MINI-REVIEW ARTICLE

Genomic and Bioinformatic Resources for Perennial Fruit Species

Jérôme Grimplet[1,2,*]

[1]*Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Unidad de Hortofruticultura, Gobierno de Aragón, Avda. Montañana, Zaragoza, Spain;* [2]*Instituto Agroalimentario de Aragón–IA2 (CITA-Universidad de Zaragoza), Calle Miguel Servet, Zaragoza, Spain*

Abstract: In the post-genomic era, data management and development of bioinformatic tools are critical for the adequate exploitation of genomics data. In this review, we address the actual situation for the subset of crops represented by the perennial fruit species. The agronomical singularity of these species compared to plant and crop model species provides significant challenges on the implementation of good practices generally not addressed in other species. Studies are usually performed over several years in non-controlled environments, usage of rootstock is common, and breeders heavily rely on vegetative propagation. A reference genome is now available for all the major species as well as many members of the economically important genera for breeding purposes. Development of pangenome for these species is beginning to gain momentum which will require a substantial effort in term of bioinformatic tool development. The available tools for genome annotation and functional analysis will also be presented.

1. INTRODUCTION

Agriculture research on perennial species faces common challenges in improving breeding strategies and cultural practices. To resolve them, genomic strategies have been developed. They are time- and funding-consuming, therefore, there is interest in comparing/extrapolating results and implementing common strategies and bioinformatics tools. Unlike annual plants, perennials have had to develop mechanisms to adapt their vegetative growth to environmental conditions during a reproductive cycle that will occur several times over their lifespan, such as bud dormancy to overwinter or vernalization at the arrival of spring [1, 2]. They must show some level of adaptability to the evolution of their environment in the longer term, such as temperature variation or water availability over various years. In addition, these species have some peculiarities, compared to annual plants, like presenting a high level of heterozygosity, a long juvenile phase, and a vegetative propagation. Therefore, these characteristics greatly impact the way of doing research, mainly due to their long life cycle, as well as the fact that the studies are commonly carried out in the field, that usually is a non-controlled environment. In the era of data science, these particularities have important consequences on data management. Samples should be described

to meet the FAIR principles [3] and also taking into consideration specific metadata such as the correct description of the genotype, environmental conditions, plant stage, genotype, *etc*. [4].

Research communities must generate and integrate heterogeneous datasets (genomics, epigenomics, transcriptomics, proteomics, metabolomics and botanical ampelography) describing genotypes, phenotypes and the environment. Integrating these multi-omics datasets will reduce the gap between data generation and the ability to analyze and understand the biological mechanisms underlying plant responses to the environment. It could therefore reduce the breeding cycles to produce new cultivars. This is achieved by applying high-throughput experimental techniques that generate large datasets (omics technologies and multiplex sampling/imaging). The resulting data of the actual scientific production are dispersed and difficult to access, hindering exploitation beyond their initial purpose. International gene data repositories do not store functional data (*e.g.*, regulatory and metabolic networks), detailed plant materials or non-molecular phenotypes. Such data for non-model species might be stored in regional or local databases but are often not easily accessible to the wider research community.

Research communities involved in fruit tree species and assimilated perennial plants (vines, shrubs or some herbaceous species such as pineapple or strawberry) are generally small, implementing resources and assuring their long-term sustainability and update can be complicated. It is advantageous to mutualize these resources because these species share many similarities, both in terms of biology and re-

*Address correspondence to this author at the Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Unidad de Hortofruticultura, Gobierno de Aragón, Avda. Montañana, Zaragoza, Spain; Instituto Agroalimentario de Aragón–IA2 (CITA-Universidad de Zaragoza), Calle Miguel Servet, Zaragoza, Spain; Tel: +34976713635;
E-mail: jgrimplet@cita-aragon.es

search priorities. In all of these species, the partthat is consumed is the fruit. Whether a species produce a fleshy fruit or a dry fruit, the primary focus of the breeding programs is its improvement. The following fruit tree characteristics can be improved in any fruit: fruit quality, fruit maturity time, disease resistance, abiotic stress tolerance, productivity, and the sustainability of cultural practices. This improvement can be carried out using the strategy that is best for each of them.

The second common feature among these species is that they are perennial plants, so the fruit can be produced only when they are old. This significantly impacts breeding and research strategies, necessitating long term planning. The majority of them are ligneous, although some, like the banana, which is a perennial herbaceous plant sometimes mistaken for a tree [5], are not actually trees. Many (and potentially all) of these crops are grown on rootstocks, which confer many advantages since the root part can be bred for vigor-conferring or root disease resistance while the aerial part can be bred for fruit properties, aerial part disease resistance or tree architecture [6]. Again, in terms of data management,

they are the only cultivated organisms composed of two different genomes, each targeted at improving complementary but unrelated traits.

2. SCOPE OF SCIENTIFIC DATA PRODUCTION IN PERENNIAL FRUIT SPECIES

The initial step is to establish the resources available for each perennial fruit species. The number of articles published for a given species was identified in the Pubmed database. It is a good estimation of the scientific production of the research community working on this species and, indirectly, the size of the available resources and the needs of data analysis tools. Economic importance is not correlated to these data because other aspects such as the status of model species and biological advantages of studying a given species (short reproductive cycle, small genome) must also be taken into account. Looking at the research articles available in Pubmed, the major species in terms of publication is by far the grapevine with 18% of the data (Fig. **1**). Other major species include the three Rosaceae members: apple, peach and pear trees and the sweet orange tree, with each of
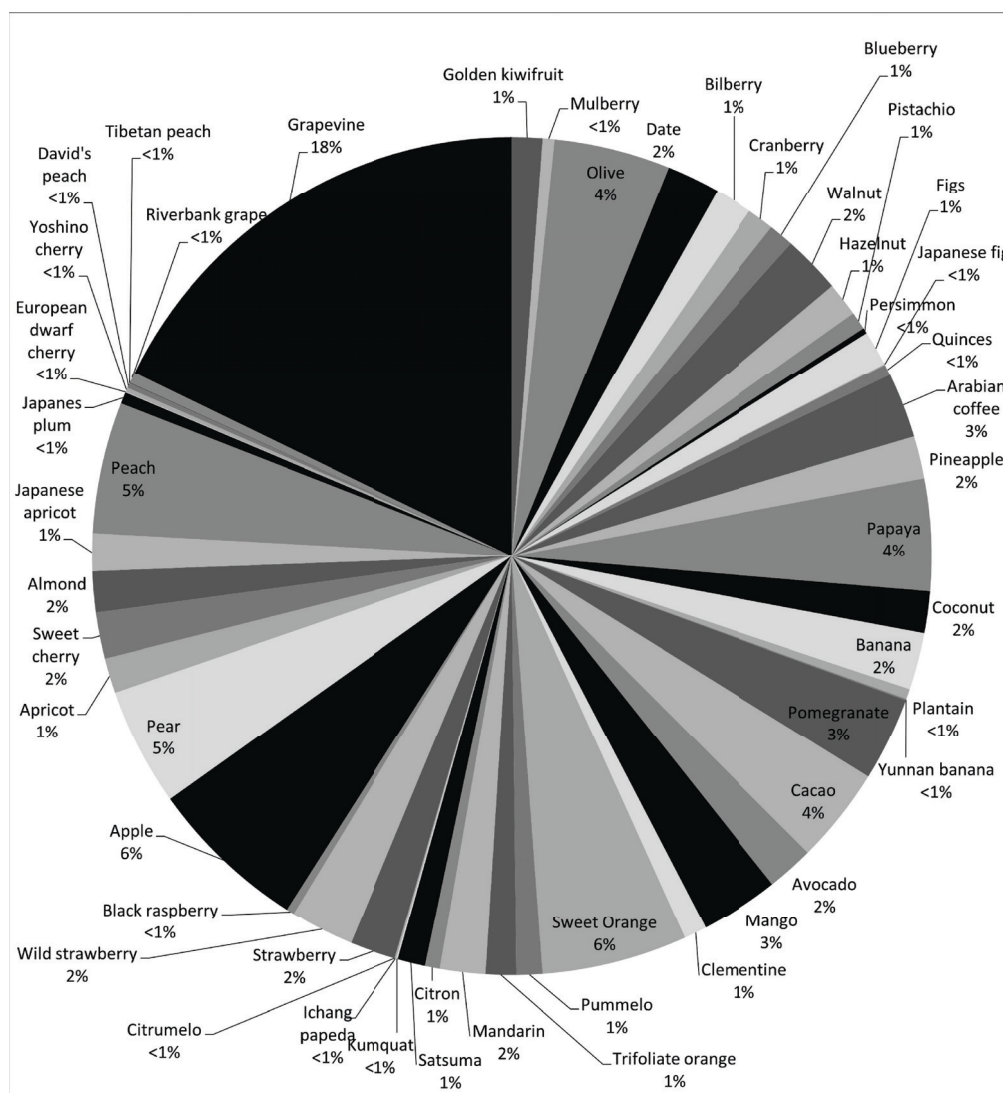

Fig. (1). Percentage of research articles per perennial fruit species. Data fetched August 2021. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

them accounting for 6% and 5% of the data. Another criterion is the number of available Biosamples for a given species. The BioSample captures the descriptive information or metadata attached to a biological material for which sequence data (from genome or transcriptome) [7] are stored in the INSDC databases [8]. They are also reciprocally linked to the BioProjects, the whole set of BioSamples participating in a single initiative. The number of BioSamples deposited in the INSDC database for each species represents the amount of sequencing data produced for a species and, therefore, the activities in terms of data release. These numbers increased more rapidly than the number of Pubmed articles (Supplementary Material **1**), but at the date the data were retrieved (August 2021), they were very similar for many perennial fruit species, hence, there are as many BioSamples as Pubmed papers. The number of BioProjects (Supplementary Material **2**) by species reveals that the species with the largest number of publications (grapevine, apple, peach, and pear) have proportionally more data per article than the others. This observation might be explained by the fact that there are more obstacles to perform genomic studies in minor species and to release the data in public databases with the proper metadata to link the samples to a particular species. These distinctions are made clear by the BioSample graph, which shows that the major species (grapevine, apple, and peach) have more BioSamples per project than the others, indicating that there are more opportunities to conduct larger sequencing studies in those species (large-scale genotypes sequencing or larger transcriptomic experiments).

Comparison of the production with model species gives another indicator of data structure and how perennial fruit research compares to other research communities. In Table **1**, all perennial fruit species have been grouped together. The total number of articles is relatively similar to those in model crop species (rice and maize) or humans and lower than in Arabidopsis, though falling within the same range. However, sequencing data production or deposition in the public database is less than half in perennial fruit species than in those other species.

These observations highlight lower dynamism in the production of data or difficulties to properly release the data in public databases. Similar conclusions are reached when comparisons are made within the perennial fruit species: more genomic data are proportionally produced in species with economical/biological interest compared to the size of its research community. However, on average, there are fewer BioSamples per BioProject in the fruits than in other plants. Individually, the research and data production for each species might remain relatively small but the sum of

the perennial fruit species-related research articles is only half of what has been performed, for instance, in Arabidopsis. These findings demonstrate the possibility of developing popular, long-lasting bioinformatic tools that have a similar impact to those developed for particular model species.

According to economic, physiological, and genetic characteristics as well as how bioinformatic technologies are used, five types of perennial fruit species can be distinguished. The tropical fruits are produced by trees or perennial herbaceous plants in tropical climate lands. They grow in wet and hot habitats with a relatively constant daylight time throughout the year. As a major feature, they do not have any phylogenetic relation, unlike the other groups (Fig. **2**). Three groups comprise species phylogenetically related (Fig. **2**), the genius Rosaceae and the genera *Vitis* and *Citrus*. A set of other species exist that cannot be classified in any of these groups (kiwi, olive tree, date, other berries and many nuts). The Rosaceae group exhibits the highest productivity overall, accounting for a third of the Pubmed papers and BioProjects, and half of the deposited data (BioSamples). The tropical fruits are the second biggest group in terms of Pubmed articles though there is more *Vitis* genomic data in the public databases.

3. REFERENCE GENOMES

Reference genomes are tools that have been developed since the beginning of genome sequencing. For perennial fruit species, they are generally built from representative individuals within a species; it can either be a commercially important cultivar or an individual that provides a significant advantage for performing the genome assembly (for example, a near homozygous grapevine individual was chosen by the consortium responsible for the sequencing project [9]).

Due to the widespread democratization of genome sequencing and assembly tools, as well as the development of pangenome and consensus genome building technologies, there has been discussion in recent years about the possible replacement of a single reference genome for those of the entire species [10]. They are still very useful, in particular for the species with a small or average size of research community. The reference genomes are used for several tasks: they provide an easy and cheap way to match sequences of a genome population with short reads and allow genome comparisons, polymorphism detection, such as GBS (Genotyping by Sequencing), or GWAS (Genome-Wide Association Study) analysis. Commercial genotyping arrays are also built from the reference genomes and molecular markers are also identified from them. They enable the specification of the species'

Table 1. Scientific and genomic production of perennial fruits compared to major model species.

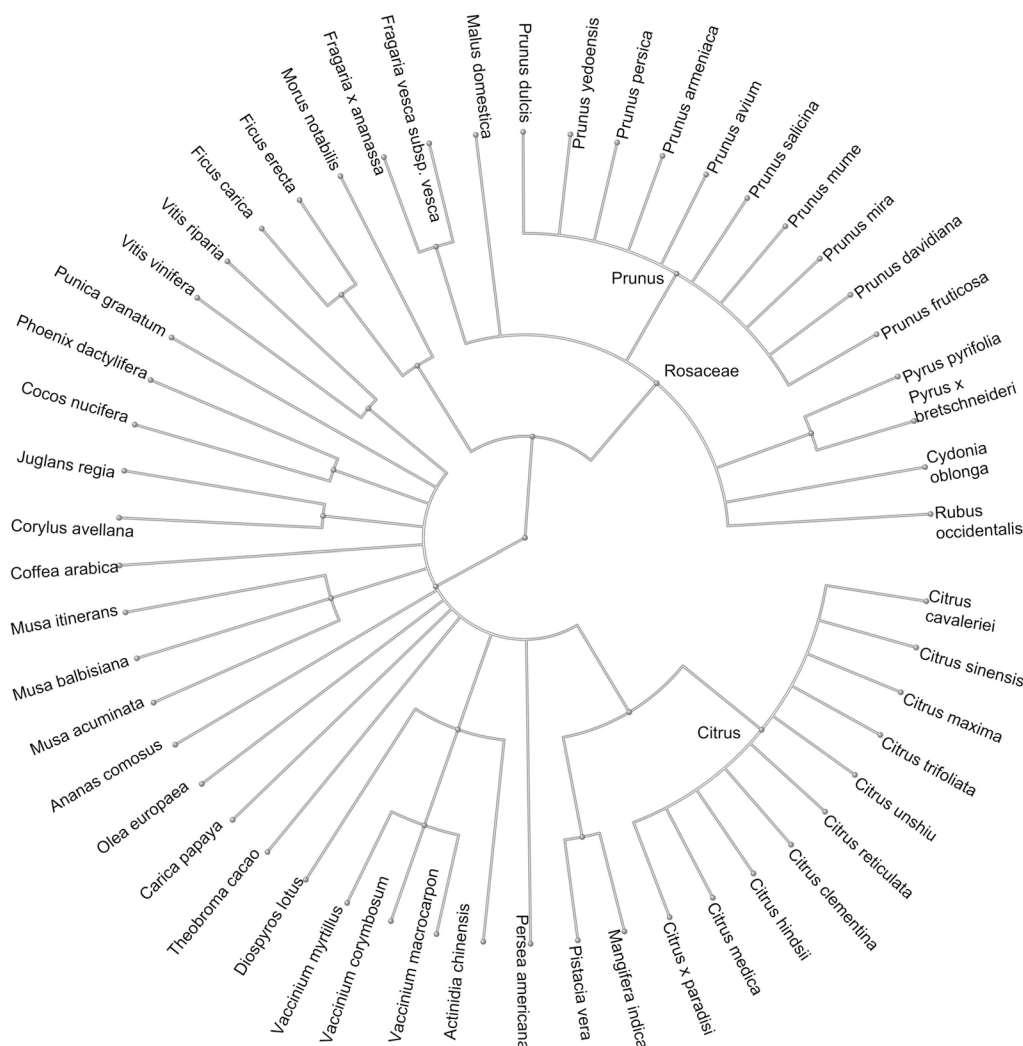| -	**Perenial Fruits**	**Arabidopsis**	***Homo sapiens***	**Rice**	**Maize**
PubMed Central	61,100	109,822	44,318	29,854	25,239
Bio Project	2,482	6,873	73,559	6,387	4,831
Bio Sample	47,251	125,842	6,318,974	96,672	75,649

Fig. (2). Phylogenetic tree of perennial fruit species. Tree was built using the NCBI taxa tool and NCBI treeview. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

gene annotation standards, and it is then also possible to manually curate the gene annotation. Quantitative data such as those from transcriptomic analyses are also generally mapped on these reference genomes.

The genomes of 53 perennial fruit species have been sequenced and are available in the genome section of the NCBI database (Table **2**). There are still a few fruit tree species to sequence amongst the fruits present in the FAO stats database (http://www.fao.org/faostat); the cashew (*Anacardium occidentale* L.) is the major species yet to be sequenced, the other ones are the areca (*Areca catechu* L.), all the currants (*Ribes* sp.), the kapok [*Ceiba pentandra* (L.) Gaertn], the nutmeg (*Myristica fragrans* Houtt.), the Brazil nut (*Bertholletia excelsa* Humb. & Bonpl.), and the carob (*Ceratonia siliqua* L.). However, the FAO stats database only gives a partial view of the potential perennial fruit species that could be sequenced since there are many other local species of insignificant economic importance but with breeding potential for relevant traits. For 22 species, a genome was integrated into the RefSeq genome database, which also implies that a gene annotation was performed by the NCBI team using the NCBI Eukaryotic Genome Anno-

tation Pipeline. For the remaining species, except for the raspberry, a representative genome assembly has been flagged by NCBI, but they have not been attributed a Refseq genome assembly ID and its corresponding annotation. There are three levels of quality for assembly, the chromosome, the scaffold, and the contig level. Only genomes assembled at the chromosome and scaffold level can be included in the RefSeq database. With the incorporation of newer and better-quality genomes, the representative genome for a species can be updated or changed based on quality data such as the scaffold N50 or a lower number of scaffolds. In the present state of the database, for some species, the current representative genome is assembled at the contig level while a recent assembly has been submitted to the database, which will modify the reference genome ultimately. While a genome assembled at the chromosome level was submitted in 2021 [11] and used here for comparison because the scaffolds number and N50 were available, the representative genome for apricot is assembled at the contig level (Table **2**). For seven species, only assembly at the contig level is available and there is no assembly at all for raspberry, for which only raw reads were deposited.

Table 2. Genomes available in perennial fruit species and sequencing data quality.

Species	Common Name	Total Ungapped Length	Number of Scaffolds	Scaffold N50	Scaffold L50	Number of Chromosomes	Date	Assembly Level	Genome Coverage	Reference
Non-group species										
Actinidia chinensis Planch.	Golden kiwifruit	6.54E+08	2366	1433402	130	29	14-11-19	Chr	63x	[13]
Morus notabilis C.K. Schneid.	Mulberry	3.04E+08	31301	405448	232	0	07-08-13	Sca	236x	[14]
Olea europaea L.	Olive	1.03E+09	41226	12567911	23	24	03-11-17	Chr	220x	[15]
Phoenix dactylifera L.	Date	7.73E+08	2391	4728343	19	20	04-11-19	Chr	100x	[16]
Vaccinium myrtillus L.	Bilberry	5.23E+08	1418	37643989	7	12	23-02-21	Chr	100x	[17]
Vaccinium macrocarpon Aiton	Cranberry	4.15E+08	200203	4291	28049	2	04-09-14	Sca	20x	[18]
Vaccinium corymbosum L.	Blueberry	2.86E+08	13757	145010	760	1	09-09-20	Sca	40x	[19]
Juglans regia L.	Walnut	5.72E+08	24757	37114715	7	17	08-07-20	Chr	155x	[20]
Corylus avellana L.	Hazelnut	3.68E+08	11	36653616	5	11	25-11-20	Chr	117x	[21]
Pistacia vera L.	Pistachio	6.7E+08	1865	948305	218	1	23-09-19	Sca	374x	[22]
Diospyros lotus L.	Persimmon	6.3E+08	42	42671757	7	15	21-09-20	Chr	167x	[23]
Ficus carica L.	Figs	3.33E+08	511	19846527	7	13	18-12-19	Chr	74x	[24]
Ficus erecta Thunb.	Japanese fig	5.96E+08	NA	NA	NA	NA	18-09-19	Con	62x	[25]
Cydonia oblonga Mill.	Quinces	4.88E+08	NA	NA	NA	NA	01-12-20	Con	108x	[26]
Tropical fruits										
Coffea arabica L.	Arabian coffee	1.09E+09	2833	42465768	11	23	08-11-18	Chr	160x	none
Ananas comosus L. Merr	Pineapple	3.75E+08	3129	11759267	13	26	14-03-16	Chr	400x	[27]
Carica papaya L.	Papaya	2.72E+08	17766	1089885	91	11	06-05-08	Sca	NA	[28]
Cocos nucifera L.	Coconut	2.1E+09	7998	570487	771	0	10-06-19	Sca	50x	[12]
Musa acuminata Colla	Banana	3.9E+08	7512	1311088	65	11	14-11-12	Chr	20x	[29]
Musa balbisiana Colla	Plantain	4.91E+08	2590	41415055	6	11	09-07-19	Chr	113x	[30]
Musa itinerans Chessman	Yunnan banana	4.15E+08	28415	195772	555	0	21-05-16	Sca	92x	[31]

Species	Common Name	Total Ungapped Length	Number of Scaffolds	Scaffold N50	Scaffold L50	Number of Chromosomes	Date	Assembly Level	Genome Coverage	Reference
Tropical fruits										
Punica granatum L.	Pomegranate	3.2E+08	474	39957110	4	9	22-10-19	Chr	60x	[32]
Theobroma cacao L.	Cacao	3.06E+08	431	36364294	5	11	09-07-16	Chr	16.5x	[33]
Persea americana Mill.	Avocado	7.49E+08	NA	NA	NA	NA	19-05-21	Con	48x	[34]
Mangifera indica L.	Mango	3.93E+08	252	17652500	10	22	10-03-20	chr	240x	[35]
Citrus										
Citrus clementina Hort. Ex. Tanaka	Clementine	2.95E+08	1398	31410901	4	0	08-11-13	Sca	7x	[36]
Citrus sinensis (L.) Osbeck	Sweet orange	3.01E+08	4995	1778813	51	10	12-12-12	chr	214x	[37]
Citrus maxima Merr.	Pummelo	3.45E+08	1612	4210623	27	9	07-03-17	Chr	426x	[38]
Citrus trifoliata L.	Trifoliate orange	3.03E+08	707	1174493	64	9	13-05-21	Chr	91x	[39]
Citrus reticulata Blanco	Mandarin	3.36E+08	67725	1288159	67	NA	20-06-18	Sca	200x	[40]
Citrus medica L.	Citron	3.7E+08	32732	369527	304	NA	24-01-18	Sca	171x	[38]
Citrus unshiu Yu. Tanaka ex Swingle	Satsuma	3.31E+08	20876	386404	203	NA	21-12-17	Sca	144x	[41]
Citrus hindsii (Champ. ex Benth.) Govaerts	Hong Kong kumquat	3.73E+08	NA	NA	NA	NA	22-04-19	Con	145x	[42]
Citrus cavaleriei H. Lév.	Ichang papeda	3.35E+08	14916	501435	193	NA	24-01-18	Sca	164x	[38]
Citrus x paradisi Macfad.	Citrumelo	2.66E+08	NA	NA	NA	NA	03-01-17	Con	24x	[43]
Rosaceae										
Fragaria x ananassa Duchesne	Strawberry	8.06E+08	204	27686486	13	28	24-06-21	Chr	77x	[44]
Fragaria vesca L.	Wild strawberry	2.02E+08	3048	27879571	4	8	24-02-11	Chr	49x	[45]
Rubus occidentalis L.	Black raspberry	2.91E+08	NA	NA	NA	7	19-01-18	raw read submitted		[46]
Malus domestica Borkh.	Apple	6.26E+08	807	37631755	9	18	28-04-17	Chr	700x	[47]
Pyrus x bretschneideri Rehder	Pear	4.98E+08	2182	535028	277	NA	03-12-12	Sca	8x	[48]

(Table 2) contd….

Species	Common Name	Total Ungapped Length	Number of Scaffolds	Scaffold N50	Scaffold L50	Number of Chromosomes	Date	Assembly Level	Genome Coverage	Reference
Rosaceae										
Prunus armeniaca L.	Apricot	2.49E+08	103	30876794	4	8	05-10-21	Chr	470x	[11]
Prunus avium L.	Sweet cherry	2.47E+08	10148	219566	316	NA	12-06-17	Chr	327x	[49]
Prunus dulcis (Mill.) D.A. Webb	Almond	2.24E+08	692	24375383	4	9	10-10-19	Chr	800x	[50]
Prunus mume (Siebold) Siebold & Zucc.	Japanese apricot	2.17E+08	8626	636887	106	9	28-02-14	Chr	180x	[51]
Prunus persica (L.) Batsch	Peach	2.25E+08	192	27368013	4	9	02-02-17	Chr	9x	[52]
Prunus salicina Lindl.	Japanese plum	2.84E+08	110	30743302	4	8	05-10-21	Chr	480x	[11]
Prunus fruticosa Pall.	European dwarf cherry	3.75E+08	57	43818497	4	8	03-06-21	Chr	97x	preprint
Prunus yedoensis Matsum.	Yoshino cherry	6.9E+08		NA	NA	NA	18-03-19	Con	54x	[53]
Prunus davidiana Carr.	David's peach	2.43E+08	127	28110464	4	8	05-10-21	Chr	390x	[11]
Prunus mira Koehne	Tibetan peach	2.39E+08	49	239877647	4	8	05-10-21	Chr	520x	[11]
Pyrus pyrifolia (Burm.) Nak.	Japanese pear	5.04E+08	NA	NA	NA	NA	23-12-20	Con	136x	[54]
Vitis										
Vitis riparia Michx.	Riverbank grape	4.94E+08	175	23918791	9	20	16-04-19	Chr	335x	[55]
Vitis vinifera L.	Grapevine	4.71E+08	2061	3426264	42	21	07-12-09	Chr	12X	[9]

NA: not available. Chr: chromosome. Sca: scaffold. Con: contig

The largest fruit tree genome is by far the one from the coconut tree, with 2.1 Gbp. As reported by some authors, its greater size could be explained by at least three rounds of duplications of the entire genomes [12]. The genomes of olive and coffee trees also exceed the size of 1 Gbp, but for the vast majority (39/53) of perennial fruit species, genomes are between 200 and 500 Mbp.

The size of the genome and number of predicted transcripts or coding sequences (CDS) do not show any correspondence when comparing a wide range of different organisms [56]. Here, when comparing perennial fruit species, CDS numbers are relatively similar between species as observed for the genome size (Table **3**). Out of the fifty-two species, forty-four have between 20,000 and 50,000 transcripts (it was not possible to retrieve the number of transcripts for Citrumelo). The density of CDS per Mbp is between 50 and 150 for almost all of them, except for the coconut tree, with a low density (16 transcripts per Mbp), and strawberry (180 transcripts per Mbp) and blueberry (246 transcripts per Mbp),

with high density. The average density is 99 ± 41 transcripts per Mbp. These numbers are generally relatively similar; the species with the largest size comprise the higher number of predicted transcripts, except for the coconut tree, which comprises a medium number of proteins. However, gene prediction methods are highly different between projects and only the predictions with the same pipeline can be compared, such as the RefSeq predictions [57].

These high similarities between genomes, genome compositions, and research priorities lines highlight the usefulness of sharing experiences, tools, and strategies among these species. Additionally, the research communities are relatively small, which makes it more difficult to leverage resources individually.

4. GENOME REPOSITORY AND DATABASES

Even if the data are deposited in public databases, the genomes are also available in community-specific databases, which allow the development of specific structural or

Table 3. Genome accession, references, and number of CDS of perennial fruit species.

Species	Infraspecific Name	BioProject	Assembly	Number of CDS
Actinidia chinensis	Hongyang	PRJNA549770	GCA_009663005.1	40,464*
Morus notabilis	Unknown	PRJNA202089	GCF_000414095.1	27,648
Olea europaea	Var. silvestris	PRJNA350614	GCF_002742605.1	58,334
Phoenix dactylifera	Barhee BC4	PRJNA322046	GCF_009389715.1	48,801
Vaccinium myrtillus	NK2018	PRJNA672146	GCA_016920895	36,404*
Vaccinium macrocarpon	Ben Lear	PRJNA245813	GCA_000775335	36,364*
Vaccinium corymbosum	W8520	PRJNA638649	GCA_014504835	70,581*
Juglans regia	Chandler	PRJNA291087	GCF_001411555	45,874
Corylus avellana	Tombul	PRJEB31933	GCA_901000735	27,270*
Pistacia vera	Batoury	PRJNA526975	GCF_008641045	41,217
Diospyros lotus	Yz01	PRJNA646788	GCA_014633365	22,844*
Ficus carica	Dottato	PRJNA565858	GCA_009761775	37,840*
Ficus erecta	FE-Hiroshima-1	PRJDB8644	GCA_008635985	51,806*
Cydonia oblonga	Quince A	PRJNA675337	GCA_015708375	30,684*
Coffea arabica	Caturra red	PRJNA497895	GCF_003713225.1	67,222
Ananas comosus	F153	PRJNA305080	GCF_001540865.1	35,775
Carica papaya	SunUp	PRJNA20267	GCF_000150535.2	25,980
Cocos nucifera	Catigan green dwarf	PRJNA483845	GCA_006176705.1	34,958*
Musa acuminata	Doubled-haploid Pahang	PRJEA82777	GCF_000313855.2	47,707
Musa balbisiana	DH-PKW	PRJNA432894	GCA_004837865	35,148*
Musa itinerans	HN9	PRJNA312694	GCA_001649415	32, 456*
Punica granatum	Tunisia	PRJNA324150	GCF_007655135.1	36,608
Theobroma cacao	B97-61/B2	PRJEB14326	GCF_000208745.1	30,773
Persea americana	Hass	PRJNA694184	GCA_018408905	24,616
Mangifera indica	Alphonso	PRJNA487154	GCA_011075055.1	41,251*
Citrus clementina	Clementine	PRJNA223006	GCF_000493195.1	32,586
Citrus sinensis	Valencia	PRJNA86123	GCF_000317415.1	39,056
Citrus maxima	Pummelo	PRJNA318855	GCA_002006925	42,886*
Citrus trifoliata	ZK8	PRJNA554539	GCA_018350135	39,675*
Citrus reticulata	Mangshan	PRJNA388397	GCA_003258625	42,653*
Citrus medica	XZ	PRJNA320023	GCA_002013955	47,506*
Citrus unshiu	Miyagawa wase	PRJDB5882	GCA_002897195	37,970*
Citrus hindsii	Hongkong kumquat	PRJNA487160	GCA_004802465	32,257*
Citrus cavaleriei	XJC	PRJNA321657	GCA_002013975	43,103*
Citrus x paradisi	Unknown	PRJNA224728	GCA_001929425	unknown
Fragaria x ananassa	Wongyo 3115	PRJNA662854	GCA_019022445	151,892*

(Table 3) contd….

Species	Infraspecific Name	BioProject	Assembly	Number of CDS
Fragaria vesca	Hawaii	PRJNA60037	GCF_000184155	24,056
Rubus occidentalis	ORUS 4115-3	PRJNA430858		34,545*
Malus domestica	Golden Delicious Doubled haploid	PRJNA534520	GCF_002114115.1	52,039
Pyrus x bretschneideri	Unknown	PRJNA259338	GCF_000315295.1	47,086
Prunus armeniaca	GSYX	PRJNA655343	GCA_020226305.1	23,445*
Prunus avium	Satonishiki	PRJDB4877	GCF_002207925.1	35,009
Prunus dulcis	Texas	PRJEB32994	GCF_902201215.1	33,326
Prunus mume	Unknown	PRJNA171605	GCF_000346735.1	29,621
Prunus persica	Lovell	PRJNA241430	GCF_000346465.2	32,595
Prunus salicina	Sanyueli	PRJNA655343	GCA_020226455.1	26,815*
Prunus fruticosa	Harmashatarhegy	PRJNA727075	GCA_018703695	58,880*
Prunus yedoensis	Somei-yoshino	PRJDB7997	GCA_005406145.1	94,776*
Prunus davidiana	ST	PRJNA655343	GCA_020226225.1	27,604*
Prunus mira	GHT	PRJNA655343	GCA_020226265.1	25,953*
Pyrus pyrifolia	Nijisseiki	PRJDB10856	GCA_016587475	44,876*
Vitis riparia	Gloire de Montpellier	PRJNA512170	GCF_004353265.1	39,906
Vitis vinifera	PN40024	PRJEA18785	GCF_000003745.3	41,050

* number of CDS was retrieved from the research article elsewhere the number is from the RefSeq database

functional genomic tools for each species or group of species. For two groups of species (Rosaceas and *Citrus*), specific websites have been developed to aggregate the species' genomes and other data. For the tropical fruits and *Vitis* communities, the resources are not centralized in a single database, a portal acts as a central hub towards various multi-polarized resources.

4.1. Genome Databases for Rosaceae, *Citrus* Genome Database, and *Vaccinium* Genome Database

The Genome Database for Rosacea has been developed since 2003 [58]. Tripal and its modules are used across the website architecture [59] for the implementation of many tools. Amongst them, it is possible to access whole genome sequence, gene annotations, synteny between species, genetic maps, genetic diversity data, markers and trait loci or phenotypic data. *De novo* transcriptome assembly is also available for some species. Further plans are aimed to provide additional analysis tools through implementing a galaxy platform into GDR. The databases for the *Citrus* genome [60] and *Vaccinium* [61] genomes use the same structure as the Rosacea database and are developed by the same team at Washington State University with same tools available. In these databases, the genome repositories are more extensive than what has been deposited at the NCBI, mainly because of the presence of projects yet under review.

4.2. Genome Databases for Tropical Fruits

South Green [62] is a portal used as a public information system that acts as a hub redirecting towards various tropi-

cal crop resources developed in the frame of a federated network of French institutes (Alliance Bioversity CIAT, CIRAD, INRAE, and IRD). South Green ensures the development and hosting of original information systems and analysis tools, such as GreenPhyl [63], SNiPlay [64], Gigwa [65], AgroLD [66] or the Genome Hubs of specific species, as well as offers sequencing data analysis pipelines through two workflow managers: Galaxy [67] and TOGGLe [68]. The genome hubs are independent websites dedicated to specific species, with many of them, such as the banana, cocoa, and coffee genome hubs, developed using Tripal following the Rosacea model.

4.3. The Case of *Vitis*

There is no centralized resource database for grapevine data but rather a constellation of repositories and tools. However, all the available tools are indexed on the resource page of the Integrape website [69], which acts as a central hub for accessing data tools and guidelines for experiment description or data submission to public databases. The grapevine research community heavily relies and promotes publishing data in the major international public multi-species repositories as the primary sources of data access rather than a grapevine specific database managed by a single or a small number of groups depending on public funding [70]. For developing interactions between these tools, this strategy is highly dependent on compliance with the FAIR principles [3] for each of the tools and on allowing correct interoperability. For that reason, several guidelines for correct and standardized metadata for the samples have been drawn with the participation of the whole community.

5. PANGENOME, CONSENSUS GENOME, AND COMPARATIVE GENOMIC DATABASES

Recent developments in sequencing technologies, such as Pacbio and Oxford Nanopore, have permitted to produce longer reads than the Illumina sequencing technologies. Longer reads facilitate the *de novo* assembly of genomes of new cultivars for a species, allowing the reconstruction of the genomes without using a reference genome as the backbone of the assembly. The first major benefit is that these technologies democratize the sequencing and the assembly of genomes that have allowed in recent years to sequence at least one genome for nearly all the fruit tree species with a peak in 2019 (Fig. **3**). All new genomes published in 2021 have used long-read technologies.

The second benefit of these technologies is the possibility to *de novo* sequence additional cultivars for each species. So, genome areas absent from the reference genome can be analyzed and compared between cultivars or closely related species. This trend is only starting to be observed in perennial fruits since, for most of the species, only one or two genomes from different cultivars have been sequenced *de novo*. The major species, grapevine, orange, peach and apple, have more than five cultivars already submitted to the NCBI, but the exact numbers have to be estimated, and sometimes, the cultivar information is missing or not easily identifiable. Clear genotype nomenclature is a point that can be improved. As an example of good practice, the grapevine community recommends a 3-field code using the FAO WIEWS code for holding institution IDs, the IDs and nomenclature from the Vitis International Variety Catalogue (VIVC [71]) for cultivars metadata and a local accession ID for the accession/clone description. As an example, FRA038_VIVC10077_274Col49 would correspond to the clone number 49 of Riesling available at INRAE Colmar.

The pan-genome is defined as the combination of a "core" genome containing the genes or regions of the genome that are present in all individuals/varieties of a species and are essential, with "dispensable" parts that may be absent in one or more varieties. Re-sequencing projects are now being initiated for a large number of varieties of crop species such as rice [72], wheat [73] maize [74], which in

the medium term could provide a more complete map of the allelic variations found in these species, broader than only the genome parts that overlap the reference genomes.

With the availability of sequencing, genome comparison and pangenome studies within and between species have been emerging in recent years. Several studies have been performed on crops. However, the main challenge for any species is the integration of all of these genomes into a single pangenome for each taxon and, more importantly, browsers that replace the reference genome with a pangenome. These integrated pangenomes must be able to provide a user friendly way to present gene annotation, mapping accessory transcriptome and epigenome that would serve as a one-stop shop for genome analysis [75]. This task is complex because it requires working with nonlinear data structure and has to be convenient for researchers and breeders. However, tools for comparing pangenomes at the gene level are available for some perennial fruit species.

GreenPhylDB is a comparative database for plant genomes [63] that has been developed since 2007. It is part of the South Green platform. Since the latest iteration (v5), it has integrated pangenomes for 19 species. This tool allows identifying the individual genomes used to build the species pangenome, containing a particular gene. The homologous genes are grouped into functional families. Among the 19 species pangenomes available so far, 3 are perennial fruits (banana, cacao and grapevine). Some other tools for the functional analysis of the families are also available for apple, orange, strawberry, coconut, coffee pomelo, and citron.

Gramene [76] is a tool jointly developed by the Cold Spring Harbor Laboratory, Oregon State University and the EMBL-EBI. It is a curated, open-source, integrated data resource for comparative functional genomics in crops and model plant species. It also includes tools for comparative genomics between species, allowing the identification of a member of gene families and orthologs within species. Pangenome tools are currently in development. The gene tree is a tool that enables comparative analysis of genes against their paralogs in other species as well as their orthologs in the sequenced genotypes within a species. In Fig. (**4**), the grapevine reference gene Vitvi02g01015 has been identified
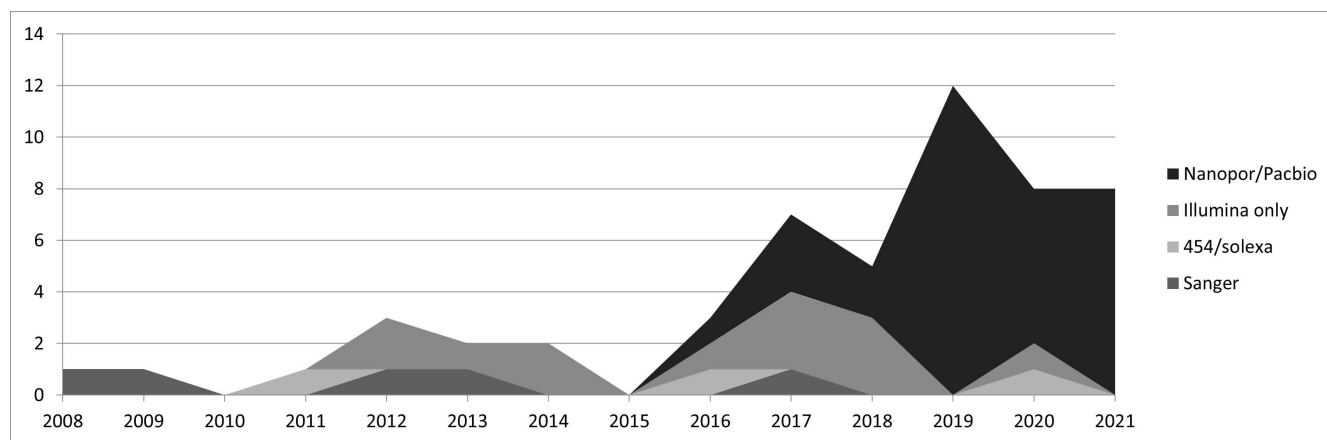

Fig. (3). Chronogram of sequencing technologies used for constructing the reference genomes of perennial fruit species. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).
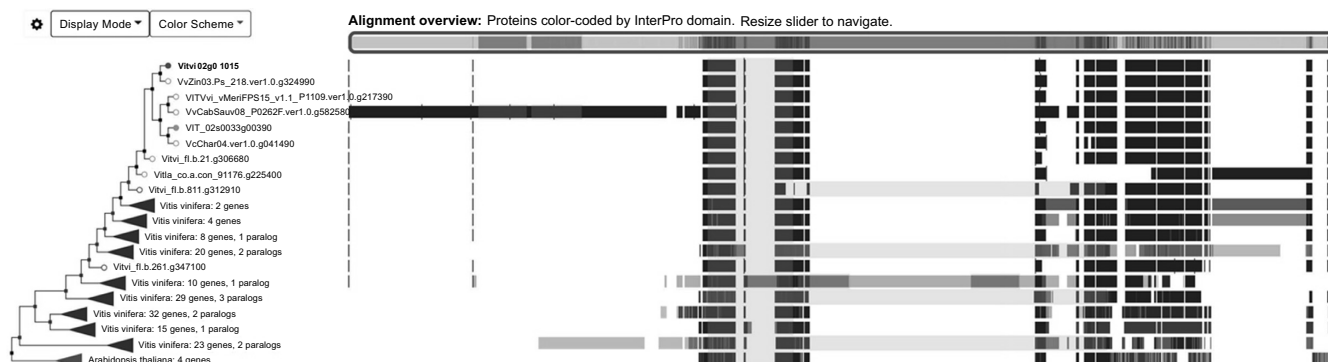
Fig. (4). Example of gramene tree for the gene Vitvi02g01015. Image was obtained from the following direction: https://vitis-ensembl.gramene.org/Vitis_viniferapn40024/Gene/Compara_Tree?collapse=5649520,5650646,5651058,5647056, 5647223,5651685,5648148,5652 258;db=core;g=Vitvi02g01015;r=2:14296841-14298234;t=Vitvi02g01015_t001.

in eight cultivars and hybrids as well as within two versions of the reference genome. It can be observed that this gene has been duplicated in the Flame Seedless cultivar and that, among sequences, the exon/intron structure is rarely identical, only 'Merlot' and 'Zinfandel' show the same structure as the reference PN40024. All three are wine grapes. The strong point of Gramene is the presence of a team of curators as well as a scientific community involved in curation workshops.

Comparative analyses among genomes of closely related species using the pangenome approach have been performed, to some extent, on some perennial fruits like *Malus* [77], *Prunus* [78], Musaceae [79], and *Juglans* [80]. These studies allowed the identification of thousands of genes per species not represented in the initial reference genomes. Six different species were assembled for the *Juglans*, and annotated genes were grouped together with orthofinder [81]. Around half of the orthogroups have gene members from all the six species (core genes) while, on the other end, between 3% and 8% of the genes are specific to one species. In the *Malus* study, in addition to comparing species, re-sequenced varieties were mapped to each *de novo* assembled genomes from three different species. For each species, the fraction of the core genome was between 81.3–87.3%, which is relatively high compared to annual plants. When orthologous genes were found, the core genes were clearly more conserved between species than the variable genes. The Musaceae study was performed at the family taxon level. As a result, the proportion of variable genes was much higher than the genes present in the core genome. In the *Prunus* study, the goal was to identify common characters from cultivars of different *Prunus* species grown at high altitudes compared to the reference genomes for these species. The core group contained 1/3 of the genes from ten Prunus genomes and the variable group 2/3. It was similar to what was observed in Musaceae, where the core gene proportion was lower than in *Malus*, though the observed core genome size is highly dependent on the number of genomes analyzed [82].

6. GENE ANNOTATION

Gene models and their functional predictions can easily be obtained from genomes with computer methods with decent but not perfect accuracy and could largely benefit from manual review [83]. Annotation errors, when evaluated, greatly vary between gene families. It is common to find around 30% of the genes affected by annotation errors though it can even reach 80% [84]. Improved accuracy of the structure of gene models largely relies on transcriptomic data availability. Therefore, the genes with lower expression are more error-prone because of lower sequencing depth and in addition, splicing variants are harder to discriminate. The full-length RNA sequencing technologies are particularly useful in that aspect but not available for all species with a sufficient tissue-representation to capture the whole transcriptome. For functional annotation, even though more sequence data are available, proportion of miss-annotation have increased over time [84]. The more common errors are related to overprediction of the function because of low cut-off or because too specific function is associated with a protein only based on homology with a well-described protein from the same family. The other cases are related to inaccurate structural annotation with important functional motifs being misannotated. Computational solutions are hard to come by because the correct decision on the cutoffs varies among protein families and the ones generically used are selected to allow the most optimal number of correct annotations. Computational methods are essential to gather the maximum information on a protein function and structure, but the decision-making process should include manual validation and eventual correction. Recommended annotation pipelines integrate tools for prediction performed *ab initio* combined to sequence comparison from other species but also complemented by visualization and edition tools for community annotation [85, 86].

As popular automatic annotation tools can be applied for any species, the current bottleneck for perennial fruit species clearly resides in community annotation and quality controls. To be useful for breeders, essential genes (involved in biological processes such as cell division, elongation and differentiation, photosynthesis, respiration, circadian rhythm, flowering, *etc*.) and adaptation genes (responses to biotic and abiotic environmental stresses) must be curated and identified on chromosomes. From the four models of manual annotation described in the literature [87], the "industry cottage" and "party" models are the more appropri-

ate, though, in this multispecies context, the update of these two models named "gatekeeper" [88] is particularly well-suited. It consists of complementing jamboree events held on different species by expert curation. Another alternative that is gaining popularity recently is student annotation [89]. So far, very few community infrastructures with the necessary tools are available for perennial fruit species. Nevertheless, genome manual annotations were carried over by small groups from strawberry [90], kiwi [91], and grapevine [92, 93] researchers. The grapevine community, however, is in the process of releasing tools such as an Apollo web browser that will be interconnected with Gramene visualization tools already used on maize [86] using a similar approach to the grapevine specific nomenclature [94] and annotation protocols available at the Integrape website [95]. The Grape Gene Reference Catalogue [96] released published gene annotations for numerous families, and curators aggregated them into the Grapevine Apollo browser.

7. FUNCTIONAL ANALYSIS DATABASES AND TOOLS

The electronic Fluorescent Pictograph – or eFP – Browse [97], already available [98], allows the exploration of transcriptomic data for hypothesis generation. Typically, it would represent the expression of a gene for a species through a collection of tissues (atlas) in a visual manner. So far, it is available for grapevine and kiwi because gene expression atlas has been previously published for these species [99, 100]. This tool is extremely useful to rapidly cross-check the location of the expression of the gene of interest.

7.1. Transcriptome Database

Most of the initiatives to develop databases centralizing the production of transcriptomic data for a perennial fruit species and allowing easy comparative studies have been performed on the grapevine. There are fewer tools for other species and the existing ones are not periodically updated. For instance, TRANSNAP [101] is a comprehensive database about transcriptomics data for pear. It integrates microarray data published in GEO for 20 samples, including 7 varieties or crossing material, with various tissues from reproductive organs [102]. Functional descriptions of the genes are very extensive (GO, Kegg pathways, domains) but do not include RNAseq data. ReprOlive [103] allows the visualization of transcriptomic data in olive reproductive cells and organs (pollen and pistils) but also data from leaves, roots and radicles. The database is available [104] and presents data from a specific experiment. The VESPUCCI [105] compendium is a complete and comprehensive database of nearly all transcriptomic experiments performed on grapevines using both microarray and RNA-seq data. It is organized as a single coherent expression matrix in which each row represents a gene, and each column represents a condition. VESPUCCI is available online [106]. VTC-Agg (v1.1) [107] is a tool to visualize co-expressed genes from lists of the gene of interest. The co-expression networks were constructed from 1,359 microarray samples (33 experiments). The top 300 co-expressed genes, their associated functional annotation (*i.e.*, Mapman BIN category, VitisNet annotation, and homology to *Arabidopsis* genes), as well as enrichment of BIN and *cis*-regulatory

elements (when available), can be downloaded for further analysis. Bulk downloads of the entire datasets are also available for local query and are available online [108]. The "GRapevine Expression Atlas" is an available [109] all-in-one curated database of grapevine transcriptomic data that contains a web application to analyze public RNAseq data. The database is periodically updated to include the latest released expression data.

CONCLUSION

Abundant genomic data have been produced for the perennial fruit species and now a reference genome is available for almost all of them, which allows the future development of molecular tools for breeding and improving farming practices. Raw data are generally available in public databases, but the description of metadata and the compliance with the FAIR principles need to improve and facilitate the development of performance analysis tools and data sharing.

CONSENT FOR PUBLICATION

Not applicable.

CONFLICT OF INTEREST

The author declares no conflict of interest, financial or otherwise.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's website along with the published article.

REFERENCES

[1]   Fadón, E.; Fernandez, E.; Behn, H.; Luedeling, E. A conceptual framework for winter dormancy in deciduous trees. *Agronomy (Basel),* **2020**, *10*(2), 241.
      http://dx.doi.org/10.3390/agronomy10020241

[2]   Brunner, A.M.; Evans, L.M.; Hsu, C-Y.; Sheng, X. Vernalization and the chilling requirement to exit bud dormancy: Shared or separate regulation? *Front. Plant Sci.,* **2014**, *5*, 732.
      http://dx.doi.org/10.3389/fpls.2014.00732 PMID: 25566302

[3]   Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.W.; da Silva Santos, L.B.; Bourne, P.E.; Bouwman, J.; Brookes, A.J.; Clark, T.; Crosas, M.; Dillo, I.; Dumon, O.; Edmunds, S.; Evelo, C.T.; Finkers, R.; Gonzalez-Beltran, A.; Gray, A.J.; Groth, P.; Goble, C.; Grethe, J.S.; Heringa, J.; 't Hoen, P.A.; Hooft, R.; Kuhn, T.; Kok, R.; Kok, J.; Lusher, S.J.; Martone, M.E.; Mons, A.; Packer, A.L.; Persson, B.; Rocca-Serra, P.; Roos, M.; van Schaik, R.; Sansone, S.A.; Schultes, E.; Sengstag, T.; Slater, T.; Strawn, G.; Swertz, M.A.; Thompson, M.; van der Lei, J.; van Mulligen, E.; Velterop, J.; Waagmeester, A.; Wittenburg, P.; Wolstencroft, K.; Zhao, J.; Mons, B. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data,* **2016**, *3*(1), 160018.

http://dx.doi.org/10.1038/sdata.2016.18 PMID: 26978244

[4]   Wegrzyn, J.L.; Staton, M.A.; Street, N.R.; Main, D.; Grau, E.; Herndon, N.; Buehler, S.; Falk, T.; Zaman, S.; Ramnath, R.; Richter, P.; Sun, L.; Condon, B.; Almsaeed, A.; Chen, M.; Mannapperuma, C.; Jung, S.; Ficklin, S. Cyberinfrastructure to improve forest health and productivity: The role of tree databases in connecting genomes, phenomes, and the environment. *Front. Plant Sci.,* **2019**, *10*, 813.
http://dx.doi.org/10.3389/fpls.2019.00813 PMID: 31293610

[5]   Karamura, D.; Karamura, E.; Blomme, G. General Plant Morphology of Musa.*Banana Breeding*; CRC Press: USA, **2011**.
http://dx.doi.org/10.1201/b10514-2

[6]   Donadio, L.C.; Lederman, I.E.; Roberto, S.R.; Stucchi, E.S. Dwarfing-canopy and rootstock cultivars for fruit trees. *Rev. Bras. Frutic.,* **2019**, *41*(3), e-997.
http://dx.doi.org/10.1590/0100-29452019997

[7]   Barrett, T.; Clark, K.; Gevorgyan, R.; Gorelenkov, V.; Gribov, E.; Karsch-Mizrachi, I.; Kimelman, M.; Pruitt, K.D.; Resenchuk, S.; Tatusova, T.; Yaschenko, E.; Ostell, J. Bioproject and biosample databases at NCBI: Facilitating capture and organization of metadata. *Nucleic Acids Res.,* **2012**, *40*, D57-D63.
http://dx.doi.org/10.1093/nar/gkr1163 PMID: 22139929

[8]   Blaxter, M.; Danchin, A.; Savakis, B.; Fukami-Kobayashi, K.; Kurokawa, K.; Sugano, S.; Roberts, R. J.; Salzberg, S. L.; Wu, C.-I. Reminder to deposit DNA sequences. *Science (80-.),* **2016**, *352*(6287), 780.
http://dx.doi.org/10.1126/science.aaf7672

[9]   Jaillon, O.; Aury, J-M.; Noel, B.; Policriti, A.; Clepet, C.; Casagrande, A.; Choisne, N.; Aubourg, S.; Vitulo, N.; Jubin, C.; Vezzi, A.; Legeai, F.; Hugueney, P.; Dasilva, C.; Horner, D.; Mica, E.; Jublot, D.; Poulain, J.; Bruyère, C.; Billault, A.; Segurens, B.; Gouyvenoux, M.; Ugarte, E.; Cattonaro, F.; Anthouard, V.; Vico, V.; Del Fabbro, C.; Alaux, M.; Di Gaspero, G.; Dumas, V.; Felice, N.; Paillard, S.; Juman, I.; Moroldo, M.; Scalabrin, S.; Canaguier, A.; Le Clainche, I.; Malacrida, G.; Durand, E.; Pesole, G.; Laucou, V.; Chatelet, P.; Merdinoglu, D.; Delledonne, M.; Pezzotti, M.; Lecharny, A.; Scarpelli, C.; Artiguenave, F.; Pè, M.E.; Valle, G.; Morgante, M.; Caboche, M.; Adam-Blondon, A.F.; Weissenbach, J.; Quétier, F.; Wincker, P. French-Italian Public Consortium for Grapevine Genome Characterization. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature,* **2007**, *449*(7161), 463-467.
http://dx.doi.org/10.1038/nature06148 PMID: 17721507

[10]   Ballouz, S.; Dobin, A.; Gillis, J.A. Is it time to change the reference genome? *Genome Biol.,* **2019**, *20*(1), 159.
http://dx.doi.org/10.1186/s13059-019-1774-4 PMID: 31399121

[11]   Tan, Q.; Li, S.; Zhang, Y.; Chen, M.; Wen, B.; Jiang, S.; Chen, X.; Fu, X.; Li, D.; Wu, H.; Wang, Y.; Xiao, W.; Li, L. Chromosome-level genome assemblies of five *Prunus* species and genome-wide association studies for key agronomic traits in peach. *Hortic. Res.,* **2021**, *8*(1), 213.
http://dx.doi.org/10.1038/s41438-021-00648-2 PMID: 34593767

[12]   Lantican, D.V; Strickler, S.R.; Canama, A.O.; Gardoce, R.R.; Mueller, L.A.; Galvez, H.F. *De novo* genome sequence assembly of dwarf coconut (Cocos Nucifera L. 'Catigan Green Dwarf') provides insights into genomic variation between coconut types and related palm species. *G3 Genes Genom Genet,* **2019**, *9*(8), 2377-2393.
http://dx.doi.org/10.1534/g3.119.400215

[13]   Wu, H.; Ma, T.; Kang, M.; Ai, F.; Zhang, J.; Dong, G.; Liu, J. A high-quality *Actinidia chinensis* (kiwifruit) genome. *Hortic. Res.,* **2019**, *6*(1), 117.
http://dx.doi.org/10.1038/s41438-019-0202-y PMID: 31645971

[14]   He, N.; Zhang, C.; Qi, X.; Zhao, S.; Tao, Y.; Yang, G.; Lee, T-H.; Wang, X.; Cai, Q.; Li, D.; Lu, M.; Liao, S.; Luo, G.; He, R.; Tan, X.; Xu, Y.; Li, T.; Zhao, A.; Jia, L.; Fu, Q.; Zeng, Q.; Gao, C.; Ma, B.; Liang, J.; Wang, X.; Shang, J.; Song, P.; Wu, H.; Fan, L.; Wang, Q.; Shuai, Q.; Zhu, J.; Wei, C.; Zhu-Salzman, K.; Jin, D.; Wang, J.; Liu, T.; Yu, M.; Tang, C.; Wang, Z.; Dai, F.; Chen, J.; Liu, Y.; Zhao, S.; Lin, T.; Zhang, S.; Wang, J.; Wang, J.; Yang, H.; Yang, G.; Wang, J.; Paterson, A.H.; Xia, Q.; Ji, D.; Xiang, Z. Draft genome sequence of the mulberry tree *Morus notabilis. Nat. Commun.,* **2013**, *4*(1), 2445.
http://dx.doi.org/10.1038/ncomms3445 PMID: 24048436

[15]   Unver, T.; Wu, Z.; Sterck, L.; Turktas, M.; Lohaus, R.; Li, Z.; Yang, M.; He, L.; Deng, T.; Escalante, F.J.; Llorens, C.; Roig, F.J.; Parmaksiz, I.; Dundar, E.; Xie, F.; Zhang, B.; Ipek, A.; Uranbey, S.; Erayman, M.; Ilhan, E.; Badad, O.; Ghazal, H.; Lightfoot, D.A.; Kasarla, P.; Colantonio, V.; Tombuloglu, H.; Hernandez, P.; Mete, N.; Cetin, O.; Van Montagu, M.; Yang, H.; Gao, Q.; Dorado, G.; Van de Peer, Y. Genome of wild olive and the evolution of oil biosynthesis. *Proc. Natl. Acad. Sci. USA,* **2017**, *114*(44), E9413-E9422.
http://dx.doi.org/10.1073/pnas.1708621114 PMID: 29078332

[16]   Hazzouri, K.M.; Gros-Balthazard, M.; Flowers, J.M.; Copetti, D.; Lemansour, A.; Lebrun, M.; Masmoudi, K.; Ferrand, S.; Dhar, M.I.; Fresquez, Z.A.; Rosas, U.; Zhang, J.; Talag, J.; Lee, S.; Kudrna, D.; Powell, R.F.; Leitch, I.J.; Krueger, R.R.; Wing, R.A.; Amiri, K.M.A.; Purugganan, M.D. Genome-wide association mapping of date palm fruit traits. *Nat. Commun.,* **2019**, *10*(1), 4680.
http://dx.doi.org/10.1038/s41467-019-12604-9 PMID: 31615981

[17]   Wu, C.; Deng, C.; Hilario, E.; Albert, N.W.; Lafferty, D.; Grierson, E.R.P.; Plunkett, B.J.; Elborough, C.; Saei, A.; Günther, C.S. A chromosome-scale assembly of the bilberry genome identifies a complex locus controlling berry anthocyanin composition. *Mol. Ecol. Resour.,* **2021**, *2021*, 13467.
http://dx.doi.org/10.1111/1755-0998.13467

[18]   Polashock, J.; Zelzion, E.; Fajardo, D.; Zalapa, J.; Georgi, L.; Bhattacharya, D.; Vorsa, N. The American cranberry: First insights into the whole genome of a species adapted to bog habitat. *BMC Plant Biol.,* **2014**, *14*(1), 165.
http://dx.doi.org/10.1186/1471-2229-14-165 PMID: 24927653

[19]   Gupta, V.; Estrada, A.D.; Blakley, I.; Reid, R.; Patel, K.; Meyer, M.D.; Andersen, S.U.; Brown, A.F.; Lila, M.A.; Loraine, A.E. RNA-Seq analysis and annotation of a draft blueberry genome assembly identifies candidate genes involved in fruit ripening, biosynthesis of bioactive compounds, and stage-specific alternative splicing. *Gigascience,* **2015**, *4*(1), 5.
http://dx.doi.org/10.1186/s13742-015-0046-9 PMID: 25830017

[20]   Martínez-García, P.J.; Crepeau, M.W.; Puiu, D.; Gonzalez-Ibeas, D.; Whalen, J.; Stevens, K.A.; Paul, R.; Butterfield, T.S.; Britton, M.T.; Reagan, R.L.; Chakraborty, S.; Walawage, S.L.; Vasquez-Gross, H.A.; Cardeno, C.; Famula, R.A.; Pratt, K.; Kuruganti, S.; Aradhya, M.K.; Leslie, C.A.; Dandekar, A.M.; Salzberg, S.L.; Wegrzyn, J.L.; Langley, C.H.; Neale, D.B. The walnut (*Juglans regia*) genome sequence reveals diversity in genes coding for the biosynthesis of non-structural polyphenols. *Plant J.,* **2016**, *87*(5), 507-532.
http://dx.doi.org/10.1111/tpj.13207 PMID: 27145194

[21]   Lucas, S.J.; Kahraman, K.; Avşar, B.; Buggs, R.J.A.; Bilge, I. A chromosome-scale genome assembly of European hazel (*Corylus avellana* L.) reveals targets for crop improvement. *Plant J.,* **2021**, *105*(5), 1413-1430.
http://dx.doi.org/10.1111/tpj.15099 PMID: 33249676

[22]   Zeng, L.; Tu, X-L.; Dai, H.; Han, F-M.; Lu, B-S.; Wang, M-S.; Nanaei, H.A.; Tajabadipour, A.; Mansouri, M.; Li, X-L.; Ji, L.L.; Irwin, D.M.; Zhou, H.; Liu, M.; Zheng, H.K.; Esmailizadeh, A.; Wu, D.D. Whole genomes and transcriptomes reveal adaptation and domestication of pistachio. *Genome Biol.,* **2019**, *20*(1), 79.
http://dx.doi.org/10.1186/s13059-019-1686-3 PMID: 30999938

[23]   Mao, W.; Yao, G.; Wang, S.; Zhou, L.; Chen, G.; Dong, N.; Hu, G. Chromosome-level genomes of seeded and seedless date plum based on third-generation DNA sequencing and Hi-C analysis. *For. Res.,* **2021**, *1*(1), 1-9.
http://dx.doi.org/10.48130/FR-2021-0009

[24]   Usai, G.; Mascagni, F.; Giordani, T.; Vangelisti, A.; Bosi, E.; Zuccolo, A.; Ceccarelli, M.; King, R.; Hassani-Pak, K.; Zambrano, L.S.; Cavallini, A.; Natali, L. Epigenetic patterns within the haplotype phased fig (*Ficus carica* L.) genome. *Plant J.,* **2020**, *102*(3), 600-614.
http://dx.doi.org/10.1111/tpj.14635 PMID: 31808196

[25]   Shirasawa, K.; Yakushiji, H.; Nishimura, R.; Morita, T.; Jikumaru, S.; Ikegami, H.; Toyoda, A.; Hirakawa, H.; Isobe, S. The Ficus erecta genome aids Ceratocystis canker resistance breeding in common fig (*F. carica*). *Plant J.,* **2020**, *102*(6), 1313-1322.
http://dx.doi.org/10.1111/tpj.14703 PMID: 31978270

[26]   Soyturk, A.; Sen, F.; Uncu, A.T.; Celik, I.; Uncu, A.O. *De novo* assembly and characterization of the first draft genome of quince (*Cydonia oblonga* Mill.). *Sci. Rep.,* **2021**, *11*(1), 3818.
http://dx.doi.org/10.1038/s41598-021-83113-3 PMID: 33589687

[27]   Ming, R.; VanBuren, R.; Wai, C.M.; Tang, H.; Schatz, M.C.; Bowers, J.E.; Lyons, E.; Wang, M-L.; Chen, J.; Biggers, E.; Zhang, J.; Huang, L.; Zhang, L.; Miao, W.; Zhang, J.; Ye, Z.; Miao, C.; Lin, Z.; Wang, H.; Zhou, H.; Yim, W.C.; Priest, H.D.; Zheng, C.; Woodhouse, M.; Edger, P.P.; Guyot, R.; Guo, H.B.; Guo, H.; Zheng, G.; Singh, R.; Sharma, A.; Min, X.; Zheng, Y.; Lee, H.; Gurtowski, J.; Sedlazeck, F.J.; Harkess, A.; McKain, M.R.; Liao, Z.; Fang, J.; Liu, J.; Zhang, X.; Zhang, Q.; Hu, W.; Qin, Y.; Wang, K.; Chen, L.Y.; Shirley, N.; Lin, Y.R.; Liu, L.Y.; Hernandez, A.G.; Wright, C.L.; Bulone, V.; Tuskan, G.A.; Heath, K.; Zee, F.; Moore, P.H.; Sunkar, R.; Leebens-Mack, J.H.; Mockler, T.; Bennetzen, J.L.; Freeling, M.; Sankoff, D.; Paterson, A.H.; Zhu, X.; Yang, X.; Smith, J.A.; Cushman, J.C.; Paull, R.E.; Yu, Q. The pineapple genome and the evolution of CAM photosynthesis. *Nat. Genet.,* **2015**, *47*(12), 1435-1442.
        http://dx.doi.org/10.1038/ng.3435 PMID: 26523774

[28]   Ming, R.; Hou, S.; Feng, Y.; Yu, Q.; Dionne-Laporte, A.; Saw, J.H.; Senin, P.; Wang, W.; Ly, B.V.; Lewis, K.L.T.; Salzberg, S.L.; Feng, L.; Jones, M.R.; Skelton, R.L.; Murray, J.E.; Chen, C.; Qian, W.; Shen, J.; Du, P.; Eustice, M.; Tong, E.; Tang, H.; Lyons, E.; Paull, R.E.; Michael, T.P.; Wall, K.; Rice, D.W.; Albert, H.; Wang, M.L.; Zhu, Y.J.; Schatz, M.; Nagarajan, N.; Acob, R.A.; Guan, P.; Blas, A.; Wai, C.M.; Ackerman, C.M.; Ren, Y.; Liu, C.; Wang, J.; Wang, J.; Na, J.K.; Shakirov, E.V.; Haas, B.; Thimmapuram, J.; Nelson, D.; Wang, X.; Bowers, J.E.; Gschwend, A.R.; Delcher, A.L.; Singh, R.; Suzuki, J.Y.; Tripathi, S.; Neupane, K.; Wei, H.; Irikura, B.; Paidi, M.; Jiang, N.; Zhang, W.; Presting, G.; Windsor, A.; Navajas-Pérez, R.; Torres, M.J.; Feltus, F.A.; Porter, B.; Li, Y.; Burroughs, A.M.; Luo, M.C.; Liu, L.; Christopher, D.A.; Mount, S.M.; Moore, P.H.; Sugimura, T.; Jiang, J.; Schuler, M.A.; Friedman, V.; Mitchell-Olds, T.; Shippen, D.E.; dePamphilis, C.W.; Palmer, J.D.; Freeling, M.; Paterson, A.H.; Gonsalves, D.; Wang, L.; Alam, M. The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature,* **2008**, *452*(7190), 991-996.
        http://dx.doi.org/10.1038/nature06856 PMID: 18432245

[29]   D'Hont, A.; Denoeud, F.; Aury, J-M.; Baurens, F-C.; Carreel, F.; Garsmeur, O.; Noel, B.; Bocs, S.; Droc, G.; Rouard, M.; Da Silva, C.; Jabbari, K.; Cardi, C.; Poulain, J.; Souquet, M.; Labadie, K.; Jourda, C.; Lengellé, J.; Rodier-Goud, M.; Alberti, A.; Bernard, M.; Correa, M.; Ayyampalayam, S.; Mckain, M.R.; Leebens-Mack, J.; Burgess, D.; Freeling, M.; Mbéguié-A-Mbéguié, D.; Chabannes, M.; Wicker, T.; Panaud, O.; Barbosa, J.; Hribova, E.; Heslop-Harrison, P.; Habas, R.; Rivallan, R.; Francois, P.; Poiron, C.; Kilian, A.; Burthia, D.; Jenny, C.; Bakry, F.; Brown, S.; Guignon, V.; Kema, G.; Dita, M.; Waalwijk, C.; Joseph, S.; Dievart, A.; Jaillon, O.; Leclercq, J.; Argout, X.; Lyons, E.; Almeida, A.; Jeridi, M.; Dolezel, J.; Roux, N.; Risterucci, A.M.; Weissenbach, J.; Ruiz, M.; Glaszmann, J.C.; Quétier, F.; Yahiaoui, N.; Wincker, P. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature,* **2012**, *488*(7410), 213-217.
        http://dx.doi.org/10.1038/nature11241 PMID: 22801500

[30]   Wang, Z.; Miao, H.; Liu, J.; Xu, B.; Yao, X.; Xu, C.; Zhao, S.; Fang, X.; Jia, C.; Wang, J.; Zhang, J.; Li, J.; Xu, Y.; Wang, J.; Ma, W.; Wu, Z.; Yu, L.; Yang, Y.; Liu, C.; Guo, Y.; Sun, S.; Baurens, F.C.; Martin, G.; Salmon, F.; Garsmeur, O.; Yahiaoui, N.; Hervouet, C.; Rouard, M.; Laboureau, N.; Habas, R.; Ricci, S.; Peng, M.; Guo, A.; Xie, J.; Li, Y.; Ding, Z.; Yan, Y.; Tie, W.; D'Hont, A.; Hu, W.; Jin, Z. Musa balbisiana genome reveals subgenome evolution and functional divergence. *Nat. Plants,* **2019**, *5*(8), 810-821.
        http://dx.doi.org/10.1038/s41477-019-0452-6 PMID: 31308504

[31]   Wu, W.; Yang, Y-L.; He, W-M.; Rouard, M.; Li, W-M.; Xu, M.; Roux, N.; Ge, X-J. Whole genome sequencing of a banana wild relative Musa itinerans provides insights into lineage-specific diversification of the Musa genus. *Sci. Rep.,* **2016**, *6*(1), 31586.
        http://dx.doi.org/10.1038/srep31586 PMID: 27531320

[32]   Luo, X.; Li, H.; Wu, Z.; Yao, W.; Zhao, P.; Cao, D.; Yu, H.; Li, K.; Poudel, K.; Zhao, D.; Zhang, F.; Xia, X.; Chen, L.; Wang, Q.; Jing, D.; Cao, S. The pomegranate (*Punica granatum* L.) draft genome dissects genetic divergence between soft- and hard-seeded cultivars. *Plant Biotechnol. J.,* **2020**, *18*(4), 955-968.
        http://dx.doi.org/10.1111/pbi.13260 PMID: 31549477

[33]   Argout, X.; Salse, J.; Aury, J-M.; Guiltinan, M.J.; Droc, G.; Gouzy, J.; Allegre, M.; Chaparro, C.; Legavre, T.; Maximova, S.N.; Abrouk, M.; Murat, F.; Fouet, O.; Poulain, J.; Ruiz, M.;

        Roguet, Y.; Rodier-Goud, M.; Barbosa-Neto, J.F.; Sabot, F.; Kudrna, D.; Ammiraju, J.S.; Schuster, S.C.; Carlson, J.E.; Sallet, E.; Schiex, T.; Dievart, A.; Kramer, M.; Gelley, L.; Shi, Z.; Bérard, A.; Viot, C.; Boccara, M.; Risterucci, A.M.; Guignon, V.; Sabau, X.; Axtell, M.J.; Ma, Z.; Zhang, Y.; Brown, S.; Bourge, M.; Golser, W.; Song, X.; Clement, D.; Rivallan, R.; Tahi, M.; Akaza, J.M.; Pitollat, B.; Gramacho, K.; D'Hont, A.; Brunel, D.; Infante, D.; Kebe, I.; Costet, P.; Wing, R.; McCombie, W.R.; Guiderdoni, E.; Quetier, F.; Panaud, O.; Wincker, P.; Bocs, S.; Lanaud, C. The genome of *Theobroma cacao. Nat. Genet.,* **2011**, *43*(2), 101-108.
        http://dx.doi.org/10.1038/ng.736 PMID: 21186351

[34]   Rendón-Anaya, M.; Ibarra-Laclette, E.; Méndez-Bravo, A.; Lan, T.; Zheng, C.; Carretero-Paulet, L.; Perez-Torres, C.A.; Chacón-López, A.; Hernandez-Guzmán, G.; Chang, T-H.; Farr, K.M.; Barbazuk, W.B.; Chamala, S.; Mutwil, M.; Shivhare, D.; Alvarez-Ponce, D.; Mitter, N.; Hayward, A.; Fletcher, S.; Rozas, J.; Sánchez Gracia, A.; Kuhn, D.; Barrientos-Priego, A.F.; Salojärvi, J.; Librado, P.; Sankoff, D.; Herrera-Estrella, A.; Albert, V.A.; Herrera-Estrella, L. The avocado genome informs deep angiosperm phylogeny, highlights introgressive hybridization, and reveals pathogen-influenced gene space adaptation. *Proc. Natl. Acad. Sci. USA,* **2019**, *116*(34), 17081-17089.
        http://dx.doi.org/10.1073/pnas.1822129116 PMID: 31387975

[35]   Wang, P.; Luo, Y.; Huang, J.; Gao, S.; Zhu, G.; Dang, Z.; Gai, J.; Yang, M.; Zhu, M.; Zhang, H.; Ye, X.; Gao, A.; Tan, X.; Wang, S.; Wu, S.; Cahoon, E.B.; Bai, B.; Zhao, Z.; Li, Q.; Wei, J.; Chen, H.; Luo, R.; Gong, D.; Tang, K.; Zhang, B.; Ni, Z.; Huang, G.; Hu, S.; Chen, Y. The genome evolution and domestication of tropical fruit mango. *Genome Biol.,* **2020**, *21*(1), 60.
        http://dx.doi.org/10.1186/s13059-020-01959-8 PMID: 32143734

[36]   Wu, G.A.; Prochnik, S.; Jenkins, J.; Salse, J.; Hellsten, U.; Murat, F.; Perrier, X.; Ruiz, M.; Scalabrin, S.; Terol, J.; Takita, M.A.; Labadie, K.; Poulain, J.; Couloux, A.; Jabbari, K.; Cattonaro, F.; Del Fabbro, C.; Pinosio, S.; Zuccolo, A.; Chapman, J.; Grimwood, J.; Tadeo, F.R.; Estornell, L.H.; Muñoz-Sanz, J.V.; Ibanez, V.; Herrero-Ortega, A.; Aleza, P.; Pérez-Pérez, J.; Ramón, D.; Brunel, D.; Luro, F.; Chen, C.; Farmerie, W.G.; Desany, B.; Kodira, C.; Mohiuddin, M.; Harkins, T.; Fredrikson, K.; Burns, P.; Lomsadze, A.; Borodovsky, M.; Reforgiato, G.; Freitas-Astúa, J.; Quetier, F.; Navarro, L.; Roose, M.; Wincker, P.; Schmutz, J.; Morgante, M.; Machado, M.A.; Talon, M.; Jaillon, O.; Ollitrault, P.; Gmitter, F.; Rokhsar, D. Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nat. Biotechnol.,* **2014**, *32*(7), 656-662.
        http://dx.doi.org/10.1038/nbt.2906 PMID: 24908277

[37]   Xu, Q.; Chen, L-L.; Ruan, X.; Chen, D.; Zhu, A.; Chen, C.; Bertrand, D.; Jiao, W-B.; Hao, B-H.; Lyon, M.P.; Chen, J.; Gao, S.; Xing, F.; Lan, H.; Chang, J.W.; Ge, X.; Lei, Y.; Hu, Q.; Miao, Y.; Wang, L.; Xiao, S.; Biswas, M.K.; Zeng, W.; Guo, F.; Cao, H.; Yang, X.; Xu, X.W.; Cheng, Y.J.; Xu, J.; Liu, J.H.; Luo, O.J.; Tang, Z.; Guo, W.W.; Kuang, H.; Zhang, H.Y.; Roose, M.L.; Nagarajan, N.; Deng, X.X.; Ruan, Y. The draft genome of sweet orange (*Citrus sinensis*). *Nat. Genet.,* **2013**, *45*(1), 59-66.
        http://dx.doi.org/10.1038/ng.2472 PMID: 23179022

[38]   Wang, X.; Xu, Y.; Zhang, S.; Cao, L.; Huang, Y.; Cheng, J.; Wu, G.; Tian, S.; Chen, C.; Liu, Y.; Yu, H.; Yang, X.; Lan, H.; Wang, N.; Wang, L.; Xu, J.; Jiang, X.; Xie, Z.; Tan, M.; Larkin, R.M.; Chen, L.L.; Ma, B.G.; Ruan, Y.; Deng, X.; Xu, Q. Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. *Nat. Genet.,* **2017**, *49*(5), 765-772.
        http://dx.doi.org/10.1038/ng.3839 PMID: 28394353

[39]   Huang, Y.; Xu, Y.; Jiang, X.; Yu, H.; Jia, H.; Tan, C.; Hu, G.; Hu, Y.; Rao, M.J.; Deng, X.; Xu, Q. Genome of a citrus rootstock and global DNA demethylation caused by heterografting. *Hortic. Res.,* **2021**, *8*(1), 69.
        http://dx.doi.org/10.1038/s41438-021-00505-2 PMID: 33790260

[40]   Wang, L.; He, F.; Huang, Y.; He, J.; Yang, S.; Zeng, J.; Deng, C.; Jiang, X.; Fang, Y.; Wen, S.; Xu, R.; Yu, H.; Yang, X.; Zhong, G.; Chen, C.; Yan, X.; Zhou, C.; Zhang, H.; Xie, Z.; Larkin, R.M.; Deng, X.; Xu, Q. Genome of wild mandarin and domestication history of mandarin. *Mol. Plant,* **2018**, *11*(8), 1024-1037.
        http://dx.doi.org/10.1016/j.molp.2018.06.001 PMID: 29885473

[41]   Shimizu, T.; Tanizawa, Y.; Mochizuki, T.; Nagasaki, H.; Yoshioka, T.; Toyoda, A.; Fujiyama, A.; Kaminuma, E.; Nakamura, Y. Draft sequencing of the heterozygous diploid genome of satsuma

(*Citrus unshiu* Marc.) using a hybrid assembly approach. *Front. Genet.,* **2017**, *8*, 180.
http://dx.doi.org/10.3389/fgene.2017.00180 PMID: 29259619

[42] Zhu, C.; Zheng, X.; Huang, Y.; Ye, J.; Chen, P.; Zhang, C.; Zhao, F.; Xie, Z.; Zhang, S.; Wang, N.; Li, H.; Wang, L.; Tang, X.; Chai, L.; Xu, Q.; Deng, X. Genome sequencing and CRISPR/Cas9 gene editing of an early flowering Mini-Citrus (*Fortunella hindsii*). *Plant Biotechnol. J.,* **2019**, *17*(11), 2199-2210.
http://dx.doi.org/10.1111/pbi.13132 PMID: 31004551

[43] Zhang, Y.; Barthe, G.; Grosser, J.W.; Wang, N. Transcriptome analysis of root response to citrus blight based on the newly assembled Swingle citrumelo draft genome. *BMC Genomics,* **2016**, *17*(1), 485.
http://dx.doi.org/10.1186/s12864-016-2779-y PMID: 27391971

[44] Lee, H-E.; Manivannan, A.; Lee, S.Y.; Han, K.; Yeum, J-G.; Jo, J.; Kim, J.; Rho, I.R.; Lee, Y-R.; Lee, E.S.; Kang, B.C.; Kim, D.S. Chromosome level assembly of homozygous inbred line 'Wongyo 3115' facilitates the construction of a high-density linkage map and identification of QTLs associated with fruit firmness in octoploid strawberry (*Fragaria × ananassa*). *Front. Plant Sci.,* **2021**, *12*, 696229.
http://dx.doi.org/10.3389/fpls.2021.696229 PMID: 34335662

[45] Shulaev, V.; Sargent, D.J.; Crowhurst, R.N.; Mockler, T.C.; Folkerts, O.; Delcher, A.L.; Jaiswal, P.; Mockaitis, K.; Liston, A.; Mane, S.P.; Burns, P.; Davis, T.M.; Slovin, J.P.; Bassil, N.; Hellens, R.P.; Evans, C.; Harkins, T.; Kodira, C.; Desany, B.; Crasta, O.R.; Jensen, R.V.; Allan, A.C.; Michael, T.P.; Setubal, J.C.; Celton, J.M.; Rees, D.J.; Williams, K.P.; Holt, S.H.; Ruiz Rojas, J.J.; Chatterjee, M.; Liu, B.; Silva, H.; Meisel, L.; Adato, A.; Filichkin, S.A.; Troggio, M.; Viola, R.; Ashman, T.L.; Wang, H.; Dharmawardhana, P.; Elser, J.; Raja, R.; Priest, H.D.; Bryant, D.W., Jr; Fox, S.E.; Givan, S.A.; Wilhelm, L.J.; Naithani, S.; Christoffels, A.; Salama, D.Y.; Carter, J.; Lopez Girona, E.; Zdepski, A.; Wang, W.; Kerstetter, R.A.; Schwab, W.; Korban, S.S.; Davik, J.; Monfort, A.; Denoyes-Rothan, B.; Arus, P.; Mittler, R.; Flinn, B.; Aharoni, A.; Bennetzen, J.L.; Salzberg, S.L.; Dickerman, A.W.; Velasco, R.; Borodovsky, M.; Veilleux, R.E.; Folta, K.M. The genome of woodland strawberry (*Fragaria vesca*). *Nat. Genet.,* **2011**, *43*(2), 109-116.
http://dx.doi.org/10.1038/ng.740 PMID: 21186353

[46] VanBuren, R.; Wai, C.M.; Colle, M.; Wang, J.; Sullivan, S.; Bushakra, J.M.; Liachko, I.; Vining, K.J.; Dossett, M.; Finn, C.E.; Jibran, R.; Chagné, D.; Childs, K.; Edger, P.P.; Mockler, T.C.; Bassil, N.V. A near complete, chromosome-scale assembly of the black raspberry (*Rubus occidentalis*) genome. *Gigascience,* **2018**, *7*(8), giy094.
http://dx.doi.org/10.1093/gigascience/giy094 PMID: 30107523

[47] Daccord, N.; Celton, J-M.; Linsmith, G.; Becker, C.; Choisne, N.; Schijlen, E.; van de Geest, H.; Bianco, L.; Micheletti, D.; Velasco, R.; Di Pierro, E.A.; Gouzy, J.; Rees, D.J.G.; Guérif, P.; Muranty, H.; Durel, C.E.; Laurens, F.; Lespinasse, Y.; Gaillard, S.; Aubourg, S.; Quesneville, H.; Weigel, D.; van de Weg, E.; Troggio, M.; Bucher, E. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nat. Genet.,* **2017**, *49*(7), 1099-1106.
http://dx.doi.org/10.1038/ng.3886 PMID: 28581499

[48] Wu, J.; Wang, Z.; Shi, Z.; Zhang, S.; Ming, R.; Zhu, S.; Khan, M.A.; Tao, S.; Korban, S.S.; Wang, H.; Chen, N.J.; Nishio, T.; Xu, X.; Cong, L.; Qi, K.; Huang, X.; Wang, Y.; Zhao, X.; Wu, J.; Deng, C.; Gou, C.; Zhou, W.; Yin, H.; Qin, G.; Sha, Y.; Tao, Y.; Chen, H.; Yang, Y.; Song, Y.; Zhan, D.; Wang, J.; Li, L.; Dai, M.; Gu, C.; Wang, Y.; Shi, D.; Wang, X.; Zhang, H.; Zeng, L.; Zheng, D.; Wang, C.; Chen, M.; Wang, G.; Xie, L.; Sovero, V.; Sha, S.; Huang, W.; Zhang, S.; Zhang, M.; Sun, J.; Xu, L.; Li, Y.; Liu, X.; Li, Q.; Shen, J.; Wang, J.; Paull, R.E.; Bennetzen, J.L.; Wang, J.; Zhang, S. The genome of the pear (*Pyrus bretschneideri* Rehd.). *Genome Res.,* **2013**, *23*(2), 396-408.
http://dx.doi.org/10.1101/gr.144311.112 PMID: 23149293

[49] Shirasawa, K.; Isuzugawa, K.; Ikenaga, M.; Saito, Y.; Yamamoto, T.; Hirakawa, H.; Isobe, S. The genome sequence of sweet cherry (*Prunus avium*) for use in genomics-assisted breeding. *DNA Res.,* **2017**, *24*(5), 499-508.
http://dx.doi.org/10.1093/dnares/dsx020 PMID: 28541388

[50] Alioto, T.; Alexiou, K.G.; Bardil, A.; Barteri, F.; Castanera, R.; Cruz, F.; Dhingra, A.; Duval, H.; Fernández I Martí, Á.; Frias, L.; Galán, B.; García, J.L.; Howad, W.; Gómez-Garrido, J.; Gut, M.;

Julca, I.; Morata, J.; Puigdomènech, P.; Ribeca, P.; Rubio Cabetas, M.J.; Vlasova, A.; Wirthensohn, M.; Garcia-Mas, J.; Gabaldón, T.; Casacuberta, J.M.; Arús, P. Transposons played a major role in the diversification between the closely related almond and peach genomes: Results from the almond genome sequence. *Plant J.,* **2020**, *101*(2), 455-472.
http://dx.doi.org/10.1111/tpj.14538 PMID: 31529539

[51] 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. The genome of *Prunus mume*. *Nat. Commun.,* **2012**, *3*(1), 1318.
http://dx.doi.org/10.1038/ncomms2290 PMID: 23271652

[52] Verde, I.; Abbott, A.G.; Scalabrin, S.; Jung, S.; Shu, S.; Marroni, F.; Zhebentyayeva, T.; Dettori, M.T.; Grimwood, J.; Cattonaro, F.; Zuccolo, A.; Rossini, L.; Jenkins, J.; Vendramin, E.; Meisel, L.A.; Decroocq, V.; Sosinski, B.; Prochnik, S.; Mitros, T.; Policriti, A.; Cipriani, G.; Dondini, L.; Ficklin, S.; Goodstein, D.M.; Xuan, P.; Del Fabbro, C.; Aramini, V.; Copetti, D.; Gonzalez, S.; Horner, D.S.; Falchi, R.; Lucas, S.; Mica, E.; Maldonado, L.; Lazzari, B.; Bielenberg, D.; Pirona, R.; Miculan, M.; Barakat, A.; Testolin, R.; Stella, A.; Tartarini, S.; Tonutti, P.; Arús, P.; Orellana, A.; Wells, C.; Main, D.; Vizzotto, G.; Silva, H.; Salamini, F.; Schmutz, J.; Morgante, M.; Rokhsar, D.S. International Peach Genome Initiative. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat. Genet.,* **2013**, *45*(5), 487-494.
http://dx.doi.org/10.1038/ng.2586 PMID: 23525075

[53] Shirasawa, K.; Esumi, T.; Hirakawa, H.; Tanaka, H.; Itai, A.; Ghelfi, A.; Nagasaki, H.; Isobe, S. Phased genome sequence of an interspecific hybrid flowering cherry, 'Somei-Yoshino' (Cerasus × yedoensis). *DNA Res.,* **2019**, *26*(5), 379-389.
http://dx.doi.org/10.1093/dnares/dsz016 PMID: 31334758

[54] Shirasawa, K.; Itai, A.; Isobe, S. Chromosome-scale genome assembly of Japanese pear (*Pyrus pyrifolia*) variety 'Nijisseiki'. *DNA Res.,* **2021**, *28*(2), dsab001.
http://dx.doi.org/10.1093/dnares/dsab001 PMID: 33638981

[55] Girollet, N.; Rubio, B.; Lopez-Roques, C.; Valière, S.; Ollat, N.; Bert, P-F. *De novo* phased assembly of the Vitis riparia grape genome. *Sci. Data,* **2019**, *6*(1), 127.
http://dx.doi.org/10.1038/s41597-019-0133-3 PMID: 31324816

[56] van Straalen, N.M.; Roelofs, D. *An Introduction to Ecological Genomics*; Oxford University Press, USA, **2011**.
http://dx.doi.org/10.1093/acprof:oso/9780199594689.001.0001

[57] Gnomon The NCBI eukaryotic gene prediction tool. Available from: https://www.ncbi.nlm.nih.gov/genome/annotation_euk/gnomon (Accessed on: Oct 18, 2021).

[58] Jung, S.; Lee, T.; Cheng, C-H.; Buble, K.; Zheng, P.; Yu, J.; Humann, J.; Ficklin, S.P.; Gasic, K.; Scott, K.; Frank, M.; Ru, S.; Hough, H.; Evans, K.; Peace, C.; Olmstead, M.; DeVetter, L.W.; McFerson, J.; Coe, M.; Wegrzyn, J.L.; Staton, M.E.; Abbott, A.G.; Main, D. 15 years of GDR: New data and functionality in the genome database for rosaceae. *Nucleic Acids Res.,* **2019**, *47*, D1137-D1145.
http://dx.doi.org/10.1093/nar/gky1000 PMID: 30357347

[59] Spoor, S.; Cheng, C-H.; Sanderson, L-A.; Condon, B.; Almsaeed, A.; Chen, M.; Bretaudeau, A.; Rasche, H.; Jung, S.; Main, D.; Bett, K.; Staton, M.; Wegrzyn, J.L.; Feltus, F.A.; Ficklin, S.P. Tripal v3: An ontology-based toolkit for construction of FAIR biological community databases. *Database (Oxford),* **2019**, *2019*, baz077.
http://dx.doi.org/10.1093/database/baz077 PMID: 31328773

[60] Citrus genome database. Available from: https://www.citrusgenomedb.org/ (Accessed on: Oct 19, 2021).

[61] Genome database for vaccinium. Available from: https://www.vaccinium.org (Accessed on: Oct 19, 2021).

[62] South Green collaborators. The south green portal: A comprehensive resource for tropical and mediterranean crop genomics. *Curr. Plant Biol,* **2016**, *2016*, 7-8.
http://dx.doi.org/10.1016/j.cpb.2016.12.002

[63] Rouard, M.; Guignon, V.; Aluome, C.; Laporte, M.A.; Droc, G.; Walde, C.; Zmasek, C.M.; Périn, C.; Conte, M.G. GreenPhylDB v2.0: Comparative and functional genomics in plants. *Nucleic Acids Res.,* **2011**, *39*(Suppl. 1), D1095-D1102.
http://dx.doi.org/10.1093/nar/gkq811 PMID: 20864446

[64] Dereeper, A.; Homa, F.; Andres, G.; Sempere, G.; Sarah, G.; Hueber, Y.; Dufayard, J.F.; Ruiz, M. SNiPlay3: A web-based application for exploration and large scale analyses of genomic variations. *Nucleic Acids Res.,* **2015**, *43*(W1), W295-300.
http://dx.doi.org/10.1093/nar/gkv351 PMID: 26040700

[65] Sempéré, G.; Pétel, A.; Rouard, M.; Frouin, J.; Hueber, Y.; De Bellis, F.; Larmande, P. Gigwa v2-extended and improved genotype investigator. *Gigascience,* **2019**, *8*(5), giz051.
http://dx.doi.org/10.1093/gigascience/giz051 PMID: 31077313

[66] Venkatesan, A.; Tagny Ngompe, G.; Hassouni, N.E.; Chentli, I.; Guignon, V.; Jonquet, C.; Ruiz, M.; Larmande, P. Agronomic Linked Data (AgroLD): A knowledge-based system to enable integrative biology in agronomy. *PLoS One,* **2018**, *13*(11), e0198270.
http://dx.doi.org/10.1371/journal.pone.0198270 PMID: 30500839

[67] Goecks, J.; Nekrutenko, A.; Taylor, J.; Afgan, E.; Ananda, G.; Baker, D.; Blankenberg, D.; Chakrabarty, R.; Coraor, N.; Goecks, J. Galaxy Team. Galaxy: A comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.,* **2010**, *11*(8), R86.
http://dx.doi.org/10.1186/gb-2010-11-8-r86 PMID: 20738864

[68] Monat, C.; Tranchant-Dubreuil, C.; Kougbeadjo, A.; Farcy, C.; Ortega-Abboud, E.; Amanzougarene, S.; Ravel, S.; Agbessi, M.; Orjuela-Bouniol, J.; Summo, M.; Sabot, F. TOGGLE: Toolbox for generic NGS analyses. *BMC Bioinformatics,* **2015**, *16*(1), 374.
http://dx.doi.org/10.1186/s12859-015-0795-6 PMID: 26552596

[69] Integrape - Data integration to maximise the power of omics for grapevine improvement News. Available from: https://integrape.eu/ (Accessed on: Oct 19, 2021).

[70] Adam-Blondon, A-F.; Alaux, M.; Pommier, C.; Cantu, D.; Cheng, Z-M.; Cramer, G.R.; Davies, C.; Delrot, S.; Deluc, L.; Di Gaspero, G.; Grimplet, J.; Fennell, A.; Londo, J.P.; Kersey, P.; Mattivi, F.; Naithani, S.; Neveu, P.; Nikolski, M.; Pezzotti, M.; Reisch, B.I.; Töpfer, R.; Vivier, M.A.; Ware, D.; Quesneville, H. Towards an open grapevine information system. *Hortic. Res.,* **2016**, *3*(1), 16056.
http://dx.doi.org/10.1038/hortres.2016.56 PMID: 27917288

[71] Vitis international variety catalogue VIVC. Available from: https://www.vivc.de/ (Accessed on: Oct 19, 2021).

[72] Duitama, J.; Silva, A.; Sanabria, Y.; Cruz, D.F.; Quintero, C.; Ballen, C.; Lorieux, M.; Scheffler, B.; Farmer, A.; Torres, E.; Oard, J.; Tohme, J. Whole genome sequencing of elite rice cultivars as a comprehensive information resource for marker assisted selection. *PLoS One,* **2015**, *10*(4), e0124617.
http://dx.doi.org/10.1371/journal.pone.0124617 PMID: 25923345

[73] Montenegro, J.D.; Golicz, A.A.; Bayer, P.E.; Hurgobin, B.; Lee, H.; Chan, C.K.; Visendi, P.; Lai, K.; Doležel, J.; Batley, J.; Edwards, D. The pangenome of hexaploid bread wheat. *Plant J.,* **2017**, *90*(5), 1007-1013.
http://dx.doi.org/10.1111/tpj.13515 PMID: 28231383

[74] Hufford, M. B.; Seetharam, A. S.; Woodhouse, M. R.; Chougule, K. M.; Ou, S.; Liu, J.; Ricci, W. A.; Guo, T.; Olson, A.; Qiu, Y. *De novo* assembly, annotation, and comparative analysis of 26 diverse maize genomes. *Science (80-.),* **2021**, *373*(6555), 655-662.
http://dx.doi.org/10.1126/science.abg5289

[75] Jayakodi, M.; Schreiber, M.; Stein, N.; Mascher, M. Building pangenome infrastructures for crop plants and their use in association genetics. *DNA Res.,* **2021**, *28*(1), dsaa030.
http://dx.doi.org/10.1093/dnares/dsaa030 PMID: 33484244

[76] Tello-Ruiz, M.K.; Naithani, S.; Gupta, P.; Olson, A.; Wei, S.; Preece, J.; Jiao, Y.; Wang, B.; Chougule, K.; Garg, P.; Elser, J.; Kumari, S.; Kumar, V.; Contreras-Moreira, B.; Naamati, G.; George, N.; Cook, J.; Bolser, D.; D'Eustachio, P.; Stein, L.D.; Gupta, A.; Xu, W.; Regala, J.; Papatheodorou, I.; Kersey, P.J.; Flicek, P.; Taylor, C.; Jaiswal, P.; Ware, D. Gramene 2021: Harnessing the power of comparative genomics and pathways for plant research. *Nucleic Acids Res.,* **2021**, *49*(D1), D1452-D1463.
http://dx.doi.org/10.1093/nar/gkaa979 PMID: 33170273

[77] Sun, X.; Jiao, C.; Schwaninger, H.; Chao, C.T.; Ma, Y.; Duan, N.; Khan, A.; Ban, S.; Xu, K.; Cheng, L.; Zhong, G.Y.; Fei, Z. Phased diploid genome assemblies and pan-genomes provide insights into the genetic history of apple domestication. *Nat. Genet.,* **2020**, *52*(12), 1423-1432.
http://dx.doi.org/10.1038/s41588-020-00723-9 PMID: 33139952

[78] Wang, X.; Liu, S.; Zuo, H.; Zheng, W.; Zhang, S.; Huang, Y.; Pingcuo, G.; Ying, H.; Zhao, F.; Li, Y.; Liu, J.; Yi, T.S.; Zan, Y.;

Larkin, R.M.; Deng, X.; Zeng, X.; Xu, Q. Genomic basis of high-altitude adaptation in *Tibetan Prunus* fruit trees. *Curr. Biol.,* **2021**, *31*(17), 3848-3860.e8.
http://dx.doi.org/10.1016/j.cub.2021.06.062 PMID: 34314676

[79] Rijzaani, H.; Bayer, P.E.; Rouard, M.; Doležel, J.; Batley, J.; Edwards, D. The pangenome of banana highlights differences between genera and genomes. *Plant Genome,* **2021**, e20100.
http://dx.doi.org/10.1002/tpg2.20100 PMID: 34227250

[80] Trouern-Trend, A.J.; Falk, T.; Zaman, S.; Caballero, M.; Neale, D.B.; Langley, C.H.; Dandekar, A.M.; Stevens, K.A.; Wegrzyn, J.L. Comparative genomics of six Juglans species reveals disease-associated gene family contractions. *Plant J.,* **2020**, *102*(2), 410-423.
http://dx.doi.org/10.1111/tpj.14630 PMID: 31823432

[81] Emms, D.M.; Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.,* **2019**, *20*(1), 238.
http://dx.doi.org/10.1186/s13059-019-1832-y PMID: 31727128

[82] Contreras-Moreira, B.; Vinuesa, P. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl. Environ. Microbiol.,* **2013**, *79*(24), 7696-7701.
http://dx.doi.org/10.1128/AEM.02411-13 PMID: 24096415

[83] Pfeiffer, F.; Oesterhelt, D. A manual curation strategy to improve genome annotation: Application to a set of haloarchaeal genomes. *Life (Basel),* **2015**, *5*(2), 1427-1444.
http://dx.doi.org/10.3390/life5021427 PMID: 26042526

[84] Schnoes, A.M.; Brown, S.D.; Dodevski, I.; Babbitt, P.C. Annotation error in public databases: Misannotation of molecular function in enzyme superfamilies. *PLOS Comput. Biol.,* **2009**, *5*(12), e1000605.
http://dx.doi.org/10.1371/journal.pcbi.1000605 PMID: 20011109

[85] Ejigu, G.F.; Jung, J. Review on the computational genome annotation of sequences obtained by next-generation sequencing. *Biology (Basel),* **2020**, *9*(9), 295.
http://dx.doi.org/10.3390/biology9090295 PMID: 32962098

[86] Tello-Ruiz, M.K.; Marco, C.F.; Hsu, F-M.; Khangura, R.S.; Qiao, P.; Sapkota, S.; Stitzer, M.C.; Wasikowski, R.; Wu, H.; Zhan, J.; Chougule, K.; Barone, L.C.; Ghiban, C.; Muna, D.; Olson, A.C.; Wang, L.; Ware, D.; Micklos, D.A. Double triage to identify poorly annotated genes in maize: The missing link in community curation. *PLoS One,* **2019**, *14*(10), e0224086.
http://dx.doi.org/10.1371/journal.pone.0224086 PMID: 31658277

[87] Stein, L. Genome annotation: From sequence to biology. *Nat. Rev. Genet.,* **2001**, *2*(7), 493-503.
http://dx.doi.org/10.1038/35080529 PMID: 11433356

[88] Loveland, J.E.; Gilbert, J.G.R.; Griffiths, E.; Harrow, J.L. Community gene annotation in practice. *Database (Oxford),* **2012**, *2012*, bas009-bas009.
http://dx.doi.org/10.1093/database/bas009 PMID: 22434843

[89] Hosmani, P.S.; Shippy, T.; Miller, S.; Benoit, J.B.; Munoz-Torres, M.; Flores-Gonzalez, M.; Mueller, L.A.; Wiersma-Koch, H.; D'Elia, T.; Brown, S.J.; Saha, S. A quick guide for student-driven community genome annotation. *PLOS Comput. Biol.,* **2019**, *15*(4), e1006682.
http://dx.doi.org/10.1371/journal.pcbi.1006682 PMID: 30943207

[90] Liu, T.; Li, M.; Liu, Z.; Ai, X.; Li, Y. Reannotation of the cultivated strawberry genome and establishment of a strawberry genome database. *Hortic. Res.,* **2021**, *8*(1), 41.
http://dx.doi.org/10.1038/s41438-021-00476-4 PMID: 33642572

[91] Pilkington, S.M.; Crowhurst, R.; Hilario, E.; Nardozza, S.; Fraser, L.; Peng, Y.; Gunaseelan, K.; Simpson, R.; Tahir, J.; Deroles, S.C.; Templeton, K.; Luo, Z.; Davy, M.; Cheng, C.; McNeilage, M.; Scaglione, D.; Liu, Y.; Zhang, Q.; Datson, P.; De Silva, N.; Gardiner, S.E.; Bassett, H.; Chagné, D.; McCallum, J.; Dzierzon, H.; Deng, C.; Wang, Y.Y.; Barron, L.; Manako, K.; Bowen, J.; Foster, T.M.; Erridge, Z.A.; Tiffin, H.; Waite, C.N.; Davies, K.M.; Grierson, E.P.; Laing, W.A.; Kirk, R.; Chen, X.; Wood, M.; Montefiori, M.; Brummell, D.A.; Schwinn, K.E.; Catanach, A.; Fullerton, C.; Li, D.; Meiyalaghan, S.; Nieuwenhuizen, N.; Read, N.; Prakash, R.; Hunter, D.; Zhang, H.; McKenzie, M.; Knäbel, M.; Harris, A.; Allan, A.C.; Gleave, A.; Chen, A.; Janssen, B.J.; Plunkett, B.; Ampomah-Dwamena, C.; Voogd, C.; Leif, D.; Lafferty, D.; Souleyre, E.J.F.; Varkonyi-Gasic, E.; Gambi, F.; Hanley, J.; Yao, J.L.; Cheung, J.; David, K.M.; Warren, B.; Marsh, K.; Snowden, K.C.; Lin-Wang, K.; Brian, L.; Martinez-Sanchez, M.; Wang,

M.; Ileperuma, N.; Macnee, N.; Campin, R.; McAtee, P.; Drummond, R.S.M.; Espley, R.V.; Ireland, H.S.; Wu, R.; Atkinson, R.G.; Karunairetnam, S.; Bulley, S.; Chunkath, S.; Hanley, Z.; Storey, R.; Thrimawithana, A.H.; Thomson, S.; David, C.; Testolin, R.; Huang, H.; Hellens, R.P.; Schaffer, R.J. A manually annotated *Actinidia chinensis* var. chinensis (kiwifruit) genome highlights the challenges associated with draft genomes and gene prediction in plants. *BMC Genomics,* **2018**, *19*(1), 257.
http://dx.doi.org/10.1186/s12864-018-4656-3 PMID: 29661190

[92]   Grimplet, J.; Van Hemert, J.; Carbonell-Bejerano, P.; Díaz-Riquelme, J.; Dickerson, J.; Fennell, A.; Pezzotti, M.; Martínez-Zapater, J.M. Comparative analysis of grapevine whole-genome gene predictions, functional annotation, categorization and integration of the predicted gene sequences. *BMC Res. Notes,* **2012**, *5*(1), 213.
http://dx.doi.org/10.1186/1756-0500-5-213 PMID: 22554261

[93]   Canaguier, A.; Grimplet, J.; Di Gaspero, G.; Scalabrin, S.; Duchêne, E.; Choisne, N.; Mohellibi, N.; Guichard, C.; Rombauts, S.; Le Clainche, I.; Bérard, A.; Chauveau, A.; Bounon, R.; Rustenholz, C.; Morgante, M.; Le Paslier, M.C.; Brunel, D.; Adam-Blondon, A.F. A new version of the grapevine reference genome assembly (12X.v2) and of its annotation (VCost.v3). *Genom. Data,* **2017**, *14*, 56-62.
http://dx.doi.org/10.1016/j.gdata.2017.09.002 PMID: 28971018

[94]   Grimplet, J.; Adam-Blondon, A-F.; Bert, P-F.; Bitz, O.; Cantu, D.; Davies, C.; Delrot, S.; Pezzotti, M.; Rombauts, S.; Cramer, G.R. The grapevine gene nomenclature system. *BMC Genomics,* **2014**, *15*(1), 1077.
http://dx.doi.org/10.1186/1471-2164-15-1077 PMID: 25481684

[95]   Guidelines for grapevine data management. Available from: https://doi.org/https://integrape.eu/resources/data-management/

[96]   Navarro-Payá, D.; Santiago, A.; Orduña, L.; Zhang, C.; Amato, A.; D'Inca, E.; Fattorini, C.; Pezzotti, M.; Tornielli, G.B.; Zenoni, S.; Rustenholz, C.; Matus, J.T. The grape gene reference catalogue as a standard resource for gene selection and genetic improvement. *Front. Plant Sci.,* **2022**, *12*, 803977.
http://dx.doi.org/10.3389/fpls.2021.803977 PMID: 35111182

[97]   Winter, D.; Vinegar, B.; Nahal, H.; Ammar, R.; Wilson, G.V.; Provart, N.J. An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS One,* **2007**, *2*(8), e718.
http://dx.doi.org/10.1371/journal.pone.0000718 PMID: 17684564

[98]   Bio-analytic resource for plant biology. Available from: http://bar.utoronto.ca/ (Accessed on: Oct 19, 2021).

[99]   Fasoli, M.; Dal Santo, S.; Zenoni, S.; Tornielli, G.B.; Farina, L.; Zamboni, A.; Porceddu, A.; Venturini, L.; Bicego, M.; Murino,

V.; Ferrarini, A.; Delledonne, M.; Pezzotti, M. The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. *Plant Cell,* **2012**, *24*(9), 3489-3505.
http://dx.doi.org/10.1105/tpc.112.100230 PMID: 22948079

[100] Brian, L.; Warren, B.; McAtee, P.; Rodrigues, J.; Nieuwenhuizen, N.; Pasha, A.; David, K.M.; Richardson, A.; Provart, N.J.; Allan, A.C.; Varkonyi-Gasic, E.; Schaffer, R.J. A gene expression atlas for kiwifruit (*Actinidia chinensis*) and network analysis of transcription factors. *BMC Plant Biol.,* **2021**, *21*(1), 121.
http://dx.doi.org/10.1186/s12870-021-02894-x PMID: 33639842

[101] Koshimizu, S.; Nakamura, Y.; Nishitani, C.; Kobayashi, M.; Ohyanagi, H.; Yamamoto, T.; Yano, K. TRANSNAP: A web database providing comprehensive information on Japanese pear transcriptome. *Sci. Rep.,* **2019**, *9*(1), 18922.
http://dx.doi.org/10.1038/s41598-019-55287-4 PMID: 31831861

[102] Japanese pear transcriptome database. Available from: http://plantomics.mind.meiji.ac.jp/nashi (Accessed on: Oct 19, 2021).

[103] Carmona, R.; Zafra, A.; Seoane, P.; Castro, A.J.; Guerrero-Fernández, D.; Castillo-Castillo, T.; Medina-García, A.; Cánovas, F.M.; Aldana-Montes, J.F.; Navas-Delgado, I.; Alché, J.D.; Claros, M.G. ReprOlive: A database with linked data for the olive tree (*Olea europaea* L.) reproductive transcriptome. *Front. Plant Sci.,* **2015**, *6*, 625.
http://dx.doi.org/10.3389/fpls.2015.00625 PMID: 26322066

[104] ReprOlive. Available from: http://reprolive.eez.csic.es (Accessed on: Oct 19, 2021).

[105] Moretto, M.; Sonego, P.; Pilati, S.; Malacarne, G.; Costantini, L.; Grzeskowiak, L.; Bagagli, G.; Grando, M.S.; Moser, C.; Engelen, K. VESPUCCI: Exploring patterns of gene expression in grapevine. *Front. Plant Sci.,* **2016**, *7*, 633.
http://dx.doi.org/10.3389/fpls.2016.00633 PMID: 27242836

[106] VESPUCCI. Available from: http://vespucci.colombos.fmach.it/ (Accessed on: Oct 19, 2021).

[107] Wong, D.C.J. Network aggregation improves gene function prediction of grapevine gene co-expression networks. *Plant Mol. Biol.,* **2020**, *103*(4-5), 425-441.
http://dx.doi.org/10.1007/s11103-020-01001-2 PMID: 32266646

[108] Aggregate and correlation rank grapevine gene co-expression networks. Available from: https://sites.google.com/view/vtc-agg (Accessed on: Oct 19, 2021).

[109] GRapevine expression atlas. Available from: https://great.colmar.inrae.fr/ (Accessed on: Oct 19, 2021).