# Transcriptomics: Open Access

# Genome Sequence Information in Fruit Crops: Current Status

**Nimisha Sharma[1]\*, Sanjay Kumar Singh[1], Lal S[2] and Nagendra Kumar Singh[3]**

[1]Division of Fruits and Horticultural Technology, ICAR-Indian Agricultural Research Institute, New Delhi, India
[2]Central Institute of Temperate Horticulture, Srinagar, India
[3]National Research Centre on Plant Biotechnology, New Delhi, India

## Abstract

Recent technological advancement in next-generation sequencing (NGS), plants can be prioritized for sequencing in relation to their value to humans. NGS provides the possibility of cost-efficient whole genome sequencing. Notable developments pertaining to genome sequencing of fruit crops are highlighted. These sequence information has been used extensively for analyzing and understanding structural and functional genomics for various biotic and abiotic stresses.

**Keywords:** Genome sequence; Fruit crops; Next generation sequencing; Transcriptomic reads

## Short Communication

Fruit Science (Pomology) is one of the important sectors of agriculture. With the growing population, demand for fruits is gradually increasing. The most important challenge in many of the fruit species is the unavailability of well-defined molecular genetic linkage maps. Many times these species are difficult to study at genetic and molecular levels because of their perennial nature and due to this, development of mapping population and map based studies are not easy. Therefore, genome sequence is a prerequisite resource for understanding completely the roles of genes in development, driving genomics-based approaches to systems biology and efficiently exploiting the natural and induced genetic diversity of a species. Initially, only plants with relatively small genomes were selected for sequencing. A decade ago, technological limitations forced the plant biology community to select a few species as models. As a consequence of continued improvements in sequencing technologies, methods and bioinformatics capabilities, sequencing goals need no longer be limited. With the development of next-generation sequencing (NGS) technologies, this paradigm can change and plants can be prioritized for sequencing in relation to their value to humans. NGS provides the possibility of cost-efficient transcriptome profiling. First plant genome sequence, Arabidopsis (*Arabidopsis thaliana*), in 2000 spawned an expansion in genomics-based research and the exploitation of annotated genes to explore orthologous genes in other plants. It also paved the way for sequencing several other model plant genomes and a few crop genomes. Sequencing of major grain crops like rice, wheat, barley and maize were more focused. Being complex, the genomes of fruit crops would be difficult to sequence as these species *de novo* using NGS technologies. Latest genomic technologies can be effectively used in fruit crop improvement programme. Availability of NGS technologies like FLX-454, Illumina, SOLiD and Helicose have brought hope to generate genomic resources for many more fruit species in a recent years' time [1-4]. NGS, although much less costly in time and money in comparison to first-generation sequencing. However, there are still many limitations. It can cost more than $100,000 in start-up and individual sequencing reactions can cost upward of $1,000 per genome. Moreover, sequencing error in homopolymer regions on certain NGS platforms, including the Ion Torrent PGM, and short-sequencing read lengths (on average 200-500 nucleotides) can lead to inaccurate sequencing. Data interpretation is time consuming and required the person highly trained in the field of Bioinformatics. Comparative details of different NGS technologies are presented in Table 1 [3,4]. Despite these limitations, it is suggested that, fruit breeders should use a multi-disciplinary approach involving people from Molecular Biology and Bioinformatics field to make use of this extensive genome information on their varietal development

programs. Genomic research has great potential to revolutionize the molecular biology research in horticultural crops in many ways. Among fruit crops, first draft with 8× high quality grapevine sequence was released by the international grape genome project (IGGP) [5]. The second fruit crop transgenic 'SunUp' papaya was sequenced by the Hawaii Papaya Genome Project [6]. Sequenced genomes of few fruits are presented in Table 2. Status and availability of genomic resources in fruit crops can be utilized for the efficient exploitation of the current research in developing improved genotypes and also defining future goals. Recent advances in automation and high throughput techniques used in decoding plant genomes play an important role to speed up the genomic research. With the establishment of genome and transcriptome sequencing projects for several horticultural crops, the huge wealth of sequence information have been generated [7]. These sequence information has been used extensively for analyzing and understanding genome structures and complexities, comparative and functional genomics and to mine useful genes and molecular markers. Present article describes the sequenced fruit genome information in brief.

### Grapevine (*Vitis vinifera* L.)

The cultivated grape species *Vitis vinifera* has potential to become a model for fruit trees genetics. Sanger shotgun sequencing and highly efficient sequencing by synthesis (SBS), together with dedicated assembly programs, resolved a complex heterozygous genome. Cultivated *V. vinifera* is highly heterozygous. The estimated genome size of *V. vinifera* to be 504.6 Mb. Genomic sequences corresponding to 477.1 Mb were assembled in 2,093 metacontigs and 435.1 Mb were anchored to the 19 linkage groups (LGs). The number of predicted genes is 29,585, of which 96.1% were assigned to LGs [5]. Nearly 2,000,000 SNPs, 1,751,176 were mapped to chromosomes and one or more of them were identified in 86.7% of anchored genes. As a result, many of the contigs were consensus sequences derived from an alignment of the two haplotypes. The set of Pinot Noir chromosome pairs included a considerable number of haplotype specific gaps. Based on resistance domain analyses, the grape genome was found to contain

**\*Corresponding author:** Nimisha Sharma, Division of Fruits and Horticultural Technology, ICAR-Indian Agricultural Research Institute, New Delhi, India, Tel: +911125843214; E-mail: nims17sharma@gmail.com

| Platform | Roche/454 | Illumina | Ion Torrent | ABI SOLiD | Helicos BioSciences | Pacific BioScience |
|---|---|---|---|---|---|---|
| Principle | Sequence by synthesis | Sequence by synthesis | Sequence by synthesis | Sequence by ligation | Single - molecule sequencing | Single - molecule sequencing |
| Sequencing method | Pyro sequencing | Reversible dye terminators | Natural nucleotides | Cleavable probe SBL | Reversible terminators | Real-time |
| Read length | 400 bases | 150 bases | 200 bases | 50 bases | 35 bases | 1500 bases |
| Run time | 10 hours | 10 days | 2 hours | 11-12 days | 30 days | 2 hours |
| Total bases per run | 500 Mb | 20 Gb | 1Gb on 318 chip | 100 Gb | 35 Gb | 100 Mb |
| Error rate | 0.1% | 1.5% | 1.71% | 4% | 2-7% | 12% |
| Sequencing cost | $12.56 × 10-6 per base | $148 per Gb | $1000 (318 chip) per Gb | $40 × 10-9 per base | $0.45 to $0.60 per Megabase | $2000 per Gb |

**Table 1:** Comparison of Different Next Generation Sequencing Technologies.

| Fruit Crop | Genome Size | Number of genes | Sequencing methodology | Reference |
|---|---|---|---|---|
| Pear *Pyrus bretscheideri* Rehd | 512.0-Mb | 42,812 | BAC-by-BAC and next-generation sequencing | [10] |
| European Pear (*Pyrus communis* L. 'Bartlett') | 577.3 Mb | 43,419 | Roche 454 sequencing | [16] |
| Grapevine *Vitis vinifera* | 504.96 (Mb) | 29,585 | Sanger shotgun sequencing and highly efficient sequencing by synthesis | [5] |
| Papaya *Carica papaya* | 372Mbp | 28,629 | whole-genome shotgun | [6] |
| Sweet orange *Citrus sinensis* | 320.5 Mb | 29,445 | whole-genome shotgun paired-end–tag sequence reads, Illumina GAII sequencer | [11] |
| Datepalm *Phoenix dactylifera* L | 605.4 Mb | 41,660 | Shotgun library Roche/454 data, SOLiD 4 System, BAC-end sequencing | [14] |
| Apple (*Malus × domestica*). | 742.3 Mbp | 57,386 | whole-genome shotgun approach | [8] |
| Peach *Prunus persia* | 265 Mbp | 27,852 | Sanger whole-genome shotgun | [15] |
| Wild Strawberry *Fragaria vesca* | 240Mbp | 34,809 | Roche/454, Illumina/Solexa and Life Technologies/SOLiD platforms | [9] |
| Mango | 450 Mbp | | PAC Biosequencing | [17] |

**Table 2:** Genome sequence information in fruit crops.

341 NBS genes. Besides NBS genes, the grape genome contains several signalling components of plant disease response which are encoded by genes EDS1, PAD4, COI1, MPK4, JAR1, ETR1 and NDR1, known to be recruited by resistance gene products. In addition, the grape genome contains eight genes similar to the MLO gene for mildew resistance in barley, compared to the 15 MLO-like genes known for Arabidopsis. In grape, the disease-related genes represent a significant part of the genome. Gene predictions corresponding to all those genes known to encode enzymes of the pathway could now be found. These include C4H and 4CL which were not previously identified in grape. The majority of genes were organized in large (PAL, F3'5'H) or small (CHS, F3H, FLS, LAR) gene families, the remainder consisting of single copy genes (C4H, 4CL, CHI, F3'H, DFR, LDOX, ANR, UFGT). Description of the grape genome sequence opens the opportunity for molecular breeding in grape. The fertility of hybrids between wild and domesticated grape species with 19 seemingly co-linear chromosomes makes it feasible to introduce new resistance genes via traditional breeding.

**Papaya (*Carica papaya*. L)**

Papaya is an exceptionally promising system for the exploration of tropical-tree genomes and fruit-tree genomics. It has a relatively small genome of 372 megabases [6]. A total of 2.8 million whole-genome shotgun (WGS) sequencing reads were generated from a female plant of transgenic cultivar SunUp, which was developed through transformation of Sunset. Papaya offers numerous advantages as a system for fruit-tree functional genomics, and this draft genome sequence provides the foundation for revealing the basis of Carica's distinguishing morpho-physiological, medicinal and nutritional properties. The sequencing of the genome of SunUp papaya makes it the best-characterized commercial transgenic crop. Since, papaya ringspot virus is widespread in nearly all papaya-growing regions, SunUp could serve as a transgenic germplasm source that could be used to evolve suitable genotypes resistant to virus in various parts of the world. Considering the assembled genome covers 92.1% of the unigenes and 92.4% of the mapped genetic markers, the number of predicted genes in the papaya genome could be 7.9% higher, or 24,746, about 11-20% less than Arabidopsis. Comparative analysis of the papaya and Arabidopsis 5′ untranslated regions showed that only 14% of orthologous promoter pairs exhibited significantly higher levels of sequence identity than random comparisons. Global analysis of all inferred protein models from papaya, Arabidopsis, poplar, grape and rice clusters the 208,901 non-redundant protein sequences into 39,706 similarity groups, or 'tribes'19, 11,851 which contained two or more genes. Sex determination in papaya is controlled by a pair of primitive sex chromosomes, with a small male-specific region of the Y chromosome (MSY)8. Papaya and Arabidopsis, respectively, have similar numbers of genes involved in ethylene synthesis, with four each for S-adenosyl methionine synthase (SAM synthase); 8 and 13 for aminocyclopropane carboxylic acid (ACC) synthase (ACS); 8 and 12 for ACC oxidase (ACO); and 42 and 64 for ethylene-responsive binding factors (AP2/ERF).

**Apple (*Malus x domestica* Borkh.)**

The domesticated apple (Malus × domestica Borkh., family Rosaceae, tribe Pyreae) is the main fruit crop of temperate regions of the world. Velasco et al. [8] showed that a relatively recent (>50

million years ago) genome-wide duplication (GWD) has resulted in the transition from nine ancestral chromosomes to 17 chromosomes in the Pyreae. Traces of older GWDs partly support the monophyly of the ancestral paleohexaploidy of eudicots. Phylogenetic reconstruction of Pyreae and the genus Malus, relative to major Rosaceae taxa, identified the progenitor of the cultivated apple as *M. sieversii*. Expansion of gene families reported to be involved in fruit development may explain formation of the pome, a Pyreae-specific false fruit that develops by proliferation of the basal part of the sepals, the receptacle. In apple, a subclade of MADS-box genes, normally involved in flower and fruit development, is expanded to include 15 members, as are other gene families involved in Rosaceae-specific metabolism, such as transport and assimilation of sorbitol. Data from a three-way sequence alignment between predicted gene space in apple (~84 Mb) and experimentally derived EST data from pear (~14.9 Mb) and peach (~18 Mb), performed, indicates that the genetic distance, based on DNA sequence divergence per base pair between members of Rosaceae, increases from apple to pear to peach. When predicted gene spaces of apple and pear were compared, a value of 96.35% nucleotide identity was calculated between these two species of the tribe Pyreae. When grape was compared with apple and pear, nucleotide identity was estimated at 85.31%. When the frequency of transitions and transversions was considered the ratio R (transitions/transversions) was similar for apple-specific and pear-specific mutations.

## Woodland strawberry (*Fragaria vesca*)

The woodland strawberry, *Fragaria vesca* (2n=2x=14), is a versatile experimental plant system. This diminutive herbaceous perennial has a small genome (240 Mb), is amenable to genetic transformation and shares substantial sequence identity with the cultivated strawberry (Fragaria × ananassa) and other economically important rosaceous plants. Shulaev et al. [9] reported the draft *F. vesca* genome, which was sequenced to × 39 coverage using second-generation technology, assembled de novo and then anchored to the genetic linkage map into seven pseudochromosomes. Gene prediction modeling identified 34,809 genes, with most being supported by transcriptome mapping. Genes critical to valuable horticultural traits including flavor, nutritional value and flowering time were identified. Genome-wide analyses provide insight into the nature and dynamics of macro-syntenic relationships among Rosaceous taxa. Comparison of the map positions of 389 rosaceous conserved ortholog set (RosCOS) markers previously been mapped in Prunus 10 to their positions on the seven pseudochromosomes of *F. vesca* H4 × 4 revealed macro-syntenic relationships between the two genomes. Markers were deemed orthologous between the two genomes when five or more RosCOS occurred within 'syntenic blocks' shared between the two genomes. This analysis revealed remarkable genome conservation between the two taxa, with complete synteny between Prunus linkage group (PG) 2 and Fragaria chromosome (FC) 7, PG8 and a section of FC2, and PG5 with a section of FC5. Annotation coverage in the strawberry genome is equivalent to that of Arabidopsis, which has a genome of similar size. Eight plant genomes were aligned, anchored by the genome of *F. vesca*. The other seven plants represent the most closely related available genomic sequences. Results shows that *Vitis vinifera* and *Populus trichocarpa* share the most genes with *F. vesca*.

## Pear (*Pyrus bretschneideri* Rehd.)

Pear, the third most important temperate fruit species after grape and apple, belongs to the subfamily Pomoideae in the family Rosaceae. The majority of cultivated pears are functional diploids (2n=34). A 512.0 Mb sequence corresponding to 97.1% of the estimated genome size of this highly heterozygous species is assembled with 194 × coverage. High-density genetic maps comprising 2005 SNP markers anchored 75.5% of the sequence to all 17 chromosomes. The pear genome encodes 42,812 protein-coding genes, and of these, ~28.5% encodes multiple isoforms. Repetitive sequences of 271.9 Mb in length, accounting for 53.1% of the pear genome, are identified [10].

## Sweet orange (*C. sinensis* cv. Valencia)

Oranges are an important nutritional source for human health and have immense economic value. Normal sweet oranges are diploids, with nine pairs of chromosomes and an estimated genome size of ~367 Mb. The assembled sequence covers 87.3% of the estimated orange genome, which is relatively compact, as 20% is composed of repetitive elements. Number of predicted Protein-coding genes are 29,445. Sequencing evidence also suggested that sweet orange originated from a backcross hybrid between pummelo and mandarin. Focused analysis on genes involved in vitamin C metabolism showed that GalUR, encoding the rate-limiting enzyme of the galacturonate pathway, is significantly upregulated in orange fruit, and the recent expansion of this gene family may provide a genomic basis. Detailed comparative analyses of the orange genome against 21 other plant genomes and screened them by EST and peptide sequences in the public database. Of the 29,445 protein-coding genes in the orange genome, 23,804 were grouped into 14,348 gene families, with 1,691 genes being specific to citrus. Most (96%) of the citrus-specific genes are 'hypothetical' and have unknown function. Of the 58 annotated genes, the over-represented protein domains are the zinc finger transcription factor domain and the nucleotide-binding-site and leucine-rich-repeat (NBS-LRR) domain, which contain potential disease-resistance genes. From the diploid genome sequencing data, were identified. 1.06 million SNPs and 176,953 small insertions/deletions (indels). More than 80% of the SNPs and indels were anchored on the nine pseudochromosomes. The overall polymorphism density was 3.6 SNPs and 0.6 indels per kb in the genome. The RNA-Seq data derived from callus, leaf, flower and fruit and identified 2,697 (9.2%) genes that were significantly upregulated (P<0.01) in fruits. Gene Ontology analysis of upregulated genes in citrus fruits revealed that genes for oxidoreductase, oxidation reduction and vitamin binding were over-represented. This draft genome represents a valuable resource for understanding and improving many important citrus traits in the future [11].

## Date palm (*Phoenix dactylifera* L.)

Date palm (*Phoenix dactylifera* L.) is a cultivated woody monocotyledons plant species with agricultural and economic importance. There have been a very limited number of genome-wide studies on *P. dactylifera*. One is a recent report on a draft genome assembly based on data generated from the Illumina GAII sequencing platform by a research team in Qatar [12]. They estimated the genome size (658 Mb), assembled 58% of the genome (382 Mb) and predicted 25,059 genes. Another is a comparative transcriptomic study on mesocarps of both oil and date palms based on pyrosequencing data from the Roche GS FLX Titanium platform [13]. Ibrahim et al. [14] reported a genome assembly for an elite variety (Khalas), which is 605.4 Mb in size and covers 490% of the genome (671 Mb) and 496% of its genes (41,660 genes). Large-scale genomic and transcriptomic data pave the way for further genomic studies not only on *P. dactylifera* but also other Arecaceae plants.

## Peach (*Prunus persica* (L.) Batsch)

Peach (*Prunus persica* (L.) Batsch), which has been bred and cultivated for more than 4,000 years, is a highly genetically characterized

tree species whose genome is important for both fruit and forest tree research. In peach, genetic and genomic efforts have identified gene-containing intervals controlling a large number of important fruit traits. Verde et al. [15] described the high-quality genome sequence of peach obtained from a completely homozygous genotype, obtained a complete chromosome-scale assembly using Sanger whole-genome shotgun methods. They reported the high-quality whole-genome shotgun assembly of a double haploid genotype of the peach cv. Lovell (2n=2x=16) with an estimated genome size of 265 Mb. A total of 27,852 protein-coding genes and 28,689 protein-coding transcripts were predicted; of these, 24,423 have Arabidopsis homologs, 18,822 have Swiss-Prot homologs and 26,731 have TrEMBL homologs. Comparative and phylogenetic analyses carried out on the manually annotated gene families among peach and other sequenced species enabled the identification of members with specific roles in peach metabolic processes (for example, sorbitol metabolism and/or transport and aroma volatile compounds metabolism) and stressed common features with other Rosaceae species. In Rosaceae, polyol biosynthesis has a more prominent role than what is seen in other plant families. Comparison study between the grape and peach genomes using the Mercator program 40, identified segments with one-to-one orthology relationships across species rather than DNA regions having multiple syntenic partners. Notably, each grape segment, corresponding to part of one of the paralogous triplets of putative ancestor paleochromosomes, showed orthology to a single peach chromosome. This suggests that the homeologous subgenomes of grape and peach derive from the same paleohexaploid event that occurred before the emergence of Vitaceae and Rosaceae.

### European pear (*Pyrus communis* L. 'Bartlett')

Pear (Pyrus) is one of the oldest temperate tree fruit crops, having been grown since antiquity from both Europe to China. Chagné et al. [16] presented a draft assembly of the genome of European pear (Pyrus communis) 'Bartlett'. The assembly was developed employing second generation sequencing technology (Roche 454), from single-end, 2 kb, and 7 kb insert paired-end reads using Newbler (version 2.7). It contains 142,083 scaffolds greater than 499 bases (maximum scaffold length of 1.2 Mb) and covers a total of 577.3 Mb, representing most of the expected 600 Mb Pyrus genome. Gene prediction using a combined ab initio prediction and homology searching approach yielded 43,419 putative gene models. The number of predicted genes is higher than for most plant species and, 30% greater than in the strawberry genome (34,809 gene models), as might be expected due to the Pyreae whole genome duplication. A total of 5,350 protein clusters was observed as conserved across all 13 species proteomes, with 14,348 predicted European pear proteins (33% of the 43,419 total predicted protein set). Only 82 protein clusters were not found in European pear compared with all other 12 species, a value less than the number of protein clusters absent from Chinese pear (298), apple (236), strawberry (192), Arabidopsis (246), potato (437), papaya (424), grape (502) and kiwifruit (558), however similar to that of sweet orange (85), clementine (34), tomato (53) and poplar (45). A total of 829,823 putative single nucleotide polymorphisms (SNPs) were detected. A total of 2,279 genetically mapped SNP markers anchor 171 Mb of the assembled genome. Analysis of the expansin gene family provided an example of the quality of the gene prediction and an insight into the relationships among one class of cell wall related genes that control fruit softening in both European pear and apple (Malus x domestica).

### Mango (*Mangifera indica* L.)

Mango (*Mangifera indica* L.) is called "King of fruits" in India due to its sweetness, richness of taste, huge variability, large production volume and variety of end usage. It is a member of the family Anacardiaceae and is an allotetraploid (2n=40) fruit tree with small genome size of about 450 Mbp. Recently, first genomic draft assembly, a total of 412,728 (~432 Mbp) contigs were generated covering more than 95% of the genome. Total of 63,130 genes were predicted with an average gene length of 711 bp. Further, about 1,96,362 SSRs has been assembled in the draft [17].

**References**

1. Barabaschi D, Guerra D, Lacrima K, Laino P, Michelotti V, et al. (2012) Emerging knowledge from genome sequencing of crop species. Molecular Biotechnology 50: 250-266.

2. Egan AN, Schlueter J, Spooner DM (2012) Applications of next-generation sequencing in plant biology. American Journal of Botany 99: 175-185.

3. Grada A, Weinbrecht K (2013) Next-Generation Sequencing: Methodology and Application. Journal of Investigative Dermatology 133.

4. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, et al. (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics 13:341.

5. Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, et al. (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS One 2: e1326.

6. Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, et al. (2008) The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus). Nature 452: 991-996.

7. Azim KM, Khan IA, Zhang Y (2014) Characterization of mango (Mangifera indica L.) transcriptome and chloroplast genome. Plant Molecular Biology 85: 193-208.

8. Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, et al. (2010) The genome of the domesticated apple (Malus x domestica Borkh.). Nature Genetics 4: 833-839.

9. Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, et al. (2011) The genome of woodland strawberry (Fragaria vesca). Nature Genetics 43: 109-116.

10. Wu J, Wang Z, Shi Z, Zhang S, Ming R, et al. (2012) The genome of pear (Pyrus bretschneideri Rehd). Genome Research 23: 396-408.

11. Xu Q, Chen LL, Ruan X, Chen D, Zhu A, et al. (2012) The draft genome of sweet orange (Citrus sinensis). Nature Genetics 45: 59-66.

12. Al-Dous EK, George B, Al-Mahmoud ME, Al-Jaber MY, Wang H, et al. (2011) De novo genome sequencing and comparative genomics of date palm (Phoenix dactylifera). Nature Biotechnology 29: 521-527.

13. Bourgis F, Aruna Kilaru, Xia Cao, Georges FNE, Noureddine D, et al. (2011) Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. Proc. Natl Acad. Sci. USA 108: 12527-12532.

14. Ibrahim S, Al-Mssallem Songnian H, Xiaowei Z, Qiang L, Wanfei L, et al. (2013) Genome sequence of the date palm Phoenix dactylifera L. Nature Communication 4: 2274.

15. Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, et al. (2013) The high-quality draft genome of peach (Prunus persica) identifies unique patterns of genetic diversity, domestication and genome evolution. Nature Genetics 45: 487-494.

16. Chagné D, Crowhurst RN, Pindo M, Thrimawithana A, Deng C, et al. (2014) The Draft Genome Sequence of European Pear (Pyrus communis L. 'Bartlett'). PLoS One 9: 92644.

17. Singh NK, Mahato AK, Sharma N, Gaikwad K, Srivastava M, et al. (2014) A draft genome of the king of fruit, mango (Mangifera indica L.). Plant and Animal Genome XXII Conference, San Diego, CA, USA.