**Updates on Bioinformatics Database Resources for Plant Genomics**

Dr. K.S. Abbiramy1, Dr. S.Sugunakala2\*, Dr. K.Thenmozhi3, Ms. T. Akilandeswari4

1 Department of Zoology & Wild life Biology, A.V.C.College (Autonomous), Mannampandal, Mayiladuthurai – 609 305.

2\* Department of Bioinformatics, A.V.C.College (Autonomous), Mannampandal, Mayiladuthurai – 609 305.

3 Department of Zoology & Wild life Biology, A.V.C.College (Autonomous), Mannampandal, Mayiladuthurai – 609 305.

4 Department of Biotechnology, A.V.C.College (Autonomous), Mannampandal, Mayiladuthurai – 609 305.

**ABSTRACT**

Advancement in genomic sequencing technology leads to a radical change in plant biology research. Plant biologists now have easy access to massive genome sequencing data to study high plant densities in depth genetic deviation at the molecular level. Hence, a complete understanding and use of bioinformatics tools to manage and analyze this omics data are necessary in current plant genomic research. Many plant genomic databases have been established and have continued to expand in recent times. In the meantime, analytical methods based on bioinformatics are also well developed in many aspects of plant genomics research including gene function prediction by using comparative genomic analysis, evolutionary analysis and whole genome association studies. However, the constant upgrade of IT infrastructure, such as high capacity and high data storage analysis software, is the existent challenge for plant genomic research. This chapter focuses on the existing database resources for various plants that will potentially helps the plant scientists in the current era of plant genomics.

**Keywords:** Plant genomics, Databases for plant genomes, Bioinformatics, Comparative genomics, GWAS, sequencing technologies, Plant transcriptome databases.

1. **INTRODUCTION**

The researchers found that there are 4, 35,000 distinctive species of land plants and they discovered that 36.5 % of plant species are especially rare plants. Majority of plant genomes are unusually diverse compared to those of other kingdoms, varying in size from ∼10 Mb to over 100 GB. In common, the plants are vital kingdom for the earth’s living systems, as they are the major sources for nutritional food, animal feed, medicines, etc., [1, 2] and they have been chosen as an model organisms to study the epigenetic role and heterochromatin elements [3]. An appreciative knowledge of variety of species genomes especially in plant genomes are primarily lagged, but have significantly increased with the advent of new technologies in DNA sequencing and leads to an exponential growth of genomic data in wide variety of biological systems.

In modern research, sequencing of a plant species genome is considered to be an important turning point. The modernism of high-throughput sequencing techniques helps the researchers to take advantage of exploring the structure of the plant genetic material at the molecular level i.e. known as “Plant Genomics”. Large volume of biological data especially plant genomic data obtained from various sequencing technologies will be effectively handled and significant findings will be extracted by applying various resources of bioinformatics that helps the scientists to have insight on the structural organization, positional arrangement, and functionality of genes in a genome unambiguously. In 2000, the first plant *Arabidopsis thaliana* genome was sequenced and from this date, an approximate of 600 plant species genomes were completely sequenced and published [4] and it also includes many economically important crops genomes. These sequential data have been properly compiled and deposited in databases and are widely used in identification, breeding and conservation purposes. In recent years, numbers of reviews had focused the updating of the plant sequencing projects and the development of crop plant database [5 - 9]. Interestingly, the availability of notable resources of bioinformatics databases and tools facilitates to grow a fruitful environment for plant research [10].

In this chapter, various sequencing technologies in plant genomic research, features and applications of plant genomic databases, most accepted plant genomic resources will be elaborated. Currently, there are notable diversity of bioinformatics resources are available for plant research. Attention was primarily given to the databases which are hosted by government organizations and academic research centers and non commercial (freely available) analysis tools and software widely used in plants research.

1. **SEQUENCING TECHNOLOGIES IN PLANT RESEARCH**

In the past decade, almost all DNA sequencing process were restricted with some conventional methods of sequencing namely capillary based, modified and partially automated Sanger’s methods of sequencing [11 – 13]. These methods have been revitalized and flourished because of various scientific innovations which ultimately lead to the development of novel experiments [14]. It includes that the release of high - throughput sequencing – next generation sequencing (NGS) technologies in 2005. They are extensivelyused to assess wide range of biological phenomenon and are employed in resequencing and whole sequencing of genomes, RNA and ChIP sequencing, etc., The more sophisticated methods of NGS technologies are sequencing by synthesis (pyro 454 sequencing, Illumina), sequencing-by-ligation (SOLiD), non optical sequencing (ion torrent semi conductor), SMRT (HeliScope Single Molecule Real Time) and DNA nanoball, nanopore sequencing etc., [15, 16]. They are all well suited for the present and future large scale sequencing demands.

1. **Sequencing by Synthesis (SBS)**
2. Pyro sequencing

Make use of DNA polymerase to lengthen various DNA strands in parallel.During the integration of nucleotide, pyrophosphate is released which is detected by the light signal emitted from the cleavage of oxyluciferin by an enzyme luciferase [17] The commercially available platform is Genome Sequencer from Roche /454 pyrosequencing.

1. llumina Sequencing (SOLEXA)

The principle involved in this sequencing technique is “Sequencing by Synthesis” (SBS) and clonal amplification. In this method, modified deoxy nucleotide triphosphates (dNTPs) i.e consisting of terminator is used, which stop further polymerization.Once, the modified dNTPs are incorporated, extension proceeds with one single base in each growing DNA copy strand [18]. Various upgraded versions of this technologyconsist of Miniseq series, Mi seq series, Nextseq series and Novaseq series [19].

1. **Sequencing-by-Ligation (SBL)**

SOLiD Sequencing

Since 2006, it has been commercially available from Applied Biosystems. DNA ligase is used to generate sequential ligation of dye labeled oligonucleotides and facilitate parallel sequencing of clonally amplified DNA strands [20]. The inconsistent nature of these amplified DNA helps to determine the concealed sequence of the target.

1. **Non Optical Sequencing**

Ion semiconductor – Sense the hydrogen ions that are released during DNA polymerization. Commercially available sequencer is Ion Torrent from Life Science, Inc.

1. **Single Molecule Sequencing (SMS)**

HeliScope Single Molecule Real Time –an enzymatic degradation of fluorescently labeled single DNA molecule releases the monomer molecules with respect to their sequential order and they will be detected and identified [21, 22].The available commercial platform is from Helicos Bioscience Corporation (True single molecule sequencing), PacBio RS sequencers from Pacific Biosciences (Single molecule, realtime sequencing) [23].

1. **Nanopore sequencing**

Using membrane lined protein nanopore, the individual nucleotides (one base at a time) will be identified by means of variation in the ion current [24]. The available commercial platform is GridION and miniaturized MinION sequencers from Oxford Nanopore Technologies (Nanopore sequencing).

Operation cost, sequence read length and specific model based error are the major differences between these systems [25 – 27]. However, the final results proved that, the sequencing data generated from these methods are similar. Hence, depends upon the research goal, the researcher can choose a suitable sequencing methods. These technologies are well applied in various plant genomic researches. For example, by exon sequencing method is well suited to study the quantitative trait loci (QTL) / genetic transmission of alleles. Further, it also facilitate to understand the biodiversity, to investigate the interactions between host – pathogens, to predict the crops evolution, to tests the genetic markers inheritance and to identify the genes involved in symbiotic systems [16]. Similarly, the single base sequencing technique provides epigenome of *A. thaliana* i.e. there is a relationship between the cytosine methylation and their abundance of sRNA targets [28]. Other important application of these high through – put and inexpensive methodsof sequencing technologies is genotyping which is especially used in the construction of genomic map and to identify the single nucleotide polymorphism (SNPs) [29]. For example, GBS analysis of 2815 maize inbred, resulted that the 6, 81,257 SNP markers were in positive association with trait related genes [30].With these novel genomic sequencing technologies, there is a rapid growth in the genomic data availability. The development of numbers of databases will effectively help the scientists to store, handle and analyze these huge data.

1. **Applications of Plant Genomes**

Drastic changes in the history of plant genomics has been observed with the advent in the field of high-throughput gene sequencing technologies. Following the first release of complete genomic sequence of Arabidopsis thaliana in 2000, the Rice genome was successfully sequenced in 2002 [31, 32]. Merging of traditional breeding process and novel *insilico* or computational technologies resulted with the enhancement in the crops quality [33]. According to the new release (Release 7.1, April 2019) of Plant DNA-C values database, around 12,273 species genomes were successfully sequenced [34]. Knowledge on genes involved in regulatory mechanisms and genome dynamics in the whole plant, influencing factors in plant growth, plant community, ecosystem dynamics, evolutionary path and interactions between plant and other species are thoroughly analyzed [35 – 38]. The plant genomes used for various studies are listed in Table1.

**Table1: Examples for plant genomes used in various studies**

|  |  |  |
| --- | --- | --- |
| **S.No** | **Plant genomes** | **uses** |
|  | Arabidopsis [31] | Genomic models for drought tolerant plants [40] |
|  | Physcomitrella patens (moss) [39] |
|  | Setaria italic (millet) [41] |
|  | Brachypodiumdistachyon (grass) [42] |
|  | Oryza sativa (rice) [43] |
|  | Populustrichocarpa (poplar) [45] | Functional Models in crops [44] |
|  | Zea mays (maize) [46] |
|  | Utriculariagibba (bladderwort) [47] | Models to construct gene map, genome size variation and to analyze the evolutionary trajectories. [44] |

Similarly, genomes of many horticultural plants were sequenced which consists of vegetables, fruits and beverage producing plants, ornamental plants, medicinal herbs etc., These plants have great impact in human health and lives globally by satisfying basic food demands, decorating the urban and rural background and enhancing personal esthetics. It also plays a role in balancing and improving of our biological environment by means of giving oxygen [48] which are given in Table 2.

**Table 2: List of database resources for horticulture plants**

|  |  |  |
| --- | --- | --- |
| **Common name** | **Botanical name** | **URL** |
| Garden asparagus | *Asparagus officinalis* | phytozome.jgi.doe.gov |
| White Guinea yam | *Dioscorearotundata* | genomevolution.org/CoGe; plants.ensembl.org |
| Jiaobai | *Zizania latifolia* | plants.ensembl.org |
| Cassava | *Manihot esculenta* | [www.plantgdb.org](http://www.plantgdb.org); phytozome.jgi.doe.gov; plants.ensembl.org; genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza |
| Cucumber, Silver-seed gourd, Winter squash, Pumpkin, Summer squash | *Cucumissativus, Cucurbita argyrosperma, Cucurbita maxima, Cucurbita moschata, Cucurbita pepo* | cucurbitgenomics.org |
| Bottle gourd | *Lagenaria siceraria* | genomevolution.org; cucurbitgenomics.org |
| Pigeon pea | *Cajanuscajan* | brassicadb.org/brad; genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza |
| Chick pea | *Cicer arietinum; Cicerreticulatum* | genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza; www.coolseasonfoodlegume.org |
| *Glycine max* | Soybean | plants.ensembl.org; www.plantgdb.org |
| Barrel clover | *Medicagotruncatula* | plants.ensembl.org; www.plantgdb.org |
| Common bean | *Phaseolus vulgaris* | plants.ensembl.org; phytozome.jgi.doe.gov |
| Mung bean | *Vigna radiata* | plants.ensembl.org |
| Carrot | *Daucus carota* | bioinformatics.psb.ugent.be/plaza; plants.ensembl.org; phytozome.jgi.doe.gov |
| Juhuanao | *Chrysanthemum nankingense* | https://genomevolution.org/coge/; |
| Cardoon | *Cynara cardunculus* | www.artichokegenome.unito.it |
| Lettuce | *Lactuca sativa* | https://phytozome-next.jgi.doe.gov/ |
| *Brassica juncea* | Zhacai | http://brassicadb.cn/ |
| Cabbage | *Brassica oleracea* | http://brassicadb.cn/; https://genomevolution.org/coge/; https://bioinformatics.psb.ugent.be/plaza/ |
| Chinese cabbage | *Brassica rapa* | plants.ensembl.org; genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza |
| Shepherd’s purse | *Capsella bursa-pastoris* | genome.ccbr.utoronto.ca/cgi-bin/hgGateway |
| Red shepherd’s purse | *Capsella rubella* | genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza; |
| Radish | *Raphanus sativus* | radish.kazusa.or.jp |
| Field pennycress | *Thlaspiarvense* | pennycress.umn.edu |
| Moringa | *Moringa oleifera* | bioinformatics.psb.ugent.be/plaza |
| Sugar beet | *Beta vulgaris* | bioinformatics.psb.ugent.be/plaza |
| Spinach | *Spinacia oleracea* | spinachbase.org |
| Spanish pepper | *Capsicum annuum* | bioinformatics.psb.ugent.be/plaza |
| Berry-like pepper | *Capsicum baccatum* | genomevolution.org/CoGe |
| Bonnet pepper | *Capsicum chinense* | www.pepperpan.org:8012 |
| Tomato | *Solanum lycopersicum* | bioinformatics.psb.ugent.be/plaza; phytozome.jgi.doe.go |
| Eggplant | *Solanum melongena* | solgenomics.net; genomevolution.org/CoGe |
| Currant tomato | *Solanum pimpinellifolium* | solgenomics.net |
| Potato | *Solanum tuberosum* | www.plantgdb.org; phytozome.jgi.doe.gov; solgenomics.net |
| Garden asparagus | *Asparagus officinalis* | phytozome.jgi.doe.gov |
| White Guinea yam | *Dioscorearotundata* | genomevolution.org/CoGe; plants.ensembl.org |
| Coconut palm | *Cocos nucifera* | gigadb.org |
| Date palm | *Phoenix dactylifera* | drdb.big.ac.cn |
| Pineapple | *Ananas comosus* | phytozome.jgi.doe.gov; genomevolution.org/CoGe; pineapple.angiosperms.org |
| Yunnan banana | *Musa itinerans* | banana-genome-hub.southgreen.fr |
| Muskmelon | *Cucumis melo* | cucurbitgenomics.org/; bioinformatics.psb.ugent.be/plaza |
| Watermelon | *Citrulluslanatus* | www.coolseasonFoodlegume.org; bioinformatics.psb.ugent.be/plaza |
| Monk fruit | *Siraitiagrosvenorii* | herbalplant.ynau.edu.cn |
| Chinese chestnut | *Castaneamollissima* | genomevolution.org/CoGe |
| Eastern black walnut | *Juglans nigra* | www.hardwoodgenomics.org |
| Common walnut | *Juglans regia* | www.hardwoodgenomics.org |
| Iron walnut | *Juglans sigillata* | www.hardwoodgenomics.org |
| Chinese walnut | *Juglanscathayensis* | www.hardwoodgenomics.org |
| Northern California walnut | *Juglanshindsii* | www.hardwoodgenomics.org |
| Texas black walnut | *Juglansmicrocarpa* | www.hardwoodgenomics.org |
| Macadamia nut | *Macadamia integrifolia* | www.hardwoodgenomics.org |
| Breadnut | *Artocarpuscamansi* | sites.northwestern.edu/zerega-lab/research/artocarpus-genomics |
| Jujube | *Ziziphus jujuba* | genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza |
| Nogo strawberry, Japanese strawberry, Tibet strawberry, Eastern strawberry, Woodland strawberry, Strawberry | *Fragariaiinuma, Fragarianipponica, FragarianubicolaFragariaorientalis, Fragariavesca, Fragaria* × *ananassa* | strawberry-garden.kazusa.or.jp |
| Apple | *Malus domestica* | genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza |
| Mulberry | *Morus notabilis* | morus.swu.edu.cn |
| Sweet cherry | *Prunus avium* | www.rosaceae.org |
| Peach | *Prunus persica* | genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza |
| Chinese pear | *Pyrusbretschneideri* | bioinformatics.psb.ugent.be/plaza |
| European pear | *Pyrus communis* | www.rosaceae.org |
| Black raspberry | *Rubus occidentalis* | www.rosaceae.org |
| Papaya | *Carica papaya* | genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza |
| Kiwifruit | *Actinidia chinensis* | bdg.hfut.edu.cn/kir; genomevolution.org/coge |
| Blueberry | *Vacciniumcorymbosum* | www.vaccinium.org |
| Date-plum | *Diospyros lotus* | gigadb.org |
| American cranberry | *Vaccinium macrocarpon* | gigadb.org |
| Clementine citrus | *Citrus clementina* | genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza; phytozome.jgi.doe.gov |
| Pummelo, Ichangpapeda,*C*itrumelo, Mandarin orange, Sweet orange, Cold hardy mandarin, | *Citrus grandis, Citrus ichangensis, Citrus paradisi* × *Poncirus trifoliate, Citrus reticulate, Citrus sinensis, Citrus unshiu* | www.citrusgenomedb.org |
| Longan | *Dimocarpuslongan* | gigadb.org |
| Grape | *Vitis vinifera* | phytozome.jgi.doe.gov; genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza |
| Pomegranate | *Punica granatum* | [www.hardwoodgenomics.org](http://www.hardwoodgenomics.org/) |
| Tianma | *Gastrodiaelata* | herbalplant.ynau.edu.cn |
| Cockspur grass | *Echinochloa crus-galli* | horticulture.eplant.org |
| Ethiopian banana | *Enseteventricosum* | horticulture.eplant.org |
| Tall-stilt mangrove | *Rhizophoraapiculata* | genomevolution.org/coge |
| Monk fruit | *Siraitiagrosvenorii* | herbalplant.ynau.edu.cn |
| Chinese liquorice | *Glycyrrhizauralensis* | ngs-data-archive.psc.riken.jp |
| Red clover | *Trifolium pratense* | [http://www.cacaogenomedb.org](http://www.cacaogenomedb.org/); bioinformatics.psb.ugent.be/plaza |
| Plume poppy, Opium poppy, | *Macleayacordata, Papaver somniferum* | herbalplant.ynau.edu.cn; genomevolution.orgauth.iplantc.org |
| Colorado blue columbine | *Aquilegia coerulea* | genome.jgi.doe.gov; genomevolution.org/CoGe; phytozome.jgi.doe.gov |
| Hemp | *Cannabis sativa* | genome.ccbr.utoronto.ca |
| Caoyeshanhuangma | *Parasponiaandersonii* | [www.bioinformatics.nl/parasponia](http://www.bioinformatics.nl/parasponia) |
| Indian charcoal tree | *Tremaorientalis* | [www.bioinformatics.nl/parasponia](http://www.bioinformatics.nl/parasponia) |
| Asian ginseng, Sanchi ginseng | *Panax ginseng, Panaxnotoginseng* | herbalplant.ynau.edu.cn |
| Sweet wormwood | *Artemisia annua* | herbalplant.ynau.edu.cn |
| Horseweed | *Conyza canadensis* | genomevolution.org/CoGe |
| Chinese fleabane | *Erigeron breviscapus* | [www.ncbi.nlm.nih.gov/genome/?term=Eleusine+coracana](http://www.ncbi.nlm.nih.gov/genome/?term=Eleusine+coracana) |
| Maca | *Lepidiummeyenii* | maca.eplant.org |
| Common primrose | *Primula vulgaris* | phytozome.jgi.doe.gov |
| Madagascar periwinkle | *Catharanthus roseus* | genomevolution.org/CoGe |
| Horse mint | *Mentha longifolia* | phytozome.jgi.doe.gov |
| Holy basil | *Ocimum sanctum* | caps.ncbs.res.in/Ote |
| Sweet osmanthus | *Osmanthus fragrans* | sweetolive.eplant.org |
| Chang shuohuang ma | *Corchorusolitorius* | bioinformatics.psb.ugent.be/plaza |
| Jiu bing le | *Atalantiabuxifolia* | [www.citrusgenomedb.org](http://www.citrusgenomedb.org/) |
| Citron | *Citrus medica* | [www.citrusgenomedb.org](http://www.citrusgenomedb.org/) |
| Tibetan Rhodiola | *Rhodiolacrenulata* | gigadb.org |
| Prairie dodder | *Cuscutacampestris* | plabipd.de/project\_cuscuta2/start.ep |
| Umbrella liverwort | *Marchantia polymorpha* | bioinformatics.psb.ugent.be/plaza; phytozome.jgi.doe.gov |
| Ginkgo tree | *Ginkgo biloba* | gigadb.org/site/index |
| Jointfir | *Gnetummontanum* | genomevolution.org/coge |
| Resuscitation moss | *Selaginellalepidophylla* | plantgdb.org/SmGDB/ |
| Spikemoss, Little club moss | *Selaginellamoellendorffii, Selaginellatamariscina* | phytozome.jgi.doe.gov; genomevolution.org/CoGe |
| Japanese lawn grass, Manila grass, Mascarene grass | *Zoysia japonica, Zoysiamatrella, Zoysiapacifica* | zoysia.kazusa.or.jp |
| Aphrodite's phalaenopsis, Horse phalaenopsis | *Phalaenopsis Aphrodite, Phalaenopsis equestris* | genomevolution.org; chibba.agtec.uga.edu/duplication; orchidstra2.abrc.sinica.edu.tw |
| Perennial ryegrass | *Loliumperenne* | pgsb.helmholtz-muenchen.de |
| Wild banana | *Musa acuminate, Musa balbisiana* | chibba.agtec.uga.edu/duplication; plants.ensembl.org; banana-genome-hub.southgreen.fr |
| Chinese tulip tree | *Liriodendron chinense* | [www.hardwoodgenomics.org](http://www.hardwoodgenomics.org/) |
| Eastern redbud | *Cercis canadensis* | genomevolution.orgauth.iplantc.org |
| Narrow-leaved lupine | *Lupinusangustifolius* | plants.ensembl.org |
| Sensitive plant | *Mimosa pudica* | [www.medicagogenome.org](http://www.medicagogenome.org/) |
| Australian pine tree | *Casuarina equisetifolia* | hardwoodgenomics.org |
| Sacred lotus | *Nelumbo nucifera* | bioinformatics.psb.ugent.be/plaza |
| California poppy | *Eschscholziacalifornica* | eschscholzia.kazusa.or.jp |
| Mei | *Prunusmume* | genomevolution.org/CoGe |
| Yoshino cherry | *Prunusyedoensis* | [www.rosaceae.org](http://www.rosaceae.org/) |
| Damask rose, Chinese rose, Many-flowered rose, Chestnut rose | *Rosa* × *damascena, Rosa chinensis, Rosa multiflora, Rosa roxburghii* | [www.rosaceae.org](http://www.rosaceae.org/) |
| Spider flower | *Tarenayahassleriana* | genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza |
| Prince’s feather | *Amaranthushypochondriacus* | phytozome.jgi.doe.gov; genomevolution.org/CoGe; bioinformatics.psb.ugent.be/plaza |
| Saguaro cactus | *Carnegiea gigantea* | phytozome.jgi.doe.gov |
| Carnation | *Dianthus caryophyllus* | carnation.kazusa.or.jp |
| Happy tree | *Camptothecaacuminata* | [www.plantkingdomgdb.com](http://www.plantkingdomgdb.com/); genomevolution.org/CoGe |
| Cowslip | *Primula veris* | plantgenie.org |
| Pink trumpet tree | *Handroanthusimpetiginosus* | [www.hardwoodgenomics.org](http://www.hardwoodgenomics.org/) |
| Scarlet sage | *Salvia splendens* | gigadb.org |
| European ash | *Fraxinus excelsior* | [www.hardwoodgenomics.org](http://www.hardwoodgenomics.org/) |
| Seep monkeyflower | *Mimulusguttatus* | phytozome.jgi.doe.gov; [www.plantgdb.org](http://www.plantgdb.org/) |
| Lavender-scallops | *Kalanchoefedtschenkoi* | phytozome.jgi.doe.gov |
| Japanese morning glory | *Ipomoea nil* | viewer.shigen.info/asagao |
| Flowering tobacco | *Nicotiana sylvestris* | solgenomics.net |

The plant genomic sequence data enhances the process of recognition, classification and exploring of specific alleles as well as significant molecular markers were screened in order to enrich the plant breeding with desired features [49].

**Resources for Plant Genomic Research**

The historical backdrop of plant genomics has been changed decisively by the production of Expressed Sequence Tags (EST) sequencing, a high-throughput quality disclosure strategy [50] and the publication of *A.thaliana* and Rice whole genomic sequence data in 2000 and 2002 respectively. These occasions have made strong waves on both plant biotechnology and bioinformatics. For the headway of learning, additional sequencing projects on imperative plant species have been done by joining novel *in silico* advancements from genomic research with customary rearing plans for additional improving the nature of harvests. Over the years, these genes have provided many valuable tools for plant research in the era of modern molecular genomics. Based on these bases, the genetic/biological activities of many genes and significant pathways have been discovered. The availability of complete genome sequences, as well and the explosion of data systems, leading to accidents the need for thorough cataloging and specification of DNA sequence databases. The freely available sequencing projects, various plant genomic databases along with their URL were listed in Table 3.

|  |  |  |
| --- | --- | --- |
| **S.No** | **Sequencing projects** | **URL** |
| 1. | Tomato genome sequencing project [51] | www.sgn.cornell.edu/about/tomato |
| 2. | Potato genome sequencing consortium [52] | www.potato genome.net |
| 3. | Papaya genome sequencing project [53] | www.asgpb.mhpcc.hawaii.edu/papaya/ |
| 4. | Grape genome sequencing project[54] | www.vitaceae.org |
| 5. | Floral genome sequencing project[55] | www.fgp.bio.psu.edu/ |
| 6. | plant Genomic databases [56] | http://www.plantgdb.org/ |
| 7. | GOLD [57] | https://gold.jgi.doe.gov/ |
| 8. | NCBI Genomes [58] | http://www.ncbi.nlm.nih.gov/genome/ |
| 9. | NCBI Assembly – [59] | http://www.ncbi.nlm.nih.gov/assembly/ |
| 10. | CoGepedia (plant genomes) [60] | https://genomevolution.org |
| 11. | plaBi | http://plabipd.de/ |
| 12. | Ensembl Plants [60] | http://plants.ensembl.org/ |
| 13. | Gramene[61] | http://www.gramene.org/ |
| 14. | PlantsDB[62] | http://pgsb.helmholtz-muenchen.de/ |
| 15. | Phytozome[63] | http://phytozome.jgi.doe.gov/ |
| 16. | PLAZA [64] | http://bioinformatics.psb.ugent.be/plaza/ |
| 17. | GreenPhylDB | http://www.greenphyl.org/ |
| 18. | PlantOrDB [65] | http://bioinfolab.miamioh.edu/plantordb/ |
| 19. | SALAD [66] | http://salad.dna.affrc.go.jp/ |
| 20. | PlantTribes [67] | http://fgp.bio.psu.edu/tribedb/ |
| 21. | PlantGenIE.org [68] | http://plantgenie.org/ |
| 22. | POGs2 [69] | http://pogs.uoregon.edu/ |
| 23. | GenomicusPlants [70] | http://www.genomicus.biologie.ens.fr/ |
| 24. | PIECE [71] | http://wheat.pw.usda.gov/piece/ |
| 25. | PlantSEED [72] | http://bioseed.mcs.anl.gov/ |
| 26. | PGDBj[73, 74] | http://pgdbj.jp/ |
| 27. | ChloroplastDB [75] | http://chloroplast.cbio.psu.edu/ |
| 28. | KEGG [76] | http://www.genome.jp/kegg/ |
| **Plant species specialized genome databases** | | |
| 29. | TAIR (Arabidopsis) [77] | http://www.arabidopsis.org/ |
| 30. | SIGnAL (Arabidopsis) | http://signal.salk.edu/ |
| 31. | RARGE II(Arabidopsis) [78] | http://rarge.psc.riken.jp/ |
| 32. | RARGE II [79] | http://rarge-v2.psc.riken.jp/ |
| 33. | RGAP v.7 (Rice) | http://rice.plantbiology.msu.edu/ |
| 34. | RAP-DB (Rice) | http://rapdb.dna.affrc.go.jp/ |
| 35. | SGN (Solanaceae) | http://solgenomics.net/solanaceae-project/index.pl |
| 36. | GrainGenes (Triticeae and Avena) | http://wheat.pw.usda.gov/GG3/ |
| 37. | SoyBase (Soybean) | http://soybase.org/ |
| 38. | MaizeGDB (Maize) | http://www.maizegdb.org |
| 39. | CyanoBase (Cyanobacteria) | http://genome.microbedb.jp/cyanobase/ |
| 40. | GDR (Rosaceae) | https://www.rosaceae.org/ |
| 41. | Brassica Genome Gateway (Brassica) | http://brassica.nbi.ac.uk/ |
| 42. | Cucurbit Genomics Database (Cucurbitaceae) | http://www.icugi.org/cgi-bin/ICuGI/index.cgi |
| **Databases for comparative Genomics** | | |
| 44. | Golm transcriptome db | http://csbdb.mpimp-golm.mpg.de/csbdb/dbxp/ath/ath\_xpmgq.html |
| 45. | ATTED-II | http://atted.jp/ |
|  | Tools Resources for Plant Genome Analysis |  |
|  | Galaxy. | http://galaxyproject |
|  | Sanger institute | http://www.sanger.ac.uk/ |
|  | GSAP | http://www.broadinstitute.org/ |

Available information in these public domain genome sequencing projects are considered as precious resources for comparative, evolutionary and functional studies [80]. For example, numbers of stress responsive factors of plants like soybean, sorghum, barley, maize and wheat were correctly predicted by comparative genomic approach in which the stress responsive transcription factors of Arabidopsis and Rice were selected as reference plant genomes [81, 82]. Likewise, numbers of functional genes and novel biosynthetic genes were discovered and annotated in newly sequenced plants by comparative genomics and co expression analysis [83, 84].

**Plant Transcriptomic Databases**

In general, the characterization and prediction of plant gene function is considered to be a difficult task because of availability of limited number of model plants, complexity in cultivation and lack of plants genomic transformation information. But, the available plant genome sequencing transcriptome data derived from different experiments and experimental treatments that capture the gene expression in tissues, organs and development phases and many bioinformatics approaches are considered to be a powerful resources for plant research scientific community. Understanding of the plant genes and their expressions will enhance the process of selection of desired plant traits [85, 86] and functional annotation of a novel gene. These gene expression data has immense values which provides significant information on the active state of the gene. Further, the expression data are used to predict gene function by comparative analysis [87]. With an exponential growing rate of available plant gene expression data, databases with different functionalities were emerged as a valuable contribution to the scientific community. Though there are huge numbers of online gene expression databases are available, only those databases used to predict gene function with unique approaches are given in Table 4.

|  |  |  |
| --- | --- | --- |
| **S.No** | **Databases** | **URL** |
| **Databases for analysis of Expression** | | |
|  | Arabidopsis | http://www.ag.arizona.edu/microarray |
|  | Barley | http://www.barleybase.org |
|  | Rice | http://www.ricearray.org |
|  | Maize | http://www.maizearray.org |
|  | Soybean | http://soybeangenomics.cropsci.uiuc.edu |
|  | Tomato | http://ted.bti.cornell.edu |
|  | GENEVESTIGATOR | https://genevestigator.com/gv/start/start.jsp |
|  | ePlant by BAR | http://bar.utoronto.ca/ |
|  | CoNekT-Plants | https://evorepro.sbs.ntu.edu.sg/ |
|  | Expression Atlas | https://www.ebi.ac.uk/gxa/home |
| **Databases for analysis of Co-Expression** | | |
|  | Expression Angler by BAR | http://bar.utoronto.ca/ |
|  | CoNekT-Plants | https://evorepro.sbs.ntu.edu.sg/ |
|  | ATTED-II | https://atted.jp/ |
| **Databases for analysis of Differentially Expressed Genes (DGE)** | | |
|  | AtCAST | http://atpbsmd.yokohama-cu.ac.jp/cgi/atcast/home.cgi |
|  | Expression Atlas | https://www.ebi.ac.uk/gxa/home |
|  | GENEVESTIGATOR | https://genevestigator.com/gv/start/start.jsp |
| **Databases for analysis of Gene expression-specificity/stability** | | |
|  | Rice Expression Database | http://expression.ic4r.org/ |
|  | CoNekT-Plants | https://evorepro.sbs.ntu.edu.sg/ |
|  | GENEVESTIGATOR | https://genevestigator.com/gv/start/start.jsp |
| **Databases for analysis of Experiment correlation** | | |
|  | AtCAST | http://atpbsmd.yokohama-cu.ac.jp/cgi/atcast/home.cgi |

**Resources for Genome – wide association studies in Plants**

Genome-wide association studies (GWAS) are a popular method to link genetic variation to phenotypic variation and are therefore important for plant genetics. The increasing wealth of publicly available genomic sequence information for crop plants provides an unprecedented opportunity to use GWAS for the identification of genes controlling a plethora of agronomic traits. However, lack of technical expertise and IT infrastructure still hinders the ability of plant breeders to conduct GWAS independently. However, more than 1300 different genes were genotyped and 107 phenotyped with respect to 2, 50,000 single nucleotide polymorphisms (SNPs) in a classic model plant Arabidopsis thaliana [88] With this ground work, there were frequent success in performing GWAS on some other qualities of interest in Arabidopsis, such as tolerance in heavy metal, salt, flowering time and heat resistance etc., [89]. Likewise, in Rice, another significant crop, mapping efforts were carried out intensively to understand the genetic variation responsible for starch quality, grain size, stress tolerance with respect to their ancestors [90, 91] In addition, the GWAS studies were performed on some other important crops for example Maize and Soybean [92, 93]. For better breeding, in GWAS, estimation of epistatic i.e. gene – gene interactions and gene environment interactions is considered to be an important task. Currently, the use of new bioinformatics approach, i.e. mixed linear model approach is widely used in the detection of epistatic and gene – gene interactions in Rice and the obtained results were also remarkable [94]. Similarly, to manage and make use of genetic association studies in plants, a special resource namely GnpIS – ASSo is available. Currently, this database dealt with GWAS in Tomato and Maize. It provides tools to study the linkage between traits and markers and the results were graphically visualized with devoted plots. Once, the best marker assisted the trait was selected, their location on chromosomes, their neighboring genes etc., were easily studied [95]. However, there is a need to enhance the speed, memory capacity and application of statistical test to make accurate results while using GWAS assisted tools.

**CONCLUSIONS**

Recent advances in the application of bioinformatics for plants Genomes not only offer enormous potential for large-scale genomics research among plant species, but also offer several technical challenges. NGS technologies and platforms will make the plant genetic data will be abundant in the coming years. With these accessible genomic data, the development of tools for managing and analyzing this data are becoming more and more important. Indeed, there is more and more genome continuously established databases of plant species fusion with different methods of analysis. Comparative genomics, the analysis provides specific information on functional genes within and between plant species. Phylogenomic results show stronger evidence for evolution studies and the hypothesized function of genes in the plant. GWAS, what is currently used in plant research, successfully detects loci and allelic variations related to valuable traits. On the contrary, one of the main challenges facing plant genomics researchers is in the great demand for knowledge and expertise in bioinformatics and computer science in order to intensely manage and manipulate the results of growth of large-scale plant genomic data. Also, from the information on the high density genotype quickly exploited, high performance Phenotype is urgently necessary to provide plant genomic analysis in high resolution.

In summary, the recent wealth of plant genomic resources, coupled with advances in bioinformatics, they have enabled plants research scientists to achieve a essential and systematic understanding of economically important plants and plant processes, essential to advance in crop improvement. While these exciting results, remains a fundamental need for effective tools and methodologies to advance plant biotechnology, to deal with problems that are difficult to solve using current approaches, and to facilitate the translation of this innovative discovery of the knowledge to get better plant productivity.

**References**

G. Eason, B. Noble, and I.N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of

Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955. (references)

1. C. Butler, and E. Millstone, “Environmental challenges. *Atlas of food: who eats what, where and why?”* Earthscan, pp.16-18, 2003.
2. J. Mann, “Natural products in cancer chemotherapy: past, present and future”. Nat. Rev. Cancer,  *2*(2), pp.143-148, 2002.
3. Z. Lippman, A.V. Gendrel, M. Black, M.W. Vaughn, N. Dedhia, W. Richard McCombie, K. Lavine, V. Mittal, B. May, K.D. Kasschau, and J.C. Carrington, “Role of transposable elements in heterochromatin and epigenetic control”. Nature, vol. *430*(6998), pp.471-476, 2004.
4. P.J. Kersey, “Plant genome sequences: past, present, future”. Curr. Opin. Plant Biol., 2019, *48*, pp.1-8.
5. C. Feuillet, J.E. Leach, J. Rogers, P.S. Schnable, and K. Eversole, “Crop genome sequencing: lessons and rationales”. Trends plant sci, vol. 16(2), pp.77-88, 2011.
6. J.P. Hamilton, and C. Robin Buell,” Advances in plant genome sequencing”. Plant J, *70*(1), pp.177-190, 2012.
7. C.N. Hirsch, and C. Robin Buell, “Tapping the promise of genomics in species with complex, nonmodel genomes”. Annu. Rev. Plant Biol,  vol. 64, pp.89-110, 2013.
8. T.P. Michael, and R. VanBuren, “Progress, challenges and the future of crop genomes”. Curr. Opin. Plant biol, vol. *24*, pp.71-81, 2015.
9. A.P. Dhanapal, and M. Govindaraj, “Unlimited thirst for genome sequencing, data interpretation, and database usage in genomic era: the road towards fast-track crop plant improvement”. Genet Res, *2015*.
10. J. Raes, and P. Bork, “Molecular eco-systems biology: towards an understanding of community function”. Nat. Rev. Microbiol, vol.*6*(9), pp.693-699, 2008.
11. S.A. Langeveld, A.D.M. Van Mansfeld, P.D. Baas, H.S. Jansz, G.A. Van Arkel, and P.J.Weisbeek, Nucleotide sequence of the origin of replication in bacteriophage ΦX174 RF DNA. Nature, vol. *271*(5644), pp.417-420, 1978.
12. H. Swerdlow, S. Wu, H. Harke, and N.J.Dovichi, “Capillary gel electrophoresis for DNA sequencing: laser-induced fluorescence detection with the sheath flow cuvette”. J. Chromatogr. A, vol. *516*(1), pp.61-67, 1990.
13. T. Hunkapiller, R.J. Kaiser, B.F. Koop, and L. Hood, “Large-scale and automated DNA sequence determination”. Science,  vol. *254*(5028), pp.59-67, 1991.
14. J. Shendure, R.D.Mitra, C. Varma, and G.M. Church, “Advanced sequencing technologies: methods and goals”. Nat Rev. Genet,  vol. *5*(5), pp.335-344, 2004.
15. J. Shendure, and H. Ji, “Next-generation DNA sequencing”, Nat. Biotech, *26*(10), pp.1135-1145, 2008.
16. D. Singh, P.K. Singh, S. Chaudhary, K. Mehla, and S. Kumar, “Exome sequencing and advances in crop improvement”, Advances in genetics, *79*, pp.87-121, 2012.
17. A. Ahmadian, B. Gharizadeh, A.C. Gustafsson, F. Sterky, P. Nyrén, M. Uhlén, and J. Lundeberg, “Single-nucleotide polymorphism analysis by pyrosequencing”. Anal Biochem, vol. 280(1), pp.103-110, 2000.
18. C.W. Fuller, L.R. Middendorf, S.A. Benner, G.M. Church, T. Harris, X. Huang, S.B. Jovanovich, J.R. Nelson, J.A.Schloss, D.C. Schwartz, and D.V.Vezenov, “The challenges of sequencing by synthesis”. Nat Biotech,  *27*(11), pp.1013-1023, 2009.
19. M.A. Quail, M. Smith, P. Coupland, T.D. Otto, S.R. Harris, T.R. Connor, A. Bertoni, H.P. Swerdlow, and Y. Gu, “A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers”. BMC genomics,  vol. 13(1), pp.1-13, 2012.
20. S. Myllykangas, J. Buenrostro and H.P. Ji, “Overview of sequencing technology platforms”, Bioinformatics for high throughput sequencing, pp. 11-25, 2012.
21. K. Dorre, S. Brakmann, M. Brinkmeier, K.T. Han, K. Riebeseel, P. J. Schwille, Stephan, T. Wetzel, M. Lapczyna, M. Stuke, and R. Bader, Techniques for single molecule sequencing. *Bioimaging*, *5*(3), pp.139-152, 1997.
22. E.R. Mardis, “Next-generation DNA sequencing methods”. Annu. Rev Genomics Hum Genet, vol. *9*(1), pp.387-402, 2008.
23. A. Rhoads, and K.F. Au, “PacBio sequencing and its applications”. Genomics, proteomics & bioinformatics, vol. *13*(5), pp.278-289, 2015.
24. D. Branton, D.W. Deamer, A. Marziali, H. Bayley, S.A. Benner, T. Butler, M. Di Ventra, S. Garaj, A. Hibbs, X. Huang, and S.B. Jovanovich, “The potential and challenges of nanopore sequencing”. Nat Biotech, vol. *26*(10), pp.1146-1153, 2008.
25. M.L. Metzker, “Sequencing technologies—the next generation”. Nat Rev. Genet, vol. *11*(1), pp.31-46, 2010.
26. S. Suzuki, N. Ono, C. Furusawa, B.W. Ying and T. Yomo, “Comparison of sequence reads obtained from three next-generation sequencing platforms”. PloS one, vol. 6(5), p.e19534, 2011.
27. C. Luo, D. Tsementzi, N. Kyrpides, T. Read and K.T. Konstantinidis, “Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample” PloS one, 7(2), p.e30087, 2012.
28. R. Lister, R.C. O'Malley, J. Tonti-Filippini, B.D. Gregory, C.C. Berry, A.H. Millar, and J.R. Ecker, “Highly integrated single-base resolution maps of the epigenome in Arabidopsis”. Cell, vol. 133(3), pp.523-536, 2008.
29. T.M. Beissinger, C.N. Hirsch, R.S. Sekhon, J.M. Foerster Johnson, G. Muttoni, B. Vaillancourt, C.R. Buell, S.M. Kaeppler, and N. de Leon, “Marker density and read depth for genotyping populations using genotyping-by-sequencing”. Genet, vol. *193*(4), pp.1073-1081, 2013.
30. M.C. Romay, M.J. Millard, J.C.Glaubitz, Peiffer, J.A., Swarts, K.L., Casstevens, T.M., Elshire, R.J., Acharya, C.B., Mitchell, S.E., Flint-Garcia, S.A. and McMullen, M.D., 2013. Comprehensive genotyping of the USA national maize inbred seed bank. *Genome biology*, *14*(6), pp.1-18.
31. Kaul, S., Koo, H.L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L.J., Feldblyum, T., Nierman, W., Benito, M.I., Lin, X. and Town, C.D., 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, *408*(6814), pp.796-815.
32. Edwards, D. and Batley, J., 2004. Plant bioinformatics: from genome to phenome. *TRENDS in Biotechnology*, *22*(5), pp.232-237.
33. Metzker, M.L., 2005. Emerging technologies in DNA sequencing. *Genome research*, *15*(12), pp.1767-1776.
34. Pellicer, J. and Leitch, I.J., 2019. The Plant DNA C‐values database (release 7.1): an updated online repository of plant genome size data for comparative studies. *[" New Phytologist"]*.
35. Dodsworth, S., Leitch, A.R. and Leitch, I.J., 2015. Genome size diversity in angiosperms and its influence on gene space. *Current opinion in genetics & development*, *35*, pp.73-78.
36. Suda, J., Meyerson, L.A., Leitch, I.J. and Pyšek, P., 2015. The hidden side of plant invasions: the role of genome size. *New Phytologist*, *205*(3), pp.994-1007.
37. Simonin, K.A. and Roddy, A.B., 2018. Genome downsizing, physiological novelty, and the global dominance of flowering plants. *PLoS biology*, *16*(1), p.e2003706.
38. Guignard, M.S., Crawley, M.J., Kovalenko, D., Nichols, R.A., Trimmer, M., Leitch, A.R. and Leitch, I.J., 2019. Interactions between plant genome size, nutrients and herbivory by rabbits, molluscs and insects on a temperate grassland. *Proceedings of the Royal Society B*, *286*(1899), p.20182619.
39. Rensing, S.A., Lang, D., Zimmer, A.D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P.F., Lindquist, E.A., Kamisugi, Y. and Tanahashi, T., 2008. The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. *Science*, *319*(5859), pp.64-69.
40. Langridge, P. and Reynolds, M.P., 2015. Genomic tools to assist breeding for drought tolerance. *Current opinion in biotechnology*, *32*, pp.130-135.
41. Bennetzen, J.L., Schmutz, J., Wang, H., Percifield, R., Hawkins, J., Pontaroli, A.C., Estep, M., Feng, L., Vaughn, J.N., Grimwood, J. and Jenkins, J., 2012. Reference genome sequence of the model plant Setaria. *Nature biotechnology*, *30*(6), pp.555-561.
42. Vogel, J.P.; Garvin, D.F.; Mockler, T.C.; Schmutz, J.; Rokhsar, D.; Bevan, M.W.; Barry, K.; Lucas, S.; Harmon-Smith, M.; Lail, K. Genome sequencing and analysis of the model grass Brachypodiumdistachyon. Nature, 2010, 463 (7282), 763-768.
43. Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H. and Hadley, D., 2002. A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). *Science*, *296*(5565), pp.92-100.
44. Michael, T.P. and VanBuren, R., 2015. Progress, challenges and the future of crop genomes. *Current opinion in plant biology*, *24*, pp.71-81.
45. Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A. and Schein, J., 2006. The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). *science*, *313*(5793), pp.1596-1604.
46. Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T.A. and Minx, P., 2009. The B73 maize genome: complexity, diversity, and dynamics. *science*, *326*(5956), pp.1112-1115.
47. Leushkin, E.V., Sutormin, R.A., Nabieva, E.R., Penin, A.A., Kondrashov, A.S. and Logacheva, M.D., 2013. The miniature genome of a carnivorous plant Genliseaaurea contains a low number of genes and short non-coding sequences. *BMC genomics*, *14*(1), pp.1-11.
48. Chen, F., Song, Y., Li, X., Chen, J., Mo, L., Zhang, X., Lin, Z. and Zhang, L., 2019. Genome sequences of horticultural plants: past, present, and future. *Horticulture research*, *6*.
49. Gupta, P.K., Langridge, P. and Mir, R.R., 2010. Marker-assisted wheat breeding: present status and future possibilities. *Molecular Breeding*, *26*(2), pp.145-161.
50. Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F. and Kerlavage, A.R., 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. Science, 252(5013), pp.1651-1656.
51. Consortium, T.G. The tomato genome sequence provides insights into fleshy fruit evolution. Nature, 2012, 485 (7400), 635-641.
52. Consortium, P.G.S. Genome sequence and analysis of the tuber crop potato. Nature, 2011, 475 (7355), 189-195.
53. Yu, Q., Tong, E., Skelton, R.L., Bowers, J.E., Jones, M.R., Murray, J.E., Hou, S., Guan, P., Acob, R.A., Luo, M.C. and Moore, P.H., 2009. A physical map of the papaya genome with integrated genetic map and genome sequence. *BMC genomics*, *10*(1), pp.1-12.
54. Welter, L.J., Göktürk-Baydar, N., Akkurt, M., Maul, E., Eibach, R., Töpfer, R. and Zyprian, E.M., 2007. Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (Vitis vinifera L). *Molecular Breeding*, *20*(4), pp.359-374.
55. Soltis, D.E., Ma, H., Frohlich, M.W., Soltis, P.S., Albert, V.A., Oppenheimer, D.G., Altman, N.S., depamphilis, C. and Leebens-Mack, J., 2007. The floral genome: an evolutionary history of gene duplication and shifting patterns of gene expression. *Trends in plant science*, *12*(8), pp.358-367.
56. Dong, Q., Schlueter, S.D. and Brendel, V., 2004. PlantGDB, plant genome database and analysis tools. *Nucleic acids research*, *32*(suppl\_1), pp.D354-D359.
57. Reddy, T.B., Thomas, A.D., Stamatis, D., Bertsch, J., Isbandi, M., Jansson, J., Mallajosyula, J., Pagani, I., Lobos, E.A. and Kyrpides, N.C., 2015. The Genomes OnLine Database (GOLD) v. 5: a metadata management system based on a four level (meta) genome project classification. *Nucleic acids research*, *43*(D1), pp.D1099-D1106.
58. NCBI Resource Coordinators, Database resources of the National Center for Biotechnology Information. Nucleic Acids Res., 2015, 43 (Database issue), D6-17.
59. Kitts, P.A., Church, D.M., Thibaud-Nissen, F., Choi, J., Hem, V., Sapojnikov, V., Smith, R.G., Tatusova, T., Xiang, C., Zherikov, A. and DiCuccio, M., 2016. Assembly: a resource for assembled genomes at NCBI. *Nucleic acids research*, *44*(D1), pp.D73-D80.
60. Bolser, D., Staines, D.M., Pritchard, E. and Kersey, P., 2016. Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomics data. In *Plant bioinformatics* (pp. 115-140). Humana Press, New York, NY.
61. Monaco, M.K., Stein, J., Naithani, S., Wei, S., Dharmawardhana, P., Kumari, S., Amarasinghe, V., Youens-Clark, K., Thomason, J., Preece, J. and Pasternak, S., 2014. Gramene 2013: comparative plant genomics resources. *Nucleic acids research*, *42*(D1), pp.D1193-D1199.
62. Spannagl, M., Nussbaumer, T., Bader, K.C., Martis, M.M., Seidel, M., Kugler, K.G., Gundlach, H. and Mayer, K.F., 2016. PGSB PlantsDB: updates to the database framework for comparative plant genome research. *Nucleic acids research*, *44*(D1), pp.D1141-D1147.
63. Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. and Rokhsar, D.S., 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic acids research*, *40*(D1), pp.D1178-D1186.
64. Proost, S., Van Bel, M., Vaneechoutte, D., Van de Peer, Y., Inzé, D., Mueller-Roeber, B. and Vandepoele, K., 2015. PLAZA 3.0: an access point for plant comparative genomics. *Nucleic acids research*, *43*(D1), pp.D974-D981.
65. Li, L., Ji, G., Ye, C., Shu, C., Zhang, J. and Liang, C., 2015. PlantOrDB: A genome-wide ortholog database for land plants and green algae. *BMC plant biology*, *15*(1), pp.1-11.
66. Mihara, M., Itoh, T. and Izawa, T., 2010. SALAD database: a motif-based database of protein annotations for plant comparative genomics. *Nucleic Acids Research*, *38*(suppl\_1), pp.D835-D842.
67. Wall, P.K., Leebens-Mack, J., Müller, K.F., Field, D., Altman, N.S. and Depamphilis, C.W., 2007. PlantTribes: a gene and gene family resource for comparative genomics in plants. *Nucleic acids research*, *36*(suppl\_1), pp.D970-D976.
68. Sundell, D., Mannapperuma, C., Netotea, S., Delhomme, N., Lin, Y.C., Sjödin, A., Van de Peer, Y., Jansson, S., Hvidsten, T.R. and Street, N.R., 2015. The plant genome integrative explorer resource: PlantGenIE. org. *New Phytologist*, *208*(4), pp.1149-1156.
69. Tomcal, M., Stiffler, N. and Barkan, A., 2013. POGs2: a web portal to facilitate cross-species inferences about protein architecture and function in plants. *PLoS One*, *8*(12), p.e82569.
70. Louis, A., Murat, F., Salse, J. and Roest Crollius, H., 2015. GenomicusPlants: a web resource to study genome evolution in flowering plants. *Plant and Cell Physiology*, *56*(1), pp.e4-e4.
71. Wang, Y., You, F.M., Lazo, G.R., Luo, M.C., Thilmony, R., Gordon, S., Kianian, S.F. and Gu, Y.Q., 2013. PIECE: a database for plant gene structure comparison and evolution. *Nucleic acids research*, *41*(D1), pp.D1159-D1166.
72. Seaver, S.M., Gerdes, S., Frelin, O., Lerma-Ortiz, C., Bradbury, L.M., Zallot, R., Hasnain, G., Niehaus, T.D., El Yacoubi, B., Pasternak, S. and Olson, R., 2014. High-throughput comparison, functional annotation, and metabolic modeling of plant genomes using the PlantSEED resource. *Proceedings of the National Academy of Sciences*, *111*(26), pp.9645-9650.
73. Asamizu, E., Ichihara, H., Nakaya, A., Nakamura, Y., Hirakawa, H., Ishii, T., Tamura, T., Fukami-Kobayashi, K., Nakajima, Y. and Tabata, S., 2014. Plant Genome DataBase Japan (PGDBj): a portal website for the integration of plant genome-related databases. *Plant and Cell Physiology*, *55*(1), pp.e8-e8.
74. Nakaya, A., Ichihara, H., Asamizu, E., Shirasawa, S., Nakamura, Y., Tabata, S. and Hirakawa, H., 2017. Plant genome database japan (PGDBj). In *Plant Genomics Databases* (pp. 45-77). Humana Press, New York, NY.
75. Cui, L., Veeraraghavan, N., Richter, A., Wall, K., Jansen, R.K., Leebens-Mack, J., Makalowska, I. and depamphilis, C.W., 2006. ChloroplastDB: the chloroplast genome database. *Nucleic acids research*, *34*(suppl\_1), pp.D692-D696.
76. Kanehisa, M. and Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, *28*(1), pp.27-30.
77. Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M. and Karthikeyan, A.S., 2012. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic acids research*, *40*(D1), pp.D1202-D1210.
78. Sakurai, T., Satou, M., Akiyama, K., Iida, K., Seki, M., Kuromori, T., Ito, T., Konagaya, A., Toyoda, T. and Shinozaki, K., 2005. RARGE: a large-scale database of RIKEN Arabidopsis resources ranging from transcriptome to phenome. *Nucleic acids research*, *33*(suppl\_1), pp.D647-D650.
79. Akiyama, K., Kurotani, A., Iida, K., Kuromori, T., Shinozaki, K. and Sakurai, T., 2014. RARGE II: an integrated phenotype database of Arabidopsis mutant traits using a controlled vocabulary. *Plant and Cell Physiology*, *55*(1), pp.e4-e4.
80. Sonah, H., Deshmukh, R.K., Singh, V.P., Gupta, D.K., Singh, N.K. and Sharma, T.R., 2011. Genomic resources in horticultural crops: status, utility and challenges. *Biotechnology advances*, *29*(2), pp.199-209.
81. Tran, L.S.P. and Mochida, K., 2010. Identification and prediction of abiotic stress responsive transcription factors involved in abiotic stress signaling in soybean. Plant Signaling & Behavior, 5(3), pp.255-257.
82. Mochida, K.E.I.I.C.H.I., Yoshida, T.A.K.U.H.I.R.O., Sakurai, T.E.T.S.U.Y.A., Yamaguchi-Shinozaki, K.A.Z.U.K.O., Shinozaki, K.A.Z.U.O. and Tran, L.S.P., 2011. In silico analysis of transcription factor repertoires and prediction of stress-responsive transcription factors from six major gramineae plants. DNA research, 18(5), pp.321-332.
83. Usadel, B., Obayashi, T., Mutwil, M., Giorgi, F.M., Bassel, G.W., Tanimoto, M., Chow, A., Steinhauser, D., Persson, S. and Provart, N.J., 2009. Co‐expression tools for plant biology: opportunities for hypothesis generation and caveats. Plant, cell & environment, 32(12), pp.1633-1651.
84. Bradbury, L.M., Niehaus, T.D. and Hanson, A.D., 2013. Comparative genomics approaches to understanding and manipulating plant metabolism. Current opinion in biotechnology, 24(2), pp.278-284.
85. Nowicka, B., Ciura, J., Szymańska, R. and Kruk, J., 2018. Improving photosynthesis, plant productivity and abiotic stress tolerance–current trends and future perspectives. Journal of plant physiology, 231, pp.415-433.
86. Niazian, M., 2019. Application of genetics and biotechnology for improving medicinal plants. Planta, 249, pp.953-973.
87. Rhee, S.Y. and Mutwil, M., 2014. Towards revealing the functions of all genes in plants. Trends in plant science, 19(4), pp.212-221.
88. Atwell, S., Huang, Y.S., Vilhjálmsson, B.J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A.M., Hu, T.T. and Jiang, R., 2010. Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature, 465(7298), pp.627-631.
89. Li, Y., Huang, Y., Bergelson, J., Nordborg, M. and Borevitz, J.O., 2010. Association mapping of local climate-sensitive quantitative trait loci in Arabidopsis thaliana. Proceedings of the National Academy of Sciences, 107(49), pp.21199-21204.
90. Xuehui, H., Xinghua, W., Tao, S., Qiang, Z., Qi, F., Yan, Z., Canyang, L., Chuanrang, Z., Tingting, L., Zhiwu, Z. and Meng, L., 2010. Genome-wide association studies of 14 agronomic traits in rice landraces. Nature genetics, 42(11), pp.961-967.
91. Famoso, A.N., Zhao, K., Clark, R.T., Tung, C.W., Wright, M.H., Bustamante, C., Kochian, L.V. and McCouch, S.R., 2011. Genetic architecture of aluminum tolerance in rice (Oryza sativa) determined through genome-wide association analysis and QTL mapping. PLoS genetics, 7(8), p.e1002221.
92. Tian, F., Bradbury, P.J., Brown, P.J., Hung, H., Sun, Q., Flint-Garcia, S., Rocheford, T.R., McMullen, M.D., Holland, J.B. and Buckler, E.S., 2011. Genome-wide association study of leaf architecture in the maize nested association mapping population. Nature genetics, 43(2), pp.159-162.
93. Hwang, E.Y., Song, Q., Jia, G., Specht, J.E., Hyten, D.L., Costa, J. and Cregan, P.B., 2014. A genome-wide association study of seed protein and oil content in soybean. BMC genomics, 15, pp.1-12.
94. Xu, H., Jiang, B., Cao, Y., Zhang, Y., Zhan, X., Shen, X., Cheng, S., Lou, X. and Cao, L., 2015. Detection of epistatic and gene-environment interactions underlying three quality traits in rice using high-throughput genome-wide data. BioMed Research International, 2015.
95. Wang, J., Joshi, T., Valliyodan, B., Shi, H., Liang, Y., Nguyen, H.T., Zhang, J. and Xu, D., 2015. A Bayesian model for detection of high-order interactions among genetic variants in genome-wide association studies. Bmc Genomics, 16(1), pp.1-20.