CrossMark

OPINION PAPER

A standard nomenclature for gene designation in the Rosaceae

Sook Jung[1] · Carole Bassett[2] · Douglas G. Bielenberg[3] · Chun-Huai Cheng[1] ·
Chris Dardick[2] · Dorrie Main[1] · Lee Meisel[4] · Janet Slovin[5] · Michela Troggio[6] ·
Robert J. Schaffer[7,8]

Abstract The plant family Rosaceae is comprised of a highly diverse group of plants, with a wide range of ornamental and agricultural uses throughout the world. Comparative genomic analyses between rosaceous species are increasingly being used to assign putative function. With the wealth of new data, there is a need to standardize nomenclature to ensure uniformity and clarity across research groups. The Rosaceae Gene Name Standardization Committee was formed to develop a naming guideline for genes in Rosaceae family members. This manuscript details the issues involved with naming genes and proposes a common nomenclature system. The Genome Database for Rosaceae (GDR) (www.rosaceae.org) has developed a gene database that will support user-submitted gene names and annotations in addition to the predicted genes from whole genome sequences. GDR also hosts user-curated lists of gene class symbols. To facilitate standardization of gene names and avoid misinterpretation, the committee recommends that the GDR serve as the database of record for the names of Rosaceae genes and that gene names, gene class symbols, and gene annotation be registered in GDR prior to publication in peer reviewed journals.

Keywords Rosaceae · Gene naming · Gene database · Gene families

✉ Sook Jung
sook_jung@wsu.edu

✉ Robert J. Schaffer
robert.schaffer@plantandfood.co.nz

1 Department of Horticulture, Washington State University, Pullman, WA 99164-6414, USA

2 USDA-ARS-Appalachian Fruit Research Laboratory, Kearneysville, WV 25430, USA

3 Department of Biological Sciences, Clemson University, Clemson, SC 29634-0314, USA

4 Universidad de Chile, Instituto de Nutrición y Tecnología de los Alimentos (INTA), El Líbano 5524, 7830490 Macul, Santiago, Chile

5 USDA/ARS Genetic Improvement of Fruits and Vegetables Laboratory, Beltsville, MA 20705, USA

6 Istituto Agrrario di San Michele All'Adige, Fondazione, Edmund Mach, San Michele All'adige, Italy

7 Plant and Food Research Ltd, 120 Mount Albert Road, Auckland, New Zealand

8 School of Biological Sciences, University of Auckland, Auckland, New Zealand

Introduction

As the amount of gene-related data grows, a standard nomenclature for gene names should be adopted to ensure uniformity. In the absence of a universally accepted naming system, researchers have typically tried to follow naming precedents

set out in published literature. For example, a two-letter species abbreviation is typically used as a prefix in gene names to identify the species of origin. For *Arabidopsis thaliana*, genes carry the prefix *At*. This species prefix is generally not required when presenting results from a single species but helps interpretation when genes from multiple species are being compared. While the use of precedent has helped reporting results more clearly, the need for naming conventions has become widely recognized especially as knowledge of genes and gene families has rapidly grown.

In response, a number of model plant communities such as *A. thaliana* (http://www.arabidopsis.org/portals/ nomenclature/guidelines.jsp), tomato (http://solgenomics.net/ static_content/solanaceae-project/docs/tomato-standards.pdf), and rice (McCouch 2008) have independently suggested naming methodologies. These were developed to address a number of issues around consistency, such as cases where the same gene has been given different names. For example, *ETHYLENE INSENSITIVE 5* has also been published as *AIN1* and *XRN4* (Olmedo et al. 2006). In tomato, the *ARF* gene family was independently named by two different research groups, resulting in each gene having two names (Kumar et al. 2011; Wu et al. 2011). whereon the other hand, in some cases, the same three letter abbreviation has been given to different genes, such as *CELL NUMBER REGULATOR* (*CNR*) in cherry (De Franceschi et al. 2013) and *COLORLESS NON RIPENING* (*CNR*) in tomato (Manning et al. 2006). Once a reference genome has become available, the problem can amplify, as independent groups annotate entire gene families. Although these problems have not been entirely resolved, clarity within genomic databases has been provided by the common naming convention of giving each gene a chromosomal locus identification number (for Arabidopsis AT1G54490) as part of whole genome sequencing and gene prediction. This convention was first established for Arabidopsis and subsequently followed by other model systems (McCouch 2008). The locus IDs provide a way to distinguish genes that have been given the same name, or to identify when a single gene has been given multiple names. While this remedy provides a mechanism for database management, it does not eliminate the problem of gene names and species designations not being translatable among various reports.

Genomics research in the Rosaceae research community is expected to grow due to the availability of powerful sequencing tools combined with the importance of commercially cultivated crops such as almond, apple, apricot, blackberry, cherry, peach, pear, plum, raspberry, rose, and strawberry. The opportunity now exists to provide the plant research community with a standardized naming convention that will help avoid significant confusion with regard to gene naming. With this in mind, the Rosaceae Executive Committee (RosEXEC) along with the Rosaceae International Genomics Initiative (RosIGI) formed a subcommittee to develop a standardized naming convention for naming genes among Rosaceae species. Adoption by the research community of the recommended gene naming conventions detailed here will benefit the Rosaceae research and breeding communities, as well as all associated fields of biology. We respectfully urge the research community as well as editors of peer reviewed journals to follow the naming conventions outlined here, and to work with the naming committee to address problems or make suggestions. Finally, it is worth noting that currently, gene names will be allocated on functional analysis occasionally tested in Rosaceae species but more often in model organisms. Currently predicted genes from whole genome sequencing with no predicted functions and/or phenotype would not be eligible for the gene names and gene symbols beyond the assigned genomic identifier.

Rosaceae gene naming convention—species prefix

The comparison of genes across Rosaceae species gives valuable insights into the way that different species have evolved. Whole genome sequence is available for strawberry (Shulaev et al. 2011), apple (Velasco et al. 2010), peach (Verde et al. 2013), pear (Wu et al. 2013; Chagné et al. 2014), and *Prunus mume* (Zhang et al. 2012), with more to come. Genome sequence has allowed cross-species comparisons at the genome level and within gene families. Considering the approximately 3000 species within the family Rosaceae (Hummer and Janick 2009), it is clear that two-letter species abbreviations will be insufficient to distinguish between species. Even among the species most commonly studied, conflicts have already arisen. Genes in *Prunus persica* and *Pyrus pyrifolia* as well as the model species of mosses, *Physcomitrella patens*, have already been given the same species prefix, *Pp*. To facilitate comparison, we recommend using a standard set of three-letter species prefixes for the major Rosaceae species (Table 1). The use of a three-letter code was chosen as it settled most naming conflicts without adding cumbersome length to gene names. In general, the first letter of the genus name and the first two letters of the species name should be used for the three-letter prefix. Using a three-letter prefix resulted in two conflicts for the major species: using our convention, both *Prunus cerasus* and *Prunus cerasifera* would be Pce, so we recommend Pci for *P. cerasifera*. Likewise, both *P. mume* and *Prunus munsoniana* would be Pmu, so we recommend Pmn for *P. munsoniana*.

In order to distinguish all taxonomic species within Rosaceae, including non-commercial species, abbreviations would need to be substantially longer. For these non-cultivated species, we suggest that authors take a UNIPROT approach (http://www.uniprot.org/docs/speclist) (Magrane and Consortium 2011) using five-letter abbreviations; three

Table 1 Proposed abbreviations for major Rosaceae species

Common name	Genus	Species	Prefix
	Aronia	*melanocarpa*	*Ame*
	Chaenomeles	*japonica*	*Cja*
	Cydonia	*oblonga*	*Cob*
	Eriobotrya	*japonica*	*Eja*
Strawberry	*Fragaria*	*x ananassa*	*Fan*
	Fragaria	*chiloensis*	*Fch*
Woodland strawberry	*Fragaria*	*vesca*	*Fve*
	Fragaria	*virginiana*	*Fvi*
Apple	*Malus*	*x domestica*	*Mdo*
	Malus	*pumila*	*Mpu*
	Malus	*sylvestris*	*Msy*
	Mespilus	*germanica*	*Mge*
	Prunus	*americana*	*Pam*
	Prunus	*angustifolia*	*Pan*
Apricot	*Prunus*	*armeniaca*	*Par*
Sweet cherry	*Prunus*	*avium*	*Pav*
Myrobalan plum	*Prunus*	*cerasifera*	*Pci*[a]
Tart cherry	*Prunus*	*cerasus*	*Pce*
Plum	*Prunus*	*domestica*	*Pdo*
Almond	*Prunus*	*dulcis*	*Pdu*
	Prunus	*hortulana*	*Pho*
Japanese apricot	*Prunus*	*mume*	*Pmu*
	Prunus	*munsoniana*	*Pmn*[a]
	Prunus	*nigra*	*Pni*
Peach	*Prunus*	*persica*	*Ppe*
Japanese plum	*Prunus*	*salicina*	*Psa*
Black cherry	*Prunus*	*serotina*	*Pse*
	Prunus	*simonii*	*Psi*
Sloe	*Prunus*	*spinosa*	*Psp*
	Pyrus	*calleryana*	*Pca*
Chinese white pear	*Pyrus*	*× bretschneideri*	*Pbr*
European pear	*Pyrus*	*communis*	*Pco*
Asian pear	*Pyrus*	*pyrifolia*	*Ppy*
	Pyrus	*ussuriensis*	*Pus*
Raspberry	*Rubus*	*idaeus*	*Rid*
Black raspberry	*Rubus*	*occidentalis*	*Roc*

[a] Exceptions of the three-letter prefix rule: the first letter of Genus and the first two letters of species names

for the genus name and two for the species such that *Potentilla simplex* (common cinquefoil) becomes *Potsi*, *Waldsteinia fragarioides* becomes *Walfr*, and *Malus platycarpa* (Bigfruit crab apple) becomes *Malpl*. This convention is still not sufficient to account for all members of the Rosaceae family and conflicts may still arise. In such cases, we recommend that researchers choose distinguishing nomenclature following the guidelines set forth here as closely as possible, making exceptions as needed to avoid conflicts with other species both

within and across other plant families. Furthermore, we strongly recommend that authors do not include a species prefix in the gene symbol when submitting the gene data to NCBI, GDR, or any other databases, so as to minimize the creation of duplicated names due to the species prefix. In databases, the species data are typically stored along with the gene names, so having a species prefix in gene names is unnecessary and generates additional aliases. In GDR, the gene symbol without the species prefix will be stored with already published genes names, with or without prefix, as aliases. Authors should use the prefix only in publications for comparisons between genes of different species origin.

Rosaceae gene naming conventions—gene name

When giving a gene a name/symbol, we encourage researchers to first check with the current literature, and within the gene databases (GenBank and GDR), to ensure that the gene they want to name has not already been assigned a name in a prior publication. It is not recommended that researchers rename genes that are already published unless a compelling reason exists such as a lack of clear orthology or to correct existing confusion. When dealing with species that have high heterozygosity, it is occasionally hard to know whether differences in sequence are allelic, a result of gene duplication, or genome duplication. Researchers should aim to reduce the incidences where the same gene name has been used for different genes and where a gene has been assigned different names, as has already been observed in other species. From a community perspective, we aim to achieve a unique name for each gene (where a name has not already been established in the literature), a single name for each gene, and a link between gene name and gene model number from whole genome sequencing.

Gene naming by function, mutant phenotype, and homology

When naming genes in Rosaceae species, we propose the common nomenclature that is now standard across most of the genomics community. The Arabidopsis community has set out clear guidelines for gene nomenclature which should be followed for the Rosaceae (Meinke and Koornneef 1997) (http://www.arabidopsis.org/portals/nomenclature/guidelines.jsp). It is ideal to design the name so that it can be associated with biological function or mutant phenotype. When the biological function of the gene has not been directly assayed, there is an advantage to assigning the name of the most similar, functionally tested gene. For example, *Malus × domestica ACC SYNTHASE1* (*ACS1*) would be named after *A. thaliana ACS1*. Co-naming with an *Arabidopsis* gene may often pose a challenge due to gene duplications in either

species through evolution, and unknown or unclear biological function for a given gene family member making it impossible to equate two genes other than by sequence. However, especially for regulatory genes such as transcription factors, we recommend that whenever possible, similar naming and numbering schemes are used.

The guidelines set out by the Arabidopsis community point out the pitfalls of naming a gene based solely on a mutant phenotype or allelic form without knowing detailed biological or biochemical function. To relate a gene to a gene family, the Arabidopsis guidelines recommend the gene name to end with "-like" when information is based solely on sequence homology. Many of the Rosaceae genes may not be characterized functionally in host species, however, and be named following orthologous relationship with genes functionally characterized in Arabidopsis or other model species. So, we recommend naming genes following the closest orthologs in other model species but not adding "like" at the end.

Another recommendation from the Arabidopsis community is that gene names should not be assigned unless a full-length cDNA sequence has been obtained. For the Rosaceae, it is advantageous to accommodate genes that are identified using various sequencing methods, including those identified through high throughput transcriptome sequencing, manual editing of multiple alignments, or other experimental or computational approaches. Therefore, GDR will accept gene models with these lower constraints and ask that when authors submit gene names/symbols to GDR, they also indicate the category of evidence for the gene structure. Four categories of evidence for a gene structure have currently been established: (A) cDNA sequencing, (B) transcriptome sequencing, (C) computational evidence, and (D) other. Detailed information about these categories of evidence is shown in Table 2. This classification system will allow the research community to take advantage of all available data while enabling individual datasets to be filtered as needed based on these evidence codes.

Assigning gene symbols

For the gene symbol, we encourage using a standard three-letter code class symbol for members of a gene family together with a hierarchical numbering system. When two- or five-letter codes have already been used for established gene families, authors are recommended to use these for consistency and avoid inventing new names. The class symbol should be derived from the full name of the gene. The current list of gene class symbols are available in GDR (http://www.rosaceae.org/gene_class). Gene symbol can be followed by a numeric suffix. A name without the numeric suffix is presumed to be the first gene with a particular function that has been identified and therefore is the equivalent of suffix "1." A numeric suffix greater than one should appear when the new gene has similar function or phenotype to a gene at one or more other loci. The numerical suffix from Arabidopsis will not necessarily be conserved in Rosaceae due to duplication events in the both species. As stated earlier, organism-specific prefixes are encouraged to be used only in publications for clarity and should not be part of the gene symbol.

Naming homoeologs, alleles, and splice variants

Within the same taxonomic family, it is advantageous to name genes after homoeologs that have already been named in closely related species. However, within the Rosaceae, genome duplications have occurred, resulting in nomenclature difficulties. For example, the Maloidae has undergone a genome duplication followed by a rearrangement leading to a haploid (x) chromosome number of 17, compared to 7, 8, or 9 of most other family members (Illa et al. 2011) thus presenting the possibility of having two genes that have a homoeologous relationship to a single gene in diploid strawberry or peach. Some effort has been made to give specific names to two genes derived from the whole genome duplication in apple by using related numerical suffixes. For example, two apple genes were named *ARF1* and *ARF101* following the strawberry homoeologous gene *ARF1* (Devoghalaere et al. 2012). This solution was chosen at the time as a few studies had shown homoeologous genes maintain a degree of conservation in function, for example *MYB10* (Espley et al. 2007) and *MYB110* (Chagné et al. 2013) both control the anthocyanin accumulation. Due to the practical difficulties in the identification of two genes with the same ancestral origin out of multiple homologs in a species that has undergone whole genome duplication, no specific naming convention is currently proposed. It is, however, recommended that at this time, genes arising by genome duplication be named sequentially within the gene families as is done for general homologs. When the

Table 2 Category of evidence for the gene structure

	Category name	Category
A	cDNA	Single molecule cloned cDNA or cDNA EST (single pass sequence)
B	Transcriptome	Compilation of RNA-seq, Sequencing of RT-PCRamplicons
C	Computational evidence	Computer generated annotation of a genome assembly, Manual or computer generated annotation of new genomic sequences
D	Other	Submitter must provide detailed information

gene symbols do not contain a number, a numeric suffix can be directly attached (e.g., *PG1*, *PG2*, etc.). If the gene symbol already has a number at the end, the numeric suffix should follow a period "." (e.g., *DHN3.1*, *DHN3.2*, etc.). If a gene has a published name, then these should be kept.

For naming alleles, we recommend a hyphen and a numeric suffix for the alleles (e.g., *DHN3.1-1*, *DHN3.1-2*) following the convention of the Arabidopsis community. For the multiple alleles found in diversity studies with multiple populations, authors are recommended to provide a table with alternate names for alleles. The alternate name will contain more information on the alleles, such as a suffix of database ID (e.g., NCBI accession number) instead of a numeric suffix. For example, the alternate name for *DHN2.1-1* will be specified as *DHN2.1-AB123456*. For the species that have a whole genome sequence available, the sequences of the reference genome should serve as the wild type sequence. For naming splice variants, we recommend an underscore and a numeric suffix (e.g., *DHN3.1_1*, *DHN3.1_2*).

Facilitation of gene naming standardization

Researchers are encouraged to submit their gene naming data to GDR (http://www.rosaceae.org/data/submission) in addition to NCBI prior to or concurrent with publication. While NCBI does not accept named genes that do not come from single molecule sequencing, the GDR database will (see above).

When researchers submit data to GDR, they will be asked to provide the following type of data:

1. Species: species from which the gene is sequenced.
2. Species prefix: species prefix to be used in publications and presentations following the recommendations of this manuscript.
3. Source germplasm: germplasm name and database/germplasm repository ID if available.
4. Gene symbol: gene symbol that is composed of two to five letters and a numeric suffix, derived from the gene name.
5. Gene name: gene full name derived from molecular function of the gene product, phenotype, or homology.
6. Synonym or alias: any other gene symbols that have been used for the gene.
7. Gene class symbol: the "root" of the gene symbol, composed of two to five letters, without the numeric suffix.
8. Gene class symbol full name: full name of the gene class symbol (gene name without the numeric suffix).
9. Gene model: gene model ID from the whole genome annotation if available.
10. Genbank ID: Genbank accession number if available.
11. Description: description of characteristics of the gene such as biochemical function, expression in particular tissue and/or growth stage, effects on phenotype, and location in subcellular component.
12. Submitter: person who submitted the gene name data and email contact information.

Table 3   Example of gene data

Species	*Prunus persica*	*Malus x domestica*	*Prunus persica*	*Fragaria vesca*
Species prefix	*Ppe*	*Mdo*	*Ppe*	*Fve*
Source germplasm		Granny Smith		*Hawaii4* and *YW5AF7*
Gene symbol	*DHN2*	*PG1*	*CKX3*	*HSP17.1*[a]
Gene name	*DEHYDRIN 2*	*POLYGALACTURONASE 1*	*CYTOKININ OXIDASE 3*	*HEAT SHOCK PROTEIN*
Synonym	*XERO1*; *LEA*		*PpCKX3*	
Gene class symbol	*DHN*	*PG*	*CKX*	*HSP*
Gene class symbol full name	*DEHYDRIN*	*POLYGALACTURONASE*	*CYTOKININ OXIDASE*	*Small HEAT SHOCK PROTEIN*
Gene model	*ppa011637m*	*MDP0000326734*	*ppa021417*	*mrna08959*
Genbank ID	AY465376.1	P48978	DW341844.1; BU040402.1	CX662181;CX661653; EX688229;EX673870
Description	Similar to Xero1 and Rab18 of Arabidopsis thaliana	Ripening associated polygalacturonase	Cytokinin oxidase/dehydrogenase; cytokinin catabolism; oxidoreductase FAD-binding domain	Cytoplasmic small heat shock protein with Mr=17.1 kD
Submitter	Carole Bassett	Robert Schaffer	Lee Meisel	Janet Slovin
Category of evidence	A	A	A	A, B, C
Reference		Atkinson et al. 1994, 2012	Immanen et al. 2013	
Comments				

[a] Members of the small heat shock protein family are traditionally named by their apparent molecular weight, so will need to be exceptions to the naming conventions we propose

13. Category of evidence: the category of evidence as described above. If authors chose category D, detailed information needs to be given.
14. Reference: citation of the publication where the gene has been published.
15. Comments: any other comments.
16. Sequence data and gene model structure: sequence in FASTA format and gene model structure (intron/exon, UTR, promoter, etc.) in a GFF file.

Table 3 shows examples of gene data to be submitted. Supplementary Table 1 shows the gene submission template, available also from GDR (http://www.rosaceae.org/data/submission). When new gene models, gene sequences, and/or splice variants are identified that are different from the whole genome sequence data, we will recommend that users also submit, in addition to the gene data submission template, a FASTA file and a GFF file that contains gene model structure such as intron, exon and promoter region.

Conclusions

Establishing a standardized naming convention for Rosaceae species, compatible with what has been done in other plant research communities, will enable researchers to perform comparative analyses within as well as outside the Rosaceae Guidelines for naming Rosaceae family genes have been developed and are presented here together with the invitation to store gene nomenclature as well as sequencing information in the central Rosaceae community database, GDR. We respectfully urge the plant research community to follow these suggestions and by doing so, we should all benefit from simplified literature reading, and less confusion in reporting the identification and function of individual genes.

References

Atkinson RG (1994) A cDNA clone for endopolygalacturonase from apple. Plant physiology 105(4):1437–8.

Atkinson RG, Sutherland PW, Johnston SL, Gunaseelan K, Hallett IC, Mitra D, Brummell DA, Schröder R, Johnston JW, Schaffer RJ (2012) Down-regulation of POLYGALACTURONASE1 alters firmness, tensile strength and water loss in apple (Malus x domestica) fruit. BMC plant biology 12:129

Chagné D, Lin-Wang K, Espley RV, Volz RK, How NM, Rouse S, Brendolise C, Carlisle CM, Kumar S, De Silva N (2013) An ancient duplication of apple MYB transcription factors is responsible for novel red fruit-flesh phenotypes. Plant Physiol 161:225–239

Chagné D, Crowhurst RN, Pindo M, Thrimawithana A, Deng C, Ireland H, Fiers M, Dzierzon H, Cestaro A, Fontana P, Bianco L, Lu A, Storey R, Knäbel M, Saeed M, Montanari S, Kim YK, Nicolini D, Larger S, Stefani E, Allan AC, Bowen J, Harvey I, Johnston J, Malnoy M, Troggio M, Perchepied L, Sawyer G, Wiedow C, Won K, Viola R, Hellens RP, Brewer L, Bus VGM, Schaffer RJ, Gardiner SE, Velasco R (2014) The draft genome sequence of European Pear (Pyrus communis L. 'Bartlett'). PLoS ONE 9, e92644

De Franceschi P, Stegmeir T, Cabrera A, Van Der Knaap E, Rosyara U, Sebolt A, Dondini L, Dirlewanger E, Quero-Garcia J, Campoy J (2013) Cell number regulator genes in Prunus provide candidate genes for the control of fruit size in sweet and sour cherry. Mol Breed 32:311–326

Devoghalaere F, Doucen T, Guitton B, Keeling J, Payne W, Ling TJ, Ross JJ, Hallett IC, Gunaseelan K, Dayatilake G (2012) A genomics approach to understanding the role of auxin in apple (Malus x domestica) fruit size control. BMC Plant Biol 12:7

Espley RV, Hellens RP, Putterill J, Stevenson DE, Kutty-Amma S, Allan AC (2007) Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. Plant J 49:414–427

Hummer KE, Janick J (2009) Rosaceae: taxonomy, economic importance, genomics. In: Genetics and genomics of Rosaceae. Springer, pp 1–17

Illa E, Sargent DJ, Girona EL, Bushakra J, Cestaro A, Crowhurst R, Pindo M, Cabrera A, van der Knaap E, Iezzoni A, Gardiner SE, Velasco R, Arús P, Chagné D, Troggio M (2011) Comparative analysis of rosaceous genomes and the reconstruction of a putative ancestral genome for the family. BMC Evol Biol 11:9

Immanen J, Nieminen K, Duchens Silva H, Rodríguez Rojas F, Meisel LA, Silva H, Albert VA, Hvidsten TR, Helariutta Y (2013) Characterization of cytokinin signaling and homeostasis gene families in two hardwood tree species: Populus trichocarpa and Prunus persica. BMC genomics 14:885

Kumar R, Tyagi AK, Sharma AK (2011) Genome-wide analysis of auxin response factor (ARF) gene family from tomato and analysis of their role in flower and fruit development. Mol Genet Genomics 285:245–260

Magrane M, Consortium U (2011) UniProt Knowledgebase: a hub of integrated protein data. Database 2011:bar009

Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nat Genet 38:948–952

McCouch SR (2008) Gene nomenclature system for rice. Rice 1:72–84

Meinke D, Koornneef M (1997) Community standards for Arabidopsis genetics. Plant J 12:247–253

Olmedo G, Guo H, Gregory BD, Nourizadeh SD, Aguilar-Henonin L, Li H, An F, Guzman P, Ecker JR (2006) ETHYLENE-INSENSITIVE5 encodes a 5′ → 3′ exoribonuclease required for regulation of the EIN3-targeting F-box proteins EBF1/2. Proc Natl Acad Sci 103:13286–13293

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P, Davis TM, Slovin JP, Bassil N, Hellens RP, Evans C, Harkins T, Kodira C, Desany B, Crasta OR, Jensen RV, Allan AC, Michael TP, Setubal JC, Celton J-M, Rees DJG, Williams KP, Holt SH, Rojas JJR, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin SA, Troggio M, Viola R, Ashman T-L, Wang H, Dharmawardhana P, Elser J, Raja R, Priest HD, Bryant DW, Fox SE, Givan SA, Wilhelm LJ, Naithani S, Christoffels A, Salama DY, Carter J, Girona EL, Zdepski A, Wang W, Kerstetter RA, Schwab W, Korban SS, Davik J, Monfort A, Denoyes-Rothan B, Arus P, Mittler R, Flinn B, Aharoni A, Bennetzen JL, Salzberg SL, Dickerman AW, Velasco R, Borodovsky M, Veilleux RE, Folta KM (2011) The genome of woodland strawberry (Fragaria vesca). Nat Genet 43:109–116

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagne D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouze P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel C-E, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus x domestica* Borkh.). Nat Genet 42:833–839

Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Marroni F, Zhebentyayeva T, Dettori MT, Grimwood J, Cattonaro F, Zuccolo A, Rossini L, Jenkins J, Vendramin E, Meisel LA, Decroocq V, Sosinski B, Prochnik S, Mitros T, Policriti A, Cipriani G, Dondini L, Ficklin S, Goodstein DM, Xuan P, Fabbro CD, Aramini V, Copetti D, Gonzalez S, Horner DS, Falchi R, Lucas S, Mica E, Maldonado J, Lazzari B, Bielenberg D, Pirona R, Miculan M, Barakat A, Testolin R, Stella A, Tartarini S, Tonutti P, Arus P, Orellana A, Wells C, Main D, Vizzotto G, Silva H, Salamini F, Schmutz J, Morgante M, Rokhsar DS (2013) The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nat Genet 45: 487–494

Wu J, Wang F, Cheng L, Kong F, Peng Z, Liu S, Yu X, Lu G (2011) Identification, isolation and expression analysis of auxin response factor (*ARF*) genes in Solanum lycopersicum. Plant Cell Rep 30: 2059–2073

Wu J, Wang Z, Shi Z, Zhang S, Ming R, Zhu S, Khan MA, Tao S, Korban SS, Wang H (2013) The genome of the pear (*Pyrus bretschneideri* Rehd.). Genome Res 23:396–408

Zhang Q, Chen W, Sun L, Zhao F, Huang B, Yang W, Tao Y, Wang J, Yuan Z, Fan G, Xing Z, Han C, Pan H, Zhong X, Shi W, Liang X, Du D, Sun F, Xu Z, Hao R, Lv T, Lv Y, Zheng Z, Sun M, Luo L, Cai M, Gao Y, Wang J, Yin Y, Xu X, Cheng T, Wang J (2012) The genome of *Prunus mume*. Nat Commun 3:1318