



A Chromosome-Level Genome Assembly of Garlic (*Allium sativum*) using NGS technologies.



Genomic studies are important for understanding different organisms by studying their genome, or the entirety of their genes. In plants, this information is particularly useful for examining the evolution of the genome to identify genes that can be selected to improve crop varieties. However, some plants have particularly large genomes which can make this task challenging. Next-generation sequencing (NGS) technologies are an important tool that has enabled several recent genomic discoveries, enabling us to go deeper in our understanding of an organism's traits. Novogene provides world-class NGS services that have contributed to recent advancements in our understanding of several areas of genomics. This article will review recent research that characterized the garlic genome, providing the first assembled genome of a species in the *Allium* genera.



Introduction

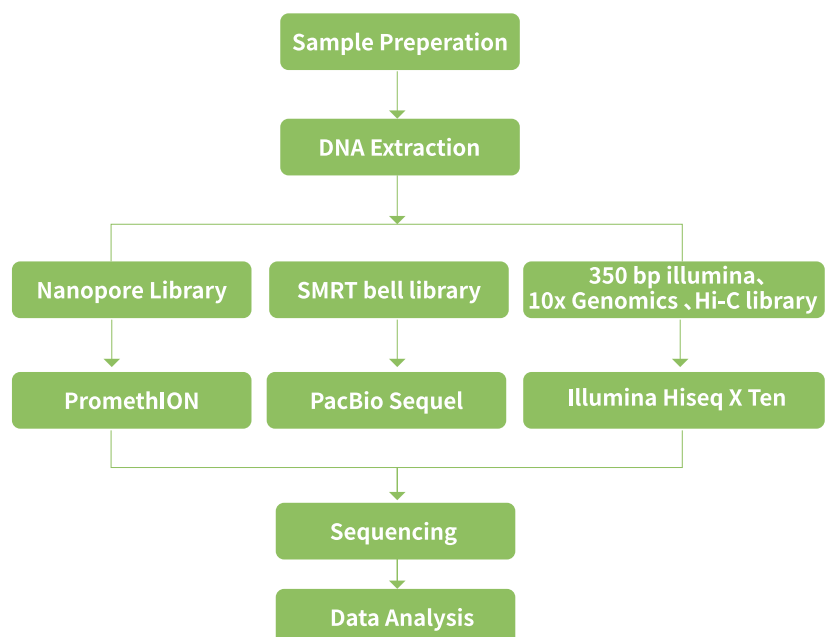
Garlic is an economically important edible bulb that has been cultivated by humans for over 5000 years. This plant is used both as a spice and a medicinal crop. The bulb of the garlic plant is very enlarged and produces distinctive organosulfur compounds and a lot of research to date has focused on the benefit that consuming garlic can have on human health and managing symptoms of various diseases. A recent study, 'A Chromosome-Level Genome Assembly of Garlic (*Allium sativum*) Provides Insights into Genome Evolution and Allicin Biosynthesis,' sequenced the garlic genome and examined the evolution of specific genes that are linked to the biosynthesis of allicin and inulin neoseries-type fructans. This work provides the first genome sequence for an *Allium* species and provides

new resources that can advance research in agriculture and crop breeding.

To construct a chromosome-level genome assembly of garlic, the study used a combination of five sequencing technologies; PacBio, Nanopore, Illumina paired-end, 10x Genomics, and Hi-C. This work represents one of the largest plant genomes sequenced to date. Novogene provides invaluable expertise and the largest sequencing capacity in the world. Researchers benefit from the latest in next-generation sequencing (NGS) technology and bioinformatics services that make research like this possible.

Materials and Methods

Genomic DNA was extracted from young fresh leaves of an Ershuizao plant, a highly adaptable garlic cultivar that is grown in China, using the DNasecure Plant Kit. This DNA was then used to construct Six PacBio libraries that were used for SMRT sequencing using the PacBio Sequel platform. A Nanopore library was also constructed for long-read sequencing which was carried out using the PromethION platform. The researchers also produced nine Illumina sequencing libraries, four 10x Genomics sequencing libraries, and



four Hi-C sequencing libraries. All three of these were sequenced using an Illumina HiSeq X Ten platform.

For the genome assembly, Pacbio SMRT reads were assembled using FALCON software and polished using Quiver. To increase the consensus accuracy of the assembly, all Illumina reads were polished using Pilon software. Following this, the assembly continuity was improved by building scaffolds using sequences produced from the 10x Genomics libraries using FragScaff. Finally, long reads generated by Nanopore sequencing were mapped to the assembly to fill in gaps between contigs, and Hi-C technology was used to assemble a chromosome-level genome.

The results from the sequencing resulted in the de novo assembly of 3.06 Tb of high-quality sequences. A genome of 16.24 GB was generated using a combination of PacBio, Illumina, and 10x Genomics libraries.

Conclusions

This study provides the first de novo genome assembly for an Allium species which was possible due to the recent advancements in NGS technology. Annotated genomes are fundamental tools that can be used to answer key questions related to the biology and evolution of crop plants. This data provides valuable information that can be used to optimize breeding and contribute to crop improvement.

This novel research provides a reference genome that will be an important resource for studying garlic biology and crop breeding. In addition, the large-scale transcriptome data produced will be valuable for future genomic and evolutionary studies.

References

Sun, X., Zhu, S., Li, N., Cheng, Y., Zhao, J., Qiao, X., ... & Liu, T. (2020). A chromosome-level genome assembly of garlic (*Allium sativum*) provides insights into genome evolution and allicin biosynthesis. *Molecular Plant*, 13(9), 1328-1339.



Follow us on LinkedIn

Novogene Co., Ltd.



en.novogene.com



Novogene Global | Novogene AMEA | Novogene America | Novogene Europe



inquiry_us@novogene.com | marketing@novogene-europe.com | marketing_amea@novogeneait.sg