Novogene

10th 2011-2021 ANNIVERSARY

Ten highly-cited Publications reflect the decade of genomic technology development.

Novogene

In 2011, life science had just entered the molecular era of quantitative research. Next-generation sequencing (NGS) technologies had been rarely applied and popularized in research fields.

However, in the past ten years, more and more NGS centers were established extensively with the scientific research and the innovation of experimental technology.

Here we selected the top 10 representative and cutting-edge articles to demonstrate the recorded the step-by-step advancement of NGS, which indicates that NGS is becoming one of leading technology in the field of life science.

And we pay tribute to every researcher who is dedicated to genetic technology development.

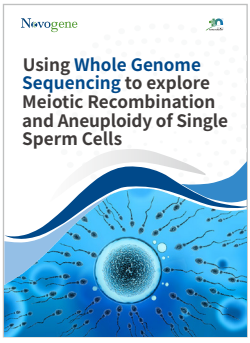
TABLE OF CONTENT

Ten Featured Publications using NGS technologies

Year published	Journal	Research Field	Novogene Service	Impact Factor	Title
2012	Science	Human Sperm	Whole Genome Sequencing	41.84	Probing Meiotic Recombination and Aneuploidy of Single Sperm Cells by Whole-Genome Sequencing
2013	Nature Genetics	Tibetan Pig	*De novo* Sequencing	27.6	Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars
2014	Nature Genetics	Snub-nosed Monkey	Whole Genome Sequencing, Assembly & Annotation Analysis	27.6	Whole-genome sequencing of the snub-nosed monkey provides insights into folivory and evolutionary history
2014	Nature Biotechnology	Soybean	*De novo* Sequecning, Pan-genome Analysis	36.6	*De novo* assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits
2015	Nature Biotechnology	Allotetraploid Cotton	*De novo* Sequencing (*Short-read + Long-read*)	36.6	Sequencing of allotetraploid cotton (*Gossypium hirsutum L. acc. TM-1*) provides a resource for fiber improvement
2017	Genome Research	Pig	*De novo* Sequencing & Assembly (*Short-read+Long-read*)	11.093	Comprehensive variation discovery and recovery of missing sequence in the pig genome using multiple de novo assemblies
2017	Nature Communications	Scallop	*De novo* Sequencing & Assembly, Transcriptome Sequencing	12.121	Scallop genome reveals molecular adaptations to semi-sessile life and neurotoxins
2018	Nature Genetics	Upland Cotton	Reseqeuncing, mRNA Sequencing, GWAS	27.6	Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield
2019	Gut	Hepatocellular Carcinomas	Whole Exome Sequencing, RNA Sequencing	17.943	Integrated multiomic analysis reveals comprehensive tumour heterogeneity and novel immunophenotypic classification in hepatocellular carcinomas
2020	Genome Biology	Influenza & gut microbes	Metagenomic Sequencing	14.028	Influenza infection elicits an expansion of gut population of endogenous Bifidobacterium animalis which protects mice against infection

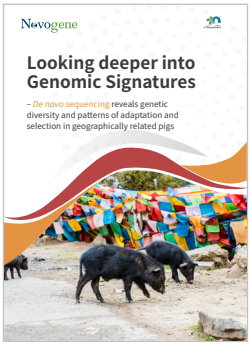
Research Highlights of the Featured Publications

***Probing Meiotic Recombination and Aneuploidy of Single Sperm Cells by Whole-Genome Sequencing*. Science, 2012. IF = 41.84**



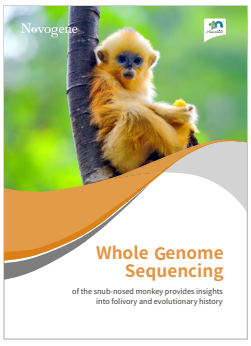
The comprehensive sequencing of human sperm revealed a new molecular mechanism, characteristic for meiosis: recombination events near the transcription start sites. Meiosis is one of the most important molecular mechanisms, allowing for the generation of gametes. A newly developed whole genome amplification method- Multiple Annealing and Looping Based Amplification Cycles (MALBAC)– integrated with whole genome sequencing (WGS) was performed in this study. MALBAC includes quasilinear preamplification when DNA fragments are used as templates. The pooling of MALBAC primers initiates the amplification, and creates semi-amplicons extended by DNA polymerases. Full-amplicons are generated by amplifying semi-amplicons with complementary ends after five cycles. The two ends will be hybridized and form looped DNA, thus the full amplicons cannot be used as template, which reduces amplification bias.The PCR products are further used for sequencing. The study provides insight and methodology for exploring the meiosis errors, infertility and human genome instability at a small scale and a large scale, respectively.

***Comprehensive variation discovery and recovery of missing sequence in the pig genome using multiple de novo assemblies*. Genome Research, 2017. IF = 11.09**



Hundreds of distinct breeds with diverse genetic variations were characterized in the agriculture and food industry previously. However, the limitation to the incomplete reference information needs to be explored. The authors performed ten *de novo* assemblies for nine pigs. The genomes were firstly sequenced using Illumina platform and whole-genome shotgun method, and then assembled with SOAPdenovo.The comparison of the assemblies to the reference genome determined novel single nucleotide polymorphisms, structural variants, and unique sequences harboring specific genes. The study has an important significance for providingdiversifying gene resources in evolution and geographical divergence of pigs.

***Whole-genome sequencing of the snub-nosed monkey provides insights into folivory and evolutionary history*. Nature Genetics, 2014. IF=27.6**



The main focus of this study was to explore the genetic adaption in endangered snub-nosed monkeys. Combining with functional experiments and multi-omics analysis, comparative genomics resequencing revealed the molecular mechanism of primate herbivorous adaptation. All samples were sequenced on Illumina HiSeq 2000 platform. Clean reads were used for further *de novo* assembly. Annotations were also added based on RNA-seq and homolog data. The study elucidates the origin and evolution of snub-nosed monkeys. The completion of the golden monkey genome map suggested that the growth, development, evolution, and origin of this species were clarified at genome

level. The findings promote the research on genetic improvement of basic molecular breeding, which is a great significance for the protection of rare animals and plants.

***De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits*. Nature Biotechnology, 2014. IF = 36.6**

Sequencing and *de novo* assembly of wild species pan-genome provided an insight for clarifying the core genome-one of the most important agricultural cultures – soybean (Glycine max). The resequencing of a single genome is insufficient to represent genomic content.. Therefore, the genome data of seven *Glycine soja* accessions were used for creating pan-genome. Results showed that the linkage-specific genes and structural variations were identified. This research provides a tremendous input for agriculture by identifying novel genes of wild relatives and improves the current domesticated soybean crops which are known to have untapped genetic diversity.

***Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement*. Nature Biotechnology, 2015. IF = 36.6**

The genome sequencing and analysis tools induce changes in agriculture, which meets the requirement of improving domesticated crops. This study focused on allotetraploid Upland cotton. By whole-genome sequencing and *de novo* assembly, the complex polyploid genome is functionally annotated. The novelty of the method is that the ultradense genetic map included assembled short reads, and it did not contain the problematic redundancy which resulted from polyploidization. The study provided an evidence on the asymmetric evolution of two subgenomes. which is helpful for improving cotton agronomic traits, such as yield and fiber quality.

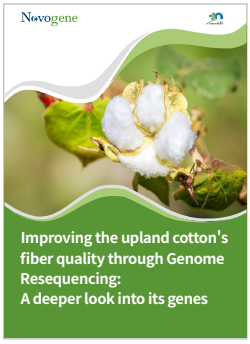
***Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars*. Nature Genetics, 2013. IF = 27.6**

The main novelty of this study is to characterize thouroughly genetic diversity and patterns of adaptation and selection. The research performed *de novo* and whole-genome sequencing of Tibetan wild boars. Different phenotypes were determined due to the impact of environmental selection. In particular, the authors compared genomes of Tibetan wild boar and domestic pigs (Duroc and Chinese domestic pigs). This enabled us to look deeper into the features, such as genomic signatures of high altitudes adaptations, as well as the possible connection of breed-specific genes to food and immunity, and reconstruction of the demographics history.

***Scallop genome reveals molecular adaptations to semi-sessile life and neurotoxins*. Nature Communications, 2017. IF = 12.121**

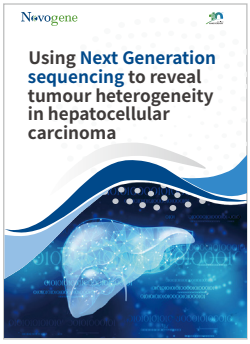
The novel genome sequencing and bioinformatics techniques help to study previous unreachable topics. The progress is particularly prominent in evolution studies of organisms with complex genomes. This study performed deep sequencing for scallop which is known to have high genome heterozygosity. SOAPdenovo was used for genome assembly. Results showed that over 28,000 protein-coding genes and about 5 million SNPs were identified. Phylogenetic analysis was performed and the evolution of scallop adductor muscle, byssal proteins, and eyes are focused in detail. In addition, the reason for bivalves tolerance of neurotoxins is due to the modification of sodium channels. The study indicates that sequencing and genetic analysis can be used to demonstrate the adaptation of the organism to the environment and the mechanisms referring to phenotypes.

***Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield*. Nature Genetics, 2018. IF = 27.6**



Cotton, one of the most widely cultivated species, is increasingly concerned by geneticists. These researchers commonly study the gene functionality referring to the improvement of crop features. Previously the lengthy and laborious selection of high quality plants was the only method to improve production. Nowadays the sequencing enables us to find and modify genes, directly impacting on the cotton quality. The researchers sequenced 419 cotton accessions, analyzed SNPs, and constructed a phylogenetic tree. The genome-wide association studies dissected cotton fiber-related traits as well as reproduction-related features. This study provides a reference of a comprehensive dataset and gene targets for crop improvements by genome sequencing.

***Integrated multiomic analysis reveals comprehensive tumour heterogeneity and novel immunophenotypic classification in hepatocellular carcinomas*. Gut, 2019. IF = 17.9**



One of the main obstacles in oncology is tumor heterogeneity, which will cause inconsistent treatment responses. In this study, integration of metabolomics and proteomics, whole exome sequencing and transcriptome sequencing revealed mutation and gene expression patterns in hepatocellular carcinoma patients as well as key pathways that related to PI3K-Akt signalling and ketone body metabolism of tumour cells. A novel classification method in tumour was also proposed. Such classification and subsequent immunotherapy have the potential to predict and improve the prognosis extensively.

***Influenza infection elicits an expansion of gut population of endogenous Bifidobacterium animalis which protects mice against infection*. Genome Biology, 2020. IF = 14**

The gut microbiome is associated with many health conditions due to its impact on immunity. This study focused on influenza and its interaction with gut microbiota, which is thought to form an immune response against this viral disease. The metagenomic sequencing analysis was used to identify the specific microorganisms in the survival and dying hosts. The diversity of gut microbes is vast. The researchers identified that the genus *Bifidobacterium*, as well as other strains, are prevalent in surviving mice. On the contrary, some strains are found to aggravate influenza symptoms. Thus, a simple dietary modification can help to prevent or relieve the burden of influenza.

References

[1]Lu S, Zong C, Fan W, et al. Probing Meiotic Recombination and Aneuploidy of Single Sperm Cells by Whole-Genome Sequencing[J]. Science, 2012. https://science.sciencemag.org/content/338/6114/1627.full
[2]Li M, Tian S, Jin L, et al. Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars[J]. Nature Genetics, 2013. https://www.nature.com/articles/ng.2811
[3]Zhou X, Wang B, Pan Q, et al. Whole-genome sequencing of the snub-nosed monkey provides insights into folivory and evolutionary history[J]. Nature Genetics, 2014. https://www.nature.com/articles/ng.3137
[4]Li Y, Zhou G, Ma J, et al. De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits[J]. Nature Biotechnology, 2014. https://www.nature.com/articles/nbt.2979
[5]Zhang T, Hu Y, Jiang W, et al. Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement[J]. Nature Biotechnology, 2015. https://www.nature.com/articles/nbt.3207
[6]Li M, Chen L, Tian S, et al. Comprehensive variation discovery and recovery of missing sequence in the pig genome using multiple de novo assemblies[J]. Genome Research, 2017. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5411780/
[7]Li Y, Sun X, Hu X, et al. Scallop genome reveals molecular adaptations to semi-sessile life and neurotoxins[J]. Nature Communications, 2017. https://pubmed.ncbi.nlm.nih.gov/29167427/
[8]Ma Z, He S, Wang X, et al. Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield[J]. Nature Genetics, 2018. https://www.nature.com/articles/s41588-018-0119-7
[9]Zhang Q, Lou Y, Yang J, et al. Integrated multiomic analysis reveals comprehensive tumour heterogeneity and novel immunophenotypic classification in hepatocellular carcinomas[J]. Gut, 2019. https://gut.bmj.com/content/68/11/2019
[10]Zhang Q, Hu J, Feng J W, et al. Influenza infection elicits an expansion of gut population of endogenous Bifidobacterium animalis which protects mice against infection[J]. Genome Biology, 2020. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-02007-1

Novogene Co., Ltd.

en.novogene.com

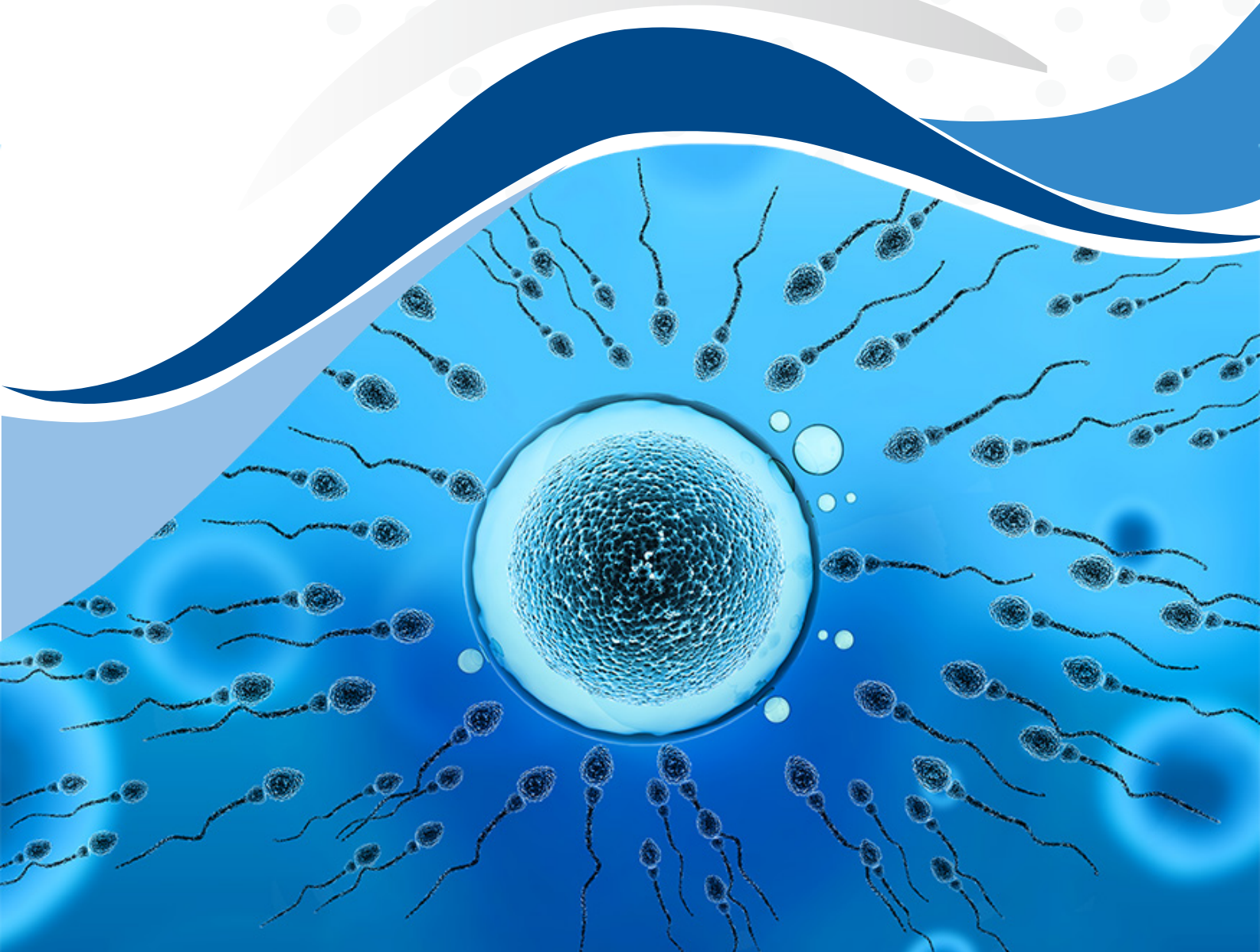
Novogene Global | Novogene AMEA | Novogene America | Novogene Europe

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

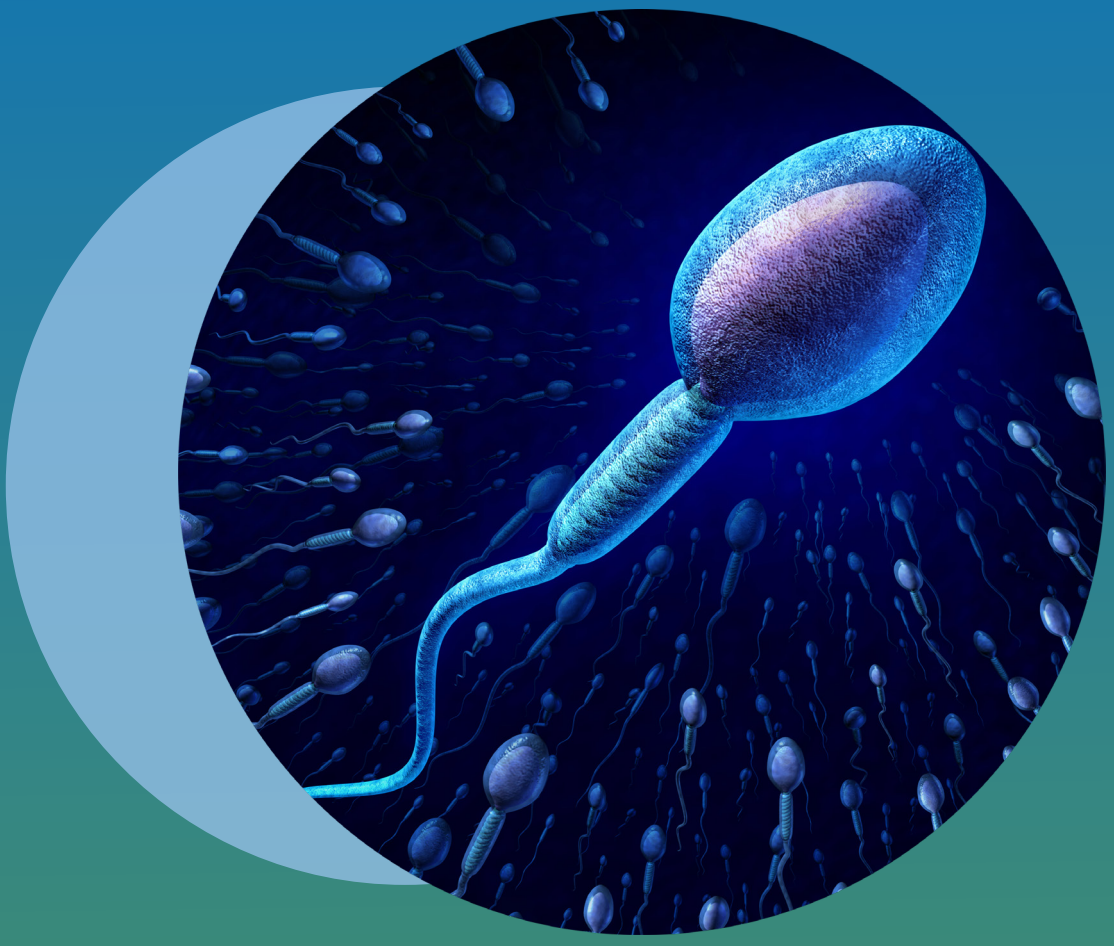
Follow us on LinkedIn

Novogene

Using Whole Genome Sequencing to explore Meiotic Recombination and Aneuploidy of Single Sperm Cells



Novogene



Meiotic recombination is an important mechanism that maximizes genetic diversity at both individual and population levels. However, variation of highly uneven recombination through genome cannot be reflected by human population average. Therefore, comparisons for crossovers among single genomes are required to achieve high resolution. Here we introduce a fascinating study led by researchers from Peking University in Beijing, China titled, '***Probing Meiotic Recombination and Aneuploidy of Single Sperm Cells by Whole Genome Sequencing*** [1].' This informative research uses a newly developed **whole genome amplification (WGA) method – Multiple Annealing and Looping Based Amplification Cycles (MALBAC)** – and sequencing to investigate meiotic recombination and aneuploidy in single sperm cells of an individual. Results showed that the researchers were able to identify recombination features based on single gametes in sperm cells. A low recombination rate was observed near the transcription start sites in individual sperm cells. Also, a different crossover mechanism between autosomes and sex chromosomes was verified. These results are helpful for studying genome instability and male infertility in clinical therapy.

Background

Meiosis is an important process that generates haploid gametes for sexual reproduction. Crossover, is also known as one type of recombination, occurs during meiosis and creates new combination of genes to help maximize genetic diversity. Abnormality in generating these crossovers is the leading cause of miscarriage and birth defects[2,3].

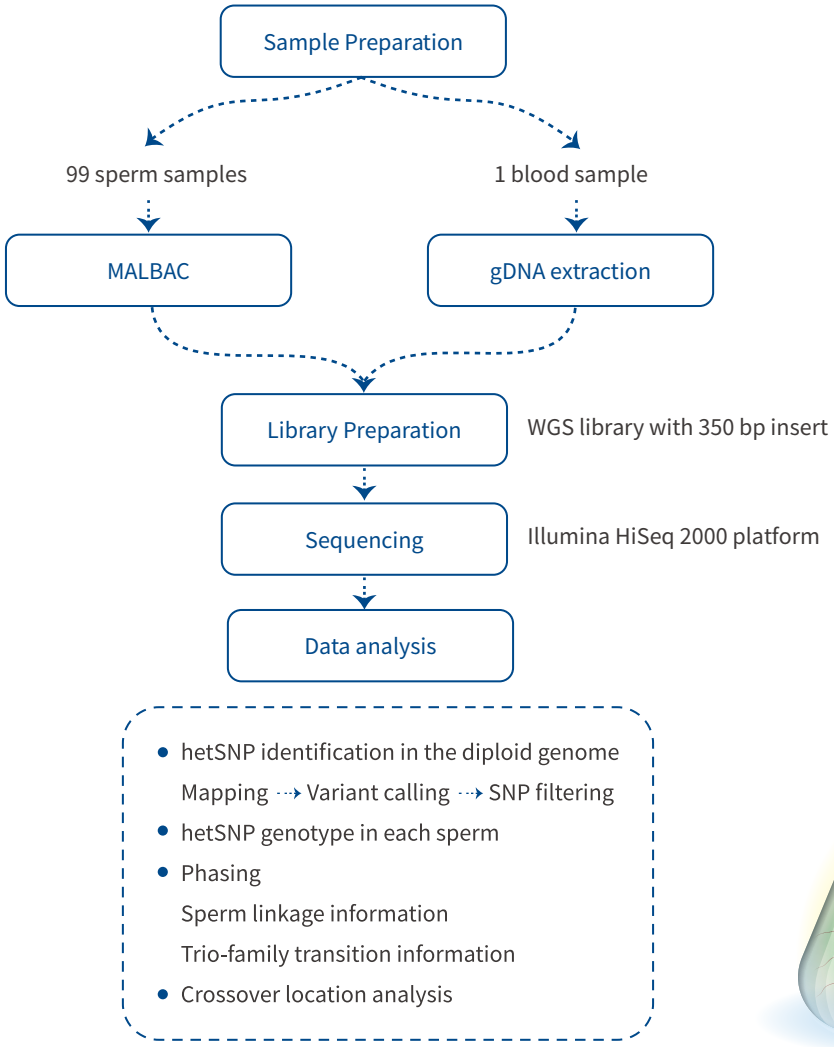
The distribution of recombination across the human genome is highly uneven and recombination active regions are not conserved among different human populations[4-6]. This suggests that these regions could be quickly evolving or might even be specific to an individual[7]. However, it is difficult to observe variation across a human population since differences may be masked by the population average.

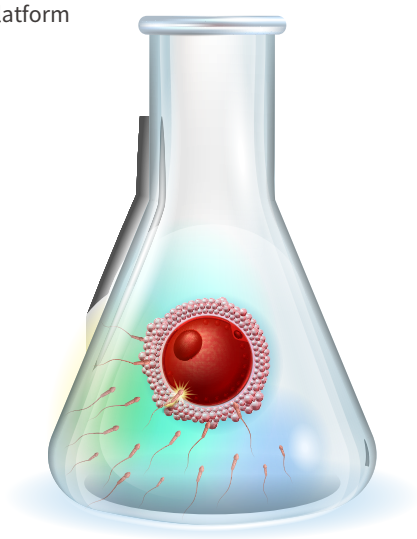
A recent study led by researchers from Peking University in Beijing studied recombination in the sperm cells of a single individual. Using MALBAC and **whole genome sequencing (WGS)**, the researchers mapped crossovers at the individual level, which could be used to provide further information for disease study.

Experiment Design and Research Pipeline

In this study, 99 grade A sperm from a de-identified Asian male were sequenced using single-cell **WGS**. More specifically, WGS was done using a newly developed amplification method called multiple annealing and looping based amplification cycles (MALBAC). The researchers sequenced the diploid genome of the individual and phased it into haplotypes of complete chromosomes. They were then able to use single nucleotide polymorphisms (SNPs) as markers to map the positions of crossovers in each sperm. WGS was carried out by **Novogene Co., Ltd..**

WGS enables researchers to construct the complete nucleotide sequence of an individual genome and identify variants such as **single-nucleotide variations (SNVs), insertions and deletions (InDels), copy number variations (CNVs)**, and large **structural variants (SVs)**. WGS has a wide range of applications, from being used in the study of genetic diseases and cancer to examining human population evolution and DNA biomarkers.





Results and Conclusions

WGS identified approximately 2.8 million SNPs in the diploid genome of the human donor. Approximately 1.4 million of these SNPs were heterozygous and were subsequently phased into chromosome level haplotypes. The increased sequencing depth is helpful to improve hetSNP phasing. By using SNP markers, 2,368 autosomal crossovers were further determined in the sperm cells. MALBAC also reached higher resolution in detecting crossovers than other WGA methods.

The researchers were then able to identify recombination features based on single gametes. A decreased crossover frequency in autosomal aneuploidy was observed on a global per-sperm basis. Recombination showed less poor rate close to genes and higher rate far away from transcription start sites (TSS) in individual sperm cells, which indicates that such a phenomenon is essential to the molecular mechanism of meiosis. .

Novogene

Novogene provides a range of services that can help researchers get high-quality, reliable, and accurate WGS result. With a wealth of experience in sequencing and analysis, Novogene could professionally recommend appropriate methods for various projects, whether researches focuse on humans, plants and animals or not. WGS has multiple applications from genetic diseases, human population evolution, to mechanisms referring conservation and natural selection.

Novogene is capable of sequencing up to 200,000 human genomes per year. Equipped with Illumina platforms, supercomputing clusters and Novogene in-house automatic pipeline, a more complete and accurate characterization of the human genome could be delivered within a fast turnaround time.

References

[1] Lu S, Zong C, Fan W, et al. Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. *Science*. 2012;338(6114):1627-1630. doi:10.1126/science.1229112.
[2] Epstein CJ. The consequences of chromosome imbalance: principles, mechanisms, and models. *Cambridge University Press.* 2007.
[3] Jeffreys AJ, May CA. Intense and highly localized gene conversion activity in human meiotic crossover hot spots. *Nature Genetics*. 2004;36:151–156. https://doi.org/10.1038/ng1287
[4] Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. A Fine-Scale Map of Recombination Rates and Hotspots Across the Human Genome. *Science*. 2005;310:321–324. doi: 10.1126/science.1117196.
[5] Paigen K, Petkov P. Mammalian recombination hot spots: properties, control and evolution. *Nature Reviews* Genetics. 2010;11:221–233.
[6] Kong A, et al. Fine-scale recombination rate differences between sexes, populations and individuals. Nature. 2010;467:1099–1103. https://doi.org/10.1038/nature09525.
[7] Hinch AG, et al. The landscape of recombination in African Americans. Nature. 2011;476:170–175. https://doi.org/10.1038/nature10336.
[8] 11. Jeffreys AJ, Neumann R. Reciprocal crossover asymmetry and meiotic drive in a human recombination hot spot. *Nature Genetics*. 2002;31:267–271. https://doi.org/10.1038/ng910

Novogene Co., Ltd.

Follow us on LinkedIn

en.novogene.com

Novogene Global | Novogene AMEA | Novogene America | Novogene Europe

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

Novogene

10th 2011-2021 ANNIVERSARY

Looking deeper into Genomic Signatures

– *De novo* sequencing reveals genetic diversity and patterns of adaptation and selection in geographically related pigs

Novogene

The advances in genomics allowed scientists to focus deeply on the evolution of species. The study on evolution and selection aids in understanding what genes are responsible for excellent adaptation to the environment, agricultural traits, and how we can utilize this knowledge for improving agricultural qualities. In the study of "**Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars**", Li et al focused on Chinese Tibetan wild boars versus domesticated pigs' comparison, asking a question about how the pigs were domesticated and how did the selection occur in different geographical locations. The authors were focusing on explaining what genetic adaptations occurred in response to high altitude regions. The study was done with the scientific and technological assistance of **Novogene Co., Ltd..**

Introduction

The difference between the habitation conditions between wild and domesticated pigs is an excellent model for learning about natural and artificial selections [1]. This topic attracted the attention of Li et al, and the main goal of the study was to identify the genetic profiles of Tibetan wild boars and Duroc pigs. The researchers, from **Sichuan Agricultural University, Peking University**, and **Novogene** worked on designing the experiments and analytical strategy, DNA library construction and sequencing, genome assembly, and annotations. Novogene technology also allowed the identification of genes of interest and identify SNPs as well as phylogenetic reconstruction.

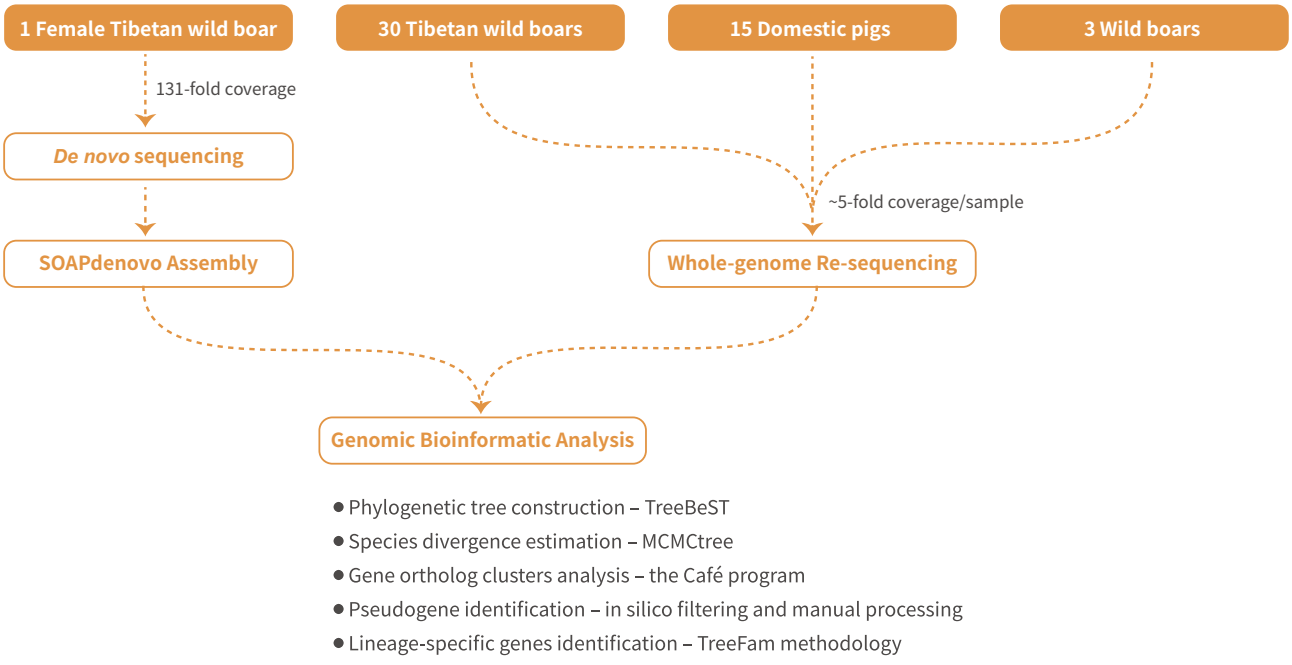


Figure 1. The Experimental Pipeline

The methods used were Illumina sequencing technology and the whole-genome shotgun strategy (*de novo* sequencing). The DNA libraries contained short- and long-inserts and were constructed with the assistance of Novogene researchers. The coverage was over 131-fold. After sequencing, the genome was assembled with SOAPdenovo. The contig sequences were constructed using reads of 180 bp and 500 bp DNA libraries [2]. The obtained contigs sizes were N50 of 1,124 bp and N90 of 252 bp [2]. Then, the reads were realigned onto the contig sequences. Finally, the authors constructed scaffolds, and the pair-end information was used to close the gaps

using the Gapcloser package [2]. The N50 size and N90 size contigs were further improved. The authors sequenced genomes of wild boars and domestic pigs from different geographic locations within China. To evaluate the quality of data, the authors analyzed 19-mer frequency, GC content, CpG frequency, heterozygosity density distribution, and divergence distribution of transposable elements. For SNPsand Indel detection, the authors used BWA, SOAPsnp, and SAMtool packages [2]. To reconstruct phylogeny and infer the genetic structure, the TreeBeST and FRAPPE program were used, respectively [2]. The bioinformatic analysis methods are based on Novogene's expertise.

Key Findings

A total of 48 wild and domesticated pigs' whole genomes were sequenced. The sequencing and genomic reconstruction resulted in 21,806 predicted protein-coding genes [2]. Tibetan wild boars have a high number of lineage-specific genes and they are enriched in specific categories. For example, Tibetan wild boars have more genes of the ferritin family, hypoxia response genes and more DNA damage repair machinery genes. The adaptations also happened in reproduction-related genes such as genes responsible for placenta development and embryonic development. On the other hand, domesticated pigs have more olfactory-related genes. This finding is interesting as pigs have long been considered an excellent olfaction study model due to high smell sensitivity and reliance on smell in their behavior [3]. The newly constructed phylogenetic tree showed that the two studied pig species diverged almost seven million years ago.

The functional genomic analysis showed that Tibetan wild boars rely more on disease-resistance genes, vascular smooth muscle contraction genes, and genes responsible for immunity. It suggests the adaptation to high-altitude harsh environments (lower oxygen and atmospheric pressure, low food diversity) on a genetic level. As compared to Duroc pigs, Tibetan wild boars have significantly fewer olfaction genes.

In summary, the study identified distinct differences between the genetic profiles of Tibetan and domesticated pigs in China. The whole-genome sequencing and SOAPdenovo reconstruction allowed to describe the genetic diversity, gene functionality, and selection patterns.

Novogene contributed greatly to the facility of the development in *de novo* assembly. The company's expertise in sequencing bioinformatics resulted in a profound depth of Tibetan and domestic pigs' analysis. The results are based on the SOAPdenovo-mediated high-quality genome reconstruction, which is Novogene's software. Additionally, the company provided cutting-edge algorithm and experiment development as well as further data processing – annotation and genes/SNPs identification. The researched genomes of Tibetan and domesticated pigs greatly improve our knowledge of selection patterns and adaptation-related gene functionality.

References

[1] Groenen, M.A. *et al.* Analyses of pig genomes provide insight into porcine demography and evolution. (2012), *Nature* **491**, 393–398. doi: https://doi.org/10.1038/nature11622
[2] Li, M., Tian, S., Jin, L. *et al.* (2013), Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars, *Nat Genet* **45**: 1431-1438. doi: https://doi.org/10.1038/ng.2811
[3] Nguyen, D.T. *et al.* (2012), The complete swine olfactory subgenome: expansion of the olfactory gene repertoire in the pig genome. BMC Genomics **13**, 584. doi: https://doi.org/10.1186/1471-2164-13-584

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

Novogene

Whole Genome Sequencing

of the snub-nosed monkey provides insights into folivory and evolutionary history

Novogene

Colobines are Old World monkeys, and unlike other Old World monkeys, their diet mainly consists of seeds and leaves. Previous studies suggest that the stomach of Colobines is similar to those of even-toed ungulates, rather than other primates. Since this species' diet is very distinct compared to the diet of closely related species, it is of high importance to study their adaptation to get better insights into the evolution of all primates. Through whole-genome sequencing and *de novo* assembly, this study sheds light on adaptation to folivory and the evolutionary history of all primates.

Introduction

Colobines are Old World monkeys whose main diet is composed of leaves and seeds, in contrast to fruits and insects of other Old World monkeys. Due to an adaptation to folivory, Colobines are here used as a model organism, to study the evolution of the primate diet. Previous research suggests that the stomachs of Colobines are compartmentalized, and symbiotic bacteria in the foregut ferment carbohydrates. Subsequently, through digesting the symbiotic bacteria, nutrients are recovered [1]. A similar strategy can also be found in other foregut fermenters, such as even-toed ungulates, a distantly related group of mammals.

Previous research has focused on sequencing several primate genomes, however, prior to this study, high-quality genome information of Asian and African Colobines was absent. This information is relevant to understand the adaptation and evolution of primates in general. This research focused on snub-nosed monkeys (Rhinopithecus species) a species of endangered Colobines [2]. Their food contains large amounts of tannins, which are difficult to digest. This study, conducted by researchers from Chinese Academy of Sciences et al. on the genomics of adaptation, hypothesizes that through genetic adaptations, easier breakdown of toxins, digestion of symbiotic microbes, and optimization of their energy metabolism are facilitated, working together with bioinformaticians at Novogene Co., Ltd..

Experiment design

The genome of a male golden snub-nosed monkey (*Rhinopithecus roxellana*) was sequenced at 146-fold coverage, followed by *de novo* assembly and an analysis of the genome. Additionally, *Rhinopithecus bieti, Rhinopithecus brelichi* and *Rhinopithecus strykeri* were re-sequenced at a 30-fold coverage.

All samples involved in *de novo* assembly and resequencing analysis were derived from the Beijing Wildlife Park (*R. roxellana*), Beijing Zoo (*R. brelichi*), Baimaxueshan National Nature Reserve (*R. bieti*) and Gaoligongshan National Nature Reserve Management Bureau (*R. strykeri*). Furthermore, tissue samples included in the transcriptome analysis originated from a dead golden snub-nosed female monkey of the Beijing Zoo.

Genomic DNA from whole blood samples was extracted, libraries constructed and sequenced on the Illumina HiSeq 2000 platform. The derived sequences were filtered for quality and consecutively assembled. The data was prepared and analyzed, focusing on the evolutionary background and the adaptation to folivory. This was done by comparing the genes and gene expression of different species with the genes of the golden snub-nosed monkey.

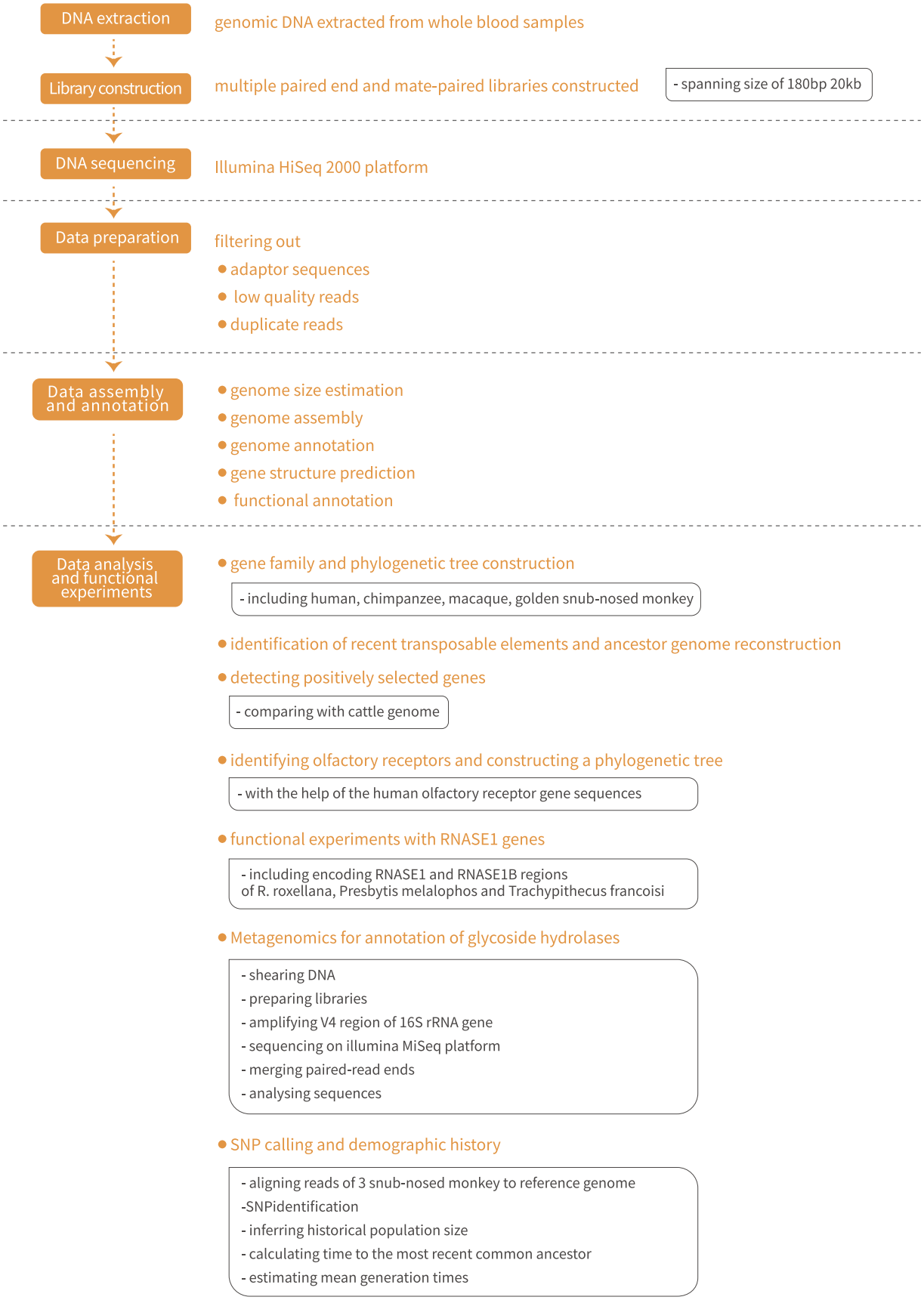


Figure 1. Pipeline for genetic analysis of this study.

Results and Conclusion

The genome of a male golden snub-nosed monkey (*Rhinopithecus roxellana*) was successfully sequenced using a whole-genome shotgun method. The research identified a total of 21.813 protein-coding genes, which were used to generate a timescale for primate evolution.

Analysis and comparisons of the genes of Asian Colobines offer new insights into their adaptation and evolutionary history. The results revealed that Colobines developed an ability to degrade xenobiotics and derive energy from fatty acids. Symbiotic microbiota in their stomach allow this species to digest cellulose. The study found evidence for functional evolution in the *RNASE1* gene. This gene is encoding a secretory RNase, which plays a key function in digesting symbiotic bacteria. Additionally, ancient effective population sizes of the snub-nosed monkey seem to be similar to those of the giant panda. Furthermore, the results of this research allow for the development of a conservation program for the endangered snub-nosed monkey.

To date, Novogene has successfully implemented next-generation sequencing of over hundreds of thousands of samples, providing rapid, efficient and reliable services. Using state-of-the-art and various platforms including Illumina NovaSeq, PacBio Sequel I, II and IIe, Oxford PromethION, and etc., Novogene guarantees excellent data quality and offer assembly solutions tailored towards different requests. Novogene' s multifaceted de novo services includes genome survey, assembly analysis and genome annotation for animals, plants, fungi and bacteria (frame map, complete map and fine map for microbes), tailored to different research needs. Overall, Novogene provides advanced and comprehensive solutions for research ranging from gene functions and profiling to structural analysis and more.

References

[1] Kay, R.N.B. & Davies, A.G. Colobine Monkeys: Their Ecology, Behaviour and Evolution (eds. Davies, A.G. & Oates, J.F.) 229–250 (Cambridge University Press, Cambridge, UK, 1994). https://doi.org/10.1002/ajpa.1331000202
[2] Li, B.G., Pan, R.L. & Oxnard, C.E. Extinction of snub-nosed monkeys in China during the past 400 years. Int. J. Primatol. 23, 1227–1244 (2002). https://doi.org/10.1023/A:1021122819845
[3] Zouh X., Wang B., Pan Qi., et al. Whole-genome sequencing of the snub-nosed monkey provides insight into folivory and evolutionary history. Nature Genetics. 46, 12, 1303-1310 (2014) https://doi.org/10.1038/ng.3137

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

Follow us on LinkedIn

Novogene

Sequencing and *de novo* assembly of pan-genome:

Giving an insight into the core genome of one of the most important agricultural cultures - soybean (*Glycine max*).

Novogene

People have been culturing soybean species for many years, and even when Sanger sequencing technology told became available, the complexity of the plant genome prevented researchers from looking into the soybean genome in detail. In the study of "***De novo* assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits**", Li et al. established a pan-genome of wild soybean (*Glycine soja*). The main benefit of using the wild soybean for the analysis is that it hasn't lost genetic diversity, as it happened with cultivated soybean. The sequencing and de novo genome assembly enabled researchers to gather knowledge about the genetic diversity of soybeans.

Research Background

The main research focus was to establish a pan-genome for soybean species, using a wild soybean as it hasn't gone through artificial selection and still contains genetic diversity [1]. The problem with the sequencing previously was dictated by high genome complexity and many types of structural variations [2]. The novel sequencing and assembly methods enabled researchers from **Chinese Academy of Agricultural Sciences (Beijing)** to characterize seven wild soybean accessions, with the support of the **Novogene**. Novogene provides assistance and expertise with SOAPdenovo assembly and genome reconstruction.

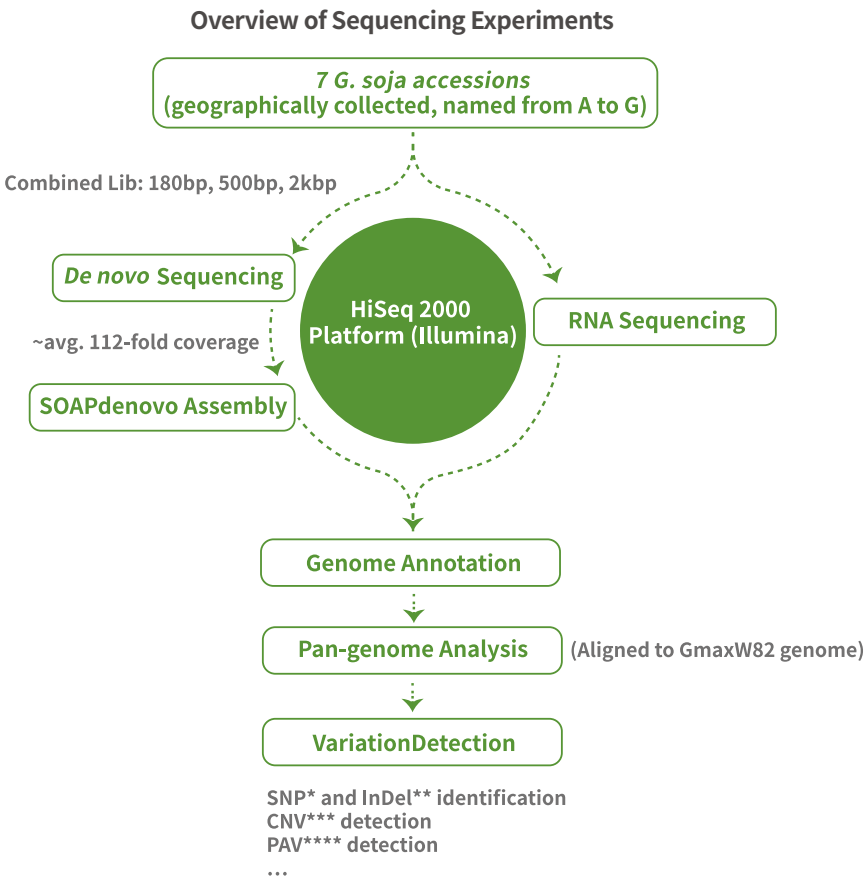


Figure 1. Schematic diagram of partial experimental pipelines

The researchers selected seven accessions of Glycine soja from different geographical regions. The sequencing was done using Illumina HiSeq2000 platform [3]. The library contained inserts of 180 bp, 500 bp, and 2Kbp [3]. Additionally, the authors extracted RNA from eight tissues, mixed it equally, and used it for constructing sequencing libraries.

The assembly was performed using SOAPdenovo. The 17-mer depth distribution of short-insert paired-end reads was used to estimate the genome sizes [3]. To remove the errors, the reads were processed with the ALLPATHS-LG module and ErrorCorrection package in SOAPdenovo [3]. The gap-filling was done with the GapCloser package from SOAPdenovo. The total amount of data from DNA and RNA sequencing was 779.2 Gb and 35.47 Gb respectively [3]. Then, the received assemblies were aligned to the domesticated soybean genome using the NUCmer program [3]. The genome was annotated for protein-coding sequences using ab initio predictions (Evidence Modeler package), mRNA expression data, and homology search integration [3]. The gene predictions were done with Augustus package and Cufflinks [3]. The other bioinformatic analyses authors also worked on gene clustering, SNP and indel identification, and CNV/PAV detection.

Table 1. Software that used in the study

· Alignment of all accessions to the reference genome	NUCmer program
· Putative gene coding region identification	Augustus package
· Protein sequences mapping	TblastN
· RNA-seq reads mapping	TopHat
· Gene structure predictions	Cufflinks, GeneWise, EvidenceModeler package
· Protein sequences alignment	BLAST
· SNPs identification	MUMer, SAMtools

Experimental Results

The core finding of the research is establishing knowledge about a core gene set and a variable set of genes. Additionally. Li *et al*. were able to reconstruct a phylogenetic tree, which demonstrated that wild soybean accessions and domesticated soybean diverged approximately 0.8 million years ago. The authors explain the genes with variations by adaptation to a variety of stresses. For example, adaptation-related genes include resistance genes with the nucleotide-binding site, nucleotide-binding-site leucine-rich domains. On the other hand, the domesticated soybean has species-specific genes such as genes involved in lipid metabolism, energy metabolism, and cell growth and death. Those are the genes are responsible for crucial agricultural features. For example, domesticated soybean is enriched in oil and fatty acid-related genes such as SAG101 (encodes triacylglycerol lipase) [3]. Some genes were shown as positive selectors for domesticated soybean: genes related to proline metabolism, nitrate transport, stomatal complex, and abiotic stress regulation. The majority of comparisons were done using a computational pipeline, which takes into account SNP，indel，CNV and PAV.

The reconstructed pan-genome undercovers many novel genes and alleles that may enrich crops, producing better agricultural traits. With the assistance of Novogene research, the reconstruction methods such as SOAPdenovo as various annotation and prediction algorithms elucidated phylogeny and gene functionality of wild and domesticated soybean species. Additionally, other available methods such as resequencing alone would not be sufficient to detect genetic variability and assemble the pan-genome of wild soybean, which is a complex genome. The findings are expected to provide significant insights into how current important agricultural corps can be further improved.

References

[1] Qiu, Lg., Xing, Ll., Guo Y. *et al*. A platform for soybean molecular breeding: the utilization of core collections for food security (2013), Plant Molecular Biology **83**, 41-50. doi: https://doi.org/10.1007/s11103-013-0076-6
[2] Lam, Hm., Xu, X., Liu, X. *et al*. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection (2010), Nature Genetics **42**, 1053-1059. doi: https://doi.org/10.1038/ng.715
[3] Li, Yh., Zhou, G., Ma, J. *et al*. *De novo* assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits (2014), *Nature Biotechnology* **32**, 1045-1052. doi: https://doi.org/10.1038/nbt.2979

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

Follow us on LinkedIn

Novogene

10th 2011-2021 ANNIVERSARY

Unraveling the allotetraploid upland cotton genome

Novogene

The Upland Cotton (*Gossypium hirsutum* L. acc. TM-1) is one of the most commonly cultivated crop plants around the world. Technological advancements and a better understanding of agricultural techniques have allowed the rise of bigger crop yields and a higher rate of desirable traits. However, some challenges still need to be solved. In the following study **'Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement'**, done by Tianzhen Zhang, Yan Hu, Wenkai Jiang, et al, affiliates of the Zhejiang University, they decided to sequence the plant' s genome, providing a reference for future projects.

Introduction

The Allotetraploid Upland Cotton accounts for more than 90% of the cotton cultivated worldwide, due its textile fiber and the oilseed crop it yields. It represents a significant income source for many countries, for example, China and the United States. Both are the largest producers and consumers of raw cotton. Not just these two, but over 80 countries are significant cotton producers. Annually, this sector yields $500 billion worldwide [1].

Most members of the Gossypium genus are diploid organisms, while only a handful are tetraploid. Hybridization of different genomes from each continent, giving rise to allotetraploid species. An allotetraploid organism has four sets of chromosomes (unlike humans and most mammals, who only have two), resulting from the crossover of different species. This process is estimated to have occurred approximately 1 to 2 million years ago.

Polyploidy is a common phenomenon in nature, particularly among plants. It occurs when an organism has more than two sets of paired chromosomes. Allopolyploidy takes place when an unreduced gametes (2n) fuse.

The unreduced gamete fuses with the unreduced game of a different plant, belonging to a different taxon. In the case of the G. hirsutum, we are talking of different clades. More specifically, the fusion of the A genome of ancestral African species with the D genome clade, found in the Americas.

Each one of these clades possess desirable traits [2]. The exact mechanisms underlying these genomic rearrangements are still poorly understood [3]. The purpose of the study was to draft the plant' s genome, allowing researchers to peer into its genetic base and construct better crop strategies [4].

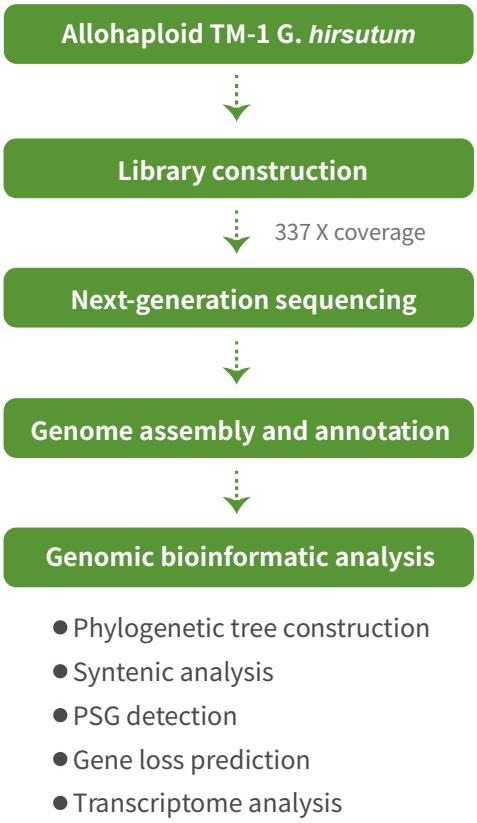
Materials and Methods

An allohaploid plant of Texas Marker-1 (TM-1), G. hirsutum, (AD)1, the genetic standard, was used. (AD)1 was isolated from the cross of TM-1 with a virescently marked semigametic line of G. barbadense (AD2). Approximately 2,400 F1 individuals, resulting from the cross, were analyzed.

The DNA was prepared using the CTAB extraction method and sheared for short-insert paired-end libraries using the Bioruptor sonication device. For mate-pair, the Hydroshear DNA Shearing Device was utilized. The libraries were subsequently sequenced at 2 x 100 bp on an Illumina HiSeq 2000 platform. 843 Gb of DNA sequenced data was generated.

The SOAPdenovo package was used to assemble the short and long paired ends, into longer sequences, organized into

scaffolds (40,407, for a total of 2.4 Gb). Gaps between the scaffolds were filled using the Gapcloser, further scaffolding was then conducted based on links between BAC-ends.



Results

Albeit the exact mechanisms (or the how's) of polyploidization are not entirely known, it' s been reported it involves genomic reorganization and massive gene loss [5]. However, gene loss in the cotton genome is rare, only a total of 228 and 141 genes were lost in the subgenome A and D, respectively. The reason is unknown, perhaps unstable allotetraploid specimens were bred out through selection, leaving a relatively stable genome.

Comparative results demonstrate, an asymmetric evolution between the A and the D subgenome, compared with their corresponding progenitor genomes. This indicates a high evolutionary rate, with significant growth (indicated by the fact the A subgenome is nearly twice the size of the D one).

Homologous genes were found predominantly found in those involved with cotton fiber elongation, as well secondary cell wall synthesis. A bigger number of transcription factor genes were found in the A homeologs, which is consistent with more significant fiber development.

Positive selected genes (the introduction of evolutionary advantageous traits into a population) for fiber improvement, as well as abiotic/biotic stress tolerance to, for example, superoxide and ethylene.

Conclusions

The domesticated cotton serves as a role model for the understanding of the role of cellulose in the biosynthesis of secondary-wall, as indicated by the presence of 23 family genes in the allotetraploid genes. Now, why go through all of this to sequence this data? To provide a genomic draft.

This draft will be used to look into the superior genes to improve positive selection, enhance the fiber production, understand better wall-cell biosynthesis, etcetera. Rapid progress in the genomic fields of agricultural sciences were largely because researchers identified genes specific to certain traits.

By manipulating the genomic lines, certain traits will surface more than others, leading to more copious crops yields, plants with stronger tolerance to endogenous and exogenous stress (temperature, pathogens such as bacteria or fungi), and so on.

Companies, such as Novogene, provide promising teams with state-of-the-art sequencing packages and software. Technological advancements have made the cost of these long processes, turning them into cost-effective scientific endeavors.

References

[1] Chen ZJ, Scheffler BE, Dennis E, et al. Toward sequencing cotton (Gossypium) genomes. Plant Physiol., 2007. doi:10.1104/pp.107.107672.
[2] Mauricio Ulloa. The Diploid D Genome Cottons (Gossypium spp.) of the New World, World Cotton Germplasm Resources, Ibrokhim Y. Abdurakhmonov, IntechOpen, 2014. DOI: 10.5772/58387.
[3] Chen ZJ, Ni Z. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. Bioessays, 2006. doi:10.1002/bies.20374.
[4] Zhang T, Hu Y, Jiang W, et al. Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement. Nat Biotechnol 33, 531–537, 2015. https://doi.org/10.1038/nbt.3207.
[5] Woodhouse MR, Schnable JC, Pedersen BS, Lyons E, Lisch D, Subramaniam S, et al. Following Tetraploidy in Maize, a Short Deletion Mechanism Removed Genes Preferentially from One of the Two Homeologs. PLoS Biol, 2010. https://doi.org/10.1371/journal.pbio.1000409

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

Novogene

Comprehensive variation discovery and recovery of missing sequence in the pig genome using multiple *De novo* assemblies

Novogene

While resequencing has limitations in terms of identifying genetic variation, and reference genomes often are incomplete, de novo assemblies are required in order to comprehensively uncover genetic variation. Here, we present the research with the title，**'Comprehensive variation discovery and recovery of missing sequence in the pig genome using multiple *de novo* assemblies'** by a team of researchers specialized in animal genetics and breeding.

This study generated *de novo* assemblies of phenotypically and geographically distinct pigs and compared these to the reference pig assembly. The results uncovered a substantial amount of variation and discovered missing sequences of the pig genome. This research shows how powerful whole genome *de novo* sequencing is in unravelling underlying genetic diversity and detecting missing sequences.

Research background

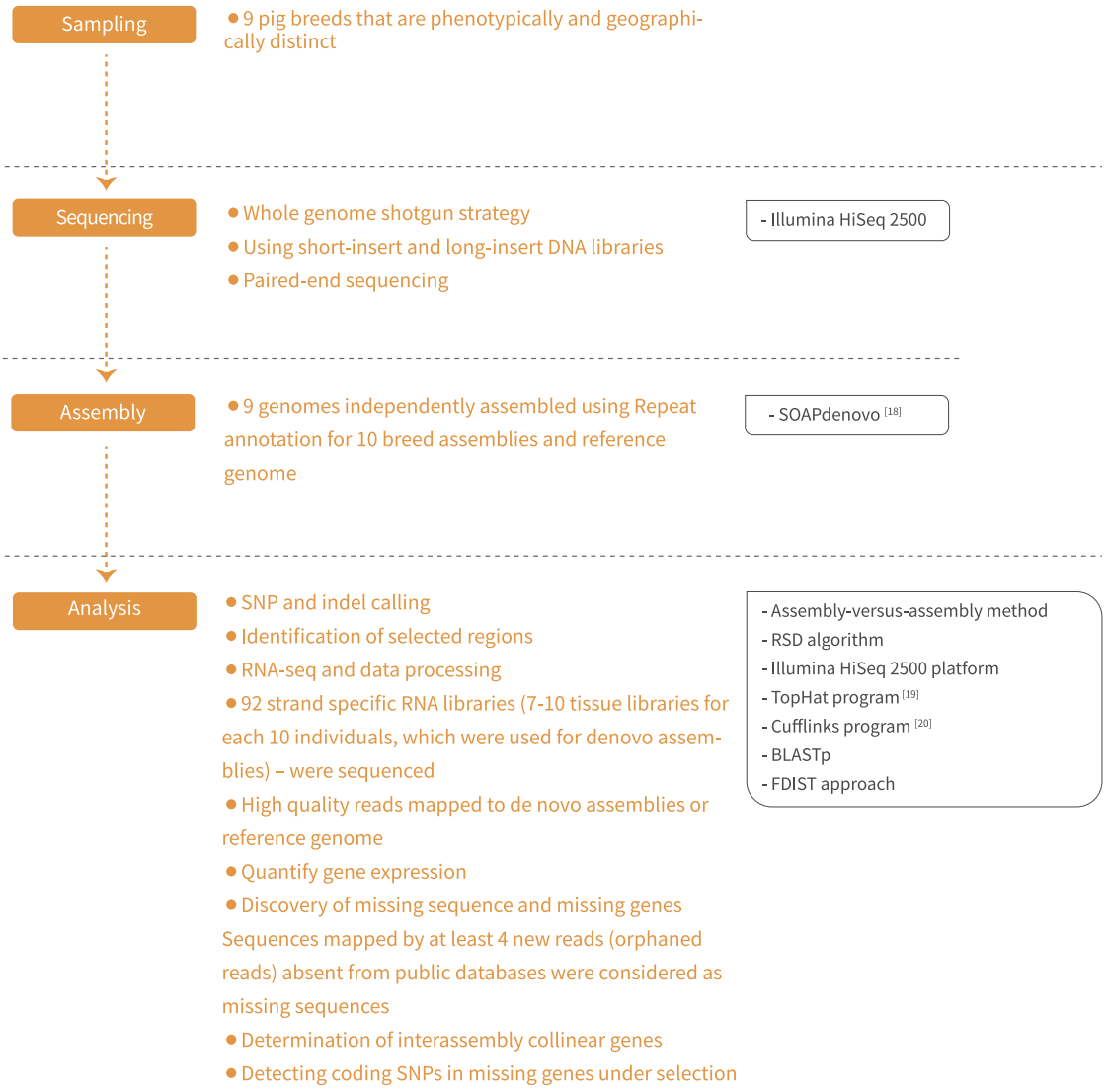
Worldwide, there are over 730 pig breeds whose different phenotypes are a combination of artificial selection and local adaptation [1, 2]. Through resequencing and using the genome of the European domestic Duroc pig as a reference genome, the genetic variation underlying the different phenotypes were previously characterized [3, 4, 5, 6]. Nevertheless, previous research shows that resequencing only partly captures genetic differences, and possibly misassigns regions of the reference genome [7].

Through *de novo* assembly, large amounts of variation in genomes of plant and animal species have been found [8, 9, 10, 11, 12, 13, 14]. To receive a complete understanding of the genetic variation of pigs, multiple de novo assemblies are necessary [15, 16]. Therefore, this research has performed nine *de novo* assemblies of phenotypically and geographically distinct individuals (of Eurasia) and compared these to a genome assembly of the Tibetan wild boar [17].

Sequencing Pipelines

This research performed whole-genome sequencing of nine different pig breeds that are phenotypically and geographically distinct. Using short-insert and long-insert DNA libraries, whole genomes were sequenced through paired-end sequencing on an Illumina HiSeq 2500 platform provided by Novogene Co., Ltd. Also, the whole-genome shotgun strategy was used. Genomes were assembled independently, using SOAPdenovo [18]. Furthermore, for ten breed assemblies and one reference genome, repeat annotations were performed.

The data analysis aimed at recovering novel SNPs, structural variants and other sequences, that are absent in the reference assembly.



Results and Conclusion

De novo assemblies of all nine pig genomes were recovered and compared to the reference pig assembly. A substantial number of structural variants, and novel single nucleotide polymorphisms, were discovered. Additionally, the results showed that hundreds of millions of base pairs, harbouring thousands of protein-coding genes were absent in the reference genome. Results suggest that these genes show variations left by selection and are hence crucial in understanding porcine evolution.

This research provides evidence that resequencing is limited to examining sequences similar to the reference genome and cannot cover the full range of genetic diversity. Hence, to identify novel genetic variation, high-quality de novo assembly of individual genomes, followed by a comparison to the reference sequence is necessary.

Novogene supplies researchers, research organizations and teams with cost-effective, high-tech, accurate, and reliable genomic profiling technologies. Novogene helps researchers achieve their goals by providing next-generation sequencing technologies for whole-genome sequencing, as demonstrated in this study. Novogene has successfully analysed hundreds of thousands of samples through next-generation sequencing, maintaining remarkably high data quality.

References

[1] Chen K, Baxter T, Muir WM, Groenen MA, Schook LB. (2007) Genetic resources, genome mapping and evolutionary genomics of the pig (Sus scrofa). Int J Biol Sci 3: 153–165. doi:10.7150/ijbs.3.153.

[2] 2] Ai H, Fang X, Yang B, Huang Z, Chen H, Mao L, Zhang F, Zhang L, Cui L, He W, et al. (2015) Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. Nat Genet 47: 217–225. doi: https://doi.org/10.1038/ng.3199

[3] Groenen MA, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens HJ, et al. (2012) Analyses of pig genomes provide insight into porcine demography and evolution. Nature 491: 393–398. doi: 10.1038/nature11622

[4] Rubin CJ, Megens HJ, Martinez Barrio A, Maqbool K, Sayyab S, Schwochow D, Wang C, Carlborg O, Jern P, Jorgensen CB, et al. (2012) Strong signatures of selection in the domestic pig genome. Proc Natl Acad Sci 109: 19529–19536. doi: 10.1073/pnas.1217149109

[5] Choi JW, Chung WH, Lee KT, Cho ES, Lee SW, Choi BH, Lee SH, LimW, Lim D, Lee YG, et al. (2015) Whole-genome resequencing analyses of five pig breeds, including Korean wild and native, and three European origin breeds. DNA Res 22: 259–267. doi: 10.1093/dnares/dsv011

[6] Moon S, Kim TH, Lee KT, Kwak W, Lee T, Lee SW, Kim MJ, Cho K, Kim N, Chung WH, et al. (2015) A genome-wide scan for signatures of directional selection in domesticated pigs. BMC Genomics 16: 130. https://doi.org/10.1186/s12864-015-1330-x

[7] Weisenfeld NI, Yin S, Sharpe T, Lau B, Hegarty R, Holmes L, Sogoloff B, Tabbaa D, Williams L, Russ C, et al. (2014) Comprehensive variation discovery in single human genomes. Nat Genet 46: 1350–1355. doi: 10.1038/ng.3121

[8] Read BA, Kegel J, Klute MJ, Kuo A, Lefebvre SC, Maumus F, Mayer C, Miller J, Monier A, Salamov A, et al. (2013) Pan genome of the phytoplankton Emiliania underpins its global distribution. Nature 499: 209–213. doi: 10.1038/nature12221

[9] Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss SJ, Osborne EJ, Sreedharan VT, et al. (2011) Multiple reference genomes and transcriptomes for Arabidopsis thaliana. Nature 477: 419–423. doi: 10.1038/nature10414

[10] Li YH, Zhou G, Ma J, Jiang W, Jin LG, Zhang Z, Guo Y, Zhang J, Sui Y, Zheng L, et al. (2014) de novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. Nat Biotechnol 32: 1045–1052. doi: 10.1038/nbt.2979

[11] Zhang QJ, Zhu T, Xia EH, Shi C, Liu YL, Zhang Y, Liu Y, Jiang WK, Zhao YJ, Mao SY, et al. (2014) Rapid diversification of five Oryza AA genomes associated with rice adaptation. Proc Natl Acad Sci 111: E4954–E4962. doi: 10.1073/pnas.1418307111

[12] Neafsey DE, Waterhouse RM, Abai MR, Aganezov SS, Alekseyev MA, Allen JE, Amon J, Arca B, Arensburger P, Artemov G, et al. (2015) Highly evolvable malaria vectors: the genomes of 16 Anopheles mosquitoes. Science 347: 1258522. doi: 10.1126/science.1258522

[13] Yan G, Zhang G, Fang X, Zhang Y, Li C, Ling F, Cooper DN, Li Q, Li Y, van Gool AJ, et al. (2011) Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques. Nat Biotechnol 29: 1019–1023. doi: https://doi.org/10.1038/ncomms6969

[14] Li R, Li Y, Zheng H, Luo R, Zhu H, Li Q, Qian W, Ren Y, Tian G, Li J, et al. (2010) Building the sequence map of the human pan-genome. Nat Biotechnol 28: 57–63. doi: 10.1038/nbt.1596

[15] Besenbacher S, Liu S, Izarzugaza JM, Grove J, Belling K, Bork-Jensen J, Huang S, Als TD, Li S, Yadav R, et al. (2015) Novel variation and de novo mutation rates in population-wide de novo assembled Danish trios. Nat Commun 6: 5969. 10.1038/nbt.1992

[16] Chaisson MJ, Wilson RK, Eichler EE. (2015) Genetic variation and the de novo assembly of human genomes. Nat Rev Genet 16: 627–640. doi: 10.1038/nrg3933

[17] Li M, Tian S, Jin L, Zhou G, Li Y, Zhang Y, Wang T, Yeung CK, Chen L, Ma J, et al. (2013) Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. Nat Genet 45: 1431–1438. doi: 10.1038/ng.2811

[18] Li R, Zhu H, Ruan J, QianW, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, et al. (2010). de novo assembly of human genomes with massively parallel short read sequencing. Genome Res 20: 265–272. doi: 10.1101/gr.097261.109

[19] Trapnell C, Pachter L, Salzberg SL. (2009) TopHat: discovering splice junctions with RNA-seq. Bioinformatics 25: 1105–1111. doi: 10.1093/bioinformatics/btp120

[20] Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7: 562–578. doi: 10.1038/nprot.2012.016

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

Novogene

Scallop genome reveals molecular adaptations to semi-sessile life and neurotoxins

Novogene

Bivalve molluscs first appeared over 500 million years ago and have survived several mass extinction events. Scallops are bivalve molluscs that primarily feed on toxic dinoflagellates and display a remarkable tolerance to neurotoxins. Researching molecular adaptations of scallops are of high importance in understanding the evolution and adaptation of organisms in general.
Here, we introduce a research with the title, 'Scallop genome reveals molecular adaptations to semi-sessile life and neurotoxins', by Yuli Li, a specialist in bioinformatics and marine genetics.

This study implemented a multi-omic approach in understanding molecular adaptations to semi-sessile life and neurotoxins of a scallop. The results suggest that simple point mutations in key genes have serious effects on an organisms' adaptation and phenotype.

Review of research background

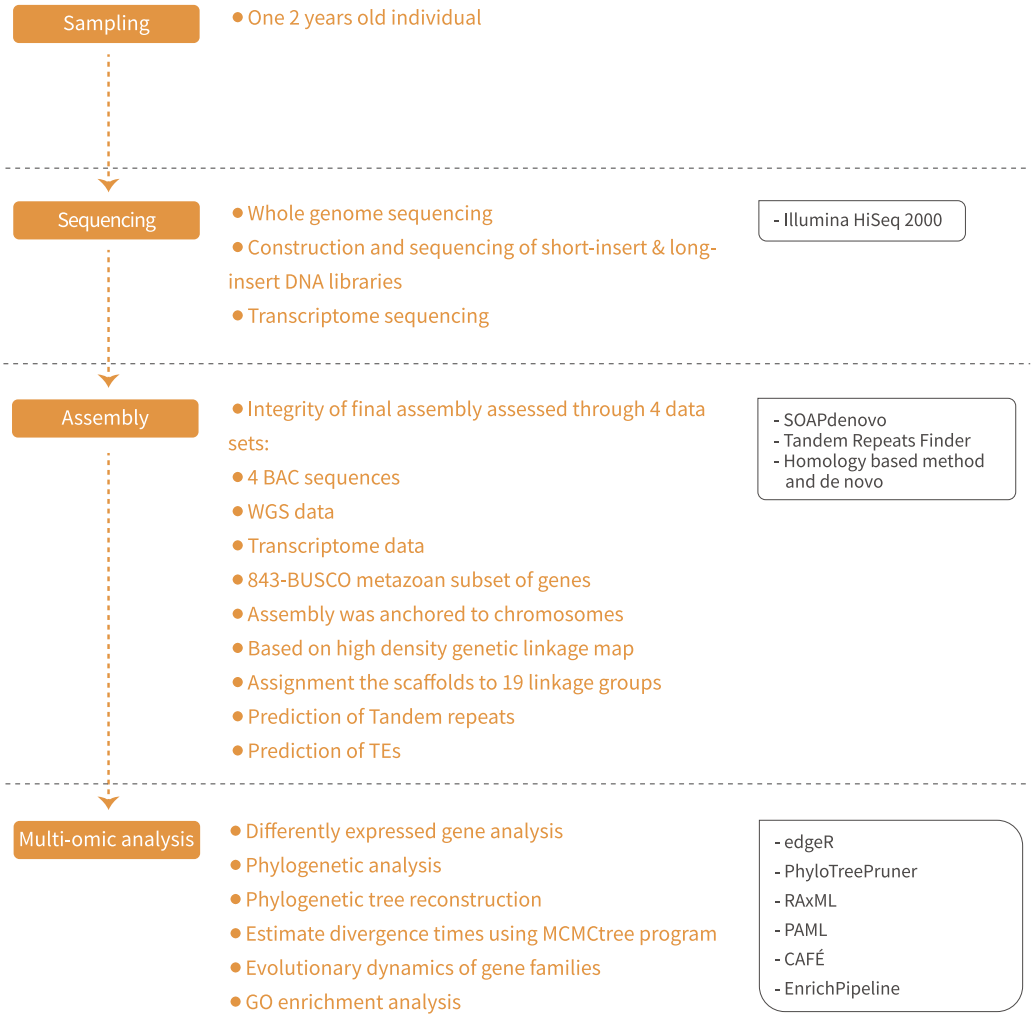
While bivalve molluscs are ecological, biological, and economical significant, their evolution, in particular molecular adaptations, are poorly understood [1, 2, 3, 4, 5]. Scallops, bivalve molluscs known for their beautiful shells, are filter feeders that primarily feed on toxic dinoflagellates. These bivalves can tolerate high levels of neurotoxins, among which are paralytic shellfish toxins (PSTs). Tolerating such toxic substances is remarkable, and understanding molecular adaptations of scallops are of high importance in grasping the evolution and adaptation of organisms.

The Zhikong scallop (*Chlamys farreri*) can tolerate and accumulate extremely high levels of PSTs. [6, 7, 8]. Previous research focused on this species in terms of PST transformation and accumulation, and ample molecular studies on this species have been performed [9, 10, 11]. Due to previous research, this species is optimal for whole-genome sequencing (WGS).

This study sequenced and analyzed the genome and 117 transcriptomes and proteomes of *C. farreri*. The multi-omic approach allows for the analysis of various organs, characteristics, and developmental stages of scallop biology.

Sequencing experiments

The whole genome of one two-year-old *C. farreri* was sequenced using the Illumina Hi Seq 2000 platform provided by **Novogene Co., Ltd**. The experimental design included the construction and sequencing of short insert and long insert DNA libraries. Additionally, the transcriptome of 13 tissues and organs were sequenced. The sequences were assembled using different methods allowing for the integrity of the final assembly. The multi-omic analysis included identifying differentially expressed genes, phylogenetic analysis, divergence time estimations, and analysis of evolutionary changes in gene families. The analysis was performed in order to understand the molecular adaptive evolution of this bivalve to neurotoxins.



Results & Conclusion

The results of this study provide evidence that *C. farreri* reached a resistance to neurotoxins through point mutations in sodium channels. The scallop uses the hepatopancreas to collect neurotoxins and the kidney to transform these toxins into high-toxicity forms. The results suggest that the transformation to highly toxic compounds in the kidney serves as a protection mechanism against predation.

The multi-omic analysis implemented in this study revealed molecular changes and novel genomic features that are related to the scallop' s adaptation to the environment. The study suggests that simple mutations and the expansion of a few key genes result in significant effects on the individuals' adaptation and phenotype. This includes the adaptation to semi-sessile life and filter-feeding, such as the remarkable resistance to neurotoxins, and the production of high toxicity in the kidney.

Novogene supports and supplies researchers with next-generation sequencing applications, such as Whole Genome Sequencing. Novogene has extensive experience in whole-genome sequencing, having successfully completed numerous projects in the field of evolutionary biology.

Additionally, Novogene ensures high-quality research results through professional material selection, library construction and sequencing. Furthermore, Novogene provides comprehensive data analysis, covering, among others, analysis on species evolution, adaptation, selection pressures, and population characteristics.

References

[1] Zhang, G. et al. (2012) The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490, 49–54 doi: 10.1038/nature11413
[2] Takeuchi, T. et al. (2016) Bivalve-specific gene expansion in the pearl oyster genome: implications of adaptation to a sessile lifestyle. Zool. Lett. 2, 3 doi: 10.1186/s40851-016-0039-2
[3] Wang, S. et al. (2017) Scallop genome provides insights into evolution of bilaterian karyotype and development. Nat. Ecol. Evol. 1, 0120. doi: 10.1038/s41559-017-0120
[4] Sun, J. et al. (2017) Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. Nat. Ecol. Evol. 1, 0121. http://dx.doi.org/10.5061/dryad.h9942
[5] Du, X. et al. (2017) The pearl oyster Pinctada fucata martensii genome and multi-omic analyses provide insights into biomineralization. GigaScience 6, 1–12. doi: 10.1093/gigascience/gix059
[6] Zhu, M., Zou, Y., Wu, R. & Hall, S. (2003) Accumulation and depuration of paralytic shellfish poisons (PSP) in Chinese scallop Chlamys farreri. Acta Oceanol. Sin. 25, 75–83. doi: 10.1016/j.toxicon.2005.07.010
[7] Takatani, T., Akaeda, H., Arakawa, O. & Noguchi, T. (1997) Occurrence of paralytic shellfish poison (PSP) in bivalves, along with mossworm adherent to their shells, collected from Fukue Island, Nagasaki, Japan during 1995 and 1996. Food Hyg. Safe Sci. (Shokuhin Eiseigaku Zasshi) 38, 430–434. DOI: 10.1016/s0041-0101(02)00290-8
[8] Tian, H. et al. (2010) Comparative study on in vitro transformation of paralytic shellfish poisoning (PSP) toxins in different shellfish tissues. Acta Oceanol. Sin. 29, 120–126. https://doi.org/10.1007/s13131-010-0015-1
[9] Wang, S. et al. (2004) AFLP linkage map of an intraspecific cross in Chlamys farreri. J. Shellfish Res. 23, 491–499
[10] Zhang, X. et al. (2011) A BAC-based physical map of Zhikong scallop (Chlamys farreri Jones et Preston). PLoS ONE 6, e27612. doi: 10.1371/journal.pone.0027612
[11] Zhan, A. et al. (2009) Construction of microsatellite-based linkage maps and identification of size-related quantitative trait loci for Zhikong scallop (Chlamys farreri). Anim. Genet. 40, 821–831.
https://doi.org/10.1111/j.1365-2052.2009.01920.x

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

Novogene

Improving the upland cotton's fiber quality through Genome Resequencing: A deeper look into its genes

Novogene

The Upland Cotton is a highly profitable and cultivated crop plant throughout the world, its industry is quite a large one. However, despite the versatility it provides, its fiber quality isn’ t top-notch, at least compared to other crop plants. Past studies identified candidate genes responsible for phenotypic fiber-related traits. The current study was conducted by Zhiying Ma, Shoupu He, Xingfen Wang et al, **'Resequencing core collection of upland cotton (*Gossypium hirsutum*) identifies genomic variation and loci influencing quality and yield of fiber.'** The results shed more light on the loci of protein-coding genes, improving the odds of effective artificial selection to bring out better fiber-related traits.

Introduction

The economic impact the Upland Cotton (*Gossypium hirsutum*) has had in the world economy is undeniable. *G. hirsutum* is an allotetraploid crop plant. The United States, China and India are the leading countries in cotton production. The USA alone produced nearly 20 million bales of cotton, representing about $7 billion in total (lint plus seed) value and a 35% of the global cotton exports, just below India and China, with 45-50 percent of the world's production, respectively [1].

G. hirsutum is a highly versatile plant, capable of enduring the abiotic and biotic stress imposed by different types of environments. For years, researchers have sought to sequence its genome, identifying the genes responsible for stress tolerance, crop yield and quality of fiber. Compared to other species of *Gossypium*, *G. hirsutum* is lacking in fiber quality. Sea Island cotton (G. Barbadense) is known for excellent fiber quality with long, strong, and fine fibers [2].

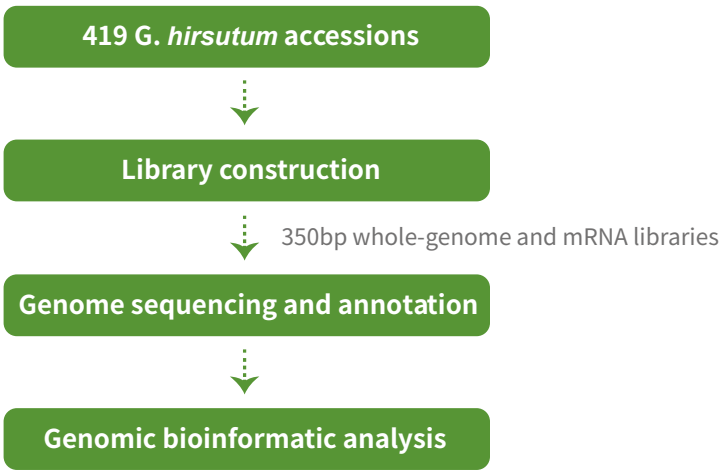
In this study, Zhiying Ma et al, affiliates to Cotton Research Programs and the Novogene Bioinformatics Institute, both located in China. They sought to identify previously undescribed causal genes related to the plant's life cycle, flowing periods, fiber length and strength, by resequencing a core collection of 419 samples [3].

Materials and Methods

G. hirsutum accessions were taken from the China National Gene Bank, Cotton Research Institute, Chinese Academy of Agriculture Sciences, Anyang, Henan Province. These specimens were selected based on the genetic diversity and, thus, the different phenotypic traits they exhibited. These come from different countries all around the world, from the United States, China, India and Brazil, to the now extinct Soviet Union.

The samples were planted and phenotyped, based on their fiber-related traits, in six different locations from different Chinese provinces. The planting was done between 2014 to 2015. Once harvested phenotypic traits (including flowering date) were recorded.

The fiber samples were taken to be examined. Subsequently, the DNA of the aforementioned samples were extracted and isolated, and the genome sequenced. For DNA isolation, Plant DNA Mini Kit (Aidlab Biotech) was used，then whole-genome libraries were constructed and sequenced using the Illumina HiSeq platform, generating 6.45 Tb of raw sequences.



Results

Researchers identified over 3 million population SNPs. Among these, 224,201 were identified in protein-coding regions, 70,959 were found in upstream or downstream regions, while others (3,369,870) were located in intergenic regions. The Upland Cotton possesses different clades, each from a different country and containing different subgenomes. The mixture of the A and the D subgenome produced the *G. hirsutum* (explaining its allotetraploid nature). The SNP distribution was 2.3 times higher in the A subgenome than the D one.

It was also found that modern varieties possessed less genomic diversity, after being compared to 20 years older variations. This phenomenon could be explained by the domestication process the upland cotton underwent.

13 fiber-related traits were studied and their coding genes identified. Each one of these phenotypic characteristics are pivotal if agricultural sciences seek to improve the fiber quality and crop yield of the *G. hirsutum* species. A total of 7,398 associated genes were detected across the 13 traits.

Genes for flowering and fiber-initiation genes, fiber-length-related genes (FL) and fiber-strength-related genes (FS) and lint-related genes (LP and LI, Lint Percentage and Lint Index, respectively) were identified. For the first, two gene candidates were put forth: *Gh_D03G0728* and *Gh_D03G0729*. The former encodes the COP1-interactive protein 1 (CIP1). In Arabidopsis, COP1 plays an important role in light-mediated development.

1,661 FL-associated SNPs were identified. The gene for fiber length was denominated *GhFL1*. Physiologically, these regions are responsible for the actin cytoskeleton organization in the plant cell. Actin is a globular protein crucial for tip growth, organelle movement, and cell elongation.

735 SNPs significantly associated with FS, 842 LP-related genes and 743 LI-related genes were located and 5,753 elite alleles were identified too. However, the genomic diversity of the *G. hirsutum* species is still too low, compared to other crop plants' accessions, such as rice, maize and soybean.

Conclusions

There is still light at the end of the tunnel. 4,820 candidate genes were associated with fiber quality and yield traits. The resequencing done in the study provides future researchers with potential candidates and genetic markers to work on. A significant number of loci regions for fiber quality were detected, much more than in previous reports [4].

These findings are promising. Increasing crop yield and fiber quality is the most imperative task that needs to be completed, if cotton is to supply the increasing global demand for it. Novogene has been highly interested in providing future researchers with state-of-the-art sequencing equipment and professional guidance.

References

[1] https://www.ers.usda.gov/topics/crops/cotton-wool/cotton-sector-at-a-glance/
[2] Avci U, et al. Cotton Fiber Cell Walls of Gossypium hirsutum and Gossypium barbadense Have Differences Related to Loosely-Bound Xyloglucan[J]. PLOS ONE, 2013. https://doi.org/10.1371/journal.pone.0056315
[3] Ma Z, He S, Wang X, et al. Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield[J]. Nat Genet, 2018.https://doi.org/10.1038/s41588-018-0119-7
[4] Fang L. et al. Genomic analyses in cotton identify signatures of selection and loci associated with fiber quality and yield traits[J]. Nat Genet, 2017. doi: 10.1038/ng.3887.

Follow us on LinkedIn

Novogene Co., Ltd.

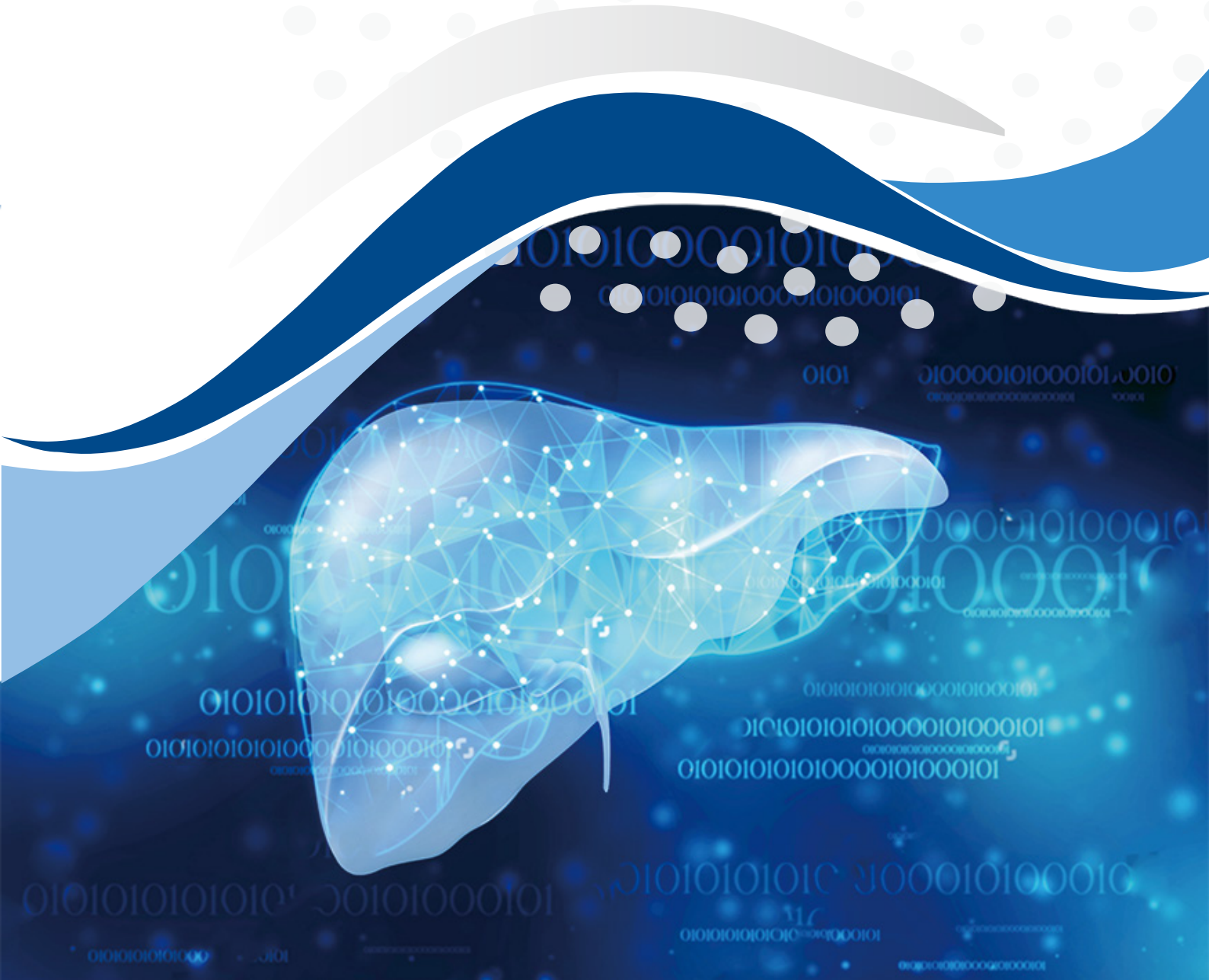
en.novogene.com

Novogene Global | Novogene AMEA | Novogene America | Novogene Europe

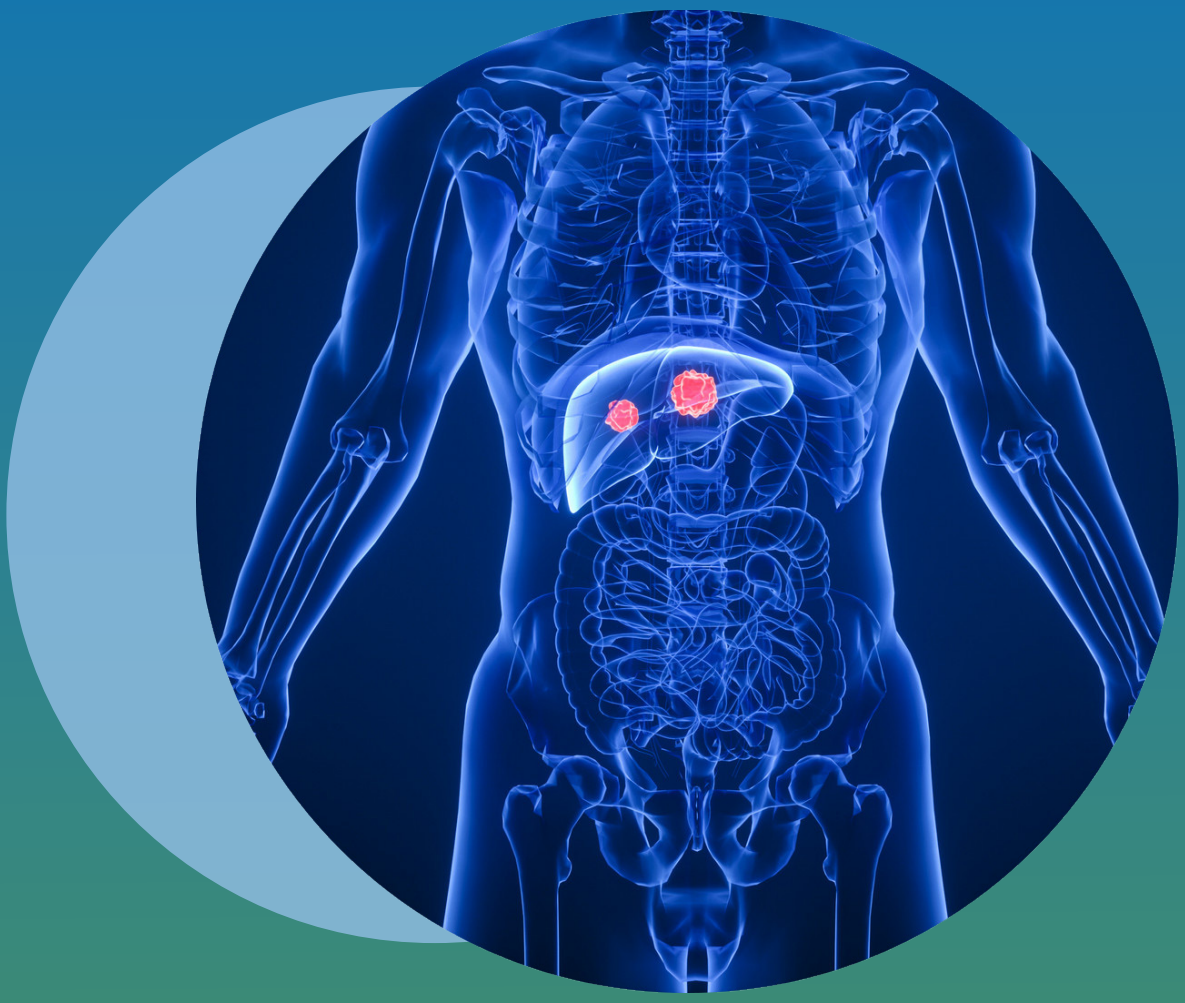
inquiry_us@novogene.com | marketing@novogene-europe.com | marketing_amea@novogeneait.sg

Novogene

Using Next Generation sequencing to reveal tumour heterogeneity in hepatocellular carcinoma

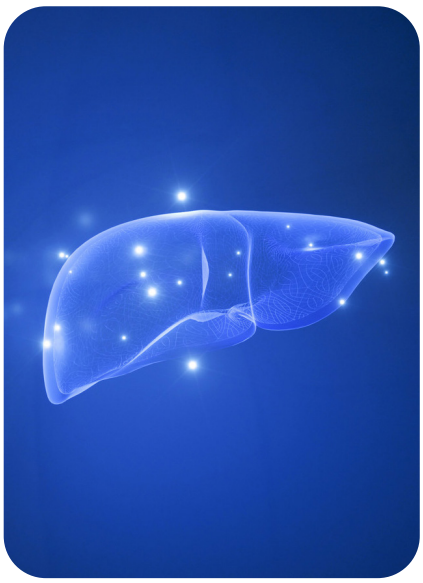


Novogene



Hepatocellular carcinoma (HCC) is a worldwide malignancy, with tumours often evolves into different subtypes which makes HCC difficult to treat. Understanding tumour heterogeneity is important for developing new treatment strategies. However, limited information is available, thereby the development of HCC novel treatments is urgently needed. Here we introduced the research of Prof. Tingbo Liang and Dr. Xueli Bai from Zhejiang University, '*Integrated multiomic analysis reveals comprehensive tumour heterogeneity and novel immunophenotypic classification in hepatocellular carcinomas*[1]' . This fascinating study used multi-omics techniques including whole exome sequencing and RNA sequencing to understand the heterogeneity of tumour cells and microenvironment in HCC. The authors further proposed a new immunophenotypic classification of HCCs that facilitates prognostic prediction and could help guide the course of treatment prescribed.

Background



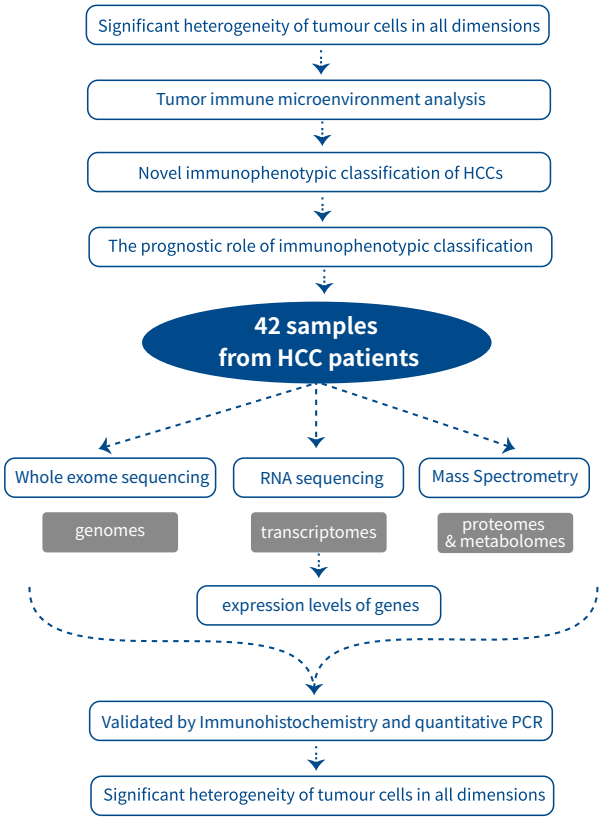
HCC is one of the leading causes of tumour-related deaths worldwide, causing viral infections and liber fibrosis that can only be cured by surgical resection or liver transplantation [2,3]. However, these strategies are only potential cures. Even then cancer still has a high rate of reoccurrence and patients could be suffered from lacking liver donors. Also, existing medicines indicate poor efficacy, which makes it urgent to find new interventions to improve the therapy[4-6].

The heterogeneity in a single tumour (Intra-tumoral heterogeneity) is one of the main reasons that tumour treatments can be ineffective including HCC. Previous research demonstrated that HCC had astonishing intratumoral heterogeneity, but unfortunately, these results failed to generate novel treatments by ignoring clinical setting[7]. Therefore, this study aims to fill this gap by carrying out a comprehensive interrogation of HCC to facilitate the development of new treatment strategies.

Research pipeline

To answer the research questions, samples were collected from patients who were diagnosed with HCC and underwent an operation during September 2018 and November 2019. The diagnosis of HCC was confirmed by pathological examinations. Samples were collected from each site and stored before proceeding **whole exome sequencing (WES) and RNA sequencing (RNA-seq)** which were carried out by **Novogene Co., Ltd**..

WES is a widely used next-generation sequencing method that involves sequencing protein-coding regions of genome. This area of the genome contains approximately 85% of known disease-related variants which makes this technique broadly applicable to a wide range of clinical studies [8].



RNA-seq uses the technology of **next-generation sequencing (NGS)** to reveal the expression profiles of mRNA in biological samples. It helps to understand the variations in the cellular transcriptome. In addition, transcriptome service has a wide range of applications from tumour-subtypes classification to clinical applications. So it could illustrate the continuous variations in the cellular transcriptome.

Results and Conclusions

This study demonstrated that tumour heterogeneity in HCC patients is considerable through genomic, transcriptomic, and various dimensions. **WES data** revealed variations of SNV, InDel and CNV in all samples, while **RNA-seq data** indicated gene expression and regulation patterns, expecially for some immune suppressive marker genes and functional genes in tumour samples.

Further integration analysis with WES and RNA-seq data identified two KEGG pathways (PI3K-Akt signalling and ketone body metabolism of tumour cells) which were found to be robustly associated with the abundance of tumour infiltrating T cells in the HCC microenvironment, providing molecular insight into tumour heterogeneity. Three subtypes of HCC and two marker genes were also found to be promising in immune treatment.

In conclusion, this study revealed tumour heterogeneity in HCC patients and proposed a novel immunophenotypic classification of HCC under genomic and transcriptome dimensions, which supports the guidance of immune therapy.

Novogene supports in this study

Increasingly, multi-omics methods, such as genome, proteome, transcriptome, epigenome, metabolome, and microbiome, are used for comprehensive research on tumours, which helps to reveal the heterogeneity of tumour cells and tumour microenvironment of individuals (indicated by this study)[1]. Based on the results of this study, it is speculated that local immunity may be a suitable target for the treatment of HCC, and a clinically practical HCC classification scheme with prognostic and potential decision-making value is proposed[1].

Novogene supplies multiple services utilized NGS-tech in multi-omics, which help researchers with high-tech, reliable, and accurate genomic profiling results. **Whole Exome Sequencing** targets specific coding regions in the human genome and is useful for genetic disease studies or cancer research (as demonstrated in this study). As a global provider of genomics solutions, Novogene also offers a variety of **Transcriptome Sequencing** services ranging from **mRNA Sequencing** to whole transcriptome analysis. Combined with multiple powerful sequencing systems, human **Whole Genome Sequencing** service allows for specific and accurate characterization of the human genome, which is very useful to detect mutational burden in small coding regions which are caused by polymorphic changes.

References

[1] Zhang Q, Lou Y, Yang J, et al. Integrated multiomic analysis reveals comprehensive tumour heterogeneity and novel immunophenotypic classification in hepatocellular carcinomas. Gut. 2019;68(11):2019-2031. doi:10.1136/gutjnl-2019-318912
[2] El-Serag HB. Hepatocellular carcinoma. N Engl J Med. 2011. 365(12):1118-27. doi: 10.1056/NEJMra1001683
[3] Li C, Li R, Zhang W. Progress in non-invasive detection of liver fibrosis. *Cancer Biol Med*. 2018;15(2):124-136. doi:10.20892/j.issn.2095-3941.2018.0018
[4] Llovet JM, Bruix J. Sorafenib in Advanced Hepatocellular Carcinoma Reply. New England Journal of Medicine. 2008. 359(23):2498-9. doi:10.1056/NEJMoa0708857
[5] Zhu AX, Finn RS, Edeline J, et al. Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. The Lancet Oncology. 2018. 19(7):940-52.. doi:10.1016/S1470-2045(18)30351-6
[6] El-Khoueiry AB, Sangro B, Yau T, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. The Lancet. 2017. 389(10088):2492-502. doi:10.1016/S0140-6736(17)31046-2
[7] Hou Y, Guo H, Cao C, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. Cell research. 2016. 26(3):304-19.doi:10.1038/cr.2016.23
[8] Van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. Trends in genetics. 2014. 30(9):418-26. https://doi.org/10.1016/j.tig.2014.07.001

Novogene Co., Ltd.

en.novogene.com

Novogene Global | Novogene AMEA | Novogene America | Novogene Europe

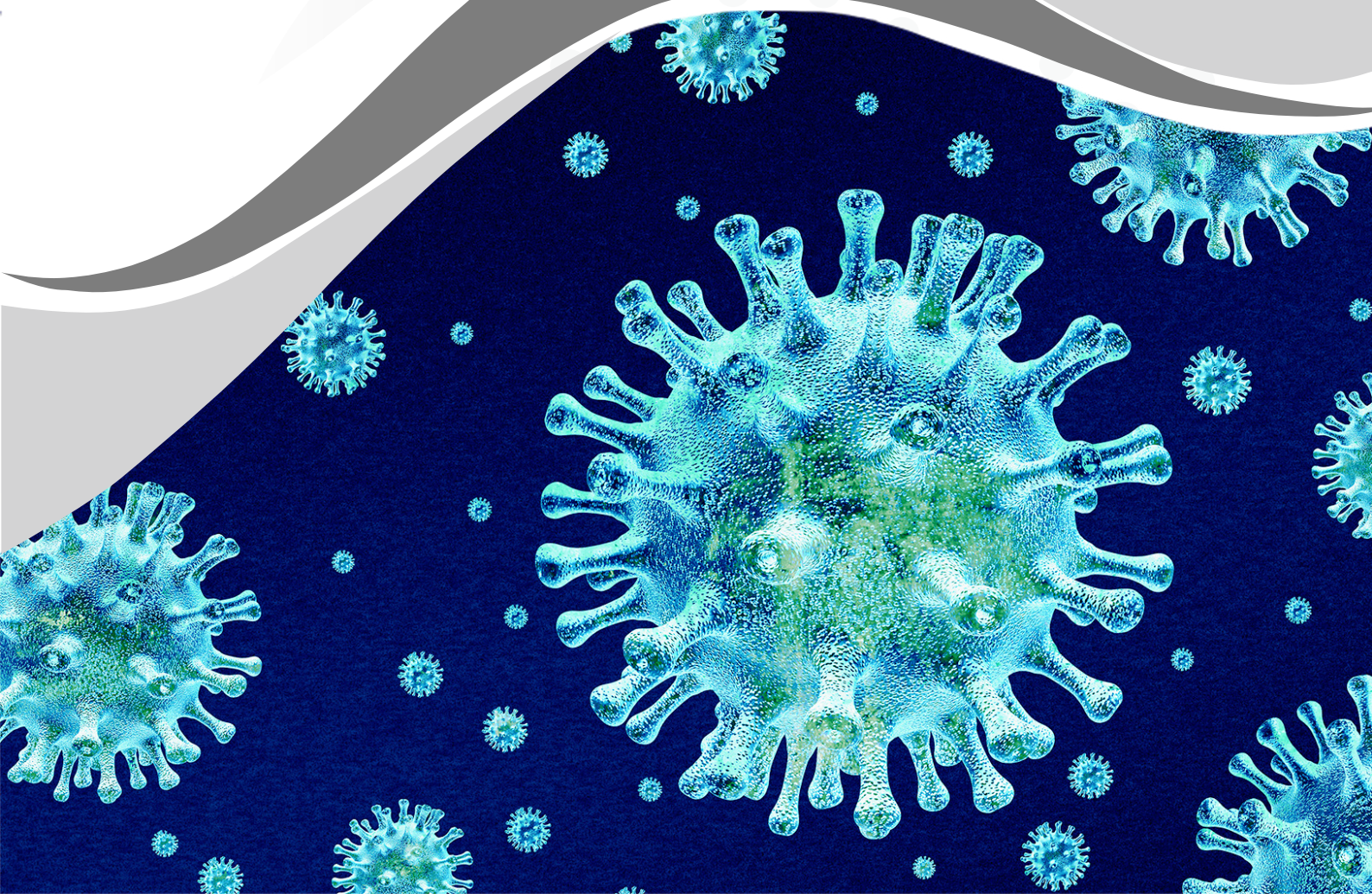
inquiry_us@novogene.com | marketing@novogene-europe.com | marketing_amea@novogeneait.sg

Follow us on LinkedIn

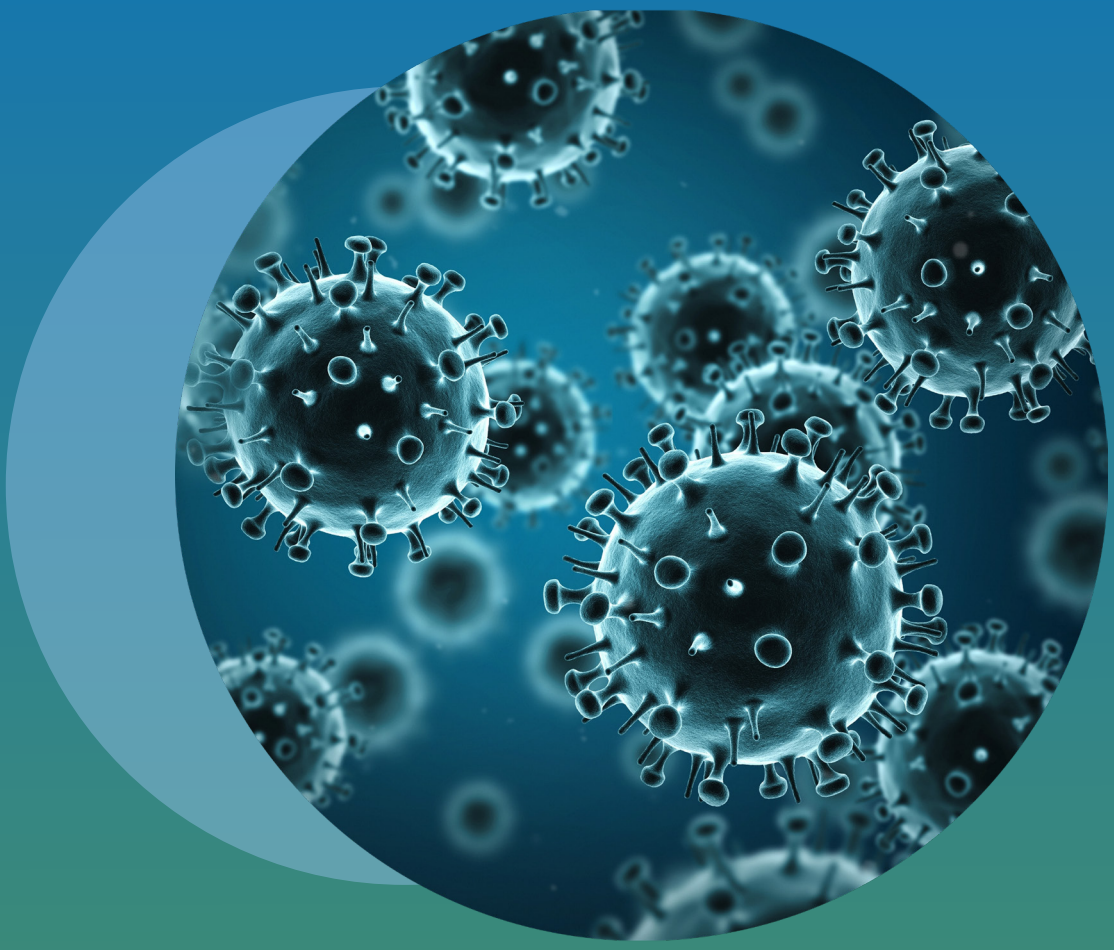
Novogene

10th 2011-2021 ANNIVERSARY

Investigating the gut microbiome using NGS technologies



Novogene



Influenza epidemics affect up to 1 billion people around the world every year, posing a serious threat to human health. Recent research has demonstrated that the **gut microbiome** may offer some protective effect against influenza infection by modulating our immune response. While several studies have demonstrated mechanisms that link the gut microbiome to the immune system, the mechanisms by which the gut microbiota protect against infections like influenza are not fully understood. Here we introduce a fascinating study led by researchers from the National Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, Wuhan, China titled, **'Influenza infection elicits an expansion of the gut population of endogenous *Bifidobacterium animalis* which protects mice against infection**[1]**.'** This informative piece uses **metagenomic sequencing technologies** to identify the gut microbiome of mice that have been exposed to the influenza infection to understand how the gut composition differs based on the severity of infection. Based on the results, the researchers show that the endogenous population of the bacteria *Bifidobacterium animalis* (*B. animalis*) increases in the gut on exposure to the influenza virus and provides protection against influenza infection.

Background

Influenza is an acute respiratory illness that can cause annual epidemics that threaten global health. This communicable illness affects the upper and lower respiratory passages and infects approximately 1 billion individuals and causes up to 500,000 deaths worldwide every year [2]. Recent research has shown that our gut microbiome modulates our immune response to protect against influenza.

The gut microbiome is made up of a complex community of gut-dwelling bacteria that are essential to human health and disease. Disruption to the normal functioning of this microbiome has been linked with a variety of diseases and disorders including inflammatory diseases, colon cancer, and autoimmunity [3]. The protective effects of gut microbiota against influenza have gained increasing interest in recent years which has led to the discovery of several novel mechanisms. For example, it has been demonstrated that microorganism within the gut can rescue the immune impairment by activating the inflammasome via Toll-like receptors [4]. However, despite these advances, the mechanisms by which the gut protects against extraintestinal infection are not fully understood [5].

This recent study increased our understanding of the link between the gut microbiome by using **metagenomic sequencing analysis** to identify specific anti-influenza gut microbes and to understand some of the mechanisms through which the gut microbiome protects against extraintestinal infection.

Sequencing

To study how gut microbiota affects the host susceptibility to influenza infection, the researchers used reverse screening for functional bacteria in mice. They examined the bacteria present in mice that had been infected with influenza and reported as either healthy, ill, or suffered mortality.

This study used metagenomics sequencing to analyse bacteria at the species level from the gut microbiome of mice after they had been infected which was carried out by **Novogene Co., Ltd..**

Metagenomic sequencing is frequently used to identify and differentiate between microbial species. There are two approaches to this type of sequencing depending upon the question that you have. **Shotgun metagenomic sequencing** sequences all given genomic DNA from an environmental sample. This approach does not require prior isolation and cultivation of individual species and has a wide range of applications from gene prediction and abundance analysis to taxonomy and function annotation. In contrast, **amplicon metagenomic** sequencing can be used to amplify short regions (<500 base pairs) of conserved genes (Amplicons). Hypervariable or intergenic regions are chosen and amplified by polymerase chain reaction (PCR) before being analysed by **next-generation**

sequencing (NGS) technologies. The sequences produced are then compared against a microbial database for identification. **16S rRNA amplicon sequencing** has a wide range of applications from identifying species in pure cultures and characterizing the microbiota of animals or plants to comparing species diversity from various environmental sources.

Results and Conclusions

The results from this study revealed that there were variations in the gut microbiota composition and metabolism in mice that responded differently to influenza infection. Interestingly, metagenomics analysis of the gut microbiome in different mouse groups revealed that the bacteria *B.animalis* may mediate a protective effect against influenza by the specific metabolic pathway.

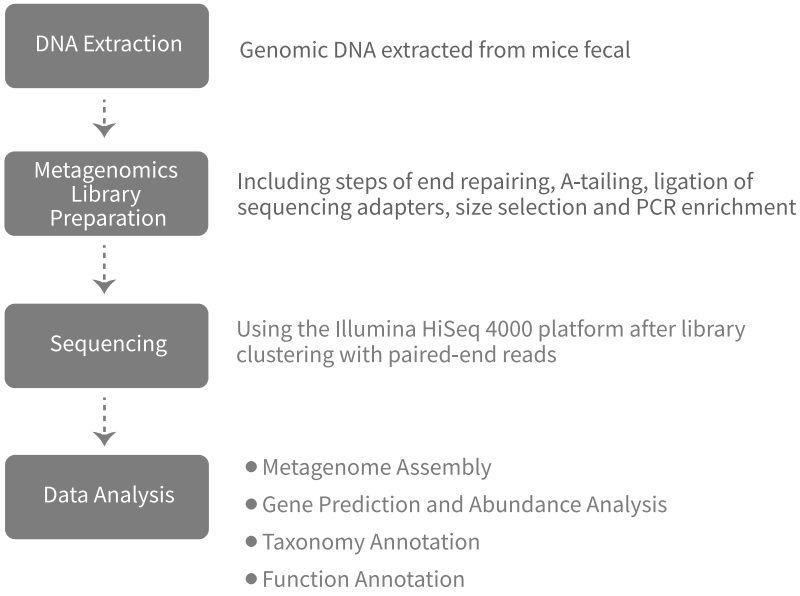
The overall findings from this study demonstrate that the severity of influence is closely related to the heterogeneous responses of the gut microbiota. Also, the researchers were able to identify that *B. animalis* has anti-influenza effects within the host and demonstrate that the endogenous gut population of *B. animalis* can increase when lethal influenza infection occurs, enhancing the ability of the host to resist influenza infection.

Novogene

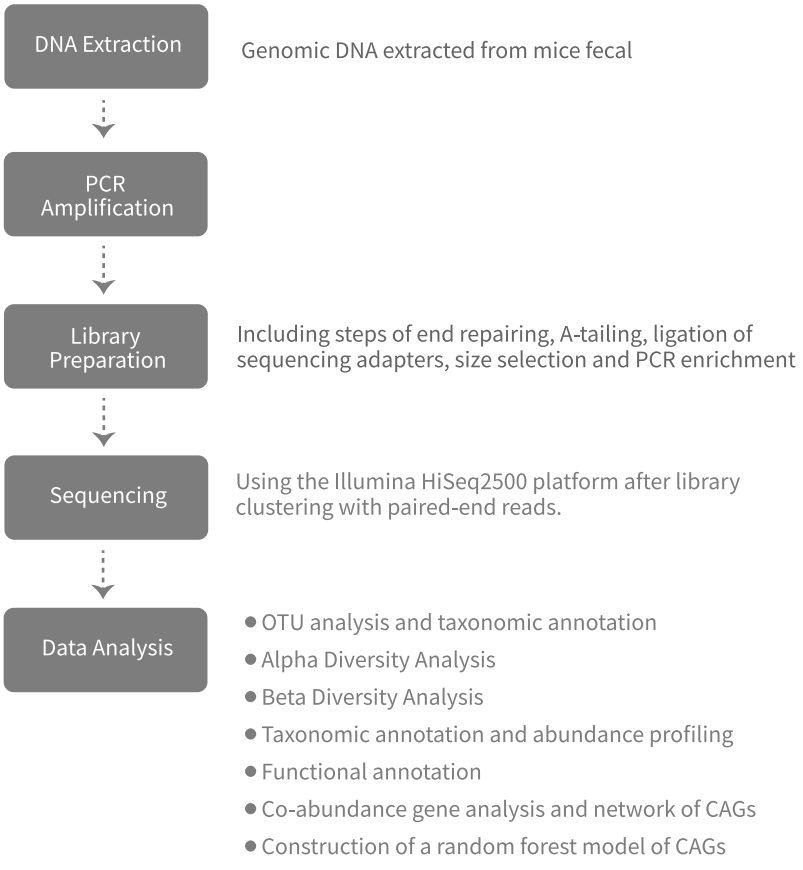
Novogene provides a range of services that can help researchers by supplying them with high-tech, reliable, and accurate genomic profiling technology. Our specialists have a wealth of experience in metagenomic sequencing, including shotgun **metagenomic sequencing** and amplicon metagenomic sequencing, and can advise you on the appropriate methods for your project, whether you wish to identify the microbiota of plants or animals (as seen in this study) or compare population structure in different geographic locations. The use of NGS technology enables the identification of hypervariable regions of conserved genes or intergenic regions in microorganisms and compare them against a database to quantify the population structure. Also, we provide cost-effective bioinformatics analysis using the latest software.

Novogene NGS technical pipelines

Shotgun Metagenomic Sequencing



Amplicon Metagenomic Sequencing



References

[1] Zhang, Q., Hu, J., Feng, JW. et al. Influenza infection elicits an expansion of gut population of endogenous *Bifidobacterium animalis* which protects mice against infection. Genome Biol 2020. **21**, 99 https://doi.org/10.1186/s13059-020-02007
[2] Davidson S. Treating Influenza Infection, From Now and Into the Future. *Frontiers in immunology*, 2018. 9, 1946. https://doi.org/10.3389/fimmu.2018.01946
[3] Cho, I., & Blaser, M. J. The human microbiome: at the interface of health and disease. *Nature reviews. Genetics*, 2012. 13(4), 260–270. https://doi.org/10.1038/nrg3182
[4] Ichinohe T, Pang IK, Kumamoto Y, et al. Microbiota regulates immune defense against respiratory tract influenza A virus infection. *Proc Natl Acad Sci U S A*. 2011;108(13):5354-5359. doi:10.1073/pnas.1019378108
[5] Brown RL, Sequeira RP, Clarke TB. The microbiota protects against respiratory infection via GM-CSF signaling. Nat Commun. 2017. 8(1):1512. doi: 10.1038/s41467-017-01803-x.

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

Follow us on LinkedIn