Understanding plant pathogens using optical mapping

Macrophomina phaseolina MS-6 is a fungal pathogen responsible for causing a plethora of despairing diseases in more than 500 host plants, such as jute. A detailed study of the organism is vital for understanding mechanisms of infection in these plants. Whole-genome sequencing can aid this process and provide a better understanding of MS-6. Previously used sequencing methods like next-generation sequencing (NGS) produced short threads which led to difficulties in assembly. Scientists of the Basic and Applied Research on Jute Project at the Bangladesh Jute Research Institute have developed optical mapping as a tool to obtain improved assembly of the genome.

Macrophomina phaseolina MS-6 is a fungal pathogen. A necrotroph, it is a parasite that feeds on dead tissue of its host. It can cause several despairing diseases in more than 500 host plants, including jute, cotton, maize, groundnut, potato and other crops. It is a soil-borne fungus and is known to decrease the yield of jute by 30%. In the soil, this fungus can live up to 15 years without attacking the host. It can survive in extreme environmental conditions, including high temperature, diverse pH, drought and low soil moisture.

Jute, an essential cash crop of Bangladesh, contributes significantly to the nation's economy. It is the country's second-largest source of foreign revenue. The adverse impact of this fungal pathogen in jute production makes it a point of interest for the research community. The researchers aimed to study the entire genome of this fungus to understand the details of *Macrophomina Phaseolina's* infection mechanism. Next-generation sequencing (NGS) was used to obtain the entire genome of the organism. However, NGS generated a large number of short reads that hampered the de novo assembly. This was due to the complex and repeated regions of the genome. The obtained data had an extensive amount of gaps and misassembles.

Edward L. Barnard, Florida Department of Agriculture and Consumer Services, CC BY 3.0, via Wikimedia Commons

Black microselerotia and orange spores of the fungi *Macrophomina phaseolina* infecting its host.

OPTICAL MAPPING - A NEW HOPE IN WHOLE-GENOME SEQUENCING

Complications exhibited by traditional whole-genome sequencing methods like NGS have escalated the need to develop novel techniques. Scientists of the Basic and Applied Research on Jute Project at the Bangladesh Jute Research Institute used an error-free tool to acquire MS-6 DNA sequences to overcome the hurdles generated by NGS. They applied optical mapping (OM) to obtain improved data of the genome. Whole-genome Optical mapping is a cutting-edge technology that can resolve several issues. It estimates the distance length between the scaffolds and combines them into much longer sequences without adding new bases. It also serves as a template for de novo genomic sequence assembly, which allows for precise detection and quantification of broad structural variations in the genome. Optical mapping can also generate high-resolution, ordered, high-throughput genomic map data providing information about the genome's structure.

THE STEPS INVOLVED IN OM ANALYSIS

M. Phaseolina MS-6 was isolated from a jute plant infected with stem rot. Spheroplasting, lysis, plug washing and restriction enzyme selection techniques were used to extract megabase size DNA. The later steps involved are - MapCard setup, optical mapping assembly, and alignment using MapSolver.

A total of 13 restriction enzymes were tested, including BamH, EcoRI, KpnI, NheI, XbaI, and others, to find the most efficient and acceptable restriction endonuclease. The Enzyme.pl script (in-house script) was used to find the best restriction endonuclease for calculating restriction fragments statistics for various restriction enzymes. The script was used to choose the best restriction enzyme based on various factors such as average fragment size (kb), fragments larger than 100 kb, maximum fragment size, and the highest percentage of average fragment size between 5 to 20 kb size fragments.

To build a consensus optical map, the mapset (total data sets created from all runs) was put together for assembly using the Argus system embedded Gentig map assembler. Integrated data from each MapCard were aligned using the MapSolver software. The software employs a dynamic programming algorithm to determine the best location in the Optical Map for each of the supplied sequence folds.

Fungal pathogens cause damage to many different plants including important crops such as jute.

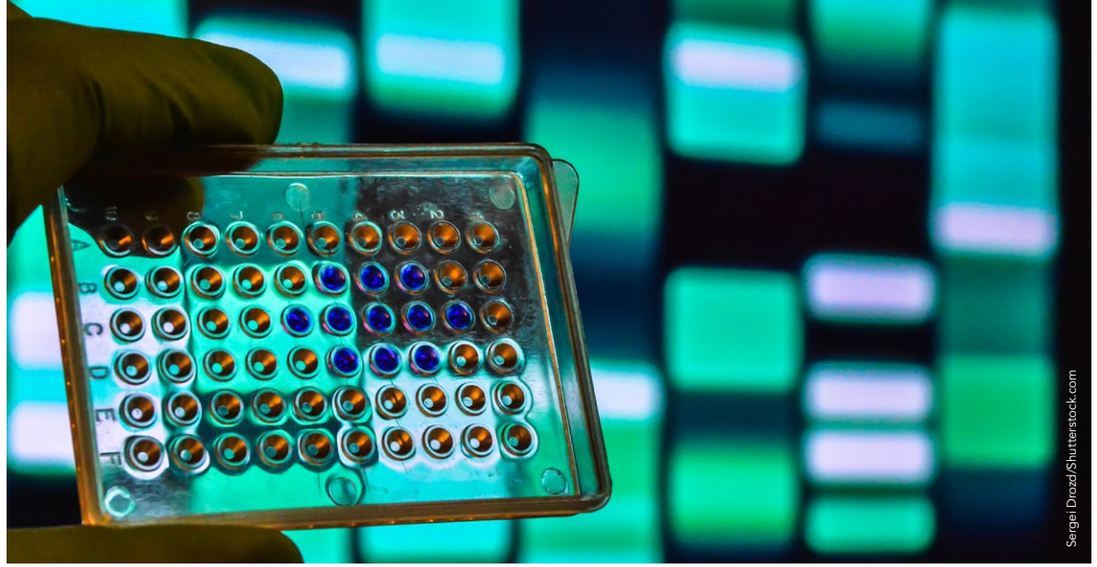
Optical mapping can also generate high-resolution, ordered, high-throughput genomic map data that provides information about the genome's structure.

A VALUABLE TOOL

KpnI, followed by NdeI and XbaI showed better results for longer fragments. KpnI stood out as the most effective and feasible restriction enzyme for optimal mapping based on the rest parameters. High molecular weight (HMW) DNA was digested with the KpnI restriction enzyme and dyed on the optical chips. It yielded 71 GB of raw data from 19 MapCards. The optical chips analysis provided a total of 270,343 Single-Molecule Restriction Fragments (SMRMs) with an average size of 263.22 kb. Furthermore, the assembly of the molecules resulted in the formation of 12 unambiguous super-scaffolds (denoted as chromosomes) with telomeric-blunt ends.

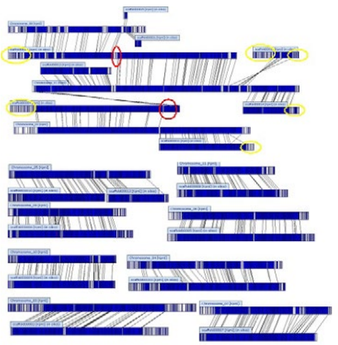
By joining all SMRMs using the ArgusTM optical mapping system, the chromosomes ranged in size from 1.6 to 6.7 Mb and spanned a total of 49.723Mb. It also decreased the number of scaffolds from 94 to 88, with the largest scaffold increasing in size by 2 Mb. N50 was

Macrophomina phaseolina growing on an agar plate.



Sergei Drozd/Shutterstock.com

Scientists of the Basic and Applied Research on Jute Project at the Bangladesh Jute Research Institute have developed optical mapping as a tool to obtain improved assembly of the genome.



Alignment between optical maps and NGS sequence generated scaffolds. The blue-shaded regions of each map represent regions of the genomes that are similar whereas white areas are different. The alignment lines (lines connecting maps) connect regions of similarity from one map to the other.

Alignment of de novo chromosomes and NGS sequence generated scaffolds. Regions encircled in red, crossed green and yellow indicate misassembled, inversion and low-quality assembly, respectively.

placed on OM scaffold number 5 rather than NGS scaffold number 6 in terms of contiguity, efficiency, and development. It grew by 4.25 Mb from NGSs 3.39 Mb. N90, meanwhile, moved from scaffold 14 to scaffold 11 by increasing its size to 2.9Mb from 1.4 Mb. The technique generated 270,343 genomic DNA molecules (>250 kb), which were assembled and aligned with NGS to retrieve sequence extension. Previous misassembles, low quality assembles, inverse sequences and other issues were resolved by combining and aligning both NGS and OM data. This method directly contributes to the improvement and validation of the NGS assembly. It also provides a valuable tool for enhancing our understanding of fungi chromosome evolution and provides an opportunity to protect them in a biotechnologically sound manner.

Jute, an essential cash crop of Bangladesh, contributes significantly to the nation's economy.

FUTURE SCOPE

Initially used in the whole-genome construction of bacteria, parasites and fungi, optical mapping has now found application in assembly validation of large-scale genome projects. This study highlights the significance of optical mapping in furnishing the MS-6 genome. It also emphasises the application of this data in developing a logical strategy that can facilitate the understanding of host-pathogen interaction. These findings can help as a foundation for future research. Studies of different pathogens and their interaction with hosts can benefit from the whole-genome optical mapping method. The cutting-edge technology exposed concordances and discordances, such as inversions, low-quality assembly, holes and overlaps. These helped in fixing the NGS misassembles. Based on the results, OM created an improved and validated assembly that can help us understand the evolution of fungi chromosomes. The furnished assembly in this research can also assist in future chromosomal re-arrangements in other fungi. It can also play a role in the development of cross-talk phenomena between host and pathogen control steps. Researchers may look to the improved non-erroneous longer scaffold-based chromosomal spanned assembly as a turning point in their search for pestilential tools and survival dimensions in various environments.

Research Objectives

To improve the NGS assembly by obviating different errors produced from the short-reads assembly, inversions, low-quality assembly, gaps, overlaps, etc. followed by development of cross-talk phenomenon between the host and pathogen for making *M. phaseolina* resistant jute plant.

References

Hossen, Q. M. M., Islam, M. S., Emdad, E. M., Haque, M. S., Alam, M. M., & Alam, M. (2019). Whole-genome optical mapping: Improving assembly of *Macrophomina phaseolina* MS-6 through spanning of twelve blunt end chromosomes by obviating all errors and misassembles. *African Journal of Biotechnology,* 18(31), 1031-1043.

Personal Response

Do you have plans to use this technique on other plant pathogens?

" Optical mapping can easily span genomic regions that are difficult to resolve by DNA sequencing. It also facilitates genome scaffolding, genome-assembly completeness validation, large-scale structural variation detection, and can be used for strain typing. Therefore we have planned to use optical mapping on the emerging disease of jute, Fusarium wilt (*Fusarium oxysporum*) in the near future. This could help us to develop cross-talk phenomenon between the pathogen and host. "

Billion Photos/Shutterstock.com