Online Available at: https://www.sciplatform.com/index.php/wjb/article/view/906

ISSN (Online) = 2522-6754 ISSN (Print) = 2522-6746
WORLD JOURNAL OF BIOLOGY AND BIOTECHNOLOGY
Research Manuscript www.sciplatform.com

Application of modern biotechnology and bioinformatics approaches in agricultural sciences; A systematic review

[a] Muhammad Taqqi Abbas, [b] Zahra Batool, [a] Chaudhry Ali Ahmad, [c] Hafiz Azhar Ali Khan, [a] Adnan Akhter [*]

[a] Department of Plant Pathology, Faculty of Agricultural Sciences, University of the Punjab, Lahore Pakistan,
[b] Department of Food Sciences, Faculty of Agricultural Sciences, University of Punjab, Lahore Pakistan,
[c] Department of Entomology, Faculty of Agricultural Sciences, University of the Punjab, Lahore Pakistan.

Authors' Contribution **Akhter, A.** & **H. A. A. Khan** conceptualized the study; **M. T. Abbas, Z. Batool** & **C. A. Ahmad** performed literature survey; **A. Akhter** & **H. A. A. Khan** draft the literature review.

[*]Corresponding Author's Email Address: adnanakhter.iags@pu.edu.pk

ABSTRACT

Review Process: Peer review

In recent times, there has been a substantial surge in biological data owing to remarkable progress in molecular biology and genomic technology. With the growing demand for data storage and management tools, coupled with the surge in genetic data, there has been a corresponding rise in the availability of tools for analysing, visualizing, modelling, and predicting large datasets. The proliferation of diverse datasets derived from a multitude of plant species has significantly broadened our comprehension, particularly within the realm of plant biotechnology. In the realm of "omics" methodology, proficient bioinformatics methodologies and tools are presently under development to facilitate expedited genome sequencing and the comprehensive analysis of plant genomes. The present investigation centres on the diverse applications of plant biotechnology bioinformatics and their potential to augment agricultural productivity. This study aims to analyse and assess the obstacles or constraints encountered in the field of plant biotechnology, which have impeded the advancement of plant genomes at a slower pace compared to animal genomes. To successfully examine the intricate details of the plant genome, bioinformatics methods must be strong enough which provides longer coverage. The area in plant biotechnology including the agricultural industry are expected to advance from the development of bioinformatics, as would mankind as a whole. The inclusion of computer based programs and data analysis tools in plant biotechnology has facilitated the in depth study of crop plants important for commerce and trade. Improved algorithm development is necessary to facilitate tasks such as data mining, analysis, comparison, and etc. Hence, the integration of novel techniques and information into bioinformatics for the advancement of plant biotechnology, the agricultural industry, and the future of humanity will heavily rely on the expertise of Bioinformaticians.

Keywords: Biotechnology, bioinformatics, plant study, modern agriculture, CRISPER/cas9.

Keywords: Antibiotic susceptibility, bacterial diversity, biochemical tests, canal water, *E. coli* O157, latex agglutination.

INTRODUCTION: The term "bioinformatics" has become increasingly prevalent across various fields of biological research in recent times. The rapid advancements and progress in the field of molecular biology necessitated the need for a more efficient and systematic computer-based approach to gather, store, regulate, and scrutinize the copious amounts of biological data generated through research across various domains (Gomez-Casati *et al.*, 2018). Bioinformatics is inter-disciplinary field which integrates CS (computer science), biology, and mathematics to analyse and interpret genomes and proteomics data (Maloy and Hughes, 2013; Tiwari *et al.*, 2020). Bioinformatics is a field that involves the development of software tools and algorithms to facilitate the interpretation of biological data. Additionally, the collection and analysis of databases are fundamental components of this field (Maloy and Hughes, 2013). The significance of bioinformatics in various fields of biology is noteworthy due to its ability to generate diverse data types such as nucleotide and amino acid sequencing, protein domains and structures, and expression patterns from multiple organisms (Tiwari *et al.*, 2020). The field in question has effectively leveraged the tools and techniques of bioinformatics, resulting in notable advantages. Similarly, the realm of plant biotechnology has also reaped significant benefits from the application of bioinformatics. The field of bioinformatics offers a wealth of genomic information on a wide range of plant species, facilitating comprehensive investigations into plants as a valuable biological resource for human use (Rhee *et al.*, 2006; Gomez-Casati *et al.*, 2018; Tiwari *et al.*, 2020). However, there are a number of challenges and limitations that must be solved despite the potential advantages of employing bioinformatics in plant biotechnology in order to reach their full potential (Gomez-Casati *et al.*, 2018). Bioinformaticians and scientists deal with a wide range of problems, which can be divided into many categories as stated in the subsections below. These difficulties accompany the rapid development of plant genomic data mining and database construction.

OBJECTIVES: The main objective of the current article is to furnish a comprehensive outline of several fundamental principles, methodologies, and implementations of bioinformatics that hold significance in the realm of plant biotechnology. The present obstacles and constraints pertaining to the advancement and continual progression of bioinformatics in the domain of plant investigation are also underscored.

Bioinformatics uses in plant biotechnology: The combined use of computational and bioinformatics methodologies in the domain of plant biology has resulted in noteworthy progressions in scientific exploration in the realm of life sciences. The utilization of sequencing technologies by plant biologists has facilitated the revelation of the genetic structure of a wide range of plant and microbial species. This includes their proteome, transcriptome, metabolome, and metabolic pathway, as noted by (Gomez-Casati *et al.*, 2018). The technique of sequence analysis is a crucial tool utilized in modern research for the purpose of extracting the entire genome sequence of an organism, which includes the DNA, RNA, and protein sequence. The utilization of whole genome sequencing allows for the identification of a species' structure and serves as a fundamental means of understanding its operation. The presence of both coding and non-coding regions within a complete sequence is a crucial requirement for the expression of the functional gene which ultimately regulates the distinct traits of living organs. The acquired sequence frequently comprises a substantial quantity of genomic data, including exons, introns, regulators, and promoters, as stated by Normand and Van den Veyver (2019). The advancement of NGS (next-generation sequencing) with other omics Plant technologies has enabled the expansion of plant genomics research.

Tools and databases: In the area of bioinformatics, there exist numerous databases and tools that can be utilized for conducting analyses pertaining to plant biotechnology. The utilization of NGS and bio-informatics in the study of genomes has resulted in the production of a significant volume of data over the course of time. The entirety of this data is transmitted to diverse online databases that are accessible to the general public. Every database possesses unique characteristics and serves a specific purpose. The Cotton Gen database is focused on collecting genetic and breeding data related to relevant cotton species, as described by Yu *et al.* (2021). The creation of a database facilitates the research endeavours of scholars engaged in genomic studies of cotton by providing a centralized repository of information, thereby obviating the need to sift through multiple datasets.

Nonetheless, certain databases are designed with the aim of catering to all plants specie, like National Canter for Biotechnology Information (NCBI) database (https://www.ncbi.nlm.nih.gov/), which currently encompasses around 21,000 accessible plant genomes (Sayers *et al.*, 2010). This type of database proves to be advantageous for conducting research that does not concentrate solely on a particular genus or species. As a consequence, accessing various types of genomic data through a unified database has become more convenient for researchers. This segment will furnish a concise synopsis of the genome databases that is publicly reachable and not restricted to a specific genus or species.

The NCBI database, a widely recognized and esteemed resource among scholars and biologists, would be given priority. NBCI has made a pledge to collect and analyse data pertaining to the fields of

Digital Object Identifier: https://doi.org/10.33865/wjb.008.02.906

genetics, biochemistry, and molecular biology. To obtain the complete genome of a plant, one may input its scientific nomenclature in the designated search bar and retrieve the relevant information from the sequence read archive (SRA) or the gene expression omnibus (GEO) repositories, which can be accessed via https://www.ncbi.nlm.nih.gov/geo/. The NCBI database contains the entire genomic data of the plant (Dhivya *et al.*, 2022). The botanical varieties that have been archived in the GEO and SRA repositories contain data that may have undergone processing or remain unprocessed for the expression desire gene to retrieve the recent datasets associated with a specific accession number, one may input the scientific name of *Rosa chinense*, commonly known as the rose plant, on the search results page. The results that researchers receive may vary and depends on the platform used in each dataset for profiling (Tan *et al.*, 2022). The study's focus can be subjected to further analysis by researchers through the utilization of relevant bioinformatics tools, including but not limited to the Basic Local Alignment Search Tool (BLAST), Gene Ontology, and [(Database for Annotation), (Visualization) and (Integration Discovery) DAVID]. These tools can aid in the comprehensive examination of the study's subject matter.

Ensembl Plants is a supplementary tool that can be employed in the aggregation of plant genome databases, and it can be accessed at https://plants.ensembl.org/index.html. Ensembl Plants is a database that is designed to provide access to plant genomes, while the NCBI database has a broader scope beyond plant genomes. Ensembl Plant is a constituent of the Ensembl program, which was founded in 1999 with the aim of automating genome annotation, merging the results with other freely accessible data, and creating an online repository which benefit the scientific community (Howe *et al.*, 2020). Subsequently, the Ensembl initiative, which also encompasses plant taxa, implemented dedicated webpages for each taxon incorporated in their undertaking. The database, known as the integrative platform, has been developed with the aim of being easily accessible to users and is frequently refreshed with recently sequenced plant genomes to enhance its range. Ensembl Plant is a comprehensive database that offers gene models as well as information on phenotypes for the relevant plant species. Bolser *et al.* (2016) and Howe *et al.* (2020) have reported that this sets it apart from the NCBI database. Ensembl Plant provides comparative genomic data for relevant plant species, as opposed to NCBI. This demonstrates how the platform offers extensive information regarding the relevant plant species that goes beyond the genomic sequence data. The time-saving benefit of conducting the study is advantageous for researchers in the field of plant bioinformatics. The level of scrutiny applied to the data by researchers may depend on the degree of thoroughness demonstrated in their study. In addition to the databases described above, which are frequently used, there are still more plant databases that should be taken into account, including Plant GDB assembled Unique Transcripts are custom transcript assemblies made available for all plants species with >10,000 sequence records, Maize DIG (Maize Database of Images and Genomes), and Phytozome (The Plant Comparative Genomics Portal of the Department of Energy's Joint Genome Institute).

Plant breeding using biotechnology and bioinformatics: According to Caligari and Brown (2017), the practice of altering or improving favorable traits in plants with the aim of producing advanced agricultural cultivars (Hybrid rice, male sterile corn, Bt-Cotton and etc.) for the betterment of mankind is referred to as plant breeding. Rani and Usha (2013) identified several benefits of genetically modified plants, such as enhanced quality, augmented nutritional value, and amplified yield. The field of plant breeding has undergone significant advancements due to the molecular biology and genomics revolution. The application of transgenic technology in contemporary agriculture involves the genetic modification of crops or plants through the introduction or alteration of foreign genes, with the aim of enhancing their characteristics and augmenting their yield (Rani and Usha, 2013; Barragán-Ocaña *et al.*, 2019). As previously mentioned, the progress of NGS and other sequencing methods has led to the generation of a substantial quantity of biological data that necessitates storage in databases. With the widespread availability of complete genome sequences in databases, it has become feasible to establish linkages between genomes through gene sequence, likely function, or genetic map position without any cost. According to Maloy and Hughes (2013) and Platten *et al.* (2019), the utilization of software can facilitate the development of prediction hypotheses and enable the incorporation of desired phenotypes into plants through a complex combination. The attainment of this objective can be facilitated by directing attention towards genetic markers that demonstrate enhanced efficacy and furnish a higher level of breeding dependability. To understand the correlation between proteomics and genomes in reflecting changes in phenotype and the precise functioning of an organism, it is essential to employ databases that incorporate information on metabolites (Gomez-Casati *et al.*, 2018). Metlin and Plant Cyc are two noticeable data-bases for plants and crops metabolomics, as reported by Almeida Filho and Bruno (2016). Metlin, accessible at http://metlin.scripps.edu, contains a vast amount of information on approximately 240,000 metabolites and nearly 72,000 high-resolution MS/MS spectra and facilitates in multiple functions. On the other hand, Plant Cyc, available at https://plantcyc.org/, is a site that provides comprehensive data on biochemical pathways, their catalytic enzymes, and the genes involved in plants. The use of RNA sequencing (RNA-seq) as a means to directly analyze the mRNA profile for the purpose of identifying established single-nucleotide polymorphisms (SNPs) through the utilization of NGS technology.

SNPs, or single nucleotide polymorphisms, are unique allelic variations found in the same species' genome that may be employed as biological markers to find the genes responsible for desirable features in plants (Hoskins *et al.*, 2001; Mammadov *et al.*, 2012). Additionally, employing NGS for transcriptome resequencing makes it possible to quickly and affordably find SNPs in the genomes of big, complicated genes with a lot of highly repeated sections, such those in wheat, maize, sugarcane, avocado, and black currant (Mammadov *et al.*, 2012).

Management of plant stress (biotic and abiotic): Understanding how plants react to stress is essential for improving agricultural breeding efforts and predicting how naturally occurring plants would fare under the abiotic changes, particularly in the modern period of ongoing climate change (Qamar *et al.*, 2016; Hassan *et al.*, 2017; Khan *et al.*, 2017; Derbyshire *et al.*, 2022). Plant stress reactions can be classified as either biotic or abiotic. Contrarily, abiotic stress refers to environmental factors that have a substantial impact on crop productivity, such as high temperatures, drought, floods, salinity, and radiation (Derbyshire *et al.*, 2022). The term "biotic stress" primarily refers to the detrimental effects brought on by living things including weeds, bacteria, fungi, insects, and viruses (Singla *et al.*, 2016). Numerous researches on plants' molecular responses to stress have been conducted as a result of NGS technology and other powerful computational tools that enabled the sequencing of the full genome and transcriptome (Maloy and Hughes, 2013; Gomez-Casati *et al.*, 2018). The enormous quantity of information about plant genomes that has been gleaned through genome sequencing enables research into the relationships in molecular foundation of living thing and their environmental adaptations (Almeida Filho and Bruno, 2016). In order to minimize the adverse impact of harsh environmental conditions on agricultural production and ensure optimal growth and development of plants and crops, it is necessary to comprehend their response to stressors (Cho *et al.*, 2019; Costa and Farrant, 2019). Hence, it is imperative to ascertain the response of the plants to stress through the use of bioinformatics modern tools. The utilization of the bioinformatics techniques may prove advantageous for the agricultural sector, as it enables the identification of essential genes within the genomes of diverse species and the determination of their role in crop development (Cho *et al.*, 2019). The plant genome sequences are vast and complex, and are accessed through genome databases that play a critical role in their exploration. Apart from data storage, some genome databases conduct gene expression profiling to determine the transcript-level expression of genes in cells or tissues. Through the utilization of in silico genomic technologies, it becomes feasible to detect the gene-enzyme responsible for disease resistance and their corresponding transcription factor. These elements play a crucial role in bolstering the body's defense system against stress, as evidenced by research conducted by Nishad *et al.* (2020). (Xu *et al.*, 2013) conducted a study on dehydration stress in *chrysanthemum* plants by sequencing the complete transcriptome of the aforementioned plants. The *chrysanthemum* (http://www.icugi.org/chrysanthemum) transcriptome were established with the primary aim of enabling the preservation and dissemination of the sequence and its corresponding analysis outcomes among the research community. *Chrysanthemum's* metabolic pathway and kinase activity under dehydration-induced stress can be evaluated by utilizing different protein databases. In 2013, Xu and colleagues have identified a total of 228 protein kinases and 306 transcription factors that play a significant role as upstream regulators in plants that are exposed to various biotic and abiotic stresses.

Researching plant pathogen resistance using bioinformatics: The issue of crop loss resulting from disease poses a significant challenge to modern agriculture in its efforts to meet the food production demands of an expanding global population. In a study Nishad *et al.* (2020) says that the discipline of plant pathogens encompasses a wide range of subjects, including the identification of pathogens, the study of disease etiology, the development of disease resistance, and the economic consequences of plant diseases. Plants utilize a sophisticated defence system to protect themselves against a variety of ailments, but not limited to insects, bacteria, fungi, and viruses. Upon detection of noxious chemicals, the mechanism of signal transduction is triggered, which prompts the activation of diverse genes and pathways within the plant's immune system, culminating in a defensive reaction (Andersen *et al.*, 2018; Abid *et al.*, 2019). The progression of molecular plant pathology has been proposed to be divided into three parts by (Schneider and Collmer, 2010). According to Schneider and Collmer (2010), the first epoch began in the early 1900s and continued until the 1980s, focusing on the examination of disease physiology. The commencement of the 3rd era of plant research in 2000 marked with advent of genome sequencing, leading to the identification of bacterial *Xylella fastidiosa* comprehensive genome. Advancements in DNA sequencing start facilitating the examination of the plant immune system at the genomic and transcriptome levels by scholars (Andersen *et al.*, 2018), Gomez-Casati *et al.* (2018) and Nishad *et al.* (2020). The study of genomics has revealed the intricate and enigmatic nature of phytopathogens, while also providing a plethora of novel insights. The utilization of transcriptome and proteome data has facilitated the examination of plant-pathogen interactions, leading to the advancement of plant resistance to microbial infections (Dong and Ronald, 2019).

PRGAB: Plant Pathogen Resistance Gene Analysis Bioinformatics Web: The resistance (R) gene serves as a mediator in the plant defense system. The R gene plays a crucial role in the defense mechanism. The protein synthesized by the organism triggers a hypersensitive reaction (HR) by activating one or more signal transduction pathways (Convergence and Divergence pathways) upon encountering specific avirulent (Avr) pathogen proteins, as reported by Nishad *et al.* (2020), Sanseverino *et al.* (2012) and Sanseverino *et al.* (2010). The precise constituents necessary for proteins to manifest their resilience, nevertheless, remain unidentified (Sanseverino *et al.*, 2010). The sequencing of plant genomes and the utilization of high-throughput genomic investigations are indispensable in the pursuit of comprehending their functions and discovering fresh R genes, with the ultimate objective of examining and identifying further innovative R genes. In 2009, the Plant Disease Resistance Gene Database (PRGdb) was launched as a database that covers numerous plant species and help in bioinformatics. Its purpose is to aid plant genome research by facilitating the identification and prediction of plant disease resistance genes which has been noted in studies conducted by Sanseverino *et al.* (2012) and (Sanseverino *et al.*, 2010) . The PRGAB 3.0 database currently provides access to 177,072 annotated potential pathogen receptor genes (PRGs) and 153 reference resistance genes. The aforementioned database serves as a crucial archive for all scholarly articles pertaining to the examination and implementation of plant resistance genes, as evidenced by the works of Osuna-Cruz *et al.* (2018).

Apart from serving as a repository for resistance genes, this extensively accessible platform encompasses several research tools important for the study and identification of R genes. Disease Resistance Analysis and Gene Orthology (DRAGO) 2.0 program has been developed to analyze resistance genes, and can be utilized to initiate any proteome. It is capable of accurately annotating and predicting PRG from DNA or amino acids, as demonstrated by Osuna-Cruz *et al.* (2018). Moreover, the capacity of PRGdb's BLAST search functionalities to contrast diverse sequences facilitates the identification of gene homology and expression analysis. Whole genome sequencing technologies have proven to be advantageous to the field of plant pathology, in addition to the utilization of databases. The integration of advanced DNA sequencing techniques helps in examination of genomes of both plant and pathogen which facilitated the implementation of many inquiries associated with them (Gomez-Casati *et al.*, 2018). As per the findings of Gomez-Casati *et al.* (2018), the process of sequencing phytopathogen genomes is anticipated to provide significant insights into the molecular mechanisms that underlie the infection of plant hosts. Additionally, this process is expected to facilitate the exploration of potential novel virulence factors by researchers.

Cas9 modification and genomics in plant biotechnology: The population of environmental microbes, specifically soil microorganisms, may have an impact on plant pathogenesis and development. Metagenomics techniques have been utilized to gain genetic insights into the physiology and disease of soil microorganisms that facilitate plant growth (Hily *et al.*, 2018; Fadiji and Babalola, 2020). Metagenomics techniques involve sequencing the entire genetic material extracted from soil, followed by data analytics to analyze the microbial community (Chaudhary *et al.*, 2021; Chukwuneme *et al.*, 2021; Zhao *et al.*, 2021). The addition of bioinformatics methodologies in metagenomics investigations has the potential to yield benefits by augmenting our understanding of plant growth, plant pathology, soil pollution, and microbial classification, consequently facilitating the identification of specific genes (Piombo *et al.*, 2021). The utilization of Next-Generation Sequencing (NGS) technology in various studies results in the production of significant amounts of genomic data. Three examples of identification methods are the implementation of UNITE for fungi identification (https://unite.ut.ee//), UNITE is a rDNA sequence database designed to provide a stable and reliable platform for sequence-borne identification of all fungal species; SILVA for 16s rRNA documentation (https://www.arb-silva.de//), SILVA is an interdisciplinary project of biologists and computer scientists to provide: fully aligned and up to date small (16S/18S, SSU) and large (23S/28S, LSU) subunit ribosomal RNA "Parc" databases on the webpage as well as ARB files and MGnify for the examination of microbiome metagenomics data (https://www.ebi.ac.uk/metagenomics), MGnify is a free to use resource for analysis, visualisation and discovery of metagenomic, metatranscriptomic, amplicon and assembly datasets. The employment of databases facilitates the retrieval and examination of relevant metagenomic sequencing data for a particular research inquiry by researchers. The use of metagenomics analysis has been shown to provide a greater level of knowledge into the interactions that occur between plants and microbes. Previous research (Musidlak *et al.*, 2014; Silva *et al.*, 2018) has revealed that genes related with plant immunity may play an important role in pathogen defense. As a result, it is fair to believe that these genes play an important role in pathogen defense. According to Feng *et al.* (2013) and Wada *et al.* (2020), the application of CRISPR gene editing technology has demonstrated potential in augmenting plant traits and imparting immunity against diseases via Cas9 modification. CRISPR/Cas9 technology is being used in the study of plant functional genomics, namely the interaction between plants and microbes. The use of Cas9 technology has accelerated gene editing by inducing a dual-stranded fracture to produce a mutant, resulting in a specific gene mutation, and finally conducting genome restoration. Several investigations, including those undertaken by Wada *et al.* (2020), Nekrasov *et al.* (2013) and Langner *et al.* (2018), have fully documented the approach in question. Zafar *et al.* (2020) used CRISPR technology to modify the OsSWEET14 genes to prevent Super Basmati Rice from *Xanthomonas oryzae pv. oryzae*. Xie and Yang (2013) discovered that deleting the OsMPK5 and OsERF922 genes using gene editing techniques results in the acquisition of rice resistance in Magnaporthe grisea and Magnaporthe oryzae. Wang *et al.* (2019) and Fister *et al.* (2018) reported that by boosting Cas9 on CsWRKY22 and TcNPR3, the defensive immunity of *Citrus sinensis* and *Theobroma cacao* was boosted via salicylic acid modulation. The use of CRISPR/Cas9 editing has the potential to be a significant scientific development in correlating metagenomic research findings on the interconnection of flora and microbes.

Complex genetic makeup of plants: In addition to the vast amount of genomic sequence data generated, the plant research community must also confront the challenging issue of the complexity of plant genetic material. Despite the advent of next-generation sequencing technology, the rate at which plants sequence DNA remains comparatively slower than that of animals and microbes. This is noteworthy that such technology has facilitated the rapid sequencing of DNA in non-model or orphan plant species. A portion

Digital Object Identifier: https://doi.org/10.33865/wjb.008.02.906

of the plant genome may exhibit polyploidy, wherein the entire genome undergoes duplication, a phenomenon that is anticipated to occur in approximately 80% of plant species (Claros *et al.*, 2012). The underlying reason for this scenario is that, at times, the genetic material of a plant may be as much as 100 times greater in size when compared to the genetic material of recently analyzed animals and microorganisms (Schatz *et al.*, 2012). According to Kyriakidou *et al.* (2018) it is believed that a certain element of the plant genome is present in approximately 80% of plant species, and it is possible for this element to be polyploid, which involves a complete duplication of the genome. The main contributing factor for this phenomenon is the necessity for a distinct assembly methodology in Next-Generation Sequencing (NGS) due to the considerably shorter sequence length compared to Sanger sequencing, as noted by Claros *et al.* (2012).

Modernization of sequencing techniques: Two main methodologies for genome association, firstly comparative genome assembly and secondly de novo genome assembly perposed by (Kyriakidou *et al.*, 2018). It is necessary to comprehend the dissimilarities between these two methodologies. The process of comparative assembly involves utilizing a genome, transcriptome, or both as a guide for assembly, in contrast to de novo assembly, which involves reconstructing a genome from species whose genomes have not yet been sequenced, as stated by Claros *et al.* (2012). (Kyriakidou et al., 2018) suggest that the lack of modern tools tailored to the distinctive and complex characteristics of plant genomes means that these two methods are not entirely incompatible with each other. The development of bioinformatics software is impeded by the challenge of creating algorithms, as noted by Mathur (2018). The computational demands of bioinformatics applications and software are significant. All comparison databases, not just the individual genome database, should be updated. Coordination of genome data resource updates across multiple plant genomic platforms enables the construction of a robust, up-to-date, and dependable database community on which all plant researchers may rely (Ong *et al.*, 2016).

CONCLUSIONS: The way researchers approach studying living things has significantly changed as a result of the application of bioinformatics in Agriculture biotechnology. Understanding plant infections and stress tolerance is crucial for improving crop breeding; these topics can also be investigated using bioinformatics. Next-generation sequencing (NGS) and other sequencing technologies have facilitated the detection of genomic variants and the anticipation structure and function of the protein, increasing plant genome data that is currently accessible in various public databases (Kyriakidou *et al.*, 2018; Mathur, 2018). Additionally, GWAS has made crop modification and improvement easier by enabling the identification of loci and allelic variation linked with desired attributes (Claros *et al.*, 2012). In short, the advancement of bioinformatics use in plant biotechnology has made it possible to conduct a thorough and fundamental study of plants that are important for commerce. Despite the promising advances in the application of bioinformatics in plant biotechnology, achieving automatic whole genome sequencing and assembly at a reasonable cost remains a distant goal (Mathur, 2018). The ability to generate longer devours with unbiased coverage is crucial for bioinformatics tools to effectively navigate the complex structure of the plant genome. Improved algorithm development is necessary to facilitate tasks such as data mining, analysis, comparison, and related endeavours. Hence, the integration of novel techniques and information into bioinformatics for the advancement of plant biotechnology, the agricultural industry, and the future of humanity will heavily rely on the expertise of Bioinformaticians and specialists possessing mathematical and coding proficiencies.

CONFLICT OF INTEREST: The authors declare no conflict of interest.

ACKNOWLEDGMENTS: The authors are thankful to the University of the Punjab, for providing resources for this study.

REFERENCES: Abid, M., M. A. ullah Khan, S. Mushtaq, M. A. Rana, S. Afzaal, M. S. J. W. J. and B. Haider and Biotechnology, 2019. A review on future of baculaviruses as a microbial bio control agent. 4(3): 1-6.

Almeida Filho, H. and O. M. Bruno, 2016. A hierarchical model of the metabolic machinery based on the kcore decomposition of plant metabolic networks. Livro de Resumos.

Andersen, E. J., S. Ali, E. Byamukama, Y. Yen and M. P. Nepal, 2018. Disease resistance mechanisms in plants. Genes, 9(7): 339.

Barragán-Ocaña, A., G. Reyes-Ruiz, S. Olmos-Peña and H. Gómez-Viquez, 2019. Transgenic crops: Trends and dynamics in the world and in latin america. Transgenic research, 28: 391-399.

Bolser, D., D. M. Staines, E. Pritchard and P. Kersey, 2016. Ensembl plants: Integrating tools for visualizing, mining, and analyzing plant genomics data. Plant bioinformatics: Methods and protocols: 115-140.

Caligari, P. and J. Brown, 2017. Plant breeding, practice.

Chaudhary, P., P. Khati, A. Chaudhary, D. Maithani, G. Kumar and A. Sharma, 2021. Cultivable and metagenomic approach to study the combined impact of nanogypsum and pseudomonas taiwanensis on maize plant health and its rhizospheric microbiome. PLoS one, 16(4): e0250574.

Cho, K. T., J. L. Portwood, J. M. Gardiner, L. C. Harper, C. J. Lawrence-Dill, I. Friedberg and C. M. Andorf, 2019. Maizedig: Maize database of images and genomes. Frontiers in plant science, 10: 1050.

Chukwuneme, C. F., A. S. Ayangbenro and O. O. Babalola, 2021. Metagenomic analyses of plant growth-promoting and carbon-cycling genes in maize rhizosphere soils with distinct land-use and management histories. Genes, 12(9): 1431.

Claros, M. G., R. Bautista, D. Guerrero-Fernández, H. Benzerki, P. Seoane and N. Fernández-Pozo, 2012. Why assembling plant genome sequences is so challenging. Biology, 1(2): 439-459.

Costa, M.-C. D. and J. M. Farrant, 2019. Plant resistance to abiotic stresses. MDPI: pp: 553.

Derbyshire, M. C., J. Batley and D. Edwards, 2022. Use of multiple 'omics techniques to accelerate the breeding of abiotic stress tolerant crops. Current Plant Biology: 100262.

Dhivya, V., V. Abilash, N. Arul, C. Lalchhandama, V. Balachandar and N. S. Kumar, 2022. Big data in tribal healthcare and biomedical research. In: Big data applications in industry 4.0. Auerbach Publications: pp: 259-295.

Dong, O. X. and P. C. Ronald, 2019. Genetic engineering for disease resistance in plants: Recent progress and future perspectives. Plant physiology, 180(1): 26-38.

Fadiji, A. E. and O. O. Babalola, 2020. Metagenomics methods for the study of plant-associated microbial communities: A review. Journal of microbiological methods, 170: 105860.

Feng, Z., B. Zhang, W. Ding, X. Liu, D.-L. Yang, P. Wei, F. Cao, S. Zhu, F. Zhang and Y. Mao, 2013. Efficient genome editing in plants using a crispr/cas system. Cell research, 23(10): 1229-1232.

Fister, A. S., L. Landherr, S. N. Maximova and M. J. Guiltinan, 2018. Transient expression of crispr/cas9 machinery targeting tcnpr3 enhances defense response in theobroma cacao. Frontiers in plant science, 9: 268.

Gomez-Casati, D. F., M. V. Busi, J. Barchiesi, D. A. Peralta, N. Hedin and V. Bhadauria, 2018. Applications of bioinformatics to plant biotechnology. Current issues in molecular biology, 27(1): 89-104.

Hassan, Z. U., A. Komaki and B. Jafar, 2017. Virus-induced manipulation paradigms and alterations in host plants and their vectors. World journal of biology biotechnology, 2(1): 143-149.

Hily, J. M., S. Demaneche, N. Poulicard, M. Tannières, S. Djennane, M. Beuve, E. Vigne, G. Demangeat, V. Komar and C. Gertz, 2018. Metagenomic-based impact study of transgenic grapevine rootstock on its associated virome and soil bacteriome. Plant biotechnology journal, 16(1): 208-220.

Hoskins, J., W. E. Alborn Jr, J. Arnold, L. C. Blaszczak, S. Burgett, B. S. DeHoff, S. T. Estrem, L. Fritz, D.-J. Fu and W. Fuller, 2001. Genome of the bacterium streptococcus pneumoniae strain r6. Journal of bacteriology, 183(19): 5709-5717.

Howe, K. L., B. Contreras-Moreira, N. De Silva, G. Maslen, W. Akanni, J. Allen, J. Alvarez-Jarreta, M. Barba, D. M. Bolser and L. Cambell, 2020. Ensembl genomes 2020—enabling non-vertebrate genomic research. Nucleic acids research, 48(D1): D689-D695.

Khan, F. F., K. Ahmad, A. Ahmed, S. J. and B. Haider, 2017. Applications of biotechnology in agriculture-review article. World journal of biology and biotechnology, 2(1): 139-142.

Kyriakidou, M., H. H. Tai, N. L. Anglin, D. Ellis and M. V. Strömvik, 2018. Current strategies of polyploid plant genome sequence assembly. Frontiers in plant science, 9: 1660.

Langner, T., S. Kamoun and K. Belhaj, 2018. Crispr crops: Plant genome editing toward disease resistance. Annual review of phytopathology, 56: 479-512.

Maloy, S. and K. Hughes, 2013. Brenner's encyclopedia of genetics. Academic Press.

Mammadov, J., R. Aggarwal, R. Buyyarapu and S. Kumpatla, 2012. Snp markers and their impact on plant breeding. International

Digital Object Identifier: https://doi.org/10.33865/wjb.008.02.906

journal of plant genomics, 2012.

Mathur, M., 2018. Bioinformatics challenges: A review. Bioinformatics, 3(6):11-13.

Musidlak, O., W. Buchwald and R. Nawrot, 2014. Plant defense responses against viral and bacterial pathogen infections. Focus on rna-binding proteins (rbps). Herba polonica, 60(4):12-15.

Nekrasov, V., B. Staskawicz, D. Weigel, J. D. Jones and S. Kamoun, 2013. Targeted mutagenesis in the model plant nicotiana benthamiana using cas9 rna-guided endonuclease. Nature biotechnology, 31(8): 691-693.

Nishad, R., T. Ahmed, V. J. Rahman and A. Kareem, 2020. Modulation of plant defense system in response to microbial interactions. Frontiers in Microbiology, 11: 1298.

Normand, E. A. and I. B. Van den Veyver, 2019. Next-generation sequencing for gene panels and clinical exomes. Human Reproductive and Prenatal Genetics: 553-575.

Ong, Q., P. Nguyen, N. Phuong Thao and L. Le, 2016. Bioinformatics approach in plant genomic research. Current genomics, 17(4): 368-378.

Osuna-Cruz, C. M., A. Paytuvi-Gallart, A. Di Donato, V. Sundesha, G. Andolfo, R. Aiese Cigliano, W. Sanseverino and M. R. Ercolano, 2018. Prgdb 3.0: A comprehensive platform for prediction and analysis of plant disease resistance genes. Nucleic acids research, 46(D1): D1197-D1201.

Piombo, E., A. Abdelfattah, S. Droby, M. Wisniewski, D. Spadaro and L. Schena, 2021. Metagenomics approaches for the detection and surveillance of emerging and recurrent plant pathogens. Microorganisms, 9(1): 188.

Platten, J. D., J. N. Cobb and R. E. Zantua, 2019. Criteria for evaluating molecular markers: Comprehensive quality metrics to improve marker-assisted selection. PloS one, 14(1): e0210529.

Qamar, M. I., S. Batool, W. Aurangzeb, R. Zainab and S. Menghwar, 2016. Different techniques for diagnostic of potato viruses: A brief review. World journal of biology and biotechnology, 1(3): 123-128.

Rani, S. J. and R. Usha, 2013. Transgenic plants: Types, benefits, public concerns and future. Journal of Pharmacy Research, 6(8): 879-883.

Rhee, S. Y., J. Dickerson and D. Xu, 2006. Bioinformatics and its applications in plant biology. Annu. Rev. Plant Biol., 57: 335-360.

Sanseverino, W., A. Hermoso, R. D'Alessandro, A. Vlasova, G. Andolfo, L. Frusciante, E. Lowy, G. Roma and M. R. Ercolano, 2012. Prgdb 2.0: Towards a community-based database model for the analysis of r-genes in plants. Nucleic acids research, 41(D1): D1167-D1171.

Sanseverino, W., G. Roma, M. De Simone, L. Faino, S. Melito, E. Stupka, L. Frusciante and M. R. Ercolano, 2010. Prgdb: A bioinformatics platform for plant resistance gene analysis. Nucleic acids research, 38(suppl_1): D814-D821.

Sayers, E. W., T. Barrett, D. A. Benson, E. Bolton, S. H. Bryant, K. Canese, V. Chetvernin, D. M. Church, M. DiCuccio and S. Federhen, 2010. Database resources of the national center for biotechnology information. Nucleic acids research, 39(suppl_1): D38-D51.

Schatz, M. C., J. Witkowski and W. R. McCombie, 2012. Current challenges in de novo plant genome sequencing and assembly. Genome biology, 13(4): 1-7.

Schneider, D. J. and A. Collmer, 2010. Studying plant-pathogen interactions in the genomics era: Beyond molecular koch's postulates to systems biology. Annual review of phytopathology, 48: 457-479.

Silva, M. S., F. B. M. Arraes, M. de Araújo Campos, M. Grossi-de-Sa, D. Fernandez, E. de Souza Cândido, M. H. Cardoso, O. L. Franco and M. F. Grossi-de-Sa, 2018. Potential biotechnological assets related to plant immunity modulation applicable in engineering disease-resistant crops. Plant Science, 270: 72-84.

Singla, J., S. Krattinger and C. Wrigley, 2016. Encyclopedia of food grains.

Tan, Y. C., A. U. Kumar, Y. P. Wong and A. P. K. Ling, 2022. Bioinformatics approaches and applications in plant biotechnology. Journal of genetic engineering and biotechnology, 20(1): 1-13.

Tiwari, A., P. Singh and S. Kumawat, 2020. Applications of bioinformatics in plant breeding system. International jurnal of current microbialogical application in sciences, 11: 2825-2831.

Wada, N., R. Ueta, Y. Osakabe and K. Osakabe, 2020. Precision genome editing in plants: State-of-the-art in crispr/cas9-based genome engineering. BMC Plant Biology, 20: 1-12.

Wang, L., S. Chen, A. Peng, Z. Xie, Y. He and X. Zou, 2019. Crispr/cas9-mediated editing of cswrky22 reduces susceptibility to xanthomonas citri subsp. Citri in wanjincheng orange (*Citrus sinensis* (L.) osbeck). Plant Biotechnology Reports, 13: 501-510.

Xie, K. and Y. Yang, 2013. Rna-guided genome editing in plants using a crispr–cas system. Molecular plant, 6(6): 1975-1983.

Xu, Y., S. Gao, Y. Yang, M. Huang, L. Cheng, Q. Wei, Z. Fei, J. Gao and B. Hong, 2013. Transcriptome sequencing and whole genome expression profiling of chrysanthemum under dehydration stress. BMC genomics, 14: 1-15.

Yu, J., S. Jung, C.-H. Cheng, T. Lee, P. Zheng, K. Buble, J. Crabb, J. Humann, H. Hough and D. Jones, 2021. Cottongen: The community database for cotton genomics, genetics, and breeding research. Plants, 10(12): 2805.

Zafar, K., M. Z. Khan, I. Amin, Z. Mukhtar, S. Yasmin, M. Arif, K. Ejaz and S. Mansoor, 2020. Precise crispr-cas9 mediated genome editing in super basmati rice for resistance against bacterial blight by targeting the major susceptibility gene. Frontiers in plant science, 11: 575.

Zhao, J., J. Ma, Y. Yang, H. Yu, S. Zhang and F. Chen, 2021. Response of soil microbial community to vegetation reconstruction modes in mining areas of the loess plateau, china. Frontiers in microbiology, 12: 714967.