*), *CONSTANS* (*CO*), *APETALA1* (*AP1*), *APETALA2* (*AP2*), *APETALA3* (*AP3*), *DELLA*, and *SLEEPY1* (*SLY*1).

Materials and Methods

Plant materials and total RNA preparation. Stem cuts of sweet potato [*I. batatas* L. (Lam.) cv. Xushu 18] were planted in May 2010 and grown under natural conditions during the first 20 d and then under drought stress during the next 20 d in Chengdu, Sichuan Province of China. After successfully inducing flowering, the whole opened flowers, including petals, stigmas, styles, anthers, filaments, ovules and ovaries, were collected and immediately snap-frozen in nitrogen until needed for further processing. Total RNAs were extracted by using the Trizol® reagent (Invitrogen, USA) according to the manufacturer's instructions. Genomic DNA was removed by DNase I (Fermentas, USA). The RNA's concentration, purity and integrity number (RIN) were assessed by SMA3000 and/or an Agilent 2100 bioanalyzer.

RNA sequencing by NGS. Qualified total RNA was submitted to Beijing Genomics Institute (BGI)-Shenzhen, Shenzhen, China (http://www.genomics.cn). A fragmented cDNA library was constructed by using an RNA-Seq 8-Sample Prep kit (Illumina, USA) according to the manufacturer's instructions, before the library was sequenced on the Illumina HiSeq 2000 platform. Raw sequence data were generated by the Illumina pipeline and are available in the NCBI Short Read Archive (SRA) database (http://www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?) under accession number SRA043584.

De novo *assembly and evaluation.* Raw paired-end (PE) reads were assessed by NGS: QC and manipulation tools on the Galaxy website

† To whom correspondence should be addressed. Tel/Fax: +86-28-85412738; E-mail: hayawang@scu.edu.cn

Abbreviations: DGE, digital gene expression profiling; F-SET, putative floral-specific expressed transcript; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; NGS, next-generation sequencing technology; NR, NCBI non-redundant protein database; RNA-Seq, RNA-sequencing; V-SET, putative vegetative-specific expressed transcript; TPM, the number of transcripts per million clean tags; RPKM, reads per kilobase per million mapped reads

(http://main.g2.bx.psu.edu/).[8,9] Sequence assembly was run on a 64-bit Linux computer system (Ubuntu 10.10), using Edena v2.1.1,[10] SOAPdenovo v1.3,[11] Oases v0.1.20[12,13] and CAP3[14] under various parametric choices by an integrated strategy.[6] Statistical data for each assembly were generated according to the assembled contigs by common perl scripts. The five best assemblies were chosen for each assembler, including assemblies produced by parametric choices of m57, m59, m61, m63 and m65 for Edena, K35, K37, K39, K41 and K43 for SOAPdenovo, and K27, K29, K31, K33 and K35 for Oases. All of these 15 assemblies were pooled together with the published vegetative transcriptome of Xushu 18 and then reassembled by CAP3 through three steps under the parametric choices described in our previous study.[6]

Statistical data for the final assembly were generated by common perl scripts. Open reading frames (ORFs) of each transcript were scanned by the EMBOSS package[15] and then submitted to calculate the proportion of long ORF-containing transcripts to the corresponding length transcripts. For example, if these are 10,000 transcripts ≥900 bp, with 6,000 of them possessing ORF ≥900 bp, this proportion is 60%.

Expression profiling. We investigated the expression level of each transcript in different samples by using the PE reads of two previously published transcriptomes and the PE reads of the floral transcriptome for a mapping analyses by using Bowtie,[16] allowing no more than two nucleotide mismatch. The reads per kilobase per million mapped reads (RPKM)[17] was calculated and used for quantifying each transcript. All 21 bp DGE tags were extracted, filtered by custom perl scripts and transformed into clean tags as previously described.[18] These clean tags were also used to map all transcripts allowing no more than one nucleotide mismatch, then the number of transcripts per million clean tags (TPM)[19] was calculated for each transcript.

Functional annotation. A sequence similarity search was conducted by Blast2GO v2.4,[20] using BLASTX against the NCBI's non-redundant (NR) protein database with an e-value ≤1e-3, and the enzyme codes and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways[21] were retrieved from the KEGG web server (http://www.genome.jp/kegg/). The Gene Ontology (GO)[22] results were visualized by using the WEGO tools.[23]

Results

De novo *assembly and quality assessment*

The RNA-Seq reads were assembled by the integrated strategy described in a previous study.[6] We constructed a comprehensive transcriptome for sweet potato containing transcripts expressed in the leaves, stems, roots and flowers. There were 70,412 contigs with length ≥200 bp (Table 1), and 47,676 contigs with length ≥300 bp (Table 2). We predicted ORF from each contig to assess the quality of this transcriptome. Compared with the previous vegetative transcriptome for sweet potato cv. Xushu 18, we obtained more long transcripts (Table 1) and more long ORF-containing transcripts (Table 2) by integrating the vegetative and floral transcriptomes in the assembly step. For example, there were 23,505 contigs with length ≥600 bp, 61.3% of them containing ORF ≥600 bp. This ORF-containing rate for contigs with length ≥900 bp is 56.0%, and 50.2% for contigs with length ≥1200 bp. All of these sequences have been deposited in NCBI SRA databases under the PRJNA187421accession number.

Table 1. Assembly Quality Statistics of the Vegetative and the Comprehensive Transcriptome

VEG, vegetative transcriptome of Xushu 18; VEG/FLO, comprehensive transcriptome including the floral and the vegetative ones.

	VEG	VEG/FLO
≥2000 bp	1072	2407
≥1000 bp	7667	12,434
Average length	567	628
Max length	5466	7619
N50 length	740	895
Number of contig	55,468	70,412

Functional annotation and expression analyses

Functional annotation of contigs was conducted by using Blast2GO v2.4.[20] Based on a BLASTX search against the NR database, 45,698 of the 70,412 contigs (64.90%) had BLASTX hits (Table S1; see *Biosci. Biotechnol. Biochem.* Web site). Annotation information could be obtained for 91.48% of contigs ≥500 bp, this value increasing to 98.77% for contigs ≥1,000 bp (Fig. 1).

The expression level for each contig was calculated according to the mapping results of the PE reads. We standardized the expression level by the RPKM[17] algorithm for each contig to eliminate the bias introduced by RNA composition, sequencing depth and transcript length. Of the 45,698 annotated transcripts, the transcripts encoding the following functional critical proteins were highly expressed in the flowers: the tonoplast intrinsic protein (Contig_10688, 2,067 RPKM), lipid transfer protein (Contig_9199, 2,080 RPKM), putative ripening protein (Contig_14314, 3,349 RPKM), proline-rich protein (Contig_6428, 6,691 RPKM), 1-aminocyclopropane-1-carboxylate oxidase (Contig_5857, 1,505 RPKM), osmotin-like protein (Contig_12307, 1,633

Table 2. ORF Prediction of the Assembly

VEG, vegetative transcriptome of Xushu 18; VEG/FLO, comprehensive one.

		VEG	VEG/FLO
Length ≥300 bp	Total seqs	37,179	47,676
	ORFs	25,771	35,770
Length ≥600 bp	Total seqs	16,781	23,505
	ORFs	9933	14,398
Length ≥900 bp	Total seqs	9243	14,498
	ORFs	5091	8115
Length ≥1,200 bp	Total seqs	5237	9107
	ORFs	2649	4572
Length ≥1,500 bp	Total seqs	2971	5765
	ORFs	1695	2494

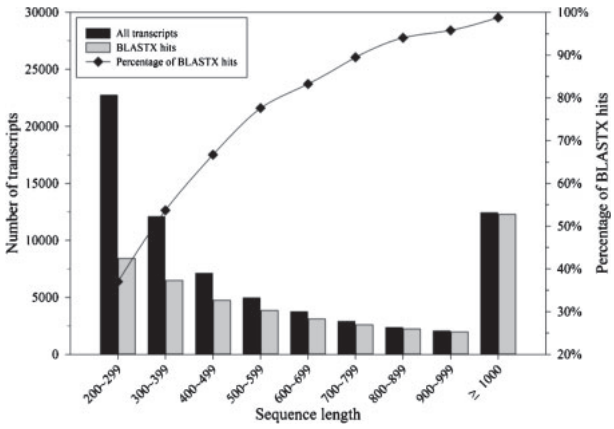

Fig. 1. Distribution of Transcripts in the Length and Percentage of Transcripts with BLASTX Hit.

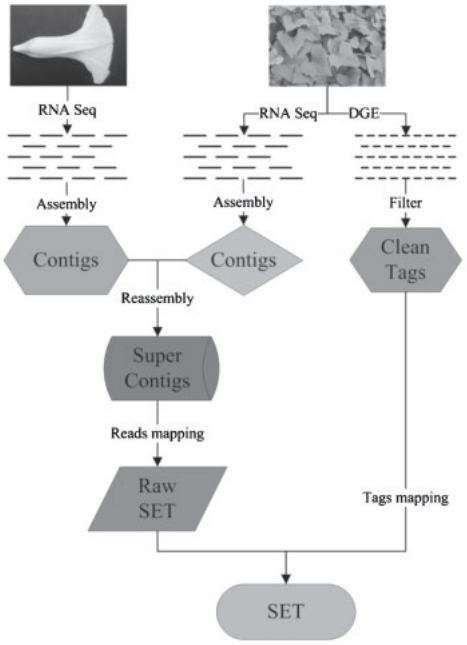

Fig. 2. Pipeline for Identifying Specific Expressed Transcripts.
The Illumina short-reads were assembled by using an integrated strategy.[6] 34 contig sets obtained from vegetative and/or floral transcriptome were pooled together and reassembled by CAP3.[14] To identify the putative specific expressed transcripts (SETs), all of the reads from the floral and vegetative transcriptome of Xushu 18[6] and the root transcriptome of Guangshu 87[7] were used for a mapping analysis by the method of Bowtie.[16] SETs identified by mapping the reads were further verified by mapping the DGE clean tags.

RPKM), xyloglucan endotransglucosylase/hydrolase 2 (Contig_42616, 1,520 RPKM), and pollen-specific protein C13 precursor (Contig_40477, 830 RPKM) (Table S1).

Putative floral- and vegetative-specific transcripts

Specifically expressed transcripts usually play a vital role in organ identity or maintaining a specific function. All of the previously published NGS data of sweet potato were employed for a specific expression analysis to identify some putative floral- and vegetative-specific transcripts in this study (Fig. 2). Those transcripts whose mapped PE reads and DGE tags numbers were zero in the vegetative organs but not in the flowers were defined as putative floral-specific expressed transcripts (F-SETs), while the opposite types were defined as putative vegetative-specific expressed transcripts (V-SETs). We identified 2595 F-SETs and 2928 V-SETs. Among these, we obtained annotation information for 1345 F-SETs and 1429 V-SETs (Table S2). The average lengths of the annotated F-SETs and V-SETs were 382 bp and 336 bp, while the average respective lengths of the 1250 no-hit F-SETs and 1499 no-hit V-SETs were 271 bp and 283 bp. ORF scanning indicated that only 35 no-hit F-SETs and 13 no-hit V-SETs had long ORFs ($\geq$300 bp). Similar to the results described in our previous study,[6] most of the F-SETs showed the highest homology to the protein sequences from *Vitis vinifera* (237 F-SETs, corresponding to 17.6% of the 1345 annotated F-SETs), *Ricinus communis* (131, 9.7%) and *Populus trichocarpa* (121, 9.0%). Similar results were also apparent for the functional annotation of the V-SETs (respectively

Table 3. KEGG Pathway Enrichment Analyses of Putative Specific Expressed Transcripts

VEG, vegetative transcriptome of Xushu 18; F-SETs, putative floral-specific expressed transcripts; V-SETs, putative vegetative-specific expressed transcripts; padj, adjusted *p*-value. All F-SETs and V-SETs were used for KEGG annotation and a hypergeometric test was applied to get enriched pathway compared with reference transcriptome.

F-SETs

Pathway	VEG	F-SETs	padj
Cutin, suberine and wax biosynthesis	0	5	0.0000
Nitrogen metabolism	119	31	0.0000
Pentose and glucuronate interconversions	254	46	0.0000
Oxidative phosphorylation	265	35	0.0000
Starch and sucrose metabolism	922	61	0.0000
Arachidonic acid metabolism	58	10	0.0000
Phenylpropanoid biosynthesis	215	18	0.0007
Retinol metabolism	29	6	0.0012
Anthocyanin biosynthesis	41	7	0.0015
Steroid hormone biosynthesis	41	7	0.0015
Linoleic acid metabolism	56	8	0.0022
Methane metabolism	418	26	0.0024
Drug metabolism-other enzymes	140	13	0.0031
Metabolism of xenobiotics by cytochrome P450	48	7	0.0050
Glycolysis/Gluconeogenesis	273	19	0.0057
Other glycan degradation	135	12	0.0089

V-SETs

Pathway	VEG	V-SETs	padj
Purine metabolism	945	44	0.0000
Retinol metabolism	29	7	0.0000
Steroid hormone biosynthesis	41	7	0.0000
Metabolism of xenobiotics by cytochrome P450	48	7	0.0001
Carotenoid biosynthesis	16	4	0.0006
Aminobenzoate degradation	143	10	0.0029
Linoleic acid metabolism	56	6	0.0045
Arachidonic acid metabolism	58	6	0.0056
Sesquiterpenoid and triterpenoid biosynthesis	5	2	0.0060

35.8%, 14.1% and 14.6% of the 1429 annotated V-SETs for *Vitis vinifera*, *Ricinus communis* and *Populus trichocarpa*). Further investigations based on the KEGG annotation were made to analyze the metabolic pathways involving the F-SETs. The result demonstrated that 418 F-SETs were annotated as 127 enzymes distributed in 91 pathways. Interestingly, there were 5 F-SETs, all possessing transferase activity, that were annotated as enzymes involving the cutin, suberine and wax biosynthesis pathway, while no transcripts of the vegetative transcriptome were involved in this pathway.[6] Significant flower-enriched pathways were identified by applying the hypergeometric test,[24] the results showing that the F-SETs were enriched in the cutin, suberine and wax biosynthesis, nitrogen metabolism, pentose and glucuronate interconversion, oxidative phosphorylation and some other pathways, while the V-SETs were enriched in the purine metabolism, retinol metabolism, and steroid hormone biosynthesis pathways (Table 3).

GO mapping was performed by using Blast2GO v2.4.[20] Of the 1345 annotated F-SETs, the primary metabolic (502 F-SETs), cellular metabolic (395 F-SETs), and biosynthetic (273 F-SETs) were the most representative biological processes, while the cell part

Table 4. Expression Analyses of Some Flowering Regulatory Genes in Sweet Potato
VEG, vegetative transcriptome of Xushu 18; FLO, floral transcriptome of Xushu 18; GD, roots transcriptome of Guangshu 87.

Gene	Seq_ID	Sequence blast		Expression level (RPKM)		
		Hit-length (AA)	Similarity (%)	VEG	FLO	GD
IbFLC	Contig_8531	241	90	0.13	77.79	0.34
	Contig_10213	238	90	0.00	22.61	0.08
	Contig_15784	196	96	0.14	58.23	0.02
IbCO	Contig_19454	121	96	10.17	3.81	1.18
	Contig_23179	116	90	5.21	4.10	1.20
	Contig_36840	129	84	0.36	24.98	1.24
IbSOC1	Contig_7067	220	100	44.44	18.54	29.95
	Contig_5676	220	100	39.72	14.98	30.28
	Contig_14166	206	82	6.16	0.53	3.44
IbFT	Contig_78915	58	94	1.47	0.69	7.71
	Contig_89908	48	92	0.93	1.41	5.53
IbLFY	Contig_14318	159	98	0.49	9.33	0.31
	Contig_34149	134	86	0.00	3.35	0.00
IbAP1	Contig_8530	247	81	10.55	15.29	18.86
	Contig_11985	252	97	0.00	34.98	0.05
	Contig_8592	256	82	0.08	16.79	0.20
	Contig_13186	255	95	0.73	13.33	14.62
	Contig_9310	240	100	4.75	9.37	5.61
IbDELLA	Contig_1888	592	80	9.19	19.59	9.54
	Contig_3300	562	81	12.14	153.38	31.14
	Contig_21075	136	58	1.98	11.93	3.98
	Contig_32087	148	88	2.01	5.00	6.17
	Contig_33965	137	91	3.81	3.92	3.91
	Contig_42156	85	89	5.00	6.84	9.75
	Contig_45065	105	89	0.15	5.36	1.46
IbGID1	Contig_5874	345	92	56.62	48.48	30.55
	Contig_8007	262	92	51.65	53.01	73.27
	Contig_9335	310	68	0.93	293.92	16.04
	Contig_28648	119	64	5.65	0.61	14.62
IbSLY1	Contig_1053	601	94	31.48	42.55	27.08

(548 F-SETs), intracellular (467 F-SETs) and intracellular organelle (360 F-SETs) were the most representative cellular components. Hydrolase activity (245 F-SETs), protein binding (180 F-SETs) and nucleotide binding (176 F-SETs) were the most representative molecular functions.

Identification and preliminary expression profiling of some candidates genes for flowering regulation

Unlike other model plants, the flowering control mechanism for sweet potato remains to be revealed. To identify the transcripts putatively involved in the flowering time, flower meristem identity and flower organ identity of sweet potato, the complete coding sequences of some flowering control genes, which were known to be involved in the flower development of *Arabidopsis thaliana* or other model plant species, were downloaded from the NCBI database (http://www.ncbi.nlm.nih.gov/). A local tBLASTX similarity search was performed to match these sequences against the assembled transcripts. Combined with the Blast2GO[20] annotation information, we found a large number of transcripts had high similarity to such generally accepted flowering control related genes as *AG1*, *AP1*, *AP2*, *AP3*, *CO*, *FLC*, *LFY*, *SOC1*, *DELLA*, and *SLY1* (Table S1). We also further investigated the expression abundance of these functional transcripts (Table 4), and found that three FLC-encoding transcripts were all up-regulated in the flowers when compared with the vegetative organs. *SOC1*, whose expression could be repressed by FLC,[25] was down-regulated. Similarly, for *LFY*, the expression level was significantly up-regulated in the flowers. Among three CO-encoding transcripts, Contig_19454 showed a high expression level in the vegetative organs, while Contig_23179 showed almost equal expression abundance between the flowers and vegetative organs, although the highest expression level (Contig_36840) was up-regulated in the flowers. A high expression of *CO* would usually activate *SOC1* and *FT*, but our results demonstrate that *SOC1* was repressed at the transcriptional level in the sweet potato flowers. Another flower meristem identity gene, *AP1*, was also identified in this study. Five AP1-encoding transcripts showed different expression patterns, four of them being more highly expressed in the flowers than in the vegetative organs.[7] We further analyzed the expression patterns of the GA-GID-DELLA mechanism-related protein-encoding transcripts and found that most *DELLA* and *GID1* were up-regulated in the flowers, while no significant difference was apparent for *SLY1* among the vegetative organs and flowers.

Discussion

The most important step in a *de novo* RNA-Seq study is *de novo* assembly. We adopted in this study an integrated *de novo* assembly strategy similar to that described in our previous study.[6] The vegetative transcriptome was integrated with the floral transcriptome to construct an improved transcriptome for sweet potato. The final assembly produced by CAP3[14] was assessed by ORF prediction methods (Table 2). The final assem-

bly was found to be the most comprehensive sequence set in comparison with the published data,[6,7,26,27] as this database contains almost all transcripts expressed in the roots, stems, leaves and flowers. The comprehensive transcriptome had a greater N50 length and average length, and more long transcripts than the previously published data.[6,7,26,27] It improved the transcriptome and genomic resources of this crop, and could be a good genetic resource for further studies to identify the functional genes.

In present study, we used the RNA-Seq PE reads and DGE tags from the two sweet potato cultivars grown in China, Xushu18[6] and Guangshu87,[7] to identify the putative floral- and vegetative-specific expressed transcripts. A total of 2595 F-SETs and 2928 V-SETs were respectively detected. The specificity was verified twice by high-throughput sequencing methodology involving RNA-Seq and DGE. About half of these SETs were annotated by Blast2GO,[20] while the others may have come from new transcripts, non-coding RNAs or some short sequences. The SETs identified in this study may play important roles in organ definition, identity and the maturation process. Some transcripts elicited by drought stress were also detected and defined as such putative floral-specific transcripts as heat shock protein and cytochrome p450.

Plants usually initiate the flowering process by coordinating and integrating the perception of such environmental cues as the day length, light condition, salinity, water content and temperature with some endogenous factors.[25] The flower developmental switch in *Arabidopsis thaliana* involves chromatin modification, translation and post-translation regulation; these can be categorized by genetics and physiology into several flowering pathways, including the photoperiod, vernalization, autonomous and gibberellin pathways.[5] Sweet potato generally rarely blooms under natural growth conditions. We employed in this study drought stress during the growth period to induce sweet potato blooming and successfully identified several transcripts as candidates for flowering regulation. The results indicate that some key genes in the flowering regulation network for *Arabidopsis thaliana* are also present in sweet potato and exhibit similar expression patterns. The expression levels of *FLC* and *CO* were up-regulated in the flowers than in the vegetative organs, (Table 4). As a MADS-box transcription factor, FLC can integrate signals from the autonomous and vernalization pathways and suppress a set of common downstream target genes such as flowering-time integrators and/or florigens SOC1 and FT.[28,29] *SOC1* and *FT* had lower expression abundance in sweet potato flowers than that in vegetative organs, while the flower meristem identity genes, *LFY* and *AP1*,[30,31] had a higher expression level in the flowers which could directly affect the initiation of the flowering process. We further analyzed the expression patterns of the DELLA plant growth inhibitor,[32,33] GIBBERELLIN INSENSITIVE DWARF1 (GID1) GA receptor,[34] and SLY1 DELLA destruction-related protein[35] which are all involved in the GA-GID1-DELLA flowering regulatory mechanism.[32] We found that five of seven DELLA and most expressed GID1 were significantly up-regulated in the flowers, while no increase was apparent for SLY1. Taking account of no changes in the photoperiod and temperature being employed in this study when comparing with the non-flowering control, the flower development of drought-treated sweet potato would have been controlled by the autonomous pathway or gibberellin pathway. As previous studies had found that LFY, AP1 and FT could be horizontally transferred from stock to scion and then induce *Arabidopsis* scion flowering,[36,37] findings described in this study also help to reveal the flower development mechanism for grafted sweet potato. This study forms a preliminary investigation of this biological process, and further studies are urgently needed to reveal the function and regulation mechanism for these putative flowering-related genes. These candidate genes may be a valuable resource for further research into sweet potato molecular breeding.

References

1) Loebenstein G and Thottappilly G, "The Sweetpotato," Springer, Berlin, pp. 9–12 (2009).
2) Varshney RK, Glaszmann JC, Leung H, and Ribaut JM, *Trends Biotechnol.*, **28**, 452–460 (2010).
3) Lardizabal RD and Thompson PG, *HortScience*, **23**, 993–995 (1988).
4) Suge H, *Jpn. J. Breed.*, **27**, 251–256 (1977).
5) Fornara F, de Montaigu A, and Coupland G, *Cell*, **141**, 550–550, e-2 (2010).
6) Tao X, Gu YH, Wang HY, Zheng W, Li X, Zhao CW, and Zhang YZ, *PLoS ONE*, **7**, e36234 (2012).
7) Wang ZY, Fang BP, Chen JY, Zhang XJ, Luo ZX, Huang LF, Chen XL, and Li YJ, *BMC Genomics*, **11**, 726 (2010).
8) Goecks J, Nekrutenko A, and Taylor J, *Genome Biol.*, **11**, R86 (2010).
9) Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, and Taylor J, *Curr. Protoc. Mol. Biol.*, **Chapter 19**, Unit 19 10 1–21 (2010).
10) Hernandez D, Francois P, Farinelli L, Osteras M, and Schrenzel J, *Genome Res.*, **18**, 802–809 (2008).
11) Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Yang H, and Wang J, *Genome Res.*, **20**, 265–272 (2010).
12) Zerbino DR and Birney E, *Genome Res.*, **18**, 821–829 (2008).
13) Schulz MH, Zerbino DR, Vingron M, and Birney E, *Bioinformatics*, **28**, 1086–1092 (2012).
14) Huang X and Madan A, *Genome Res.*, **9**, 868–877 (1999).
15) Rice P, Longden I, and Bleasby A, *Trends Genet.*, **16**, 276–277 (2000).
16) Langmead B, Trapnell C, Pop M, and Salzberg SL, *Genome Biol.*, **10**, R25 (2009).
17) Mortazavi A, Williams BA, Mccue K, Schaeffer L, and Wold B, *Nat. Methods*, **5**, 621–628 (2008).
18) Hao DC, Ge G, Xiao P, Zhang Y, and Yang L, *PLoS ONE*, **6**, e21220 (2011).
19) Morrissy AS, Morin RD, Delaney A, Zeng T, McDonald H, Jones S, Zhao Y, Hirst M, and Marra MA, *Genome Res.*, **19**, 1825–1835 (2009).
20) Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, and Robles M, *Bioinformatics*, **21**, 3674–3676 (2005).
21) Kanehisa M, Goto S, Kawashima S, Okuno Y, and Hattori M, *Nucleic Acids Res.*, **32**, D277–D280 (2004).
22) Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, and Sherlock G, *Nat. Genet.*, **25**, 25–29 (2000).
23) Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, and Bolund L, *Nucleic Acids Res.*, **34**, W293–W297 (2006).
24) Benjamini Y and Hochberg Y, *J. R. Stat. Soc. B Meth.*, **57**, 289–300 (1995).

25) He Y and Amasino RM, *Trends Plant Sci.*, **10**, 30–35 (2005).
26) Xie F, Burklew CE, Yang Y, Liu M, Xiao P, Zhang B, and Qiu D, *Planta*, **236**, 101–113 (2012).
27) Schafleitner R, Tincopa LR, Palomino O, Rossel G, Robles RF, Alagon R, Rivera C, Quispe C, Rojas L, Pacheco JA, Solis J, Cerna D, Kim JY, Hou J, and Simon R, *BMC Genomics*, **11**, 604 (2010).
28) Michaels SD and Amasino RM, *Plant Cell*, **11**, 949–956 (1999).
29) Hepworth SR, Valverde F, Ravenscroft D, Mouradov A, and Coupland G, *EMBO J.*, **21**, 4327–4337 (2002).
30) Boss PK, Bastow RM, Mylne JS, and Dean C, *Plant Cell*, **16 Suppl**, S18–31 (2004).
31) Putterill J, Laurie R, and Macknight R, *Bioessays*, **26**, 363–373 (2004).
32) Harberd NP, Belfield E, and Yasumura Y, *Plant Cell*, **21**, 1328–1339 (2009).
33) Peng J and Harberd NP, *Plant Cell*, **5**, 351–360 (1993).
34) Ueguchi-Tanaka M, Ashikari M, Nakajima M, Itoh H, Katoh E, Kobayashi M, Chow TY, Hsing YI, Kitano H, Yamaguchi I, and Matsuoka M, *Nature*, **437**, 693–698 (2005).
35) Fu X, Richards DE, Fleck B, Xie D, Burton N, and Harberd NP, *Plant Cell*, **16**, 1406–1418 (2004).
36) Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, and Coupland G, *Science*, **316**, 1030–1033 (2007).
37) Sessions A, Yanofsky MF, and Weigel D, *Science*, **289**, 779–782 (2000).