Transcriptional Profiling of Rapidly Growing Cucumber Fruit by 454-Pyrosequencing Analysis

Kaori Ando and Rebecca Grumet¹

Department of Horticulture and Graduate Program in Plant Breeding, Genetics, and Biotechnology, Michigan State University, East Lansing, MI 48824

ADDITIONAL INDEX WORDS. *Cucumis sativus*, fruit development, fruit expansion, fruit growth, transcriptome, RNA-seq

ABSTRACT. Fruit development proceeds from cell division to expansion, maturation, and ripening. Expansion is critical for size, yield, and quality; however, this period of development has received little attention. We used 454-pyrosequencing to develop a cucumber (*Cucumis sativus*) fruit transcriptome, identify highly expressed transcripts, and characterize key functions during exponential fruit growth. The resulting 187,406 expressed sequence tags (ESTs) were assembled into 13,878 contigs. Quantitative real-time polymerase chain reaction (qRT-PCR) verification of differentially expressed genes from fruit of different ages, and high correlation in transcript frequency between replicates, indicated that number of reads/contig reflects transcript abundance. Putative homologs were identified in *Arabidopsis thaliana* for 89% of the contigs represented by at least 10 ESTs; another 4% had homologs in other species. The remainder had homologs only in cucurbit species. The most highly expressed contigs were strongly enriched for growth (aquaporins, vacuolar ATPase, phloem proteins, tubulins, actins, cell wall-associated, and hormone-related), lipid, latex, and defense-related homologs. These results provide a resource for gene expression analysis in cucumber, profile gene expression in rapidly growing fruit, and shed insight into an important, but poorly characterized, developmental stage influencing fruit yield and quality.

Cucumber fruit develop from an enlarged inferior ovary formed from three fused carpels (Robinson and Decker-Walters, 1997). After pollination, the developing fruit follow the typical sequence of cell division, exponential enlargement, maturation, and ripening (Gillaspy et al., 1993). Cell division, which occurs most rapidly before anthesis, slows at anthesis pending pollination, and then resumes for the first 0 to 5 d post-pollination (DPP) (Boonkorkaew et al., 2008; Fu et al., 2008). Rapid cell expansion and fruit enlargement follow cell division. For cucumber fruit, these first two stages of development are largely completed during the first 2 weeks post-pollination, with some variation depending on cultivar and season. Unlike fruit that are consumed ripe, cucumbers are typically harvested at the mid to late phase of rapid fruit enlargement.

Fruit enlargement is clearly an important phase of development for all fruit; however, it is less well studied than the later stages of fruit development, especially ripening, or the initial stages of fruit set. In addition to the importance of fruit expansion for ultimate fruit size, shape, and yield, the expansion phase is also becoming increasingly recognized as an important determinant of fruit quality traits such as accumulation of sugars and organic acids and the establishment of cell wall and cuticle characteristics (Lemaire-Chamley et al., 2005; Mounet et al., 2009; Schlosser et al., 2008). Studies of tomato (*Solanum lycopersicum*) fruit development emphasized our lack of knowl-

edge of early fruit growth; about one-third of the expressed sequence tags (ESTs) identified from tomato fruit at 8, 12, and 15 d post-anthesis showed no homology with previously noted tomato ESTs (Lemaire-Chamley et al., 2005). The proportion of newly identified genes markedly increased with younger fruit samples (8 vs. 12 vs. 15 d post-anthesis). Transcripts that showed increased expression at the transition from cell division to cell expansion include genes associated with cell wall modification, sugar and organic acid transport, and water uptake (aquaporins) (Lemaire-Chamley et al., 2005; Pascual et al., 2007). Metabolic changes during tomato fruit expansion included accumulation of sugars, organic acids, and some amino acids, especially during the later phase of cell expansion (Mounet et al., 2009). In apple (*Malus domestica*), predominant categories of gene expression 3 weeks after bloom included photosynthetic genes, protein biosynthesis, cell cycle (histones and ubiquitin), and cell enlargement (tonoplast and cell wall proteins) genes (Janssen et al., 2008; Lee et al., 2007). Recent work with grape (*Vitis vinifera*) berries showed that high expression of cell wall- and water transport-associated genes coincided with periods of rapid cell expansion (Schlosser et al., 2008).

Cucurbit fruit can grow extremely rapidly to produce immense fruit. Pumpkins (*Cucurbita pepo*) have been recorded in excess of 400 kg, and watermelons (*Citrullus lanatus*) as large as 100 kg (Robinson and Decker-Walters, 1997). During peak growth, commercial watermelon fruit growth may be as much as 0.5 kg·d⁻¹ (A. Levi, personal communication), and cell size in rapidly expanding cucumber fruit increased 10-fold in 4 d (Ando, 2009), making these fruit a good subject for analysis of the exponential growth stage. However, as summarized below, there is relatively little information on gene expression for the cucurbit fruit, and virtually none for cucumber.

Suyama et al. (1999) cloned three highly expressed genes from young cucumber fruit, *Csfl-3* (predicted functions: extensin, major latex protein, and ribosomal protein L3, respectively). Later studies found that cell division post-anthesis was

Received for publication 26 Apr. 2010. Accepted for publication 28 May 2010. We thank Sue Hammar and Matt Duncan for greenhouse assistance, Victor Kaytser for cell size measurements, Tony Schillmiller (Department of Biochemistry and Molecular Biology) for advice on the cDNA preparation protocol, and Shari Tjugum-Holland and Kevin Carr of the Michigan State University Research Technology Support Facility for pyrosequencing and contig assembly, respectively. We also thank Cornelius Barry and Curtis Wilkerson for critical reviews of the manuscript.

This work was in part supported by USDA-SCRI Award No. 2008-51180-04881, the MSU Rackham Foundation, and the PPI Agricultural Research Fund.

¹Corresponding author. E-mail: grumet@msu.edu.

associated with increased expression of several cell cycle genes (e.g., CycA, CycB, CycD, and CDKB) (Fu et al., 2008). In melon (*Cucumis melo*), 14 clones were identified in association with initial fruit set (24 h post-pollination), and nine were identified in 9- and 18 DPP fruit (Choi et al., 2004; Nagasawa et al., 2005). More recently, an $\approx 33,000$ -EST database derived from a variety of melon tissues (roots, leaves, and fruit at 15 and 46 DPP) was used to assemble an oligonucleotide-based microarray to study changes in gene expression in melon fruit during ripening and pathogen attack (Gonzalez-Ibeas et al., 2007; Mascarell-Creus et al., 2009). Recent microarray analyses of developing and ripening watermelon fruit identified 335 ESTs that were differentially regulated by at least 2-fold in watermelon fruit during early, ripening/maturation, or mature stages (12, 24, or 36 d post-anthesis, respectively) when compared with leaf (Wechter et al., 2008). Of the annotated, upregulated genes, 23 were uniquely expressed during early growth.

The development of next-generation sequencing technologies has dramatically improved sequencing capabilities with respect to amount of data, time, and cost, making it possible to address questions of global gene expression based on the frequency of transcript appearance in the sequenced population and to address new questions in less well-characterized species (Droege and Hill, 2008; Mardis, 2008; Margulies et al., 2005). These methods allow for very large-scale, unbiased analyses that do not depend on prior gene identification or assembly onto microarrays. Correlation of transcript frequencies observed by 454-pyrosequencing analyses with microarrays, northern blot analysis, or quantitative real-time polymerase chain reaction (qRT-PCR) assays indicate high reproducibility and suitability for gene expression analysis (Ohtsu et al., 2007; Torres et al., 2008; Weber et al., 2007). In this study, we used 454-pyrosequencing to characterize gene expression during the phase of exponential fruit expansion in cucumber. Our objectives were to develop a cucumber fruit transcriptome, to assess gene expression levels, and to identify highly expressed transcripts during peak fruit expansion in cucumber.

Materials and Methods

PLANT MATERIAL AND FRUIT GROWTH. Cucumber plants ['Vlaspek' pickling type (Seminis Vegetable Seeds, Oxnard, CA)] were grown in the greenhouse in 3.78-L plastic pots filled with BACCTO medium (Michigan Peat, Houston) and were fertilized once per week. Temperature was kept between 21 and 25 °C and supplemental lights were used to provide an 18-h light period. Pest control was performed according to standard management practices in the greenhouse. Only one to two fruit were set per plant to limit interfruit competition effects on fruit development. For fruit growth analysis, length and diameter of 10 fruit were measured daily for 34 DPP. Fruit cell size was measured from sets of 15 fruit harvested at 0, 4, 8, 12, 16, 20, 26, and 32 DPP. Sections from the middle of the fruit were dissected into $\approx 1\text{-cm}^3$ pieces and were fixed in formalin-alcohol-acetic acid (FAA) solution as per Olmstead et al. (2007). Free-hand, thin sections (≈ 1 mm thick) of transverse and cross sectional tissues of fruit from five fruit of each sample age were isolated by razor blade. Each sample was measured for five neighboring consecutive cells to obtain mean cell size and were observed at three locations by light microscope at 100 $\times$ magnification with acridine orange dye (0.1 mg·mL⁻¹). Impressions of intact 8 and 16 DPP epidermal tissues were

prepared for microscopy by applying a thin layer of clear nail polish to the fruit surface as was described in Iwano et al. (1997).

CDNA LIBRARY PRODUCTION AND 454 SEQUENCING. Twenty fruit were harvested at 8 DPP and were ranked by size; the middle 10 fruit were selected for RNA extraction. Pericarp samples consisting of exocarp, mesocarp, and placenta tissue, but no seeds, were isolated from the middle of fruit by razor blade, immediately frozen by liquid nitrogen, and kept at -80 °C until RNA was isolated. RNA and cDNA sample preparation were based on the procedures of Schillmiller et al. (2009). Samples from 10 fruit were pooled for RNA extraction; two samples were prepared. Briefly, total RNA was isolated by TRIzol reagent (Invitrogen, Carlsbad, CA), followed by DNase treatment, and were cleaned up by RNeasy column (Qiagen, Valencia, CA). RNA was quantified using a spectrophotometer (ND-1000; NanoDrop Technologies, Wilmington, DE) and were quality assessed by formaldehyde agarose gel electrophoresis. mRNA was isolated using the Qiagen oligotex kit and was quantified by Nanodrop analysis; 200 ng was used for cDNA synthesis. First-strand cDNA was synthesized using the Creator SMART cDNA library construction kit (Clontech, Mountain View, CA). Following second-strand cDNA synthesis, PCR was performed for 15 to 27 cycles to determine the optimal number of PCR cycles as assessed by 1.1% TAE agarose gel. Once the optimum cycle was determined, second-strand synthesis followed by PCR was performed, and quality was assessed by 1.1% TAE agarose gel. To increase the quantity of cDNA, secondary PCR was performed for 10 cycles. PCR products were purified with the Wizard SV Gel and PCR Clean-Up column to remove fragments smaller than 100 bp (Promega, Madison, WI). Purified PCR products were digested with *Sfi*I to remove primers and purified by Wizard SV Gel and PCR Clean-Up system. Final concentration was assessed by the nanodrop ND-1000 method. Subsequent steps for 454 FLX pyrosequencing analysis were performed by the Michigan State University (MSU) Research Technology Support Facility (RTSF). Each sample was loaded on a one-quarter plate Pico TiterPlate (454 Life Sciences, Branford, CT).

CONTIG ASSEMBLY AND GENE ANNOTATION. Contigs were assembled by the MSU RTSF Bioinformatics Group. Reads were processed through The Institute for Genomic Research (TIGR) SeqClean pipeline to trim residual sequences from the cDNA preparation, poly(A) tails, and other low-quality or low-complexity regions. Trimmed sequences were assembled into contigs using the TIGR Gene Indices Clustering Tools (TIGCL) (Pertea et al., 2003). Stringent clustering and alignment parameters were used to limit the size of clusters for assembly. Contigs from the first pass of assembly were then combined and subjected to a second assembly pass with CAP3 (Huang and Madan, 1999). Less stringent alignment parameters were used for this pass to allow for minor sequencing errors or allelic differences in the cDNA sequence. To estimate relative expression, the number of reads originating from each cDNA library were counted for each contig and were reported relative to the total number of reads generated for that library as transcripts per thousand (TPT). The final contigs were subjected to BLASTX search against the green plant subdivision of the National Center for Biotechnology Information (NCBI) nr protein database and/or the *Arabidopsis thaliana* (arabidopsis) protein (TAIR8) databases to search for homology to previously identified genes and to assign possible gene functions. BLASTN analysis was

performed for highly expressed contigs for which homologs were not identified by BLASTX searches.

TRANSCRIPTOME ANALYSIS. To assign gene ontology (GO) categories, contigs were subjected to BLASTX analysis against the Arabidopsis Information Resource (TAIR V7.0) database to search for putative homologs in arabidopsis. Subsequently, The Classification SuperViewer Tool w/Bootstrap web database (Provart and Zhu, 2003) was used for GO categorization, determination of normalized frequencies relative to arabidopsis, and calculation of bootstrap standard deviations. Highly expressed transcripts (>0.2% frequency) were also subjected to BLAST analysis against the International Cucumber Genomic Initiative EST database [International Cucumber Genomic Initiative (ICuGI), 2010] to search for presence or absence in current libraries from cucumber flower buds, flowers, and fruit, melon fruit, and watermelon fruit.

qRT-PCR. Total RNA was prepared from samples of 4, 12, 20, and 32 DPP fruit as described above. RT reactions were performed using the High Capacity RNA-to-cDNA kit (Applied Biosystems, Foster City, CA). Gene-specific primers were designed using Primer Express software (Applied Biosystems). The ABI Prism 7900HT Sequence Detection System was used for qRT-PCR analysis. Power SYBR Green PCR Master Mix (Applied Biosystems) was used for PCR quantification. Actin from *C. sativus* (NCBI DQ641117) was used as an endogenous control for normalization. PCR products were quantified with reference to corresponding standard curves.

Results and Discussion

Fruit growth

The increase in fruit size for ‘Vlaspik’ cucumber began shortly after pollination with the exponential increase in length occurring from ≈4 to 12 DPP (Fig. 1A). This rapid increase in fruit length, ≈6-fold in 8 d, corresponded to a rapid increase in cell size between 4 and 12 DPP (Fig. 1B). The mesophyll cells became filled with a large central vacuole, while growth in the epidermal cells was accompanied by extensive thickening of the cell walls. The increase in fruit diameter, which was somewhat delayed relative to increase in fruit length, occurred primarily between 4 and 16 DPP, at which time fruit growth was largely completed. The fruit matured over the next 2 weeks; seeds are typically ready for harvest at 30 to 35 DPP. To perform transcriptional analysis during the phase of rapid fruit expansion, fruit samples were collected at 8 DPP.

Cucumber fruit transcriptome analysis by 454 sequencing

454-Pyrosequencing analysis was performed on cDNA libraries prepared from cucumber fruit pericarp and placenta samples harvested during midexponential growth stage, at 8 DPP. Two replicate samples were prepared, and each was run on one-quarter of an FLX Pico Titer pyrosequencing plate (454 Life Sciences). These samples generated about 187,000 clean reads with mean and median read lengths of 222 and 235 bases (Table 1). The 454 sequence data were deposited to NCBI GenBank as SCRA study SPR001601. Nearly 90% of the reads (175,350 ESTs) could be assembled into 13,878 contigs with a mean length of 416 bp, and range of 52 to 2824 bp; an additional 12,056 ESTs remained as singletons, for a total of 25,934 unigenes.

The number of the reads per contig ranged from two to more than 5000, with a mean of 13 reads per contig. There was excellent

Fig. 1. Cucumber fruit and cell size in relation to fruit age. (A) Fruit length and diameter were measured daily for 34 d post-pollination (DPP). Each data point is the mean of 10 fruit ± SE. (B) Fruit cell size was measured from 0, 4, 8, 12, 16, 20, 26, and 32 DPP fruit as described in “Materials and Methods.” Each data point is mean ± SE of 15 fruit.

Table 1. Summary of 454 pyrosequencing results and contig assembly for transcript samples of cucumber fruit at 8 d post-pollination.

Parameter	Value
Clean reads (no.)	187,406
Clean/passed (%)	94.2
Mean read length (bases)	222
Median read length (bases)	235
Total bases (no.)	41,658,727
Contigs (no.)	13,878
Reads assembled into contigs (no.)	175,350
Reads assembled into contigs (%)	88.4
Mean contig length (bp)	416
Range of contig length (bp)	52–2,824
Mean reads per contig (no.)	13
Range of reads per contig (no.)	2–5,167

correlation ($r = 0.98$) in transcript frequency for contig groups between the two replicate samples, indicating reproducibility of results (data not shown). The shortest contigs also tended to be among the least frequently represented in the population, as might be expected, as assembly of full-length transcripts increases with more EST sequences (Fig. 2). Contigs of less than 400 bp had an average of three ESTs per contig, while those greater than 800 bp averaged 81 ESTs per contig.

To verify that the transcript read number was reflective of transcript abundance, a test 454 sequence analysis was performed

Fig. 2. (A) Number of transcripts in the 8 d post-pollination cucumber fruit sample versus contig length. (B) Relationship between contig length and number of transcripts comprising the contig.

for fruit at 0, 4, 8, 12, 16, 20, 26 and 32 d using 1/16 plate runs [total clean reads 18,349 (Ando, 2009)]. Based on these data, a set of transcripts representing differential expression [as measured by transcripts per thousand (TPT)] in young [4 DPP (exponential growth)] versus mature [20 DPP (post-rapid expansion)] fruit were selected for verification by qRT-PCR analysis (Fig. 3A). Comparison of the ratio of gene expression between the two ages as estimated by qRT-PCR closely resembled the 454 results, indicating that high relative transcript abundance obtained by 454 analysis reflects high levels of gene expression.

Analysis of putative gene function

Comparisons with the Arabidopsis TAIR v.7 database provided putative homologs for 62% of the contigs; however, the portion with matches in arabidopsis was not evenly distributed (Table 2). For those contigs represented by 10 or more ESTs, the number of hits increased to 89%. This is likely due to generally longer contig length, and therefore greater ability to identify homologs. A small portion of the contigs (4%) without homologs in arabidopsis had putative homologs in other species, especially the well-characterized species, grape and poplar (*Populus trichocarpa*). Highly abundant contigs (>0.1% frequency) that did not have homologs in the current nr protein databases were subjected to BLAST N analysis against the GenBank nucleotide and EST databases and the Cucurbit Genomics Database (ICuGI, 2010). In all cases, very strong homology (77%–97% identity) was observed with other cucurbit (e.g., melon or watermelon) ESTs. The majority had clearly recognizable open reading frames, suggesting that these contigs may represent genes that are unique to cucurbits.

Fig. 3. qRT-PCR analysis of selected, differentially expressed genes. (A) Relative transcript levels of selected highly expressed genes in fruit at 20 d post-pollination (DPP) versus 4 DPP measured by 454 sequencing (dark bars) and quantitative PCR (light bars). (B and C) qRT-PCR analysis of selected highly expressed genes in 4, 12, 20, and 32 DPP fruit. Gene expression was calculated using the relative standard curve method and was normalized with actin expression level. Putative arabidopsis homologs are indicated in parentheses. (B) Transcripts more highly expressed early in development [α -tubulin (At4g14960), chlorophyll a/b-binding protein (At2g34430), and aquaporin (At2g36830)]. (C) Transcripts more highly expressed in later development [auxin-repressed protein (At1g28330), catalase (At4g35090), metallothionein (At3g09390), and snakin (At2g18420)].

The most highly abundant transcripts (present at >0.15% of total EST reads) are listed in Table 3. The predicted functions included latex-related proteins, and proteins associated with lipids, growth, protein synthesis, defense, phloem transport, and photosynthesis. When frequency of the most highly expressed group of transcripts is normalized relative to frequency of their counterparts in the arabidopsis genome, functional category analysis shows over-representation (>2 $\times$) of ribosome-related genes, and structural categories such as structural molecule activity, cytosol, plastid, other cytoplasmic components, and other intracellular components (Fig. 4A). Genes associated with response to stress and response to abiotic or biotic stimuli also were overexpressed.

Under-represented or non-represented groups among the most highly expressed contigs tended to be associated with functions such as transcription and signaling (e.g., kinase activity, transferase activity, nucleotide binding, DNA or RNA binding, and transcription factor and transcription factor activity) (Fig. 4). The predicted functional distributions of the most highly represented contigs can be contrasted with the total set of 13,878 contigs. For the total set of contigs, the distribution of

Table 2. Transcript frequency relative to identification of putative homologs to cucumber contigs as identified by BLAST X analysis of the green plant subdivision of the National Center for Biotechnology Information non-redundant protein database.

Contig frequency (reads/contig)	Contigs (no.)	Avg length (bp)	Putative arabidopsis homologs		Other putative homologs		Homologs in cucurbit species only	
			No.	%	No.	%	No.	%
>200	117	1,102	105	89.7	4	3.4	2	1.7
100–199	138	1,132	127	92.1	3	2.2	0	0
10–99	2,071	818	1,830	88.4	60	2.9	5	0.2
2–9	12,010	328	6,595	55.6	511	4.5	45	0.4
Total	13,879		8,657	62.4	578	4.1	52	0.4

predicted functions more closely resembled the overall distribution for arabidopsis genes (Fig. 4B). For example, the total set includes 115 contigs that are putative protein kinases, 35 that are putative phosphatases, and 112 that are putative zinc-finger containing proteins (a total of 652, 211, and 1004 reads, respectively). A wide distribution of putative transcription factor homologs was also observed in the total set. Those that were represented by at least 30 ESTs include two floral-related (*AGAMOUS* and *SEPALLATA*) and an auxin-induced transcription factor (Table 4). Additional transcription factor homologs were found for *APETALA 1*, *APETALA 2*, *APETALA 3*, *ETHYLENE RESPONSE FACTOR 3*, *ETHYLENE INSENSITIVE 3 (EIN3)*, *GIBBERELLIN INSENSITIVE (GAI)*, *KNOTTED-LIKE*, *REVOLUTA*, *SCARECROW*, *SHATTERPROOF*, and *WUSCHEL*-related, as well as several leucine zipper, myb family, NAC domain, helix-loop-helix, CCAAT binding, G-box, homeobox, WRKY, and zinc finger domain factor genes. These observations suggest that the total pool of ESTs likely reflects a broad representation of the cucumber transcriptome.

Growth-associated transcripts

CELL EXPANSION. Many of the most highly expressed contigs (>100 ESTs/contig) have high homology with genes associated with cell structure and expansion, including putative homologs of 13 extensins, tubulins, and actin-associated proteins (Table 5). Homologs of α -tubulin, which is related to formation of cytoskeleton, were highly expressed in developing watermelon and apple (Janssen et al., 2008; Wechter et al., 2008). Other highly expressed contigs had homology with genes associated with cell wall structure and modification, including putative expansins that collectively contributed 166 EST reads, as well as several putative cellulose synthases (195 ESTs), xyloglucan endotransglucosylases (194 ESTs), glycosyl hydrolases (171 ESTs), pectinesterase inhibitors (178 ESTs), pectinase methyltransferase inhibitors (211 ESTs), pectate lyases (100 ESTs), pectinases (72 ESTs), and pectin methylesterases (35 ESTs). High expression of cell wall-modifying genes such as expansins, endo-1,2- β -glucanase, pectate lyases, and pectin methylesterases was also observed in expanding grape berries (Schlosser et al., 2008). Early fruit development in apple and tomato also was marked by elevated expression of tubulins, expansins, and β -glucosidases (Janssen et al., 2008; Lee et al., 2007).

An additional, very highly expressed gene in the 8 DPP fruit was highly homologous to 24-sterol C-methyltransferase 1 (*SMT1*) (76% identity, 90% similarity; E value: 1.0E-159). *SMT1* is involved in sterol biosynthesis, including the production and response to brassinosteroids (BRs), and has been associated with rapidly growing tissue (Huang et al., 2007; Shi et al., 1996). BRs have been studied extensively with respect to

plant growth and are well characterized with respect to stimulation of cell elongation (Mussig, 2005). BRs in combination with other hormones, especially auxins, act synergistically to promote growth by stimulating cell division and elongation.

Recently, BRs have been shown to be critical for cucumber fruit set and early development (Fu et al., 2008). BR treatment can substitute for pollination in inducing cucumber fruit growth, including induction of cell division and increased expression of cell cycle genes in non-parthenocarpic varieties, while inhibition of BR will inhibit fruit development in parthenocarpic varieties. Expression of *SMT* in cucumber fruit decreased for the first 2 DPP or BR treatment, followed by a steady rise from 2 to 6 d after anthesis. The expression of *SMT*, which acts early in the BR biosynthetic pathway, can be contrasted with a homolog of an enzyme involved in a later step of BR synthesis, BR-6-oxidase (BR-6-ox) (Fu et al., 2008). Expression of the *BR6ox* homolog steadily declined following pollination or BR treatment without a subsequent increase (Fu et al., 2008), and our 8 DPP fruit sample also had a very low level of expression of *BR-6-oxidase* (data not shown). This difference in expression pattern suggests possible additional functions for *SMT1*, and is consistent with mutational analyses in *Arabidopsis*, indicating that *SMT1* influences other growth and development functions not rescued by BR application (Fujioka and Yokota, 2003).

More generally, the role of hormones in fruit development is best known in association with fruit set rather than the exponential growth phase. For example, auxins and gibberellins (GAs) can induce parthenocarpic fruit growth in tomato (Vriezen et al., 2008; Wang et al., 2009), and increased production of auxin (IAA) in transgenic cucumbers resulted in parthenocarpic fruit set (Yin et al., 2006). Ethylene and abscisic acid also have been implicated in tomato fruit set (Vriezen et al., 2008). One transcript with homology to auxin-associated genes (At1g28330 and At2g33830) was highly represented (1114 reads) in the 8 DPP cucumber fruit library. Other possible hormone-related genes represented by more than 10 reads included eight putative IAA-induced or -repressed genes and five ethylene biosynthetic genes and ethylene-regulated transcription factors. Transcripts of homologs of GA synthesis genes, such as ent kaurene synthase or oxidase, were observed at low levels; however, a putative GA responsive/snakin gene (homolog of arabidopsis At2g18420) was represented by 498 ESTs.

Homologs of several genes associated with protein synthesis (e.g., ribosomal proteins and translation elongation factors) and photosynthesis (e.g., RuBisCo and chlorophyll a/b-binding proteins) were also highly expressed, as would be expected for metabolically active cells at this stage of fruit development

Table 3. Predicted function, contig length, and number of reads for highly expressed cucumber fruit expressed sequence tags (ESTs) (transcript frequency $\geq 0.15\%$).

Contig name	Contig length (bp)	Total reads (no.)	Predicted function of arabidopsis homolog or other closest homolog based on the Arabidopsis Information Resource (TAIR) 7.0 and National Center for Biotechnology Information (NCBI) databases	E value
Cs1695	1,051	5,167	MLP43 (MAJOR LATEX PROTEIN-LIKE PROTEIN 43)/Csf-2 (<i>Cucumis sativus</i>)	6.0E-47
Cs239	1,778	1,992	SMT1 (STEROL METHYLTRANSFERASE 1)	1.0E-159
Cs273	960	1,790	Lipid transfer protein, putative	2.0E-24
Cs1784	867	1,726	MLP423 (MAJOR LATEX PROTEIN-LIKE PROTEIN 423)	1.0E-49
Cs23	2,479	1,716	PRXR1 (peroxidase 42)	1.0E-155
Cs112	1,247	1,526	GAMMA-TIP [tonoplast intrinsic protein (TIP) γ]	1.0E-121
Cs154	903	1,300	PFN3/PRF3 (PROFILIN 3)	6.0E-61
Cs175	1,434	1,197	LP1 (nonspecific lipid transfer protein 1)	2.0E-23
Cs1735	654	1,187	Unknown, <i>Cucumis melo</i> ssp. <i>Piel de Sapo</i>	8.0E-83
Cs126	980	1,114	DRM1 (DORMANCY-ASSOCIATED PROTEIN 1)/auxin associated	1.0E-41
Cs2	2,209	1,059	Putative histone H2B	2.0E-17
Cs241	815	1,053	Protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	1.0E-32
Cs699	1,667	1,022	60S acidic ribosomal protein P3 (RPP3B)	3.0E-39
Cs22	590	819	Unknown, <i>C. melo</i>	4.0E-75
Cs1835	855	816	SP1L1 (SPIRAL1-LIKE1)	2.0E-36
Cs110	970	776	Unknown, <i>C. melo</i> ssp. <i>Piel de Sapo</i>	5.0E-176
Cs209	820	717	PP1 (phloem protein 1) (<i>Cucurbita maxima</i>)	1.0E-30
Cs220	2,001	711	CAT2 (CATALASE 2)	0.0
Cs299	1,219	702	60S ribosomal protein L21 (RPL21C)	9.0E-82
Cs13	696	692	60S ribosomal protein L30 (RPL30A)	2.0E-51
Cs121	1,784	663	TUA6 (tubulin α -6 chain)	0.0
Cs307	509	654	Unknown, <i>C. melo</i>	7.0E-147
Cs265	1,299	649	VPS46.1 (VACUOLAR PROTEIN SORTING 46.1)	9.0E-32
Cs172	1,340	642	CAB1 (CHLOROPHYLL A/B-BINDING PROTEIN 1)	1.0E-139
Cs1586	807	637	SAG21 (SENESCENCE-ASSOCIATED GENE 21)	5.0E-21
Cs182	904	585	MT2A (METALLOTHIONEIN 2A)	3.0E-28
Cs141	1,515	573	MAT2/SAM-2 (S-adenosylmethionine synthetase 2)	0.0
Cs692	1,025	563	Unknown; <i>C. melo</i> ssp. <i>agrestis</i>	0.0
Cs124	1,747	541	EF-1- α (elongation factor 1- α)	0.0
Cs227	857	538	AVA-P2 (vacuolar-H ⁺ -pumping ATPase 16-kDa proteolipid subunit 2)	2.0E-85
Cs168	1,271	518	RPL15A (60S ribosomal protein L15)	1.0E-109
Cs1861	760	501	Similar to unknown protein (<i>Arabidopsis thaliana</i>)	6.0E-27
Cs36	702	498	Gibberellin-regulated GASA/GAST/Snakin family protein	7.0E-32
Cs714	1,070	498	ATARFA1E (ADP-RIBOSYLATION FACTOR A1E)	1.0E-100
Cs271	2,053	498	CIPK6 (CBL-INTERACTING PROTEIN KINASE 6)	0.0
Cs1946	882	487	Unknown, <i>C. melo</i>	7.0E-147
Cs131	797	451	Proteasome maturation factor UMP1 family protein	3.0E-61
Cs128	1,025	439	DNA-binding family protein/remorin family protein	2.0E-54
Cs1554	692	430	Thioredoxin h (<i>Hevea brasiliensis</i>)	2.0E-44
Cs75	1,285	418	Latex allergen (<i>H. brasiliensis</i>)	9.0E-11
Cs58	1,413	406	GDSL-motif lipase/hydrolase family protein	1.0E-146
Cs267	1,033	392	CAM7 (CALMODULIN 7); calcium ion binding	2.0E-80
Cs85	1,039	386	LHB1B1 (photosystem II light-harvesting complex gene 1.4)	1.0E-103
Cs155	1,503	372	GRF9 (GEN REGULATORY FACTOR 9); protein phosphorylated amino acid binding	1.0E-121
Cs6	875	371	EMB2386 (EMBRYO DEFECTIVE 2386); structural constituent of ribosome	8.0E-98
Cs34	1,751	370	OASB [<i>O</i> -ACETYLSELINE (THIOL) LYASE B]; cysteine synthase	1.0E-157
Cs723	718	368	RPP2D (60S acidic ribosomal protein P2)	4.0E-40
Cs244	1,048	366	GAMMA-TIP [tonoplast intrinsic protein (TIP) γ]	1.0E-106
Cs109	1,362	365	PIP1C (PLASMA MEMBRANE INTRINSIC PROTEIN 1;3)	1.0E-148
Cs1774	1,011	362	PP1 (phloem protein 1) (<i>Cucurbita maxima</i>)	1.0E-25
Cs163	1,106	361	Ribulose biphosphate carboxylase small chain 3B/RuBisCO small subunit 3B	5.0E-81

Continued next page

Table 3. Continued.

Contig name	Contig length (bp)	Total reads (no.)	Predicted function of arabidopsis homolog or other closest homolog based on the Arabidopsis Information Resource (TAIR) 7.0 and National Center for Biotechnology Information (NCBI) databases	E value
Cs217	1,323	359	CBS domain-containing protein	3.0E-89
Cs40	1,019	351	ATPP2-A1 (<i>A. thaliana</i> phloem protein 2-A1)	3.0E-28
Cs218	913	349	ADF3 (ACTIN DEPOLYMERIZING FACTOR 3)	5.0E-63
Cs1136	839	345	CCH (COPPER CHAPERONE)	6.0E-28
Cs724	759	343	Integral membrane family protein	2.0E-37
Cs28	752	341	MLP34 (MAJOR LATEX PROTEIN-LIKE PROTEIN 34)	4.0E-17
Cs298	878	338	Gibberellin-regulated family protein	7.0E-16
Cs1130	885	337	TCTP (TRANSLATIONALLY CONTROLLED TUMOR PROTEIN)	2.0E-69
Cs88	592	329	Protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	9.0E-24
Cs71	1,823	324	Arginosuccinate synthase family	0.0
Cs117	504	319	Unknown, <i>C. melo</i>	4.0E-85
Cs51	816	311	40S ribosomal protein S23 (RPS23B)	2.0E-74
Cs12	992	305	LHB1B1 (photosystem II light-harvesting complex gene 1.4)	1.0E-136
Cs719	881	299	60S ribosomal protein L32 (RPL32A)	8.0E-45
Cs1129	954	297	Similar to four F5 protein-related	1.0E-23
Cs186	713	296	PSBR (photosystem II subunit R)	3.0E-56
Cs275	1,478	296	POM1 (POM-POM1); chitinase	1.0E-135
Cs282	971	290	HIS1-3 (HISTONE H1-3)	2.0E-35
Cs1836	442	287	40S ribosomal protein S29 (RPS29C)	2.0E-28
Cs1142	904	286	ROC3 (rotamase CyP 3)	6.0E-87
Cs727	916	283	Similar to unknown protein	2.0E-43

(Table 3). Consistent with the high transcript abundance in the 8 DPP fruit, qRT-PCR analysis of selected aquaporin, α -tubulin, and chlorophyll a/b-binding protein genes at 4, 12, 20 and 32 DPP showed a pattern of high expression in the young, rapidly growing fruit that progressively declined to low expression after reaching full size (Fig. 3B). The above genes can be contrasted with transcripts whose expression increased later in fruit development, such as homologs of auxin-repressed protein, catalase, metallothionein, and snakin (Fig. 3C).

UPTAKE OF WATER AND PHOTOASSIMILATES. A primary factor contributing to cell expansion is turgor pressure caused by uptake of water and associated increase in vacuole size as was seen in the cucumber fruit mesophyll cells (Fig. 1). Eight of the most highly expressed genes were associated with water uptake, such as those with high homology to aquaporin and tonoplast intrinsic protein genes (more than 3400 ESTs). Members of the aquaporin family are small integral membrane proteins that serve as water channels to mediate water flow across membranes, and are associated with cell expansion (Maurel et al., 2008). Aquaporins also have been observed to be highly expressed in rapidly growing grape, apple, and tomato fruit (Lemaire-Chamley et al., 2005; Janssen et al., 2008; Lee et al., 2007; Schlosser et al., 2008).

Cell expansion is associated with enlargement of the vacuole, as was seen in the cucumber fruit mesophyll cells. The vacuole serves to accumulate solutes, causing increased solute pressure, which in turn drives uptake of water into the cell (Amemiya et al., 2006). Movement of solutes into the vacuole across a concentration gradient is driven by specific transporters, including vacuolar H⁺ ATPases, which also have been implicated in fruit growth in tomato, and were highly expressed in developing watermelon fruit (Amemiya et al., 2006; Wechter et al., 2008). A

putative vacuolar H⁺ ATPase was highly expressed (538 reads) in the growing cucumber fruit.

Import of photoassimilates via the phloem is also essential for fruit growth. Angiosperm phloem sieve elements contain proteinaceous filaments formed from P proteins, which for cucurbits, are comprised of the cucurbit-specific PP1 protein, and the more widely distributed PP2 protein (Clark et al., 1997). *PPI*, which has been isolated from several cucurbit species including pumpkin, cucumber, melon, and watermelon (Turgeon and Wolf, 2009; Wechter et al., 2008), has no direct counterpart in arabidopsis; it appears to be unique to cucurbits as indicated by BLAST sequence analysis (Sept. 2009) and lack of antibody cross-reactivities (Clark et al., 1997; Turgeon and Wolf, 2009). Two genes encoding homologs of *PPI* were among the top group of transcripts in the 8 DPP cucumber fruit. *PPI* also was very highly expressed in green stage watermelon fruit (Wechter et al., 2008). Two additional highly expressed, predicted phloem protein genes (at frequencies of 242 and 219) have homologs in arabidopsis (At1g33920 and At4g19840), as does the 26-kDa phloem protein 2 (At4g19840). Cucurbits have unusual phloem anatomy. In addition to phloem located inside vascular bundles, cucurbit plants have a network of extrafascicular phloem associated with the outside of the vascular bundles (Clark et al., 1997). *PPI* is strongly expressed in both locations, and along with *PP2*, was found to be predominantly expressed in actively growing tissue.

LATEX, LIPIDS, AND DEFENSE. The single most highly expressed cucumber fruit gene was *Csf-2*, with more than 5000 reads. *Csf-2* has high homology with a major latex-like protein (MLP) gene from arabidopsis and to a family of flower- and fruit-specific genes (Table 6). Several other contigs with homology to MLP genes were among the most frequently expressed group. MLP

Fig. 4. Biological enrichment analysis of contigs with putative arabidopsis homologs in the 8 d post-pollination cucumber fruit expressed sequence tag (EST) library. (A) Most highly represented contigs (transcript frequency >0.1%; 127 contigs, representing 61,188 ESTs). (B) Total 13,878 contigs (187,406 ESTs). Functional distribution, normalized frequency, and bootstrap standard deviation (SD) of cucumber fruit contigs relative to arabidopsis was determined using the categories classification from the Classification SuperViewer from Bio-Array Resource for Arabidopsis Functional Genomics for Gene Ontology (Provart and Zhu, 2003).

homologs have been observed from fruit of other plants such as opium poppy (*Papaver somniferum*), bell pepper (*Capsicum annuum*), raspberry (*Rubus idaeus*), strawberry (*Fragaria* spp.), soybean (*Glycine max*), and tomato (e.g., Nessler, 1994; Ruperti et al., 2002). Cucumber, peach (*Prunus persica*), and soybean MLP homologs were highly expressed in immature fruit, while muskmelon, strawberry, and raspberry homologs were expressed in ripening fruit, and tobacco (*Nicotiana tabacum*) and bell pepper homologs were expressed upon wounding. Despite high levels of expression, the function of MLPs is largely unknown. There was a linear correlation between peach MLP homolog (*Pp-MLP1*) mRNA accumulation and fruit relative growth rate, suggesting that peach MLP may be associated with fruit cell expansion. Preliminary examination of *Csf2* expression showed elevated expression during peak cucumber expansion (Ando, 2009).

Latex can serve as a defense against herbivory, and coagulation of latex following injury can seal wounds, preventing

subsequent pathogen entry (Looze et al., 2009; Moutin et al., 1999). Latex production in various species such as rubber (*Hevea brasiliensis*) and papaya (*Carica papaya*) also has been associated with several kinds of wound- or infection-inducible PR proteins such as chitin-binding protein (hevein), chitinase, β -1,3-glucanase, and thaumatin (Looze et al., 2009; Moutin et al., 1999; Subruto et al., 2001). Homologs of several of these sorts of genes also are highly expressed in the rapidly growing cucumber fruit [e.g., chitinase, thionin, hevein, and snakain (Table 6)]. Other contigs represented at least 100 times in the 8 DPP cucumber fruit EST population had homology to abiotic stress-related genes such as peroxidase, catalase, thioredoxin, dehydrin, and COR (cold-responsive) genes (Table 6).

Despite extremely high levels of expression in the cucumber fruit, *Csf-2* was not seen in melon or watermelon fruit according to the Cucurbit Genomics Database (ICuGI, 2010). Although *Csf-2* homologs may not have been present in the Cucurbit

Table 4. Frequently observed transcription factor homologs in the 8 d post-pollination cucumber fruit expressed sequence tag (EST) library.

Putative arabidopsis homolog		Total reads (no.)	Contig name	Contig length (bp)	E value
AGAMOUS	At4G18960	71	Cs1196	1009	4.0E-73
AKR2B ankyrin-repeat	At2G17390	117	Cs771	1530	1.0E-120
ATHB1 leucine zipper	At3G01470	87	Cs132	1288	2.0E-40
ATHB13 leucine zipper	At1G69780	67	Cs1199	1375	1.0E-103
CCR-NOT transcription complex	At2G32070	54	Cs789	1590	1.0E-121
Helix-loop-helix family	At3G57800	52	Cs1217	1908	6.0E-91
IAA-induced	At3G04730	64	Cs355	1157	2.0E-72
Myb family/DNA binding	At4G39250	60	Cs9978	647	5.0E-25
SEPALLATA 1	At5G15800	31	Cs767	730	3.0E-69

Table 5. Highly expressed transcripts in the 8 d post-pollination cucumber fruit library predicted to be associated with import of water, photoassimilates, and cell expansion.

Contig name	Contig length (bp)	Predicted function	Arabidopsis putative homolog	E value	Total reads (no.)
Cs239	1778	24-sterol C-methyltransferase 1	At5G13710	1.0 E-159	1992
Cs112	1247	Aquaporin (tonoplast intrinsic)	At2G36830	1.0 E-121	1526
Cs154	903	Profilin/actin binding	At5G56600	6.0 E-61	1300
Cs1835	855	SP1L1 (Spiral-like 1)	At1G26355	2.0 E-36	816
Cs209	820	Phloem filament protein/PP1	(<i>Cucurbita maxima</i>)	1.0 E-30	717
Cs121	1784	α -Tubulin	At4G14960	00	663
Cs277	857	Vacuolar H ⁺ ATPase	At1G19910	2.0 E-85	538
Cs244	1048	Tonoplast intrinsic protein	At2G36830	1.0 E-106	366
Cs109	1362	Aquaporin 1 (PIP 1C)	At1G01620	1.0 E-148	365
Cs1774	1011	Phloem filament protein	(<i>C. maxima</i>)	1.0 E-25	362
Cs40	1019	26-kDa phloem protein A21	At4G19840	3.0 E-128	351
Cs218	913	Actin depolymerizing factor-like	At5G59880	5.0 E-63	349
Cs724	759	Integral membrane family protein	At4G15360	2.0 E-37	343
Additional contigs with >100 reads					
Extensin					215, 214, 174, 106
Tubulin (α and β)					160, 159, 133, 128
Aquaporin					242, 127, 121
Phloem protein					244, 219
Cell wall modifying					211, 194, 178, 171

Genomics Database due to limited EST library sizes, if they were expressed at a comparable level as in cucumber, it seems likely they would be also present in melon and watermelon EST collections. However, other MLP homologs were expressed in melon fruit (Aggelis et al., 1997; Hadfield et al., 2000), suggesting that related genes may play similar functions in the different cucurbit fruit. Furthermore, the majority of MLP homologs in melon, including the melon *Csf2* homolog, are expressed in root or phloem tissue (Aggelis et al., 1997; Hadfield et al., 2000; ICuGI, 2010). Defense systems such as latex and resin canals have been suggested to be under extensive diversification, reflective of ongoing evolutionary challenges between herbivores and their hosts (Farrell et al., 1991). The variation for MLP homologs among cucurbits may reflect such diversification.

At least four of the very highly expressed contigs had homology to lipid transfer protein (LTP) genes (Table 6). LTPs possess a hydrophobic pocket that can bind long-chain fatty acids and have been implicated in a broad range of functions, including lipid transport to extracellular surfaces or wax transport to the cuticle (Carvalho and Gomes, 2007; Samuels et al., 2008; Yeats and Rose, 2008). LTPs are hypothesized to trans-

port lipids across the cell wall based on expression in epidermal tissue and signal peptides conferring secretion into the apoplast (Samuels et al., 2008; Yeats and Rose, 2008). LTPs also have been associated with cell wall loosening and are frequently found in fruit and seeds, including elevated expression during early fruit development in grape and tomato (Janssen et al., 2008; Yeats and Rose, 2008), and later development in pear (*Pyrus communis*), strawberry, grape, and tomato (Fonseca et al., 2004; Janssen et al., 2008). In addition to a structural role associated with cuticle formation, LTPs have been demonstrated to have antimicrobial properties (Carvalho and Gomes, 2007; Yeats and Rose, 2008). LTPs are members of the pathogenesis-related (PR) class of proteins, the PR14 family. They are induced in response to infection, and have been shown to inhibit bacterial and fungal growth in vitro, possibly by promoting membrane permeabilization of the pathogen but not the host.

Conclusions

Next generation, ultra-high-throughput sequencing technologies such as 454-pyrosequencing are providing an alternative approach for gene expression analysis relative to methods such

Table 6. Highly expressed transcripts from the 8 d post-pollination cucumber fruit library predicted to be associated with latex, lipids, and defense.

Contig name	Contig length (bp)	Predicted function	Arabidopsis putative homolog	E value	Total reads (no.)
Cs1695	1051	Csf-2 latex-like protein MLP43	At1G70890	6.0 E-47	5167
Cs273	960	Lipid transfer protein	At5G01870	2.0 E-24	1790
Cs1784	867	Major latex-like protein 423	At1G24020	1.0 E-49	1726
Cs175	1434	Lipid transfer protein isoform 1	At2G38450	2.0 E-23	1197
Cs241	815	Lipid transfer family/protease inhibitor	At2G45180	1.0 E-32	1053
Cs75	1285	Latex allergen	no hit, best <i>Hevea</i>	9.0 E-11	418
Cs58	1413	Lipase	At5G33370	1.0 E-146	406
Cs28	752	Major latex-like protein 34	At1G70850	4.0 E-17	341
Cs88	592	Lipid transfer-like protein	At1G48750	9.0 E-24	329
Cs23	2479	Peroxidase	At4G21960	1.0 E-169	1716
Cs220	2001	Catalase	At4G35090	1.0 E-2018	711
Cs36	702	GA-regulated/snakin2/COR 11	At2G18420	7.0 E-32	498
Cs1554	692	Thioredoxin h	At1G11530	1.0 E-37	430
Cs275	1478	Chitinase	At1G05850	1.0 E-135	296
Cs738	1318	Thionin	At1G12663	1.0 E-12	276

as microarrays and subtraction cDNA libraries (Shendure, 2008; Vera et al., 2008; Weber et al., 2007). Advantages of 454-pyrosequencing include the ability to obtain massive numbers of gene sequences, reduced cost relative to developing microarrays for non-sequenced species, recovery of rare transcripts, and the lack of necessity to clone into vectors. RNA samples prepared from exponentially expanding cucumber fruit at 8 DPP yielded $\approx 190,000$ ESTs representing nearly 14,000 contigs, providing a valuable resource for gene expression analysis in cucumber. The contig set suggests potentially broad representation of the cucumber transcriptome and including highly expressed transcripts as well as classes of transcripts, such as transcription factors and signal molecules that may be less strongly expressed. High correlation in transcript frequency observed between replicate samples, and analyses of selected genes showing that qRT-PCR results closely resembled relative transcript levels indicated by 454 sequencing, validated that the usefulness of the 454 approach for gene expression analysis (i.e., the number of reads per contig) is reflective of transcript abundance.

BLASTX analysis of the NCBI database indicated that $\approx 95\%$ of the most highly expressed genes had homologs in other species. The remaining transcripts only matched other cucurbit clones, suggesting that there may be genes unique to cucurbits. Analysis of the draft cucumber genome sequence indicated 4362 potentially cucumber unique families, although this may be an overestimation as this group of predicted genes had a much lower rate of EST confirmation than the total set of predicted cucumber genes (Huang et al., 2009).

The group of contigs showing greatest expression (e.g., $>0.15\%$ transcript frequency) in these rapidly growing fruit was highly enriched for genes with homologs associated with cell structure, growth, or expansion, including extensive representation of putative homologs of aquaporins, vacuolar ATPase, phloem proteins, tubulins, actins, numerous cell wall-associated proteins, ribosome- and photosynthesis-related genes, as well as homologs of putative brassinosteroid, auxin, and gibberellin biosynthesis or response genes. Finally, many of the most highly expressed genes had homologs associated with lipids, latex, and

response to biotic and abiotic stresses. In summary, these results provide an extensive profile of gene expression during exponential fruit growth of cucumber and suggest important functions associated with cell expansion, interorgan transport and communication, and interface with the environment.

Literature Cited

- Aggelis, A., I. John, A. Karvouni, and D. Grierson. 1997. Characterization of two cDNA clones for mRNAs expressed during ripening of melon (*Cucumis melo* L.) fruits. *Plant Mol. Biol.* 33:313–322.
- Amemiya, T., Y. Kanayama, S. Yamaki, K. Yamada, and K. Shiratake. 2006. Fruit-specific V-ATPase suppression in antisense-transgenic tomato reduces fruit growth and seed formation. *Planta* 223:1272–1280.
- Ando, K. 2009. Evaluation of the role of plant architecture and cucumber fruit development in *Phytophthora capsici* disease development. Ph.D. Diss., Michigan State Univ., East Lansing.
- Boonkorkaew, P., S. Hikosaka, and N. Sugiyama. 2008. Effect of pollination on cell division, cell enlargement, and endogenous hormones in fruit development in a gynoeocious cucumber. *Sci. Hort.* 116:1–7.
- Carvalho, A.D. and V.M. Gomes. 2007. Role of plant lipid transfer proteins in plant cell physiology: A concise review. *Peptides* 28:1144–1153.
- Choi, J.W., G.B. Kim, Y.C. Huh, M.R. Kwon, I.G. Mok, J.W. Kim, T.S. Lee, S. Kim, and K.H. Im. 2004. Cloning of genes differentially expressed during the initial stage of fruit development in melon (*Cucumis melo* cv. *Reticulatus*). *Mol. Cells* 17:237–241.
- Clark, A.M., K.R. Jacobsen, D.E. Bostwick, J.M. Dannenhoffer, M.I. Skaggs, and G.A. Thompson. 1997. Molecular characterization of a phloem-specific gene encoding the filament protein, phloem protein 1 (PP1), from *Cucurbita maxima*. *Plant J.* 12:49–61.
- Droege, M. and B. Hill. 2008. The genome sequencer FLX™ system: Longer reads, more applications, straight forward bioinformatics and more complete data sets. *J. Biotechnol.* 136:3–10.
- Farrell, B.D., D.E. Dossourd, and C. Mitter. 1991. Escalation of plant defense: Do latex and resin canals spur plant diversification? *Amer. Nat.* 138:881–900.
- Fonseca, S., L. Hackler, A. Zvara, S. Ferreira, A. Baldé, D. Dudits, M.S. Pais, and L.G. Puskás. 2004. Monitoring gene expression along pear fruit development, ripening and senescence using cDNA microarrays. *Plant Sci.* 167:457–469.

- Fu, F.Q., W.H. Mao, K. Shi, Y.H. Zhou, T. Asami, and J.Q. Yu. 2008. A role of brassinosteroids in early fruit development in cucumber. *J. Expt. Bot.* 9:2299–2308.
- Fujioka, S. and T. Yokota. 2003. Biosynthesis and metabolism of brassinosteroids. *Annu. Rev. Plant Biol.* 54:137–164.
- Gillaspy, G., H. Ben-David, and W. Gruissem. 1993. Fruits: A developmental perspective. *Plant Cell* 5:1439–1451.
- Gonzalez-Ibeas, D., J. Blanca, C. Roig, M. Gonzalez-To, B. Pico, V. Truniger, P. Gomez, W. Deleu, A. Cano-Delgado, P. Arus, F. Nuez, J. Garcia-Mas, P. Puigdomenech, and M.A. Aranda. 2007. MEL-OGEN: An EST database for melon functional genomics. *BMC Genomics* 8:306.
- Hadfield, K.A., T. Dang, M. Guis, J. Pech, M. Bouzayen, and A.B. Bennett. 2000. Characterization of ripening-regulated cDNAs and their expression in ethylene-suppressed Charentais melon fruit. *Plant Physiol.* 122:977–983.
- Huang, I., D.P. Paudyal, S.K. Kim, and H. Cheong. 2007. Influence of the SMT2 knock-out on hypocotyl elongation in *Arabidopsis thaliana*. *Biotechnol. Bioprocess Eng.* 12:157–164.
- Huang, S., R. Li, Z. Zhang, L. Li, X. Gu, W. Fan, W.J. Lucas, X. Wang, B. Xie, P. Ni, Y. Ren, H. Zhu, J. Li, K. Lin, W. Jin, Z. Fei, G. Li, J. Staub, A. Kilian, E.A.G. van der Vossen, Y. Wu, J. Guo, J. He, J. Jia, Y. Ren, G. Tan, Y. Lu, J. Ruan, W. Qian, M. Wang, Q. Huang, B. Li, Z. Xuan, J. Cao, Asan, Z. Wu, J. Ahang, Q. Cai, Y. Bai, B. Zho, Y. Han, Y. Li, X. Li, S. Wang, Q. Shi, S. Liu, W.K. Cho, J.Y. Kim, Y. Xu, K. Heller-Uszynska, H. Miao, Z. Cheng, S. Zhang, J. Wu, Y. Yang, H. Kang, M. Li, H. Liang, X. Ren, A. Shi, M. Wen, M. Jian, H. Yang, G. Zhang, Z. Yang, R. Chen, S. Liu, J. Li, L. Ma, H. Liu, Y. Zhou, J. Zhao, X. Fang, G. Li, L. Fang, Y. Li, D. Liu, H. Zheng, Y. Zhang, N. Qin, Z. Li, G. Yang, S. Yang, L. Bolund, K. Kristiansen, H. Zheng, S. Li, X. Zhang, H. Yang, J. Wang, R. Sun, B. Zhang, S. Jiang, J. Wang, Y. Du, and S. Li. 2009. The genome of the cucumber, *Cucumis sativus* L. *Nat. Genet.* 1:10.1038/ng.475.
- Huang, X. and A. Madan. 1999. CAP3: A DNA sequence assembly program. *Genome Res.* 9:868–877.
- International Cucumber Genomic Initiative. 2010. Cucurbit Genomics Database. 28 May 2010. <<http://www.icugi.org/>>.
- Iwano, A.D., T.N. Sreenivasan, and P. Umaharan. 1997. *Phytophthora* resistance in cacao (*Theobroma cacao*): Influence of pod morphological characteristics. *Plant Pathol.* 46:557–565.
- Janssen, B.J., K. Thodey, R.J. Schaffer, R. Alba, L. Balakrishnan, R. Bishop, J.H. Bowen, R.N. Crowhurst, A.P. Gleave, S. Ledger, S. McArtney, F.B. Pichler, K.C. Snowden, and S. Ward. 2008. Global gene expression of apple fruit development from the floral bud to ripe fruit. *BMC Plant Biol.* 8:16.
- Lee, Y.P., G.H. Yu, Y.S. Seo, S.E. Han, Y.O. Choi, D. Kim, I.G. Mok, W.T. Kim, and S.K. Sung. 2007. Microarray analysis of apple gene expression engaged in early fruit development. *Plant Cell Rep.* 26:917–926.
- Lemaire-Chamley, M., J. Petit, V. Garcia, D. Just, P. Baldet, V. Germain, F. Fagard, M. Mouassite, C. Cheniclet, and C. Rothan. 2005. Changes in transcriptional profiles are associated with early fruit tissue specialization in tomato. *Plant Physiol.* 139:750–769.
- Looze, Y., P. Boussard, J. Huet, G. Vandenbussche, V. Raussens, and R. Wintjens. 2009. Purification of a wound-inducible thaumatin-like protein from the latex of *Carica papaya*. *Phytochemistry* 70:970–978.
- Mardis, E.R. 2008. The impact of next-generation sequencing technology on genetics. *Trends Genet.* 24:133–141.
- Margulies, M., M. Egholm, W.E. Altman, S. Attiya, J.S. Bader, L.A. Bernben, J. Berka, M.S. Braverman, Z. Chen, S.B. Dewell, L. Du, J.M. Fierro, X.V. Gomes, K.C. Godwin, W. He, S. Helgesen, C.H. Ho, G.P. Ircyk, S.C. Jando, N.L.I. Alenquer, T.P. Jarvie, K.B. Jirage, K.B. Kim, J.R. Knight, J.R. Lanza, J.H. Leamon, S.M. Lefkowitz, M. Lei, K.L. Lohman, H. Lu, V.B. Makhijani, K.E. McDada, M.P. McKenna, E.W. Myers, E. Nickerson, J.R. Nobile, R. Plant, B.P. Puc, M.T. Ronan, G.T. Roth, G.J. Sarkis, J.F. Simons, J.W. Simpson, M. Srinivasan, K.R. Tartaro, A. Tomasz, K.A. Vogt, G.A. Volkmer, S.H. Wang, Y. Wang, M.P. Weiner, P.G. Yu, R.F. Begley, and J.M. Rothberg. 2005. Genome sequencing in microfabricated high-density picoliter reactors. *Nature* 437:376–380.
- Mascarell-Creus, A., J. Canizares, J. Vilarrasa-Blasi, S. Mora-Garcia, J. Blanca, D. Gonzalez-Ibeas, M. Saladie, C. Roig, W. Deleu, B. Pico-Silvent, N. Lopez-Bigas, M.A. Aranda, J. Garcia-Mas, F. Nuez, P. Puigdomenech, and A.I. Cano-Delgado. 2009. An oligo-based microarray offers novel transcriptomic approaches for the analysis of pathogen resistance and fruit quality traits in melon (*Cucumis melo* L.). *BMC Genomics* 10:467.
- Maurel, C., L. Verdoucq, D.-T. Luu, and V. Santoni. 2008. Plant aquaporins: Membrane channels with multiple integrated functions. *Annu. Rev. Plant Biol.* 59:595–624.
- Mounet, F., A. Monig, V. Garcia, J. Petit, M. Maucourt, C. Deborde, S. Bernillon, G. Le Gall, I. Colquhoun, M. Defernez, J.-L. Giraudel, D. Rolin, C. Rothan, and M. Lemaire-Chamley. 2009. Gene and metabolite regulatory network analysis of early developing fruit tissues highlights new candidate genes for the control of tomato fruit composition and development. *Plant Physiol.* 149:1505–1528.
- Moutin, V., L.G. Silva, M.T.P. Lopes, G.W. Fernandes, and C.E. Salas. 1999. Spontaneous processing of peptides during coagulation of latex from *Carica papaya*. *Plant Sci.* 142:115–121.
- Mussig, C. 2005. Brassinosteroid-promoted growth. *Plant Biol.* 7:110–117.
- Nagasawa, M., H. Mori, K. Shiratake, and S. Yamaki. 2005. Isolation of cDNAs for genes expressed after/during fertilization and fruit set of melon (*Cucumis melo* L.). *J. Jpn. Soc. Hort. Sci.* 74:25–30.
- Nessler, C.L. 1994. Sequence analysis of two new members of the major latex protein gene family supports the triploid-hybrid origin of the opium poppy (*Laticifer; Papaver somniferum*). *Gene* 139:207–209.
- Ohtsu, K., M.B. Smith, S.L. Emrich, L.A. Borsuk, R. Zhou, T. Chen, X. Zhang, M.C.P. Timmermans, J. Beck, B. Buckner, D. Janick-Buckner, N. Nettleton, M.J. Scanlon, and P.S. Schnable. 2007. Global gene expression analysis of the shoot apical meristem of maize (*Zea mays* L.). *Plant J.* 52:391–404.
- Olmstead, J.W., A.F. Iezzoni, and M.D. Whiting. 2007. Genotypic differences in sweet cherry (*Prunus avium* L.) fruit size are primarily a function of cell number. *J. Amer. Soc. Hort. Sci.* 132:697–703.
- Pascual, L., J.M. Blanca, J. Canizares, and F. Nuez. 2007. Analysis of gene expression during the fruit set of tomato: A comparative approach. *Plant Sci.* 173:609–620.
- Pertea, G., X. Huang, F. Liang, V. Antonescu, R. Sultana, S. Karamycheva, Y. Lee, J. White, F. Cheung, B. Parvizi, J. Tsai, and J. Quackenbush. 2003. TIGR gene indices clustering tools (TIGCL): A software system for fast clustering of large EST databases. *Bioinformatics* 19:651–652.
- Provart, N. and T. Zhu. 2003. A browser-based functional classification SuperViewer for *Arabidopsis* genomics. *Curr. Computational Mol. Biol.* 2003:271–272.
- Robinson, R.W. and D.S. Decker-Walters. 1997. Cucurbits. CAB International, New York.
- Rupert, B., C. Bonghi, F. Ziliotto, S. Pagni, A. Rasori, S. Varotto, P. Tonutti, J.J. Giovannoni, and A. Ramina. 2002. Characterization of a major latex protein (MLP) by ethylene during peach fruitlet abscission. *Plant Sci.* 163:265–272.
- Samuels, L., L. Kunst, and R. Jetter. 2008. Sealing plant surfaces: Cuticular wax formation by epidermal cells. *Annu. Rev. Plant Biol.* 59:683–707.
- Schillmiller, A.L., I. Schauvinhold, M. Larson, M. Xu, A.L. Charbonneau, A. Schmidt, C. Wilkerson, R.L. Last, and E. Pichersky. 2009. Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. *Proc. Natl. Acad. Sci. USA.* 106:0865–10870.
- Schlosser, J., N. Olsson, M. Weis, K. Reid, F. Peng, S. Lund, and P. Bowen. 2008. Cellular expansion and gene expression in the developing grape (*Vitis vinifera* L.). *Protoplasma* 232:255–265.
- Shendure, J. 2008. The beginning of the end for microarrays? *Nat. Methods* 5:585–587.

- Shi, J.R., R.A. Gonzales, and M.K. Bhattacharyya. 1996. Identification and characterization of an S-adenosyl-L-methionine: $\delta(24)$ -sterol-C-methyltransferase cDNA from soybean. *J. Biol. Chem.* 271:9384–9389.
- Subrato, T., H. de Vries, J.J. Schuringa, U.M.S. Soedjanaatmadja, J. Hofsteenge, P.A. Jekel, and J.J. Beintema. 2001. Enzymatic and structural studies on processed proteins from the vacuolar (luteoid-body) fraction of latex of *Hevea brasiliensis*. *Plant Physiol. Biochem.* 39:1047–1055.
- Suyama, T., K. Yamada, H. Mori, K. Takeno, and S. Yamaki. 1999. Cloning cDNAs for genes preferentially expressed during fruit growth in cucumber. *J. Amer. Soc. Hort. Sci.* 124:136–139.
- Torres, T.T., M. Metta, B. Ottenwälder, and C. Schlöttere. 2008. Gene expression profiling by massively parallel sequencing. *Genome Res.* 18:172–177.
- Turgeon, R. and S. Wolf. 2009. Phloem transport: Cellular pathways and molecular trafficking. *Annu. Rev. Plant Biol.* 60:207–221.
- Vera, J.C., C.W. Wheat, H.W. Fescemyer, M.J. Frilander, D.L. Crawford, I. Hanski, and J.H. Marden. 2008. Rapid transcriptome characterization for a non-model organism using 454 pyrosequencing. *Mol. Ecol.* 17:1636–1647.
- Vriezen, W.H., R. Feron, F. Maretto, J. Keijman, and C. Mariani. 2008. Changes in ovary transcriptome demonstrate complex hormonal regulation of fruit set. *New Phytol.* 177:60–76.
- Wang, H., N. Schauer, B. Usadej, P. Frasse, M. Zourine, M. Hernould, A. Latche, J.C. Pech, A.R. Fernie, and M. Bouzayen. 2009. Regulatory features underlying pollination-dependent and -independent tomato fruit set revealed by transcript and metabolite profiling. *Plant Cell* 21:1428–1452.
- Weber, P.M., K.L. Weber, K. Carr, C. Wilkerson, and J.B. Ohlrogge. 2007. Sampling the *Arabidopsis* transcriptome with massively parallel pyrosequencing. *Plant Physiol.* 144:1–11.
- Wechter, W.P., A. Levi, K.R. Harris, A.R. Davis, Z. Fei, N. Katzir, J.J. Giovannoni, A. Salman-Minkov, A. Hernandez, J. Thimmapuram, Y. Tadmor, V. Portnoy, and T. Trebitsh. 2008. Gene expression in developing watermelon fruit. *BMC Genomics* 9:275.
- Yeats, T.H. and J.K.C. Rose. 2008. The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). *Protein Sci.* 17: 191–198.
- Yin, A.M., R. Malinowski, A. Ziolkowska, H. Sommer, W. Plader, and S. Malepszy. 2006. The *DefH9-iaaM* containing construct efficiently induces parthenocarpy in cucumber. *Cell. Mol. Biol. Lett.* 11:279–290.