FragariaCyc: A Metabolic Pathway Database for Woodland Strawberry Fragaria vesca

Sushma Naithani *, Christina M. Partipilo, Rajani Raja, Justin L. Elser and Pankaj Jaiswal

Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR, USA

FragariaCyc is a strawberry-specific cellular metabolic network based on the annotated genome sequence of Fragaria vesca L. ssp. vesca, accession Hawaii 4. It was built on the Pathway-Tools platform using MetaCyc as the reference. The experimental evidences from published literature were used for supporting/editing existing entities and for the addition of new pathways, enzymes, reactions, compounds, and small molecules in the database. To date, FragariaCyc comprises 66 super-pathways, 488 unique pathways, 2348 metabolic reactions, 3507 enzymes, and 2134 compounds. In addition to searching and browsing FragariaCyc, researchers can compare pathways across various plant metabolic networks and analyze their data using Omics Viewer tool. We view FragariaCyc as a resource for the community of researchers working with strawberry and related fruit crops. It can help understanding the regulation of overall metabolism of strawberry plant during development and in response to diseases and abiotic stresses. FragariaCyc is available online at http://pathways.cgrb.oregonstate.edu.

Keywords: FragariaCyc, Fragaria vesca, strawberry, plant pathway database, metabolic network, gene-expression analysis

INTRODUCTION

Strawberries, like other fleshy fruits, are rich in essential nutrients, vitamins, anti-oxidants, dietary fibers, and diverse set of volatile compounds—all of which offer potential benefits to human health, including protection against cancer, inflammation, coronary heart diseases, and other age-related diseases (Nijveldt et al., 2001; Boyer and Liu, 2004; Gallus et al., 2005; Liu et al., 2005; Rasmussen et al., 2005; Giampieri et al., 2015; Mazzoni et al., 2016). For that reason, strawberries have been the subject of extensive research using both chemical (Du et al., 2011; Egea et al., 2014; Prat et al., 2014; Schwieterman et al., 2014) and molecular approaches (Schwab et al., 2009; Chambers et al., 2014; Najda et al., 2014; Sanchez-Sevilla et al., 2015; Vallarino et al., 2015; Paniagua et al., 2016). In the past decade, high-throughput experimental approaches have been employed to study strawberry transcriptomes, metabolomes, and proteomes (Perez et al., 1999; Aharoni and O’connell, 2002; Wein et al., 2002; Aharoni et al., 2004; Lunkenbein et al., 2006b; Fait et al., 2008; Bombarely et al., 2010; Zhang et al., 2011). However, these omics data have not been fully utilized due to the lack of a comprehensive cellular metabolic framework for analysis.

Fragaria vesca, the diploid woodland strawberry, is an ancestral subgenome donor to the octoploid F. × ananassa, and its immediate ancestors F. chiloensis and F. virginiana (Rousseau-Gueutin et al., 2009). The F. vesca has a small stature, short life cycle, and a small sequenced genome (Shulaev et al., 2011) that is amenable to genetic manipulations (Slovin et al., 2009).
It shares synteny with many commercially important fruit crop species of the family Rosaceae, such as apple, peach, pear, plum, apricot, raspberry, blueberry, etc. (Illa et al., 2011; Shulaev et al., 2011). Therefore, *F. vesca* is emerging as an attractive model for functional genomics studies within the Rosaceae family (Xia et al., 2015).

We report the development of FragariaCyc, a curated metabolic pathway database for strawberry, based on the genome sequence of the *Fragaria vesca* L. ssp. *vesca*, accession Hawaii 4 (Shulaev et al., 2011). FragariaCyc was constructed using the reference eukaryotic metabolic network MetaCyc (Caspi et al., 2014) and Pathway-Tools software (http://bioinformatics.ai.sri.com/ptools/).

Conceptually, FragariaCyc is a repertoire of the strawberry metabolic and transport pathways depicted by a complex network of dots and connectors. The metabolic reactions show the enzymatic conversion of metabolites. The enzymes with Enzyme Commission (EC) number and associated coding genes are displayed on the top of the reactions. The transport reactions portray transportation of metabolites from one subcellular compartment to another or between the cells. The pathways, reactions, genes, enzymes, and proteins carry electronic and systematic annotations based on orthology and are supported by manually curated experimental evidences. Each component has a detailed page containing a summary, citations, and other relevant information. Currently, for analyzing genomics datasets researchers are using *F. vesca* geneIDs for various members of the *Fragaria* genus. In the wider interest of the strawberry research community, we are curating FragariaCyc at the level of the genus *Fragaria*, but whenever available, we are incorporating information relevant to species or subspecies. The evidences for the existence of various metabolites, reactions, and pathways were collected from published studies on transcriptomes (Aharoni et al., 2002; Aharoni and O’Connell, 2002; Folta et al., 2010; Shulaev et al., 2011), proteomes (Alm et al., 2007; Bianco et al., 2009), and metabolomes (Aharoni et al., 2004; Fait et al., 2008; Zhang et al., 2011) from various *Fragaria* species.

Many excellent tools such as KAAS (Moriya et al., 2007), Mercator (May et al., 2008), MapMan (Thimm et al., 2004), RAST (Meyer et al., 2008; Wilke et al., 2016) etc. allow pathway-based (species-neutral) gene annotation analysis of omics data, but these may not provide information on the species-specific pathways, reactions, and metabolite variations and may not support the import/export of network data in the standardized formats. In contrast, FragariaCyc, based on BioCyc platform (supported by Pathway-Tools software), allows import and export of network data in the standard machine-readable SBML and BioPax formats and supports data interoperability (Jaiswal and Usadel, 2016). Due to the common underlying platform, exchange of relevant information between FragariaCyc and other BioCyc databases is greatly facilitated. Also, users can make quick cross-species comparisons of pathways, compounds, reactions, genes, and gene products with other publicly available BioCyc databases including AraCyc (*Arabidopsis thaliana*) (Mueller et al., 2003), RiceCyc (*Oryza sativa*) (Dharmawardhana et al., 2013), MedicCyc (*Medicago truncatula*) (Urbanczyk-Wochniak and Sumner, 2007), MaizeCyc (*Zea mays*) (Monaco et al., 2013), PoplarCyc (*Populus trichocarpa*) (Zhang et al., 2010), VitisCyc (*Vitis vinifera*) (Naithani et al., 2014), etc. BioCyc databases also provide built-in tools for carrying out gene-expression analysis and flux balance analysis. Therefore, we chose the BioCyc platform to construct FragariaCyc.

Researchers can access various components of the FragariaCyc, and can conduct Omics Viewer analysis of expression data of their choice to assess the broader role of a gene(s), mutant and phenotype in the context of overall cellular metabolism. Here, publicly available transcriptomic data (Kang et al., 2013) was used to illustrate the functionality of Omics Viewer tool.

**MATERIALS AND METHODS**

**Annotation of *F. vesca* Proteins**

FragariaCyc was developed based on the annotations of the version 1.0 hybrid gene models identified in the sequenced genome of the *F. vesca* L. ssp. *vesca* accession Hawaii 4 (https://www.rosaceae.org/species/fragaria/fragaria_vesca/genome_v1.0) (Shulaev et al., 2011). To enrich annotations of the *F. vesca* proteins, a workflow was employed to ascertain various conserved structural-functional domains, transmembrane domains, and subcellular localization sequences. Subsequently, Gene Ontology (GO) annotations were imported for strawberry genes based on orthology to *Arabidopsis thaliana* from GO and TAIR (Poole, 2007) and *Oryza sativa* from Gramene (Tello-Ruiz et al., 2016). Using both methods, we provided preliminary annotations and GO assignments to 25,050 polypeptides. Similar methods have been employed earlier for the construction of RiceCyc (Dharmawardhana et al., 2013), MaizeCyc (Monaco et al., 2013), and VitisCyc (Naithani et al., 2014).

**Construction of FragariaCyc**

The Pathway-Tools software (Karp et al., 2002) and the MetaCyc reference database (Caspi et al., 2014) were employed to create FragariaCyc by following standard protocols supplied by software developers. In the first step, following the PathoLogic input file format, a putative protein was assigned PRODUCT-TYPE code “P” and two required attributes, NAME and FUNCTION. The values for FUNCTION were supplied from structural-functional annotations and GO terms. Subsequently, optional attributes such as gene ID, gene symbols, SYNONYMS, EC number, GO assignments, and a free text COMMENT were assigned. Then GENETIC ELEMENT (required) and FASTA sequence files (optional) were added. In the second step, the “PathoLogic” option was performed with taxonomic filtering ON to find the best matches of enzyme name, gene name and synonyms, EC numbers, reactions, and pathways in the reference database MetaCyc. Once a correspondence was detected for a given entity associated with a reaction of the known pathway, the respective pathway was created in the FragariaCyc. Following the initial assembly of FragariaCyc, standard quality control checks were performed using a built-in consistency tool (Karp et al., 2002) and then plant-specific relationships.
pathways were enriched by comparing it with PlantCyc (http://www.plantcyc.org/).

Curation and Updates of FragariaCyc
Several pathways and reactions specific to bacteria, fungi, and animals that bypassed the routine quality checks were manually removed from the FragariaCyc. Standard methods were employed to curate strawberry-specific pathways, reactions, metabolites and small molecules, genes, and literature citations. We have successfully incorporated the gene annotations based on recently published transcriptomic studies in _F. vesca_ (Darwish et al., 2013, 2015; Kang et al., 2013). We continue to add and revise gene models in FragariaCyc based on the revised annotation of _F. vesca_ Genome v1.1 (Darwish et al., 2015). The Pathway-Tool software and FragariaCyc contents are regularly updated. The current version of FragariaCyc, version 2.19, is hosted on the Pathway-Tool version 19.0.

Omics Viewer Tool and Gene Expression Analysis
The Omics Viewer tool within FragariaCyc allows analysis of large-scale expression data. To show the functionality of this tool, we analyzed a subset of publicly available RNA-Seq expression data (Kang et al., 2013) from _F. vesca_ cultivar YW5AF7. The data file (Supplementary Table 1) was uploaded to the Omics Viewer tool for visualizing the differentially expressed genes on the “Cellular Overview Diagram,” and to understand the tissue-specific regulation of metabolic pathway genes.

RESULTS
An Overview of FragariaCyc, the Cellular Metabolic Networks of _F. vesca_
Several species-specific plant metabolic pathway databases including FragariaCyc are available from our website (http://pathways.cgrb.oregonstate.edu). At present, FragariaCyc contains 66 super-pathways grouped into 488 unique pathways, 2348 enzymatic and 101 transport reactions, 3507 enzymes, 289 transporters, and 2134 metabolites and small molecules.

The FragariaCyc homepage provides hyperlinks to access a database summary, “Cellular Overview Diagram,” Omics Viewer. This homepage also facilitates “quick search” and browsing of the pathways, enzymes, genes and compounds using a hierarchical classification schema based on ontology (Figure 1). Every entity in the database has a detail page (e.g., ethylene biosynthesis pathway shown in Figure 1C) that provides a summary, literature citations, and other relevant information. “Cellular Overview Diagram” (a conceptual framework of the cellular metabolic networks) is also available from the pull-down “Metabolism” menu on the top navigation bar (Figure 1E).

Based on the ontology and enzyme function, pathways in FragariaCyc are organized into several categories: activation/inactivation/inter-conversion, biosynthesis, degradation/utilization/assimilation, detoxification, generation of precursor metabolites and energy, and transport. Typically, a pathway page provides a textbook style interactive diagram that includes hyperlinks to the detail pages of corresponding reactions, metabolites, enzymes, genes, cofactors, and small molecules. Figure 2 shows the L-homoserine biosynthesis pathway as an example of pathway detail page. The Pathway page also shows experimentally verified enzymes and associated genes in boldface. A right-hand side menu within a pathway page provides options to (i) upload expression data on a pathway diagram (Figure 2B); (ii) view a pathway in other plant species or reference database MetaCyc; and (iii) conduct pathway comparisons across two or more metabolic pathway databases (Figure 2C). For the convenience of researchers, our website hosts three reference databases, MetaCyc (Caspí et al., 2014), PlantCyc (Chae et al., 2012), and EcoCyc (Karp et al., 2000), and 13 species-specific plant pathway databases.

The “Cellular Overview Diagram” is accessible from the main page (Figure 1E). It represents a conceptual framework for metabolic networks and its various entities within a simple, non-compartmentalized _F. vesca_ cell containing various specialized clusters of related pathways: biosynthesis pathways of primary metabolites, secondary metabolites, and co-factors; catabolic pathways; glycolysis and the TCA cycle; photosynthesis and energy-related pathways, etc. (Figure 3). In general, nodes of various shapes (for example, circle, diamonds, square, etc.) depict metabolites and lines depict either a metabolic or a transport reaction. Users can zoom into detailed view (by clicking on the widget located on the top left-hand side corner) to see the names of metabolites, reactions and associated enzymes with corresponding EC numbers (Figure 3). Typically, reactions that have not yet been assigned to a pathway represent the clutter at the extreme right side of the “Cellular Overview Diagram.” Additional options such as highlighting chosen entities and uploading of user-defined data on the “Cellular Overview Diagram” are available under the “OPERATIONS” menu on the right-hand side of the page (Figure 3).

Manual Curation of FragariaCyc
We constructed a strawberry metabolic network based on extensive computational analysis, automated and manual curation. Manual curation includes (i) confirmation of computational mappings of genes to reactions, and pathways; (ii) addition/editing of strawberry-specific compounds, reactions, sub-pathways, pathways and super-pathways including assignment of genes to the relevant enzyme and pathways based on gene orthology and from published studies; and (iii) deletion of nonplant pathways.

We added 773 compounds manually to FragariaCyc based on literature survey (Aharoni et al., 2004; Lunkenbein et al., 2006a; Fait et al., 2008; Osorio et al., 2011; Zhang et al., 2011; Rohloff et al., 2012; Kim et al., 2013; Gunduz and Ozdemir, 2014; Najda et al., 2014; Schwieterman et al., 2014; Voca et al., 2014; Xu et al., 2014). At present ~50% of the compounds out of the total 2069 are supported by experimental evidences. Many of these compounds impart high nutritional value to strawberries, such as flavonols, anthocyanins, and their derivatives. We have extensively edited biosynthesis pathways of several secondary metabolites including plant hormones (e.g., auxin, abscisic acid, brassinosteroids, ethylene, gibberellins and...
gibberellic precursors, jasmonic acid, etc.), phenylpropanoids, trans-lycopene, pro-anthocyanidins, isoflavonoids, zeaxanthin, beta-carotene, plant sterols, chlorophyllide a, and mevalonate.

Recently, Darwish et al. (2015) have revised annotation of *F. vesca* genome v1.1 based on transcriptomes of 25 different tissues in combination with the MAKER2 annotation pipeline (https://www.rosaceae.org/species/fragaria_vesca/genome_v1.1.a2) that identified 2286 new gene models, and 6006 new exons. The efforts to integrate these new genes and revised gene models into FragariaCyc are ongoing. However, we have updated annotations for 6598 genes based on the recent transcriptomic studies from *F. vesca* (Darwish et al., 2013, 2015; Kang et al., 2013).

We have not curated transport reactions in FragariaCyc so far, but FragariaCyc contains 101 transport reactions based on the automated baseline projections from MetaCyc.

**Omics Viewer Analysis of RNA-Seq Data Suggests Tissue-Specific Regulation of Metabolic Pathways**

The Omics Viewer tool enables users to upload and analyze a wide array of expression data (e.g., transcriptomes, proteomes, metabolomes, reaction flux data, etc.). We chose a subset of RNA-Seq expression data from the *F. vesca* line YW5AF7 (Kang et al., 2013) representing the differential expression of 16,316 genes during berry development in three tissues, ovule, vegetative cortex, and embryo (see Supplementary Table 1).

We were able to map 16,271 out of the total 16,316 genes on FragariaCyc. The remaining 45 genes belong to newly identified *F. vesca* genes (Darwish et al., 2015) awaiting their addition into the database. We successfully mapped 2378 genes to the metabolic pathways and compared their expression...
across various samples. **Figure 3A** shows a “Cellular Overview Diagram” displaying the transcriptome of cortex4 (collected from stage-4 small green berry, cutoff value = 500). The “Cellular Overview Diagram” is an interactive platform, where users can click on reactions, compounds, and pathways to access more information about a particular entity and to navigate to respective detail pages. For example, a click on the initial reaction of the ascorbate biosynthesis I pathway opens a popup window showing the differential expression of gene20851 in samples #1-#10 (**Figure 3B**). Similarly, users can display expression profiles of the gene homologs (**Figure 3C**) that map to the same enzyme. It is evident from the expression profiles of several gene-duplicates that homologous genes are likely to respond to different signals or regulators. Additional options from the right-hand side panel allow highlighting pathways/reactions/genes within the “Cellular Overview.” Users can also use the pathway customization tool to import expression data onto desired pathway pages. For example, **Figure 4A** shows the mapping of transcriptomic data on the abscisic acid biosynthesis pathway.

Furthermore, we generated a table of differentially regulated pathways (cutoff = 200) to identify ovule-, embryo-, and cortex-specific pathways. As shown in **Figure 4**, enzymes involved in the biosynthesis pathways of ethylene, 4-coumarate, flavonoids, chorismate, and quercetin are highly expressed in the cortex and/or in ovule1 and ovule2, but are clearly repressed in the embryo. In contrast, several genes of oleate biosynthesis show preferential expression in ovule1 and ovule2, and some expression in the developing embryo, but clearly repressed in the cortex (**Figure 4C**). For example, genes involved in fatty acid activation are clearly repressed in embryo and cortex, but express in the ovule. Another gene encoding an ammonium transporter expresses in the ovule and embryo, but not in the cortex (**Figure 4C**).

We investigated changes in the expression of genes involved in the biosynthesis of various plant hormones during the early fruit development in embryo and cortex, and pre- and post-fertilized ovules. In general, auxin biosynthesis genes show the highest expression in the tissues derived from small green berries and then slowly decline as the berry grows. Many embryo-specific genes involved in indole-3-acetate biosynthesis I, such as gene11728 (encoding indole-3-pyruvate monooxygenase), gene03586 (encoding L-tryptophan aminotransferase), gene04251 (encoding carboxylesterase), and gene22572 (encoding an amidotransferase) show the highest expression embryo4 tissue (from small green berries) and then show a decline in their expression in embryo5 (from big green berry). Similarly, other auxin biosynthesis pathway genes expressed in the cortex and ovules showed higher expression in stage-4 small green berry and decreased expression in stage-5 big green berry. Gibberellins (GA) biosynthesis genes (e.g., gene19437, gene13360, gene01059, etc.) show higher expression in embryo following fertilization event. Low expression of a
FIGURE 3 | A view of the F. vesca “Cellular Overview Diagram” displaying gene expression data from cortex4 tissue sample collected from small green berries. The color scale used for depicting the expression level is shown (A). The popup windows show the tissue-specific expression profile of a gene (B) or genes mapped to the same reaction (C). Samples: #1, ovule1 (from open pre-fertilized flower); #2, ovule2 (achene collected just after fertilization containing a globular stage embryo); #3, embryo3 (heart-stage embryo from very small green berry); #4, embryo4 (torpedo-stage embryo from small green berry); #5, embryo5 (mature embryo with two cotyledons from big green berry); #6, cortex1 (from pre-fertilized flower); #7, cortex2 (from pollinated flower, 2–4 DPA); #8, cortex3 (from very small green berry); #9, cortex4 (from small green berry); #10, cortex5 (from big green berry) (Kang et al., 2013).

few GA biosynthesis genes was observed in the cortex tissue as well. In general, no significant change in the expression of GA biosynthesis gene was observed between the samples obtained from small and big green berries (Supplementary Figure 1B).

The abscisic acid biosynthesis genes show a gradual increase in their expression during berry development in the cortex tissue, eventually attaining a peak at stages 4–5 (Figure 4A). However, no significant change or very little increase in their expression was observed in the embryos. The genes involved in initial steps of ethylene biosynthesis show higher expression across all tissue types. However, gene01202 and gene07935 (encoding for the enzyme 1-aminocyclopropane-1-carboxylate oxidase catalyzing the final step) show higher expression in the cortex, compared to the embryo (Figure 4B). The majority of jasmonic acid biosynthesis genes show preferential expression in the cortex throughout the early stages of berry development, although gene16355, gene28898, and gene16355 show embryo-specific expression (Supplementary Figure 1C). The majority of brassinosteroid biosynthesis genes show high expression in the ovule, embryo and cortex tissues (Supplementary Figure 1D), although a few genes show highly tissue-specific expression profile (e.g., gene20962, gene19221, gene03747).

Overall, the Omics Viewer analysis can reveal information about subcellular/cell-/tissue-/organ-specific localization of a metabolic pathway; differential expression of gene family members mapped to the same reaction; and help to identify the rate-limiting step within a pathway—all required for the identification of candidate genes for the purpose of cultivars improvement.

DISCUSSION

To date, FragariaCyc is the only resource that provides a comprehensive knowledgebase of strawberry metabolism supported by extensive manual curation and allows analysis of omics data in the context of the cellular-level metabolic network. FragariaCyc is based on the sequenced genome of F. vesca (Shulaev et al., 2011) and shares functionalities with other BioCyc pathway databases including several species-specific plant pathway databases (Mueller et al., 2003; Urbanczyk-Wochniak and Sumner, 2007; Dharmawardhana et al., 2013; Monaco et al., 2013; Naithani et al., 2014).

The published literature and publicly available experimental data serve as a vital resource for improving the functional annotations and curation of the genes, enzymes, reactions, pathways, and compounds in FragariaCyc. We have updated gene annotations in FragariaCyc using information from several published studies (Aharoni et al., 2002; Wein et al., 2002; Bianco et al., 2009; Osorio et al., 2011; Darwish et al., 2013, 2015; Kang et al., 2013; Xu et al., 2014). In 2011, an improved version of the F. vesca genome assembly became available (https://www.rosaceae.org/species/fragaria/fragaria_vesca/genome_v1.1), but no new genes were predicted. Recently, a large-scale transcriptomic study (Darwish et al., 2015) identified 2286 new genes and 6006 new exons (by correcting the exon-intron boundaries in several genes). Our efforts to incorporate newly identified and revised genes in FragariaCyc, and their mapping to appropriate reactions and pathways are ongoing. We expect that FragariaCyc as a community resource will continue to mature.
FIGURE 4 | Visualization of differentially expressed metabolic pathways. (A) Abscisic acid biosynthesis pathway displaying expression profile of various associated genes. (B) A table of ovule or cortex-specific pathways with expression profiles of associated genes. (C) Ovule-specific pathways. By clicking on the genes and compounds more information can be accessed. The samples refer as #1, ovule1; #2, ovule2; #3, embryo3; #4, embryo4; #5, embryo5; #6, cortex1; #7, cortex2; #8, cortex3; #9, cortex4; #10, cortex5 (for details see Figure 3 legend). The common color scale shown in (A) depicts the expression level of genes and also applies to (B,C).

as more experimental data become available and integrated into it.

We anticipate that FragariaCyc will aid strawberry researchers in (i) discovering the broader role of new and known genes/molecular interactions in the context of overall cellular metabolism; (ii) understanding how a plant’s environment and genetic diversity relates to yields and the quality of its fruit (e.g., texture, taste, aroma, flavor, etc.); and (iii) exploring how the plant’s overall metabolism and its ability to adapt are affected in response to climate change. To show the utility of the Omics Viewer, we analyzed a subset of publicly available RNA-Seq data from *F. vesca* (Kang et al., 2013) to understand tissue-specific regulation of the metabolic pathways during fruit development.

Figure 3 shows visualization of RNA-Seq data over “Cellular Overview Diagram.” Users can customize the display of expression data by defining a cut-off value, and/or choosing an available color schema to depict up- and down-regulation of the genes. When data from multiple samples are uploaded, an automated animation displays painted Cellular Overview Diagrams in sequential order. Users have options to capture this animation as a movie (by using a third-party software) or to pause it at desired sample (e.g., Figure 3) for further exploration. The users can also display expression data on the pathway diagram, and generate a table of pathways in which one or more genes show differential expression (Figure 4).

The unique feature of Omics Viewer analysis is its ability to compare expression profiles of paralogs mapped to the same reaction considering extensive gene duplications in plant genomes (Shiu and Bleecker, 2003; Jaillon et al., 2007; Velasco et al., 2007; Myburg et al., 2014; Renny-Byfield and Wendel, 2014). The information about differential expression of the homologous genes can help researchers to articulate functional redundancies. Such analysis is very useful for characterizing a mutant/phenotype, and for shortlisting candidate genes for functional characterization using transient (Guidarelli and Baraldi, 2015) or stable transformation systems available for strawberry (Slovin et al., 2009). Also, the mapping of the expression profile of genes in a given pathway can identify (i) a rate limiting step in a pathway, and (ii) clusters of pathways regulated in response to a common signal/regulator.

Strawberry is not a true fruit. Indeed, the fleshy edible fruit, known as “receptacle,” develops from the stem tip instead of an ovary. The receptacle is the primary site of accumulation of various nutrients and metabolites (Fait et al., 2008; Bianco et al., 2009) that impart color, flavor and taste to the berry.
The receptacle is made of two tissues: the fleshy tissue placed immediately underneath achenes is called the cortex, and the interior tissue is known as the pith. The seed-like achenes, embedded in the fleshy tissue (receptacle) are the real fruit as they develop from an ovary and contain developing embryo inside. Thus, strawberry offers an interesting system to explore how various accessory fruits relate to each other, and to the true fleshy fruits (e.g., grape and tomato) regarding their global gene expression profile, cell signaling, and metabolic milieu.

Overall, our analysis suggests that the majority of auxin and gibberellin biosynthesis genes show preferential expression in the embryo, which is consistent with their role in driving early fruit development (Nitsch, 1950; Dreher and Poovaiah, 1982; Csukasi et al., 2011; Kang et al., 2013). The levels of free auxins in both the achene and the receptacle have been shown to rise after fertilization, peak at small green fruit stage, and then decline slowly in big green fruit (Dreher and Poovaiah, 1982; Symons et al., 2012).

In contrast, genes involved in the biosynthesis of the abscisic acid, ethylene, and jasmonic acid show higher expression in the cortex tissue when compared with embryos, and their expression peaks at the stage-4 small green or stage-5 big green berries. Abscisic acid has been shown to play a key role in the ripening of strawberry (Jia et al., 2011; Li et al., 2011; Vallarino et al., 2015) and tomato (Zhang et al., 2009), likely by triggering ethylene biosynthesis. Ethylene is known as a major driver of ripening in climacteric fruits, such as apple and tomato (Seymour et al., 2013). However, the interplay between various hormones is more likely to foster ripening transition in non-climacteric fruits, such as strawberry (Kang et al., 2013; Mcatee et al., 2013; Osorio et al., 2013; Fortes et al., 2015). Recent studies suggest some role of ethylene in non-climacteric fruit ripening in both grapes and strawberries (Merchant et al., 2013; Lopes et al., 2015). Brassinosteroids have also been proposed to positively influence ripening associated processes, whereas, jasmonic acid seems to interfere with the ripening (Jia et al., 2011; Concha et al., 2013; Merchant et al., 2013). The large-scale gene expression studies suggest similarities at the molecular level between climacteric and non-climacteric ripening process. However, the relative contribution of various hormones in fruit ripening could vary between two categories.

FragariaCyc is freely accessible online at http://pathways.cgrb.oregonstate.edu. The data content in FragariaCyc will updated once a year as long as we have some support for curation.

Afterward, an archive copy of FragariaCyc will be maintained on our website and will be deposited in a public repository. Users are encouraged to view an online video on metabolic pathway databases available at http://biocyc.org/webinar.shtml.

**AUTHOR CONTRIBUTIONS**

SN and PJ conceived the project. SN led the project and its overall curation process. SN and CP did curation. PJ and RR constructed the automated projections of the FragariaCyc. JE carried out the Inparanoid-based annotations and provided technical support. SN analyzed RNA-Seq expression data. SN and PJ wrote the manuscript.

**FUNDING**

This work was supported by the startup funds provided to SN and PJ by Oregon State University.

**ACKNOWLEDGMENTS**

We thank Peter Karp and his group (SRI International) for providing excellent support for Pathway-Tool software and Prof. Zhongchi Liu (University of Maryland) for providing expression data used in this study for Omics Viewer analysis.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00242

**Supplementary Table 1 | RNA-Seq expression data used for Omics Viewer analysis.** This is a subset of publicly available data from *F. vesca* (Kang et al., 2013) that represent differential expression of 16,316 genes across 10 samples: ovule1 and cortex1 were from open flower prior to fertilization, ovule2 and cortex2 from pollinated flowers just after fertilization, while the cortex3-5 and embryo3-5 were collected from the developing strawberry as described previously (Kang et al., 2013).

**Supplementary Figure 1 | Differential expression of genes using absolute RPKM values (cutoff =50) associated with gibberellin (B), jasmonic acid (C), and brassinosteroid (D), biosynthesis pathway during fruit development in three tissues: ovule, embryo, and cortex.** A common color scale shown (A) depicts the expression level of genes belonging to pathways (B–D). The samples refer as #1, ovule1; #2, ovule2; #3, embryo3; #4, embryo4; #5, embryo5; #6, cortex1; #7, cortex2; #8, cortex3; #9, cortex4; #10, cortex5 as described in Figure 3.

**REFERENCES**

Aharoni, A., Giri, A. P., Verstappen, F. W., Bertea, C. M., Sevenier, R., Sun, Z., et al. (2004). Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. Plant Cell 16, 3110–3131. doi: 10.1105/tpc.104.023895

Aharoni, A., Keizer, L. C., Van Den Broeck, H. C., Blanco-Portales, R., Munoz-Blanco, J., Bois, G., et al. (2002). Novel insight into vascular, stress, and auxin-dependent and -independent gene expression programs in strawberry, a non-climacteric fruit. Plant Physiol. 129, 1019–1031. doi: 10.1104/pp.003558

Aharoni, A., and O’Connell, A. P. (2002). Gene expression analysis of strawberry achene and receptacle maturation using DNA microarrays. J. Exp. Bot. 53, 2073–2087. doi: 10.1093/jxb/erf026

Alm, R., Ekefjord, A., Krogh, M., Hakkinen, J., and Emanuelsson, C. (2007). Proteomic variation is as large within as between strawberry varieties. J. Proteome Res. 6, 3011–3020. doi: 10.1021/pr0700450

Bianco, L., Lopez, L., Scalone, A. G., Di Carlì, M., Desiderio, A., Benvenuto, E., et al. (2009). Strawberry proteome characterization and its regulation during fruit ripening and in different genotypes. J. Proteomics 72, 586–607. doi: 10.1016/j.jprot.2008.11.019

Bombarely, A., Merchante, C., Csukasi, F., Cruz-Rus, E., Caballero, J. L., Medina-Escobar, N., et al. (2010). Generation and analysis of ESTs from strawberry (*Fragaria xananassa*) fruits and evaluation of their utility in genetic and molecular studies. BMC Genomics 11:503. doi: 10.1186/1471-2164-11-503

Boyer, J., and Liu, R. H. (2004). Apple phytochemicals and their health benefits. Nutr. J. 3:5. doi: 10.1186/1475-2891-3-5
Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., et al. (2014). The genome of Eucalyptus grandis. Nature 510, 356–362. doi: 10.1038/nature13308

Naithani, S., Raja, R., Waddell, E. N., Elser, J., Gouthu, S., Deluc, L. G., et al. (2014). VitisCyc: a metabolic pathway knowledgebase for grapevine (Vitis vinifera). Front. Plant Sci. 5:444. doi: 10.3389/fpls.2014.00644

Najda, A., Dyduch-Sieminska, M., Dyduch, J., and Gantner, M. (2014). Comparative analysis of secondary metabolites contents in Fragaria vesca L. fruits. Ann. Agric. Environ. Med. 21, 339–343. doi: 10.5604/1232-1966.1108601

Nijveldt, R. J., Van Nood, E., Van Hoorn, D. E., Boelen, P. G., Van Norren, K., and Van Leeuwen, P. A. (2001). Flavonoids: a review of probable mechanisms of action and potential applications. Am. J. Clin. Nutr. 74, 418–425.

Nitsch, J. (1950). [Growth and morphogenesis of strawberry the strawberry as related to auxin.] A. J. Bot 37, 211–215. doi: 10.2307/2437903

Osorio, S., Bombarely, A., Giavalisco, P., Usadel, B., Stephens, C., Aragüez, L., et al. (2011). Demethylation of oligogalacturonides by FaPEI in the fruits of the wild strawberry Fragaria vesca triggers metabolic and transcriptional changes associated with defence and development of the fruit. J. Exp. Bot. 62, 2855–2873. doi: 10.1093/jxb/erq465

Osorio, S., Scossa, F., and Fernie, A. R. (2013). Molecular regulation of fruit ripening. Front. Plant Sci. 4:198. doi: 10.3389/fpls.2013.00198

Paniagua, C., Blanco-Portales, R., Barcelo-Munoz, M., Garcia-Gago, J. A., Waldron, K. W., Quesada, M. A., et al. (2016). Antisense down-regulation of the strawberry β-galactosidase gene FaβGal during cell wall galactose levels and reduces fruit softening. J. Exp. Bot. 67, 619–631. doi: 10.1093/jxb/erv462

Perez, A. G., Olias, R., Olias, J. M., and Sanz, C. (1999). Biosynthesis of 4-hydroxy-2,5-dimethyl-3(2H)-furanone and derivatives in in vitro grown strawberries. J. Agric. Food Chem. 47, 655–658. doi: 10.1021/jf980404t

Poole, R. L. (2007). The TAIR database. Methods Mol. Biol. 406, 179–212.

Prat, L., Espinoza, M. I., Agosin, E., and Silva, H. (2014). Identification of the strawberry β-galactosidase gene FaβGal and its expression during cold stress induced ripening and development. Plant Mol. Biol. 85, 418–425. doi: 10.1007/s11103-014-9936-2

Shiu, S. H., and Bleecker, A. B. (2003). Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. Plant Physiol. 132, 530–543. doi: 10.1104/pp.103.019264

Shulaev, V., Sargent, D. J., Crowhurst, R. N., Mockler, T. C., Folkers, O., Delcher, A. L., et al. (2011). The genome of woodland strawberry (Fragaria vesca). Nat. Genet. 43, 109–116. doi: 10.1038/ng.740

Slivov, J. P., Schmitt, K., and Folta, K. M. (2009). An inbred line of the diploid strawberry Fragaria vesca L. semperflorens for genomic and molecular genetic studies in the Rosaceae. Plant Methods 5, 15. doi: 10.1186/1746-4811-5-15

Symons, G. M., Chua, Y. J., Ross, J. I., Quittenden, L. J., Davies, N. W., and Reid, J. B. (2012). Hormonal changes during non-climacteric ripening in strawberry. J. Exp. Bot. 63, 4741–4750. doi: 10.1093/jxb/ers147

Tello-Ruiz, M. K., Stein, J., Wei, S., Preece, J., Olson, A., Naithani, S., et al. (2016). Gramene 2016: comparative plant genomics and pathway resources. Nucleic Acids Res. 44, D1133–D1140. doi: 10.1093/nar/gkv1179

Thimm, O., Blasing, O., Gibson, Y., Nagel, A., Meyer, S., Kruger, P., et al. (2004). MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J. 37, 914–939. doi: 10.1111/j.1365-313X.2004.02016.x

Urbanczyk-Wochniak, E., and Sumner, L. W. (2007). Medicyc: a biochemical pathway database for Medicago truncatula. Bioinformatics 23, 1418–1423. doi: 10.1093/bioinformatics/btm040

Vallarino, J. G., Osorio, S., Bombarely, A., Casanal, A., Cruz-Rus, E., Sanchez-Sevilla, J. F., et al. (2015). Central role of FaGA MYB in the transition of the strawberry receptacle from development to ripening. New Phytol. 208, 482–496. doi: 10.1111/nph.13463

Velasco, R., Zahirikhi, A., Troggio, M., Cartwright, D. A., Cestaro, A., Pruss, D., et al. (2007). A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. Pla S ONE 2:e2362. doi: 10.1371/journal.pone.0001326

Voca, S., Zlabur, J. S., Dobricevic, N., Jakobek, L., Seruga, M., Galic, A., et al. (2014). Variation in the bioactive compound content at three ripening stages of strawberry fruit. Molecules 19, 10370–10385. doi: 10.3390/molecules190710370

Wein, M., Navid, N., Lenkenstein, B., Lewinsohn, E., Schwab, W., and Kaldenhoff, R. (2002). Isolation, cloning and expression of a multifunctional O-methyltransferase capable of forming 2,5-dimethyl-4-methoxy-3(2H)-furanone, one of the key aroma compounds in strawberry fruits. Plant J. 31, 755–765. doi: 10.1046/j.1365-313X.2002.01396.x

Wilke, A., Bischof, J., Gerlach, W., Glass, E., Harrison, T., Keegan, K. P., et al. (2016). The MG-RAST metagenomics database and portal in 2015. Nucleic Acids Res. 44, D590–D594. doi: 10.1093/nar/gkv1322

Xia, R., Ye, S., Liu, Z., Meyers, B. C., and Liu, Z. (2015). Novel and Recently Evolved MicroRNA Clusters Regulate Expansive F-BOX Gene Networks through Phased Small Interfering RNAs in Wild Diploid Strawberry. Plant Physiol. 169, 594–610. doi: 10.1093/pl Physiol.169.3.594

Xu, W., Peng, H., Yang, T., Whitaker, B., Huang, L., Sun, J., et al. (2014). Effect of calcium on strawberry fruit flavonoid pathway gene expression and anthocyanin accumulation. Plant Physiol. Biochem. 82, 289–298. doi: 10.1016/j.plaphy.2014.06.015

Zhang, J., Wang, X., Yu, O., Tang, J., Gu, X., Wan, X., et al. (2011). Metabolic profiling of strawberry (Fragariaxananassa Duch.) during fruit development and maturation. J. Exp. Bot. 62, 1103–1118. doi: 10.1093/jxb/erq343

Zhang, M., Yuan, B., and Leng, P. (2009). The role of ABA in triggering ethylene biosynthesis and ripening of tomato fruit. J. Exp. Bot. 60, 1579–1588. doi: 10.1093/jxb/erp026

Zhang, P., Dreher, K., Karthikeyan, A., Chi, A., Pujar, A., Caspi, R., et al. (2010). Generation of a genome-wide metabolic pathway database for Populus trichocarpa using a new approach for reconstruction and curation of metabolic pathways for plants. Plant Physiol. 153, 1479–1491. doi: 10.1104/pp.110.157396

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Naithani, Partipilo, Raja, Elser and Jaiswal. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided that the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.