UPDATES ON BIOINFORMATICS DATABASE RESOURCES FOR PLANT GENOMICS

Abstract

Advancement in genomic sequencing technologies causes a radical change in agricultural research. Currently, Plant biologists have simplest access to massive genome sequencing data to understand plant densities based on genetic deviation. Therefore, in contemporary plant genomic research, it is imperative to employ bioinformatics tools comprehensively to effectively manage and analyze omics data while maintaining a sympathetic approach. Numerous plant genomic databases have been acknowledged and are steadily growing in size in recent years. Simultaneously, bioinformatics-based analytical techniques have made significant advancements across various facets of plant genomic research. For example, prediction of evolutionary relationship, comparative genome based function prediction and whole genome wide association studies. However, the persistent challenge in plant genomic research remains the ongoing enhancement of IT infrastructure, including the development of high-capacity data storage and analysis software. This chapter highlights 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.

Authors

Dr. K.S. Abbiramy
Department of Zoology & Wild life Biology
A.V.C. College (Autonomous)
Mannampandal, Mayiladuthurai, India.

Dr. S. Sugunakala*
Department of Bioinformatics
A.V.C. College (Autonomous)
Mannampandal, Mayiladuthurai, India.
suguna_murali@hotmail.com

Dr. K. Thenmozhi
Department of Zoology & Wild life Biology
A.V.C. College (Autonomous)
Mannampandal, Mayiladuthurai, India

Ms. T. Akilandeswari
Department of Biotechnology A.V.C. College (Autonomous)
Mannampandal, Mayiladuthurai, India.

I. 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. The majority of plant genomes exhibit remarkable diversity, surpassing those of other biological kingdoms, with sizes ranging from approximately 10 megabytes to well over 100 gigabytes. Generally, plants play a crucial role in Earth's ecosystems as they serve as primary sources of nutrition, animal feed, medicines, and more [1, 2]. Additionally, they are frequently selected as model organisms for studying the role of epigenetics and heterochromatin elements [3]. An appreciative knowledge of variety of species genomes especially in plant genomes are primarily lagged, but have significantly improved with the advent of latest 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 advancement of automated sequencing techniques empowers researchers to delve into the molecular-level exploration of plant genetic material, a field commonly referred to 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 functions of a gene in a genome unambiguously. In the year 2000, the sequencing of the first plant genome, that of Arabidopsis thaliana, marked a significant milestone and 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, a growing number of reviews have concentrated on updating plant sequencing projects and the advancement of crop plant databases [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.

II. 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 HTS (High - Throughput Sequencing) & NGS (Next Generation Sequencing) technologies. They are extensively used 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 (SBS) such as pyro 454 sequencing &Illumina, Sequencing-By-Ligation (SBL) (i.e. SOLiD), Non Optical Sequencing (i.e. Ion Torrent Semi Conductor), SMRT (HeliScope Single Molecule Real Time), DNA Nanoball, and Nanopore Sequencing etc., [15, 16]. They are all well suited for the present and future large scale sequencing demands.

1. **Sequencing by Synthesis (SBS):**

   - **Pyro Sequencing:** Make use of DNA polymerase enzyme to lengthen various DNA fragments in parallel. In the process of nucleotide integration, the release of pyrophosphate leads to the detection of a light signal emitted due to the cleavage of oxyluciferin by the enzyme luciferase [17]. The commercially available platform is Genome Sequencer from Roche /454 pyrosequencing.

   - **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 technology consist of Miniseq series, Mi seq series, Nextseq series and Novaseq series [19].

2. **Sequencing-by-Ligation (SBL):**

   - **SOLiD Sequencing:** Starting in 2006, Applied Biosystems made this technology commercially accessible. DNA ligase is employed to enable consecutive ligation of dye-labeled oligonucleotides, facilitating the parallel sequencing of clonally amplified DNA strands [20]. The inconsistent nature of these amplified DNA helps to find out the concealed target sequence.

3. **Non Optical Sequencing:** Ion semiconductor sequencing involves the detection of hydrogen ions that are emitted during DNA polymerization. Commercially available sequencer is Ion Torrent from Life Science, Inc.

4. **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]. Commercial platforms currently accessible include those from Helicos Bioscience Corporation, which offers True single molecule sequencing, as well as the PacBio RS sequencers from Pacific Biosciences, providing Single molecule, real-time sequencing [23].

5. **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]. A commercial platform readily accessible is the GridION, along with the miniaturized MinION sequencers, both offered by Oxford Nanopore Technologies for 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 fast 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.

6. **Applications of Plant Genomes:** Drastic changes in the history of plant genomics has been observed with the advent in the field of high-throughput genome sequencing technologies. Following the first release of whole genomic sequence of Arabidopsis thaliana in 2000, the Rice whole 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]. In the latest update, as of April 2019 (Release 7.1) of the Plant DNA-C Values Database, successful sequencing has been reported for approximately 12,273 species genomes [34]. Knowledge on genes involved in regulatory mechanisms and genome dynamics in the whole plant, influencing factors in plant development, plant group, 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
1.	Arabidopsis [31]	Genomic models for drought tolerant plants [40]
2.	Physcomitrella patens (moss) [39]	
3.	Setaria italic (millet) [41]	
4.	Brachypodiumdistachyon (grass) [42]	
5.	Oryza sativa (rice) [43]	
6.	Populustrichocarpa (poplar) [45]	Functional Models in crops [44]
7.	Zea mays (maize) [46]	
8.	Utriculariagibba (bladderwort) [47]	Models to construct gene map, genome size variation and to analyze the evolutionary trajectories. [44]

Likewise, the genomes of numerous horticultural plants, encompassing vegetables, fruits, beverage-producing plants, ornamental plants, and medicinal herbs, have been successfully sequenced. 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 is given in Table 2.

Table 2: Listing of Database Resources for Horticulture Plants

S. No	Databases and their Website Address	Botanical Name	Common Name
1.	Phytozome 13 https://phytozome-next.jgi.doe.gov/	*Asparagus officinalis*	Garden asparagus
		Manihot esculenta	Cassava
		Phaseolus vulgaris	Common bean
		Manihot esculenta	Cassava
		Phaseolus vulgaris	Common bean
		Daucus carota	Carrot
		Lactuca sativa	Lettuce
		Solanum lycopersicum	Tomato
		Solanum tuberosum	Potato
		Asparagus officinalis	Garden asparagus
		Ananas comosus	Pineapple
		Citrus clementina	Clementine citrus
		Vitis vinifera	Grape
		Aquilegia coerulea	Colorado blue columbine
		Primula vulgaris	Common primrose
		Mentha longifolia	Horse mint
		Selaginellamoellendorffii,	Spikemoss
		Selaginellatamariscina	Little club moss
		Marchantia polymorpha	Umbrella liverwort
		Amaranthushypochondriacus	Prince’s feather
		Carnegiea gigantea	Saguaro cactus
		Mimulusguttatus	Seep monkeyflower
		Kalanchoefedtschenkoi	Lavender-scallops
2.	CoGe- https://genomevolution.org/coge/	*Dioscorearotundata*	White Guinea yam
		Manihot esculenta	Cassava
		Lagenaria siceraria	Bottle gourd
		Cajanuscajan	Pigeon pea
		Cicer arietinum; Cicerreticulatum	Chick pea
		Chrysanthemum nankingense	Juhuanao
		Brassica oleracea	Cabbage
		Brassica rapa	Chinese cabbage
		Capsella rubella	Red shepherd’s purse
		Capsicum baccatum	Berry-like pepper

		Solanum melongena	Eggplant
		Dioscorearotundata	White Guinea yam
		Ananas comosus	Pineapple
		Castaneamollissima	Chinese chestnut
		Ziziphus jujuba	Jujube
		Malus domestica	Apple
		Prunus persica	Peach
		Actinidia chinensis	Kiwifruit
		Citrus clementina	Clementine citrus
		Vitis vinifera	Grape
		Rhizophoraapiculata	Tall-stilt mangrove
		Aquilegia coerulea	Colorado blue columbine
		Conyza canadensis	Horseweed
		Catharanthus roseus	Madagascar periwinkle
		Gnetummontanum	Jointfir
		Phalaenopsis Aphrodite	Aphrodite's phalaenopsis
		Phalaenopsis equestris	Horse phalaenopsis
		Prunusmume	Mei
		Tarenayahassleriana	Spider flower
		Amaranthushypochondriacus	Prince’s feather
		Camptothecaacuminata	Happy tree
3.	Ensemble Plants - https://plants.ensembl.org/	*Dioscorearotundata*	White Guinea yam
		Zizania latifolia	Jiaobai
		Manihot esculenta	Cassava
		Glycine max	Soybean
		Medicagotruncatula	Barrel clover
		Phaseolus vulgaris	Common bean
		Vigna radiata	Mung bean
		Daucus carota	Carrot
		Brassica rapa	Chinese cabbage
		Dioscorearotundata	White Guinea yam
		Musa acuminate, Musa balbisiana	Wild banana
		Lupinusangustifolius	Narrow-leaved lupine
4.	Plant GDB - www.plantgdb.org;	*Manihot esculenta*	Cassava
		Glycine max	Soybean
		Medicagotruncatula	Barrel clover
		Solanum tuberosum	Potato
		Selaginellalepidophylla	Resuscitation moss
		Mimulusguttatus	Seep monkeyflower
5.	Plaza - https://bioinformatics.psb.ugent.be/plaza/	*Manihot esculenta*	Cassava
		Cajanuscajan	Pigeon pea
		Cicer arietinum; Cicerreticulatum	Chick pea

		Daucus carota	Carrot
		Brassica oleracea	Cabbage
		Brassica rapa	Chinese cabbage
		Capsella rubella	Red shepherd’s purse
		Moringa oleifera	Moringa
		Beta vulgaris	Sugar beet
		Capsicum annuum	Spanish pepper
		Solanum lycopersicum	Tomato
		Cucumis melo	Muskmelon
		Citrulluslanatus	Watermelon
		Ziziphus jujuba	Jujube
		Malus domestica	Apple
		Prunus persica	Peach
		Pyrusbretschneideri	Chinese pear
		Carica papaya	Papaya
		Citrus clementina	Clementine citrus
		Vitis vinifera	Grape
		Trifolium pratense	Red clover
		Corchorusolitorius	Chang shuohuang ma
		Marchantia polymorpha	Umbrella liverwort
		Nelumbo nucifera	Sacred lotus
		Tarenayahassleriana	Spider flower
		Amaranthushypochondriacus	Prince’s feather
6.	CuGenDB - http://cucurbitgenomics.org/	*Cucumissativus*	Cucumber
		Cucurbita argyrosperma	Silver-seed gourd
		Cucurbita maxima	Winter squash
		Cucurbita moschata	Pumpkin
		Cucurbita pepo	Summer squash
		Lagenaria siceraria	Bottle gourd
		Cucumis melo	Muskmelon
7.	Brassicaceae Database (BRAD) - http://brassicadb.org/brad	*Cajanuscajan*	Pigeon pea
		Brassica juncea	Zhacai
		Brassica oleracea	Cabbage
8.	globe artichoke genome - www.artichokegenome.unito.it	*Cynara cardunculus*	Cardoon
9.	C. sativa (Cannabis) Genome Browser Gateway- http://genome.ccbr.utoronto.ca/cgi-bin/hgGateway	*Capsella bursa-pastoris*	Shepherd’s purse
10.	Raphanus sativus Genome DataBase - http://radish.kazusa.or.jp/	*Raphanus sativus*	Radish

11.	Pennycress genomics - https://pennycress.umn.edu/	*Thlaspiarvense*	Field pennycress
12.	SpinachBase - http://spinachbase.org/	*Spinacia oleracea*	Spinach
13.	SolGenomics Network - https://solgenomics.net	*Solanum pimpinellifolium*	Currant tomato
		Solanum tuberosum	Potato
		Nicotiana sylvestris	Flowering tobacco
14.	GiGa DB - http://gigadb.org/	*Cocos nucifera*	Coconut palm
		Diospyros lotus	Date-plum
		Vaccinium macrocarpon	American cranberry
		Dimocarpuslongan	Longan
		Rhodiolacrenulata	Tibetan Rhodiola
		Ginkgo biloba	Ginkgo tree
		Salvia splendens	Scarlet sage
15.	Banana Genome Hub - https://banana-genome-hub.southgreen.fr/	*Musa itinerans*	Yunnan banana
		Musa acuminate, Musa balbisiana	Wild banana
16.	Herbalplant DB - http://herbalplant.ynau.edu.cn/	*Siraitiagrosvenorii*	Monk fruit
		Gastrodiaelata	Tianma
		Macleayacordata, ,	Plume poppy
		Papaver somniferum	Opium poppy
		Asian ginseng,	*Panax ginseng,*
		Sanchi ginseng	*Panaxnotoginseng*
		Artemisia annua	Sweet wormwood
17.	Hardwod Genomic DB - www.hardwoodgenomics. org	*Juglans nigra*	Eastern black walnut
		Juglans regia	Common walnut
		Juglans sigillata	Iron walnut
		Juglanscathayensis	Chinese walnut
		Juglanshindsii	Northern California walnut
		Juglansmicrocarpa	Texas black walnut
		Macadamia integrifolia	Macadamia nut
		Punica granatum	Pomegranate
		Liriodendron chinense	Chinese tulip tree
		Casuarina equisetifolia	Australian pine tree
		Handroanthusimpetiginosus	Pink trumpet tree
		Fraxinus excelsior	European ash
18.	*Artocarpus Genomics* - http://sites.northwestern.edu/zerega-lab/research/artocarpus-genomics	*Artocarpuscamansi*	Breadnut
19.	Strawberry GARDEN - https://strawberry-garden.kazusa.or.jp/	*Fragariaiinuma*	Nogo strawberry, , , , ,
		Fragarianipponica	Japanese strawberry
		Fragarianubicola	Tibet strawberry

		Fragariaorientalis	Eastern strawberry
		Fragariavesca	Woodland strawberry
		Fragaria × ananassa	Strawberry
20.	Morus DB – Mulberry Genome Database - https://morus.swu.edu.cn/	*Morus notabilis*	Mulberry
21.	GDR – Genome Database for Rosaceae - www.rosaceae.org	*Prunus avium*	Sweet cherry
		Pyrus communis	European pear
		Rubus occidentalis	Black raspberry
		Prunusyedoensis	Yoshino cherry
		Rosa × damascena	Damask rose
		Rosa chinensis	Chinese rose
		Rosa multiflora	Many-flowered rose
		Rosa roxburghii	Chestnut rose
22.	GDV – Genome Database for Vaccinum - www.vaccinium.org	*Vacciniumcorymbosum*	Blueberry
23.	Citrus Genome Databases - www.citrusgenomedb.org	*Citrus grandis*	Pummelo
		Citrus ichangensis	Ichangpapeda
		Citrus paradisi × Poncirus trifoliate	Citrumelo
		Citrus reticulate	Mandarin orange
		Citrus sinensis	Sweet orange
		Citrus unshiu	Cold hardy mandarin
		Atalantiabuxifolia	Jiu bingle
		Citrus medica	Citron
24.	Horticulture Plant Database - https://horticulture.eplant.org/	*Echinochloa crus-galli*	Cockspur grass
		Enseteventricosum	Ethiopian banana
25.	Next Generation Genome Database - https://ngs-data-archive.psc.riken.jp/	*Glycyrrhizauralensis*	Chinese liquorice
26.	Cacao Genome Daatabase - http://www.cacaogenomedb.org	*Trifolium pratense*	Red clover
27.	Portal for Parasponia Genomics Research - www.bioinformatics.nl/parasponia	*Parasponiaandersonii*	Caoyeshanhuangma
		Tremaorientalis	Indian charcoal tree
28.	GDPR - https://www.plabipd.de/	*Cuscutacampestris*	Prairie dodder
29.	Zoysia Genome Database - https://zoysia.kazusa.or.jp/	*Zoysia japonica*	Japanese lawn grass, ,
		Zoysiamatrella	Manila grass
		Zoysiapacifica	Mascarene grass

30.	Eschscholzia Genome DataBase - https://eschscholzia.kazusa.or.jp/	Eschscholziacalifornica	California poppy
31.	Carnation DB - https://carnation.kazusa.or.jp/	*Dianthus caryophyllus*	Carnation
32.	Orchidstra2.0 – A transcriptomic database for Orchid family - https://orchidstra2.abrc.sinica.edu.tw/	*Phalaenopsis Aphrodite*	Aphrodite's phalaenopsis
		Phalaenopsis equestris	Horse phalaenopsis
33.	Medica Go Genome - www.medicagogenome.org	*Mimosa pudica*	Sensitive plant
34.	Plant Genie Database - https://plantgenie.org/	*Primula veris*	Cowslip

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].

III. RESOURCES FOR PLANT GENOMIC RESEARCH

The landscape of plant genomics data has undergone a profound transformation due to the emergence of Expressed Sequence Tags (EST) sequencing, a high-throughput method for uncovering genetic information [50] and the publication of *A.thaliana* and Rice whole genomic sequence data in 2000 and 2002 respectively. These developments have left a lasting impact on both agricultural biotechnology and bioinformatics. In the pursuit of advancing our understanding, additional sequencing projects on vital plant species have been conducted by blending innovative *in-silico* techniques from genomic research with traditional breeding methods, further enhancing crop quality. Over time, these genes have offered invaluable tools for contemporary molecular genomics in plant research. As a result, numerous genes and important pathways' genetic and biological functions have been unveiled. With the availability of complete genome sequences and the proliferation of data systems, there has been an increasing demand for comprehensive cataloging and precise 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]	https://solgenomics.net/organism/Solanum_lycopersicum/genome
2.	Potato genome sequencing consortium [52]	https://solgenomics.net/organism/Solanum_tuberosum/genome
3.	Papaya genome sequencing project [53]	https://hawaiipapaya.com/

4.	Grape genome sequencing project[54]	http://www.grapegenomics.com/
5.	Floral genome sequencing project[55]	http://fgp.bio.psu.edu/
6.	plant Genomic databases [56]	http://www.plantgdb.org/
7.	Genomes Online Database - 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/
29.	The Arabidopsis Information Resource-TAIR (Arabidopsis) [77]	http://www.arabidopsis.org/
30.	Salk Institute Genomic Analysis Laboratory - SIGnAL (Arabidopsis)	http://signal.salk.edu/
31.	Riken Arabidopsis + Genome Encyclopedia RARGE II(Arabidopsis)[78]	http://rarge.psc.riken.jp/
32.	RARGE II [79]	http://rarge-v2.psc.riken.jp/
33.	Rice Genome Annotation Project - RGAP v.7 (Rice)	http://rice.uga.edu/
34.	Rice Annotation Project - RAP-DB (Rice)	http://rapdb.dna.affrc.go.jp/
35.	Sol Genomic Network (SGN) (Solanaceae)	https://solgenomics.net/solanaceae-project/index.pl
36.	GrainGenes - A Database for Triticeae and Avena	https://wheat.pw.usda.gov/GG3/
37.	SoyBase (Soybean)	http://soybase.org/
38.	MaizeGDB (Maize)	https://www.maizegdb.org/
39.	CyanoBase – A Database for Cyanobacteria	http://genome.microbedb.jp/cyanobase/

40.	Genome Database for Rosaceae - GDR	https://www.rosaceae.org/
41.	Brassica Genome Gateway	http://brassica.nbi.ac.uk/
42.	Cucurbit Genomics Database	http://www.icugi.org/cgi-bin/ICuGI/index.cgi
43.		
44.	Golm Metabolome Database	http://gmd.mpimp-golm.mpg.de/
45.	A Database for co regulated genes - ATTED-II	http://atted.jp/
46.	GoMapMan – Annotation of Gene functions	https://gomapman.nib.si/
47.	NPACT – Database for plant derived anticancer compounds	http://crdd.osdd.net/raghava/npact
48.	PGDD – Identification tools for synteny relationship intra genomes	http://chibba.agtec.uga.edu/duplication
49.	PlantRNA – Database consists of tRNA of photosynthetic Eukaryotes	http://seve.ibmp.unistra.fr/plantrna/
50.	PMRD: plant microRNA database	http://bioinformatics.cau.edu.cn/PMRD/
51.	SALAD - database for systematic comparison of proteome data	https://salad.dna.affrc.go.jp/salad/en/

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].

1. **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		
1	Arabidopsis	http://www.ag.arizona.edu/microarray
2	Barley	http://www.barleybase.org
3	Rice	http://www.ricearray.org
4	Maize	http://www.maizearray.org
5	Soybean	http://soybeangenomics.cropsci.uiuc.edu
6	Tomato	http://ted.bti.cornell.edu
7	GENEVESTIGATOR	https://genevestigator.com/gv/start/start.jsp
8	ePlant by BAR	http://bar.utoronto.ca/
9	CoNekT-Plants	https://evorepro.sbs.ntu.edu.sg/
10	Expression Atlas	https://www.ebi.ac.uk/gxa/home
Databases for analysis of Co-Expression		
11	Expression Angler by BAR	http://bar.utoronto.ca/
12	CoNekT-Plants	https://evorepro.sbs.ntu.edu.sg/
13	ATTED-II	https://atted.jp/
Databases for analysis of Differentially Expressed Genes (DGE)		
14	AtCAST	http://atpbsmd.yokohama-cu.ac.jp/cgi/atcast/home.cgi
15	Expression Atlas – Database of Gene expression crosswise species and biological condition	https://www.ebi.ac.uk/gxa/home
16	Genevestigator	https://genevestigator.com/gv/start/start.jsp
Databases for analysis of Gene expression-specificity/stability		
17	RED - Rice Expression Database	http://expression.ic4r.org/
18	CoNekT- Database for plant gene expression and co expression	https://evorepro.sbs.ntu.edu.sg/
19	Genevestigator	https://genevestigator.com/gv/start/start.jsp
Databases for analysis of Experiment correlation		
20	AtCAST – Tool for analysis of transcriptome data	http://atpbsmd.yokohama-cu.ac.jp/cgi/atcast/home.cgi

2. **Resources for Plant Genome – Wide Association Studies (PGWAS):** 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.

IV. CONCLUSIONS

Latest developments in bioinformatics application for plants Genomes offer enormous potential for large-scale genomics research among plant species along with several technical challenges. In near future, plant genetic data will be abundant with the application of NGS technologies and platforms. The tools development for managing and examining these data are becoming more and more highly significant with the accessible genomic data. Undoubtedly, continuous establishment of several of plant specific genome databases facilitates retrieval and analysis of plant genomes. For example, specific information on functional genes, their evolutionary relationships within and between plant species will be identified by comparative genome and phylogenetic analysis respectively. Likewise, from the genome wide association studies loci and allelic variations related to valuable traits will be detected successfully. However, the researchers those who are working in plant genomic analysis are still facing the great demand for managing and manipulating the results obtained from plant genomic data by using the available bioinformatics resources. In summing up, for essential advances in crop improvement, the recent collection of plant genomic data have enabled plants research scientists to accomplish an essential and systematic understanding of economically important plants and plant processes by utilizing advances in bioinformatics resources. Besides these exciting results, there is a basic need for efficient tools and methods to advance plant biotechnology, solve difficult problems to address with recent approaches, and smooth the progress of the implementation of this innovative discovery of knowledge to attain excellent yield from plants,

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”. Sci, 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”, Adv Genet, *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., “Comprehensive genotyping of the USA national maize inbred seed bank”, *Genome biology*, *14*(6), pp.1-18, 2013..

[31] S. Kaul, H.L. Koo, , J. Jenkins, M. Rizzo, T. Rooney, L.J. Tallon, T. Feldblyum, W. Nierman, M.I. Benito, X. Lin and C.D. Town, “Analysis of the genome sequence of the flowering plant Arabidopsis thaliana”, Nature, *408*(6814), pp.796-815, 2000.

[32] D. Edwards, D. and J. Batley, “Plant bioinformatics: from genome to phenome”, Trends Biotechnol, *22*(5), pp.232-237, 2004..

[33] M.L. Metzker, “Emerging technologies in DNA sequencing”, Genome res, *15*(12), pp.1767-1776, 2005.

[34] J. Pellicer, and I.J. Leitch, “The Plant DNA C□values database (release 7.1): an updated online repository of plant genome size data for comparative studies”. New Phytologist, 2019.

[35] S. Dodsworth, A.R. Leitch, and I.J. Leitch, “Genome size diversity in angiosperms and its influence on gene space”, Curr. Opin. Genet. Dev, *35*, pp.73-78, 2015.

[36] J. Suda, L.A. Meyerson, I.J. Leitch, and P. Pyšek, “The hidden side of plant invasions: the role of genome size”, New Phytol, *205*(3), pp.994-1007, 2015.

[37] K.A. Simonin, and A.B. Roddy, “Genome downsizing, physiological novelty, and the global dominance of flowering plants”, PLoS Biol, *16*(1), p.e2003706, 2018.

[38] M.S. Guignard, M.J. Crawley, D. Kovalenko, R.A, Nichols, M. Trimmer, A.R. Leitch, and I.J. Leitch, “Interactions between plant genome size, nutrients and herbivory by rabbits, molluscs and insects on a temperate grassland”, Proc. R. Soc. B, 286(1899), p.20182619, 2019.

[39] S.A. Rensing, D. Lang, A.D. Zimmer, A. Terry, A. Salamov, H. Shapiro, T. Nishiyama, P.F. Perroud, E.A. Lindquist, Y. Kamisugi, and T. Tanahashi, “The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants” *Science*, *319*(5859), pp.64-69, 2008.

[40] P. Langridge, and M.P. Reynolds, “Genomic tools to assist breeding for drought tolerance”, Curr. Opin. Biotechnol, *32*, pp.130-135, 2015.

[41] J.L. Bennetzen, J. Schmutz, H. Wang, R. Percifield, J. Hawkins, A.C. Pontaroli, M. Estep, L. Feng, J.N. Vaughn, J. Grimwood, and J. Jenkins, “Reference genome sequence of the model plant Setaria”, Nat. Biotechnol, *30*(6), pp.555-561, 2012.

[42] J.P.Vogel, D.F. Garvin, T.C. Mockler, J. Schmutz, D. Rokhsar, M.W. Bevan, K. Barry, S. Lucas, M. Harmon-Smith, K. Lail, “Genome sequencing and analysis of the model grass Brachypodiumdistachyon”, Nature, 463 (7282), 763-768, 2010.

[43] S.A. Goff, D. Ricke, T.H. Lan, G. Presting, R. Wang, M. Dunn, J. Glazebrook, A. Sessions, P. Oeller, H. Varma, and D. Hadley, “A draft sequence of the rice genome (Oryza sativa L. ssp. japonica)”, Science, 296(5565), pp.92-100, 2002.

[44] T.P. Michael, and R. VanBuren, “Progress, challenges and the future of crop genomes”, Curr. Opin. Plant. Biol, 24, pp.71-81, 2015.

[45] G.A.Tuskan, S. Difazio, S. Jansson, J. Bohlmann, I. Grigoriev, U. Hellsten, N. Putnam, S. Ralph, S. Rombauts, A. Salamov, and J. Schein, “The genome of black cottonwood, Populus trichocarpa (Torr. & Gray)”, Science, 313(5793), pp.1596-1604, 2006.

[46] P.S. Schnable, D. Ware, R.S. Fulton, J.C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton, T.A. Graves, and P. Minx, “The B73 maize genome: complexity, diversity, and dynamics”, Science, 326(5956), pp.1112-1115, 2009.

[47] E.V. Leushkin, R.A. Sutormin, E.R. Nabieva, A.A. Penin, A.S. Kondrashov, and M.D. Logacheva, “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, 2013.

[48] F. Chen, Y. Song, X. Li, J. Chen, L. Mo, X. Zhang, Z. Lin, and L. Zhang, “Genome sequences of horticultural plants: past, present, and future”, Hortic res, *6*, 2019.

[49] P.K. Gupta, P. Langridge, and R.R. Mir, “Marker-assisted wheat breeding: present status and future possibilities”, Mol. Breeding, 26(2), pp.145-161, 2010.

[50] M.D. Adams, J.M. Kelley, J.D. Gocayne, M. Dubnick, M.H. Polymeropoulos, H. Xiao, C.R. Merril, A. Wu, B. Olde, R.F. Moreno, and A.R, “Kerlavage, Complementary DNA sequencing: expressed sequence tags and human genome project”, Science, 252(5013), pp.1651-1656, 1991.

[51] T.G. Consortium, “The tomato genome sequence provides insights into fleshy fruit evolution”, Nature, 485 (7400), 635-641,2012.

[52] P.G.S. Consortium, “Genome sequence and analysis of the tuber crop potato”, Nature, 475 (7355), 189-195, 2011.

[53] Q. Yu, E. Tong, R.L. Skelton, J.E. Bowers, M.R. Jones, J.E. Murray, S. Hou, P. Guan, R.A. Acob, M.C. Luo, and P.H. Moore, “A physical map of the papaya genome with integrated genetic map and genome sequence”, BMC Genomics, 10(1), pp.1-12, 2009.

[54] L.J. Welter, N. Goktürk-Baydar, M. Akkurt, E. Maul, R. Eibach, R. Topfer, and E.M. Zyprian, “Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (Vitis vinifera L)”, Mol Breeding, 20(4), pp.359-374, 2007.

[55] D.E. Soltis, H. Ma, M.W. Frohlich, P.S. Soltis, V.A. Albert, D.G. Oppenheimer, N.S. Altman, C. Depamphilis, and J. Leebens-Mack, “The floral genome: an evolutionary history of gene duplication and shifting patterns of gene expression”, Trends. plant sci, 12(8), pp.358-367, 2007.

[56] Q. Dong, S.D. Schlueter, and V. Brendel, “PlantGDB, plant genome database and analysis tools”, Nucleic Acids Res, 32(suppl_1), pp.D354-D359, 2004.

[57] T.B. Reddy, A.D. Thomas, D. Stamatis, J. Bertsch, M. Isbandi, J. Jansson, J. Mallajosyula, I. Pagani, E.A. Lobos, and N.C. Kyrpides, “The Genomes OnLine Database (GOLD) v. 5: a metadata management system based on a four level (meta) genome project classification”, Nucleic Acids Res, 43(D1), pp.D1099-D1106, 2015.

[58] NCBI Resource Coordinators, Database resources of the National Center for Biotechnology Information. Nucleic Acids Res, 2015, 43 (Database issue), D6-17.

[59] P.A. Kitts, D.M. Church, F. Thibaud-Nissen, J. Choi, V. Hem, V. Sapojnikov, R.G. Smith, T. Tatusova, C. Xiang, A. Zherikov, and M. DiCuccio, “Assembly: a resource for assembled genomes at NCBI”, Nucleic Acids Res, 44(D1), pp.D73-D80, 2016.

[60] D. Bolser, D.M. Staines, E. Pritchard, and P. Kersey, “Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomics data”, J. Plant Bioinform, (pp. 115-140), 2016.

[61] M.K. Monaco, J. Stein, S. Naithani, S. Wei, P. Dharmawardhana, S. Kumari, V. Amarasinghe, K. Youens-Clark, J. Thomason, J. Preece, and S. Pasternak, “Gramene 2013: comparative plant genomics resources”, Nucleic Acids Res, 42(D1), pp.D1193-D1199.

[62] M. Spannagl, T. Nussbaumer, K.C. Bader, M.M. Martis, M. Seidel, K.G. Kugler, H. Gundlach, and K.F. Mayer, “PGSB PlantsDB: updates to the database framework for comparative plant genome research”, Nucleic Acids Res, 44(D1), pp.D1141-D1147, 2016.

[63] D.M. Goodstein, S. Shu, R. Howson, R. Neupane, R.D. Hayes, J. Fazo, T. Mitros, W. Dirks, U. Hellsten, N. Putnam, and D.S. Rokhsar, “Phytozome: a comparative platform for green plant genomics”, Nucleic Acids Res, 40(D1), pp.D1178-D1186, 2012.

[64] S. Proost, M. Van Bel, D. Vaneechoutte, Y. Van de Peer, D. Inzé, B. Mueller-Roeber, and K. Vandepoele, “PLAZA 3.0: an access point for plant comparative genomics”, Nucleic Acids Res, 43(D1), pp.D974-D981, 2015.

[65] L. Li, G. Ji, C. Ye, C. Shu, J. Zhang, and C. Liang, “PlantOrDB: A genome-wide ortholog database for land plants and green algae”, BMC Plant Biol, 15(1), pp.1-11, 2015.

[66] M. Mihara, T. Itoh, and T. Izawa, “SALAD database: a motif-based database of protein annotations for plant comparative genomics”, Nucleic Acids Res, 38(suppl_1), pp.D835-D842, 2010.

[67] P.K. Wall, J. Leebens-Mack, K.F. Müller, D. Field, N.S. Altman, and C.W. Depamphilis, “PlantTribes: a gene and gene family resource for comparative genomics in plants”, Nucleic Acids Res, 36(suppl_1), pp.D970-D976, 2007.

[68] D. Sundell, C. Mannapperuma, S. Netotea, N. Delhomme, Y.C. Lin, A. Sjodin, Y. Van de Peer, S. Jansson, T.R. Hvidsten, and N.R. Street, “The plant genome integrative explorer resource: PlantGenIE. Org”, New Phytol, 208(4), pp.1149-1156, 2015.

[69] M. Tomcal, N. Stiffler, and A. Barkan, “POGs2: a web portal to facilitate cross-species inferences about protein architecture and function in plants”, PLoS One, 8(12), p.e82569, 2013.

[70] A. Louis, F. Murat, J. Salse, and H. Roest Crollius, “GenomicusPlants: a web resource to study genome evolution in flowering plants”, Plant and Cell Physiol, 56(1), pp.e4-e4, 2015.

[71] Y. Wang, F.M. You, G.R. Lazo, M.C. Luo, R. Thilmony, S. Gordon, S.F. Kianian, and Y.Q. Gu, “PIECE: a database for plant gene structure comparison and evolution”, Nucleic Acids Res, 41(D1), pp.D1159-D1166, 2013.

[72] S.M. Seaver, S. Gerdes, O. Frelin, C. Lerma-Ortiz, L.M. Bradbury, R. Zallot, G. Hasnain, T.D. Niehaus, B. El Yacoubi, S. Pasternak, and R. Olson, “High-throughput comparison, functional annotation, and metabolic modeling of plant genomes using the PlantSEED resource”, Proc. Natl. Acad. Sci, 111(26), pp.9645-9650, 2014.

[73] E. Asamizu, H. Ichihara, A. Nakaya, Y. Nakamura, H. Hirakawa, T. Ishii, T. Tamura, K. Fukami-Kobayashi, Y. Nakajima, and S. Tabata, “Plant Genome DataBase Japan (PGDBj): a portal website for the integration of plant genome-related databases”, Plant Cell Physiol, 55(1), pp.e8-e8, 2014.

[74] A. Nakaya, H. Ichihara, E. Asamizu, S. Shirasawa, Y. Nakamura, S. Tabata, and H. Hirakawa, “Plant genome database japan (PGDBj)”, Plant Genomics Databases (pp. 45-77), 2017.

[75] L. Cui, N. Veeraraghavan, A. Richter, K. Wall, R.K. Jansen, J. Leebens-Mack, I. Makalowska, and C.W. depamphilis, “ChloroplastDB: the chloroplast genome database”, Nucleic Acids Res, 34(suppl_1), pp.D692-D696, 2006.

[76] M. Kanehisa, and S. Goto, “KEGG: kyoto encyclopedia of genes and genomes”, Nucleic Acids Res, 28(1), pp.27-30, 2000.

[77] P. Lamesch, T.Z. Berardini, D. Li, D. Swarbreck, C. Wilks, R. Sasidharan, R. Muller, K. Dreher, D.L. Alexander, M. Garcia-Hernandez, and A.S. Karthikeyan, “The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools”, Nucleic Acids Res, 40(D1), pp.D1202-D1210, 2012.

[78] T. Sakurai, M. Satou, K. Akiyama, K. Iida, M. Seki, T. Kuromori, T. Ito, A. Konagaya, T. Toyoda, and K. Shinozaki, “RARGE: a large-scale database of RIKEN Arabidopsis resources ranging from transcriptome to phenome”, Nucleic Acids Res, 33(suppl_1), pp.D647-D650, 2005.

[79] K. Akiyama, A. Kurotani, K. Iida, T. Kuromori, K. Shinozaki, and T. Sakurai, “RARGE II: an integrated phenotype database of Arabidopsis mutant traits using a controlled vocabulary”, Plant Cell Physiol, 55(1), pp.e4-e4, 2014.

[80] H. Sonah, R.K. Deshmukh, V.P. Singh, D.K. Gupta, N.K. Singh, and T.R. Sharma, “Genomic resources in horticultural crops: status, utility and challenges”, Biotechnol Adv, 29(2), pp.199-209, 2011.

[81] L.S.P. Tran, and K. Mochida, “Identification and prediction of abiotic stress responsive transcription factors involved in abiotic stress signaling in soybean”, Plant Signal Behav, 5(3), pp.255-257, 2010.

[82] K.E.I.I.C.H.I. Mochida, T.A.K.U.H.I.R.O. Yoshida, T.E.T.S.U.Y.A. Sakurai, K.A.Z.U.K.O. Yamaguchi-Shinozaki, K.A.Z.U.O. Shinozaki, and L.S.P. Tran, “In silico analysis of transcription factor repertoires and prediction of stress-responsive transcription factors from six major gramineae plants”, DNA Res, 18(5), pp.321-332, 2011.

[83] B. Usadel, T. Obayashi, M. Mutwil, F.M. Giorgi, G.W. Bassel, M. Tanimoto, A. Chow, D. Steinhauser, S. Persson, and N.J. Provart, “Co-expression tools for plant biology: opportunities for hypothesis generation and caveats”, Plant, Cell Environ, 32(12), pp.1633-1651, 2009.

[84] L.M. Bradbury, T.D. Niehaus, and A.D. Hanson, “Comparative genomics approaches to understanding and manipulating plant metabolism”, Curr. Opin. Biotechnol, 24(2), pp.278-284, 2013.

[85] B. Nowicka, J. Ciura, R. Szymańska, and J. Kruk, “Improving photosynthesis, plant productivity and abiotic stress tolerance–current trends and future perspectives”, J Plant Physiol, 231, pp.415-433, 2018.

[86] M. Niazian, “Application of genetics and biotechnology for improving medicinal plants”, Planta, 249, pp.953-973, 2019.

[87] S.Y. Rhee, and M. Mutwil, “Towards revealing the functions of all genes in plants”, . Trends plant Sci, 19(4), pp.212-221, 2014.

[88] S. Atwell, Y.S. Huang, B.J. Vilhjálmsson, G. Willems, M. Horton, Y. Li, D. Meng, A. Platt, A.M. Tarone, T.T. Hu, and R. Jiang, “Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines”, Nature, 465(7298), pp.627-631, 2010.

[89] Y. Li, Y. Huang, J. Bergelson, M. Nordborg, and J.O. Borevitz, “Association mapping of local climate-sensitive quantitative trait loci in Arabidopsis thaliana”, Proc Nat Acad Sci, 107(49), pp.21199-21204, 2010.

[90] H. Xuehui, W. Xinghua, S. Tao, Z. Qiang, F. Qi, Z. Yan, L. Canyang, Z. Chuanrang, L. Tingting, Z. Zhiwu, and L. Meng, “Genome-wide association studies of 14 agronomic traits in rice landraces”, Nat Genet, 42(11), pp.961-967, 2010.

[91] A.N. Famoso, K. Zhao, R.T. Clark, C.W. Tung, M.H. Wright, C. Bustamante, L.V. Kochian, and S.R. McCouch, “Genetic architecture of aluminum tolerance in rice (Oryza sativa) determined through genome-wide association analysis and QTL mapping”, PLoS Genet, 7(8), p.e1002221, 2011.

[92] F. Tian, P.J. Bradbury, P.J. Brown, H. Hung, Q. Sun, S. Flint-Garcia, T.R. Rocheford, M.D. McMullen, J.B. Holland, and E.S. “Buckler, Genome-wide association study of leaf architecture in the maize nested association mapping population”. Nat Genet, 43(2), pp.159-162, 2011.

[93] E.Y. Hwang, Q. Song, G. Jia, J.E. Specht, D.L. Hyten, J. Costa, and P.B. Cregan, “A genome-wide association study of seed protein and oil content in soybean”, BMC Genomics, 15, pp.1-12, 2014.

[94] H. Xu, B. Jiang, Y. Cao, Y. Zhang, X. Zhan, X. Shen, S. Cheng, X. Lou, and L. Cao, “Detection of epistatic and gene-environment interactions underlying three quality traits in rice using high-throughput genome-wide data”, BioMed Res Inter Nat, 2015.

[95] J. Wang, T. Joshi, B. Valliyodan, H. Shi, Y. Liang, H.T. Nguyen, J. Zhang, and D. Xu, “A Bayesian model for detection of high-order interactions among genetic variants in genome-wide association studies”, BMC Genomics, 16(1), pp.1-20, 2015.