Int.J.Curr.Microbiol.App.Sci (2024) *13*(01): 87-93

International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 **Volume 13 Number 1 (2024)**
Journal homepage: http://www.ijcmas.com

Review Article

https://doi.org/10.20546/ijcmas.2024.1301.011

A Genomic Approach to Understand Plants - The Pandora Box yet to be Explored

Aleena Johny*

Department of Zoology, St. Aloysius College, Elthuruth, Thrissur – 680611, India

**Corresponding author*

Keywords

Bioinformatics, Genomics, Omics, producers, food, medicine

Article Info

Received:
25 November 2023
Accepted:
26 December 2023
Available Online:
10 January 2024

A B S T R A C T

The recent advances in plant sciences, in combination with bioinformatics and analogous development in tools, software and visualisation modelling, have steered the scientific community to an aggressive argumentation of information. Even after the emergence of Omics and several such splendid bioinformatics tools, a substantial number of researchers need to become more familiar with these tools. This review focuses on the applications of various *in silico* tools and technologies that may be adopted to analyse the plant sciences. The understanding of these different technologies shall lead us to a better understanding of plant properties such as stress tolerance, pathogenic resistance nutritional improvement. We are also converging thoughts on different trials and constraints in plant sciences in connection with the bioinformatics approach.

Introduction

Plants have always played an immensely important role through the inception and development of humankind being the producers. It provided food, medicine and sometimes protection to the human population (Vadakkan, 2019; Vadakkan *et al.,* 2019a, 2019b).

Several activities medicinal and non-medicinal such as anti-cancer (Suriyakala *et al.,* 2021), anti-helminthic (Vadakkan *et al.,* 2021), anti-microbial (Vadakkan *et al.,* 2018) and larvicidal, are investigated to date. As the importance of plants improved, so as plant studies which created a vast amount of data.

As an unprecedented amount of digital data is being produced because of recent hardware and technical developments that enable both large- and nano-scale investigation of biological samples. There is a high growth in the demand for computational methods to analyse and contextualise this ocean of data because the human brain cannot handle it (Mu *et al.,* 2022).

Bioinformatics is the analysis of biological data using approaches and principles from computer science, statistics, and engineering. It fits best within computational biology and biological information management (Bayat, 2002). However, in this modern era of computational biology, In addition to genomic sequence data analysis, bioinformatics is being used for a number of significant activities., such as the investigation of gene variation and expression, analysis and modelling of the composition and function of genes and proteins, identification and forecasting of gene regulatory networks, and simulation settings (Keerthana and Gochhait, 2022; Şahin *et al.,* 2022).

Understanding of an organism or its analysis shall be

done in various layers such as genomic, transcriptomic proteomic and finally, it's metabolomic profiling. Genomics is the study of the entirety or a portion of the genetic or epigenetic sequence information of organisms to comprehend the structure and function of these sequences and the biological products derived from them (Li *et al.,* 2020). It is widely assumed that transcriptomics, the analysis of all RNA molecules in a cell, is the continuation or updated version of genomic studies. In a parallel scenario, it should be emphasised that transcriptomics is the study of expressed genes, therefore they are interrelated (Supplitt *et al.,* 2021). This gets us to the final member of the molecular central dogma, proteins, and their study, proteomics, which is the broad study of proteomes (Kwon *et al.,* 2021). In this article, we discuss various bioinformatics tools that enable us to understand the molecular and genomic integrity of a plant. We are categorized various tools based on their influence in each "Omic" level.

Genomic studies- Understanding the predetermined 'fate'

Understanding the gene composition of an organism shall be considered one of the major milestones in plant biology. Since identifying the *Arabidopsis thaliana* genome in the early 20th century, several plant genomes have been identified and analysed (Boris *et al.,* 2011). The introduction of genomic analysis and bioinformatic tools revolutionized the dimensions of plant biology as it complimented the area by plant identification methods, phenotypical mutation studies, analysis of epigenetic variations and single nucleotide polymorphisms (Rhee *et al.,* 2006). Identifying plants is considered the cornerstone of all plant-related studies wherein the bioinformatic methods provide us with the upper hand over conventional methods. Genomic methods were introduced to a plant identification fraternity where plant identification was entirely dependent upon observation methods (Salazar Robles *et al.,* 2022). A study conducted by Wittich *et al.,* (2018), to identify plant taxa on-site suggested that the utilization of bioinformatics tools enhanced the resulting outcome by providing accuracy and reliability (Wittich *et al.,* 2018).

The process of genomics can be studied under various stages such as gene sequencing, sequence alignment and analysis of sequenced data (Cheng *et al.,* 2021; Henry, 2022). Even after developing high standards of genomic analysis generating and storing and characterizing the genomic data is yet to be improvised and made more compatible (Kress *et al.,* 2022). The biggest challenge faced by genomic scientists are the under par knowledge about transposons, ecological factors and lack of genotype-phenotype correlations (Pemmasani *et al.,* 2020). Once after these challenges are addressed then genomic data carries huge promise in developing healthcare and agricultural industries. The bioinformatic tools used in genomic analysis are illustrated in figure 1.

Gene sequencing – something to start with

Genomic sequencing consists of the procedures involved in identifying nucleotide composition of a genome (Posey, 2019). Initially, genomic studies were conducted by the employment of the Sangers sequencing method developed in 1977, it was widely considered to be among the first-generation sequencing technology (Heather and Chain, 2016). The technology functions with the principle of chain termination and each step involved in this methodology were revised and improvised in the course of time. One of the major improvements in this technology was the substitution of radio labelling with fluorometric-based detection and the introduction of capillary electrophoresis, both these changes facilitated the transformation of first-generation sequencing towards automation (Slatko *et al.,* 2018). The first-generation sequencing technology was successfully utilized for the identification of plants in the initial era and mainly consisted three steps thermal cycling, sample purification, and capillary electrophoresis (Blazej *et al.,* 2006). Arabidopsis (Initiative, 2000) and rice (Goff *et al.,* 2002) were among the first to be sequenced by standard Sanger sequencing of lined genomic fragments preserved in BAC vectors. The generalized plant genome sequence workflow consists of the extraction of DNA, sequencing, assembly of sequenced reads and chromosome assembly.

The second-generation sequencing techniques embarked themselves by generating large genomes at lower cost and are divided into two classes: sequencing by hybridization and synthesis. Some well-known examples of second-generation sequencing methods are 454 pyrosequencing, Ion Torrent, Illumina sequencing etc. (Edwards and Batley, 2010). The third-generation sequence mainly focuses on identifying the long nucleotide chains whereas the second-generation sequencing methods prefer short-chain sequencing and its assembly (Xiao and Zhou, 2020). The widely used third-generation sequencing platform is PacBio which is also sometimes mentioned as SMRT (Single Molecule Real Time). This strategy empowers the identification of

larger sequences that ranges from 30–50 kb, or longer (Roberts *et al.,* 2013) and also makes the hierarchical genome-assembly process (HGAP) possible (Chin *et al.,* 2013). Even though, all these mentioned sequencing methods have their own merits and shortcoming currently we have emerged the fourth generation sequencing technology that mainly comprises of nanopore technology which gives the scientific community the power to sequence hundreds of kb long chains (Mignardi and Nilsson, 2014). So far there are two different technologies are available in this category biological membrane systems and solid-state sensor technology. The commercialization of nanopore technology was catalysed by Oxford Nanopore Technologies (ONT) (Ying *et al.,* 2022) with a MinION, benchtop GridION, and a high throughput PromethION (He *et al.,* 2021; Jain *et al.,* 2016; Wang *et al.,* 2021).

To encapsulate the bioinformatics workflow of next-generation sequencing can be divided into primary secondary and tertiary levels whereas the primary level consists the recognition and evaluation of crude data, the secondary analysis composed of assembly of reads and the final stage comprises the annotation, prioritization and visualization (Pandey *et al.,* 2016; Pereira *et al.,* 2020).

Sequence alignment and analysis of sequenced genomic data

Once the row sequence data are generated the most popular and extensively used format to store nucleic acid sequences FASTA format, that generally comes with a".fasta" or ".fa" extensions (Pearson, 2016). One FASTA file will be composed of the name of sequence and the base pairs itself it could also contain various annotations such as sequence length (Eric *et al.,* 2014). Comparing and detecting correlations among biological sequences is the process of sequence alignment. The "similarities" discovered rely on the objectives of the alignment procedure. Counting the number of nucleotides that are identical in each of the sequences been evaluated is an easy option to reduce the complexity of analysing the sequences (Wang *et al.,* 2018). The score that displays how exactly two sequences matched to one another is alluded to as the alignment score of the sequences. This score reflects the degree to which two sequences are identical to one another. Distance between sequences is typically described to as the value that is the opposite of similarity among sequences and corresponds to the level of dissimilarity between sequences (Chao *et al.,* 2022). In other words, distance between sequences measures how dissimilar one sequence is to another. The total number of characters for which there is no match is referred to as the "Hamming distance" (Thompson *et al.,* 2011).

The Basic Local Alignment Search Tool, or BLAST, is with no doubt the most utilised sequence alignment tool in bioinformatics as this tool have been designed to overcome the constraints of the existing alignment algorithms, which are often employed for quite short sequences (Dash *et al.,* 2021; McGinnis and Madden, 2004; Pertsemlidis and Fondon, 2001). The tool do also have its variants such as BLASTN and Mega BLAST, which makes it possible to align both nucleic acid and amino acid sequences and this tool is available in NCBI website (Chen *et al.,* 2015). Despite what sequencing methodology, technique, or specimen was used to generate sequencing data, quality control is a vital component of every experiment. FastQC, which can be used to check the quality of diverse sequence data, is a highly crucial tool for ensuring the quality of a project. It has the benefit that it doesn't rely on any extra information, like a reference genome, and that it does quality control primarily using a FastQC file containing sequences and the quality values that correspond to them (Al Yami and Huang, 2019; Wingett and Andrews, 2018).

Assembly and annotation of sequenced genome

Sequence assembly is the process of aligning and merging fragments of a longer DNA sequence to reconstruct the sequence in its original form (Nagarajan and Pop, 2013). By aligning and combining fragments, DNA sequence assembly recreates the original DNA structure. Due to the inability of current sequencing technology to interpret the entire genome in a single step, this step is essential to genome analysis. Due to the presence of genomic repeats and numerous sequencing inaccuracies, none of the tools discovered during the sequencing period can be considered the optimal genome assembler in every circumstance (Prjibelski *et al.,* 2018). Currently, a sequenced genome is assembled using either long or short reads, although hybrid assembly is occasionally employed. Long read assembly employs the Overlap-Layout-Consensus (OLP) principle, which involves joining overlapping reads one by one (Marx, 2021). This method is divided into three stages: first, find all possible overlaps; second, create a layout for the overlaps; and third, use the overlaps to reconstruct the

sequence. Each stage builds on the one before it. Celera is a well-known long read assembly tool for genome reconstruction which is a first generation of assembler capable of assembling the genomes of multicellular organisms (Miller *et al.,* 2017).

However, given the abundance of short reads in NGS data, the OLC method seemed inconvenient. De Bruijn graph, which was proposed in 2001, effectively addressed the issue by making it possible to assemble a genome from scratch using short reads. The reads are broken up into a collection of sequences called k-mers, each of which has a defined length of k. These sequences serve as replacements for the vertices of the de Bruijn graph. Contrary to the construction of an overlap graph, the construction of a de Bruijn graph does not require a read alignment stage, which results in a considerable reduction in the amount of time required for the execution. Because each k-mer is only saved once, the de Bruijn graph approach appears to be more memory-efficient than the OLC technique. As a result, it is suitable for high-coverage datasets that contain millions of reads. (Chaisson and Pevzner, 2008). In addition, the introduction of new sequencing methods made hybrid assembly possible. Hybrid assembly is a way of assembling genomes that uses readings from different sequencing technologies all at once. (Brown *et al.,* 2021).

The prediction of protein-coding genes and other functional genomic units, such as structural RNAs, tRNAs, short RNAs, pseudogenes, regulatory regions, direct and inverted repeats, insertion sequences, transposons, and other mobile elements, are all part of the multi-level process known as genomic annotation. (Stein, 2001). It is necessary to employ a variety of methods in order to make predictions on the activities of genes; statistical gene prediction and the search of general-purpose databases for sequence similarity are essential components of every genome project. The utilisation of genomic context analysis, in addition to domain databases like Pfam, SMART, and CDD, and genome-oriented databases like COGs, KEGG, and WIT, varies greatly from one study to the next. Other genomic databases like WIT, KEGG, and COGs are also used (Salzberg, 2019).

Comparative genomics and Phylogenomics

The objective of comparative genomics is to compare two or more genomes to find their similarities and differences and to learn more about the biology of each genome, which can be accomplished at various levels of the genomes to acquire multiple viewpoints about the organisms (Wei *et al.,* 2002). The significance of genetic analysis increased with the advent of comparative genomics, which enables us to comprehend and predict the functions of multiple unknown genes when an organism is allowed to survive in various environments (Genereux *et al.,* 2020). To obtain this data, a number of bioinformatic tools, including GolmTranscriptome DB and ATTED-II, are being used.

Although there are some effective tools in the field, storing and making the data accessible to aspiring scientists was the main issue. The majority of the time, this problem is being solved by using online data platforms like Phytosome, PLAZA, Greenphyll DB, and Plants DB (Ong *et al.,* 2016).

Understanding phylogenomics, also known as Molecular Phylogenetic Analysis, which is frequently used to discover evolutionary relationships and Phylogenetic position of a plant species, is another crucial component of plant studies (Delsuc *et al.,* 2005). The primary goal of plant phylogenomics is to find evolutionary patterns and the secondary goal is to develop new hypotheses on unknown plant genes. To obtain phylogenomic data, a variety of bioinformatical methods are applied. Hyb-Seq is a potent method used to enrich organelle genome high copy repeats and their flanking regions as well as nuclear exons. ExaML is another such technology that can be productively utilised in plant phylogenetics (Theys *et al.,* 2019).

Author Contribution

Aleena Johny: Investigation, formal analysis, writing—original draft.

Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Research Funding: Not applicable

Ethical Approval: Not applicable.

Consent to Participate: Not applicable.

Consent to Publish: Not applicable.

Conflict of Interest: The author declare no competing interests.

References

Al Yami, S. and Huang, C.-H., 2019. LFastqC: A lossless non-reference-based FASTQ compressor. PLoS One 14, e0224806. https://doi.org/10.1371/journal.pone.0224806

Bayat, A., 2002. Science, medicine, and the future: Bioinformatics. BMJ 324, 1018–1022. https://doi.org/10.1136/bmj.324.7344.1018

Blazej, R. G., Kumaresan, P., Mathies, R. A., 2006. Microfabricated bioprocessor for integrated nanoliter-scale Sanger DNA sequencing. Proc. Natl. Acad. Sci. U. S. A. 103, 7240–7245. https://doi.org/10.1073/pnas.0602476103

Boris, K. V., Ryzhova, N. N., Kochieva, E. Z., 2011. Identification and characterization of intraspecific variability of the sucrose synthase gene Sus4 of potato (*Solanum tuberosum*). Russ. J. Genet. 47, 168–175. https://doi.org/10.1134/S1022795411020074

Brown, C.L., Keenum, I.M., Dai, D., Zhang, L., Vikesland, P.J., Pruden, A., 2021. Critical evaluation of short, long, and hybrid assembly for contextual analysis of antibiotic resistance genes in complex environmental metagenomes. Sci. Rep. 11, 3753. https://doi.org/10.1038/s41598-021-83081-8

Chaisson, M.J., Pevzner, P.A., 2008. Short read fragment assembly of bacterial genomes. Genome Res. 18, 324–330. https://doi.org/10.1101/gr.7088808

Chao, J., Tang, F., Xu, L., 2022. Developments in Algorithms for Sequence Alignment: A Review. Biomolecules 12. https://doi.org/10.3390/biom12040546

Chen, Y., Ye, W., Zhang, Y., Xu, Y., 2015. High speed BLASTN: an accelerated MegaBLAST search tool. Nucleic Acids Res. 43, 7762–7768. https://doi.org/10.1093/nar/gkv784

Cheng, Q.-Q., Ouyang, Y., Tang, Z.-Y., Lao, C.-C., Zhang, Y.-Y., Cheng, C.-S., Zhou, H., 2021. Review on the Development and Applications of Medicinal Plant Genomes. Front. Plant Sci. 12, 791219. https://doi.org/10.3389/fpls.2021.791219

Chin, C.-S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E. E., Turner, S.W., Korlach, J., 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat. Methods 10, 563–569. https://doi.org/10.1038/nmeth.2474

Dash, S., Rahman, S.R., Hines, H.M., Feng, W., 2021. iBLAST: Incremental BLAST of new sequences via automated e-value correction. PLoS One 16, e0249410. https://doi.org/10.1371/journal.pone.0249410

Delsuc, F., Brinkmann, H., Philippe, H., 2005. Phylogenomics and the reconstruction of the tree of life. Nat. Rev. Genet. 6, 361–375. https://doi.org/10.1038/nrg1603

Edwards, D., Batley, J., 2010. Plant genome sequencing: applications for crop improvement. Plant Biotechnol. J. 8, 2–9. https://doi.org/https://doi.org/10.1111/j.1467-7652.2009.00459.x

Eric, S.D., Nicholas, T.K.D.D., Theophilus, K.A., 2014. Bioinformatics with basic local alignment search tool (BLAST) and fast alignment (FASTA). J. Bioinforma. Seq. Anal. 6, 1–6. https://doi.org/10.5897/ijbc2013.0086

Genereux, D.P., Serres, A., Armstrong, J., Johnson, J., Marinescu, V.D., Murén, E., Juan, D., Bejerano, G., Casewell, N.R., Chemnick, L.G., Damas, J., Di Palma, F., Diekhans, M., Fiddes, I.T., Garber, M., Gladyshev, V.N., Goodman, L., Haerty, W., Houck, M.L., Hubley, R., Kivioja, T., Koepfli, K.-P., Kuderna, L.F.K., Lander, E.S., Meadows, J.R.S., Murphy, W.J., Nash, W., Noh, H.J., Nweeia, M., Pfenning, A.R., Pollard, K.S., Ray, D.A., Shapiro, B., Smit, A.F.A., Springer, M.S., Steiner, C.C., Swofford, R., Taipale, J., Teeling, E.C., Turner-Maier, J., Alfoldi, J., Birren, B., Ryder, O.A., Lewin, H.A., Paten, B., Marques-Bonet, T., Lindblad-Toh, K., Karlsson, E.K., Consortium, Z., 2020. A comparative genomics multitool for scientific discovery and conservation. Nature 587, 240–245. https://doi.org/10.1038/s41586-020-2876-6

Goff, S.A., Ricke, D., Lan, T.-H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller, R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A., Briggs, S., 2002. A Draft Sequence of the Rice Genome (Oryza sativa L. ssp. japonica). Science (80-.). 296, 92–100. https://doi.org/10.1126/science.1068275

He, M., Chi, X., Ren, J., 2021. Applications of Oxford Nanopore Sequencing in Schizosaccharomyces pombe. Methods Mol. Biol. 2196, 97–116. https://doi.org/10.1007/978-1-0716-0868-5_9

Heather, J.M., Chain, B., 2016. The sequence of sequencers: The history of sequencing DNA. Genomics 107, 1–8. https://doi.org/10.1016/j.ygeno.2015.11.003

Henry, R.J., 2022. Progress in Plant Genome Sequencing.

Appl. Biosci. 1, 113–128. https://doi.org/10.3390/applbiosci1020008

Initiative, T.A.G., 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408, 796–815. https://doi.org/10.1038/35048692

Jain, M., Olsen, H.E., Paten, B., Akeson, M., 2016. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol. 17, 239. https://doi.org/10.1186/s13059-016-1103-0

Keerthana, P., Gochhait, S., 2022. Application of Bioinformatics in Health Care and Medicine BT - Information Retrieval in Bioinformatics: A Practical Approach, in: Dutta, S., Gochhait, S. (Eds.),. Springer Nature Singapore, Singapore, pp. 83–99. https://doi.org/10.1007/978-981-19-6506-7_5

Kress, W.J., Soltis, D.E., Kersey, P.J., Wegrzyn, J.L., Leebens-Mack, J.H., Gostel, M.R., Liu, X., Soltis, P.S., 2022. Green plant genomes: What we know in an era of rapidly expanding opportunities. Proc. Natl. Acad. Sci. U. S. A. 119, 1–9. https://doi.org/10.1073/pnas.2115640118

Kwon, Y.W., Jo, H.S., Bae, S., Seo, Y., Song, P., Song, M., Yoon, J.H., 2021. Application of Proteomics in Cancer: Recent Trends and Approaches for Biomarkers Discovery. Front. Med. 8, 1–18. https://doi.org/10.3389/fmed.2021.747333

Li, H., Yang, Y., Hong, W., Huang, M., Wu, M., Zhao, X., 2020. Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects. Signal Transduct. Target. Ther. 5, 1. https://doi.org/10.1038/s41392-019-0089-y

Marx, V., 2021. Long road to long-read assembly. Nat. Methods 18, 125–129. https://doi.org/10.1038/s41592-021-01057-y

McGinnis, S., Madden, T.L., 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res. 32, W20-5. https://doi.org/10.1093/nar/gkh435

Mignardi, M., Nilsson, M., 2014. Fourth-generation sequencing in the cell and the clinic. Genome Med. 6, 31. https://doi.org/10.1186/gm548

Miller, J.R., Zhou, P., Mudge, J., Gurtowski, J., Lee, H., Ramaraj, T., Walenz, B.P., Liu, J., Stupar, R.M., Denny, R., Song, L., Singh, N., Maron, L.G., McCouch, S.R., McCombie, W.R., Schatz, M.C., Tiffin, P., Young, N.D., Silverstein, K.A.T., 2017. Hybrid assembly with long and short reads improves discovery of gene family expansions. BMC Genomics 18, 541. https://doi.org/10.1186/s12864-017-3927-8

Mu, H., Wang, B., Yuan, F., 2022. Bioinformatics in Plant Breeding and Research on Disease Resistance. Plants 11. https://doi.org/10.3390/plants11223118

Nagarajan, N., Pop, M., 2013. Sequence assembly demystified. Nat. Rev. Genet. 14, 157–167. https://doi.org/10.1038/nrg3367

Ong, Q., Nguyen, P., Phuong Thao, N., Le, L., 2016. Bioinformatics Approach in Plant Genomic Research. Curr. Genomics 17, 368–378. https://doi.org/10.2174/1389202917666160331202956

Pandey, R.V., Pabinger, S., Kriegner, A., Weinhäusel, A., 2016. ClinQC: a tool for quality control and cleaning of Sanger and NGS data in clinical research. BMC Bioinformatics 17, 56. https://doi.org/10.1186/s12859-016-0915-y

Pearson, W.R., 2016. Finding Protein and Nucleotide Similarities with FASTA. Curr. Protoc. Bioinforma. 53, 3.9.1-3.9.25. https://doi.org/10.1002/0471250953.bi0309s53

Pemmasani, S.K., Raman, R., Mohapatra, R., Vidyasagar, M., Acharya, A., 2020. A Review on the Challenges in Indian Genomics Research for Variant Identification and Interpretation. Front. Genet. 11, 1–7. https://doi.org/10.3389/fgene.2020.00753

Pereira, R., Oliveira, J., Sousa, M., 2020. Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics. J. Clin. Med. 9. https://doi.org/10.3390/jcm9010132

Pertsemlidis, A., Fondon, J.W., 2001. Having a BLAST with bioinformatics (and avoiding BLASTphemy). Genome Biol. 2, reviews2002.1. https://doi.org/10.1186/gb-2001-2-10-reviews2002

Posey, J.E., 2019. Genome sequencing and implications for rare disorders. Orphanet J. Rare Dis. 14, 153. https://doi.org/10.1186/s13023-019-1127-0

Prjibelski, A.D., Korobeynikov, A.I., Lapidus, A.L., 2018. Sequence analysis, Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics. https://doi.org/10.1016/B978-0-12-809633-8.20106-4

Rhee, S.Y., Dickerson, J., Xu, D., 2006. Bioinformatics and its applications in plant biology. Annu. Rev. Plant Biol. 57, 335–360. https://doi.org/10.1146/annurev.arplant.56.032604.144103

Roberts, R.J., Carneiro, M.O., Schatz, M.C., 2013. The advantages of SMRT sequencing. Genome Biol. 14, 405. https://doi.org/10.1186/gb-2013-14-7-405

Şahin, E.Ç., Aydın, Y., Gilles, T., Uncuoğlu, A.A., Lucas, S.J., 2022. Chapter 28 - Concepts and applications of bioinformatics for sustainable agriculture, in: Sharma, P., Yadav, D., Gaur, R.K.B.T.-B. in A. (Eds.),. Academic Press, pp. 455–489. https://doi.org/https://doi.org/10.1016/B978-0-323-89778-5.00012-X

Salazar Robles, G., Hernández, L.R., Pedraza Pérez, Y., Juárez, Z.N., Rodríguez Acosta, M., Pérez Armendáriz, B., Bautista Rodríguez, E., El Kassis, E.G., 2022. Bioinformatic approach for the identification of plant species that accumulate palmitoleic acid. Electron. J. Biotechnol. 60, 58–69. https://doi.org/https://doi.org/10.1016/j.ejbt.2022.09.0

08

Salzberg, S.L., 2019. Next-generation genome annotation: we still struggle to get it right. Genome Biol. 20, 92. https://doi.org/10.1186/s13059-019-1715-2

Slatko, B.E., Gardner, A.F., Ausubel, F.M., 2018. Overview of Next-Generation Sequencing Technologies. Curr. Protoc. Mol. Biol. 122, e59. https://doi.org/10.1002/cpmb.59

Stein, L., 2001. Genome annotation: from sequence to biology. Nat. Rev. Genet. 2, 493–503. https://doi.org/10.1038/35080529

Supplitt, S., Karpinski, P., Sasiadek, M., Laczmanska, I., 2021. Current Achievements and Applications of Transcriptomics in Personalized Cancer Medicine. Int. J. Mol. Sci. 22. https://doi.org/10.3390/ijms22031422

Suriyakala, G., Sathiyaraj, S., Gandhi, A.D., Vadakkan, K., Mahadeva Rao, U.S., Babujanarthanam, R., 2021. Plumeria pudica Jacq. flower extract - mediated silver nanoparticles: Characterization and evaluation of biomedical applications. Inorg. Chem. Commun. 126, 108470. https://doi.org/https://doi.org/10.1016/j.inoche.2021.108470

Theys, K., Lemey, P., Vandamme, A., Baele, G., 2019. Advances in Visualization Tools for Phylogenomic and Phylodynamic Studies of Viral Diseases 7, 1–18. https://doi.org/10.3389/fpubh.2019.00208

Thompson, J.D., Linard, B., Lecompte, O., Poch, O., 2011. A Comprehensive Benchmark Study of Multiple Sequence Alignment Methods: Current Challenges and Future Perspectives. PLoS One 6, e18093. https://doi.org/10.1371/journal.pone.0018093

Vadakkan, K., 2019. Acute and sub-acute toxicity study of bacterial signaling inhibitor *Solanum torvum* root extract in Wister rats. Clin. Phytoscience 5. https://doi.org/10.1186/s40816-019-0113-3

Vadakkan, K., Cheruvathur, M.K., Chulliparambil, A.S., Francis, F., Abimannue, A.P., 2021. Proteolytic enzyme arbitrated antagonization of helminthiasis by Cinnamomum cappara leaf extract in Pheretima posthuma. Clin. Phytoscience 7, 23. https://doi.org/10.1186/s40816-021-00261-9

Vadakkan, K., Hemapriya, J., Selvaraj, V., 2019a. Quorum quenching intervened in vivo attenuation and immunological clearance enhancement by *Solanum torvum* root extract against *Pseudomonas aeruginosa* instigated pneumonia in Sprague Dawley rats. Clin Phytosci 5, 24 (2019). https://doi.org/10.1186/s40816-019-0120-4

Vadakkan, K., Vijayanand, S., Choudhury, A.A., Gunasekaran, R., Hemapriya, J., 2018. Optimization of quorum quenching mediated bacterial attenuation of *Solanum torvum* root extract by response surface modelling through Box-Behnken approach. J. Genet. Eng. Biotechnol. 16, 381–386. https://doi.org/10.1016/j.jgeb.2018.02.001

Vadakkan, K., Vijayanand, S., Hemapriya, J., Gunasekaran, R., 2019b. Quorum sensing inimical activity of Tribulus terrestris against gram negative bacterial pathogens by signalling interference. 3 Biotech 9, 163. https://doi.org/10.1007/s13205-019-1695-7

Wang, Y., Wu, H., Cai, Y., 2018. A benchmark study of sequence alignment methods for protein clustering. BMC Bioinformatics 19, 529. https://doi.org/10.1186/s12859-018-2524-4

Wang, Yunhao, Zhao, Y., Bollas, A., Wang, Yuru, Au, K.F., 2021. Nanopore sequencing technology, bioinformatics and applications. Nat. Biotechnol. 39, 1348–1365. https://doi.org/10.1038/s41587-021-01108-x

Wei, L., Liu, Y., Dubchak, I., Shon, J., Park, J., 2002. Comparative genomics approaches to study organism similarities and differences. J. Biomed. Inform. 35, 142–150. https://doi.org/10.1016/s1532-0464(02)00506-3

Wingett, S.W., Andrews, S., 2018. FastQ Screen: A tool for multi-genome mapping and quality control. F1000Research 7, 1338. https://doi.org/10.12688/f1000research.15931.2

Wittich, H.C., Seeland, M., Wäldchen, J., Rzanny, M., Mäder, P., 2018. Recommending plant taxa for supporting on-site species identification. BMC Bioinformatics 19, 190. https://doi.org/10.1186/s12859-018-2201-7

Xiao, T., Zhou, W., 2020. The third generation sequencing: the advanced approach to genetic diseases. Transl. Pediatr. 9, 163–173. https://doi.org/10.21037/tp.2020.03.06

Ying, Y.-L., Hu, Z.-L., Zhang, S., Qing, Y., Fragasso, A., Maglia, G., Meller, A., Bayley, H., Dekker, C., Long, Y.-T., 2022. Nanopore-based technologies beyond DNA sequencing. Nat. Nanotechnol. 17, 1136–1146. https://doi.org/10.1038/s41565-022-01193-2

How to cite this article:

Aleena Johny. 2024. A Genomic Approach to Understand Plants -The Pandora Box yet to be Explored. *Int.J.Curr.Microbiol.App.Sci.* 13(01): 87-93. **doi:** https://doi.org/10.20546/ijcmas.2024.1301.011