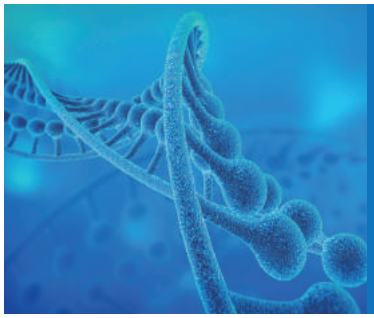
Novogene

Human Whole Genome Sequencing



Human whole genome sequencing enables researchers to catalog the genetic constitution of individuals and capture all the variants present in a single assay. It is commonly used in studies of a variety of diseases, especially cancer, as well as human population evolution, and pharmacogenomics.

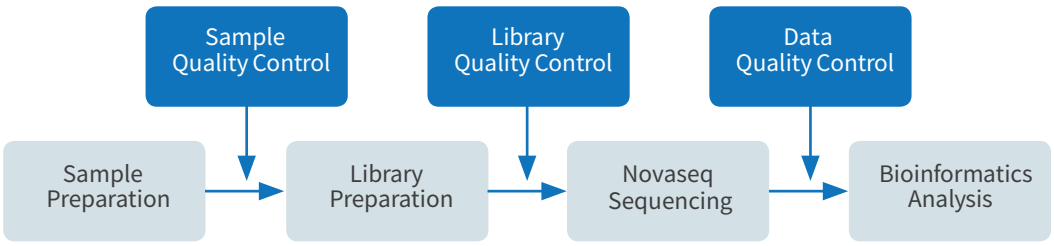
Equipped with the powerful Illumina NovaSeq 6000 system, Novogene is capable of sequencing up to 280,000 human genomes per year at the lowest cost per genome. With extensive experience in whole genome sequencing and advanced bioinformatics capabilities, Novogene is able to expertly meet customer needs for delivering large project results with quick turnaround times and the highest quality results.

Extensive experience with over 3000 projects

Data quality exceeds Illumina's official guarantee

In house pipeline to meet different analysis requirement

Project workflow



Sequencing parameter

Platform	Illumina NovaSeq 6000
Read length	Pair-end 150
Recommended Sequencing Depth	Recommended: Tumor tissues: 50~70X; Adjacent normal tissues or blood: 30X Rare diseases: 30~50X
Data quality	Guarantee Q30 ⩾80%
Turnaround time	23 working days from verification of sample quality to data delivery （<24 samples）

Samples requirement

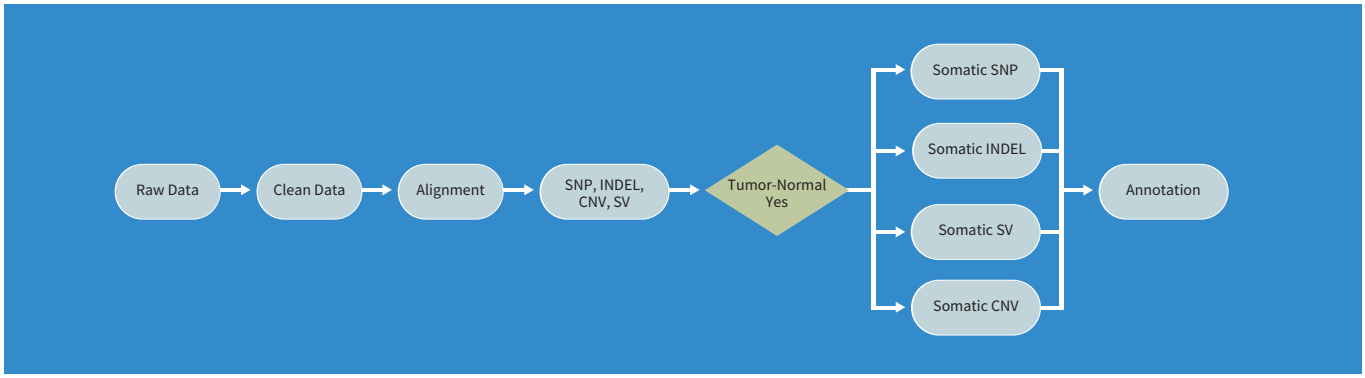
Library Type	Sample Type	Amount Required	Volume	Concentration	Purity
Short insert library (350 bp)	Genomic DNA	⩾200 ng	⩾ 20 μL	⩾ 10 ng/μL	OD260/280=1.8~2.0 No degradation, no contamination
	FFPE Genomic DNA	⩾800 ng	-	-	

NOVOGENE (UK) COMPANY LTD.
25 Cambridge Science Park
Milton Road
Cambridge, CB4 0FW
United Kingdom

Tel: +44(0)1223 628750
Eml: europe@novogene.com
Web: www.novogene.com
China · China Hong Kong · Singapore · UK · USA

Follow us on LinkedIn

Analysis Pipeline



Standard Analysis

- Data quality control
- Alignment with reference genome, statistics of sequencing depth and coverage
- SNP/InDel/SV/CNV calling, annotation and statistics
- Somatic SNP/InDel/SV/CNV calling, annotation and statistics (paired tumor samples)

Advanced Analysis

- Tumor evolution analysis (Cancer)
- Tumor neoantigen identification (Cancer)
- Candidate variant identification (Disease)
- Linkage analysis (Disease)
- CRISPR/Cas9 on-target and off-target detecting
- .....

Novogene Data

Below are Novogene data from human genome sequencing projects.

Sample name	Raw Data[1]	Effective (%) [2]	Q30 (%)[3]	Mapped (%) [4]	Average sequencing depth[5]	Genome coverage (%) [6]	Percentage of Genome with ⩾ 4X coverage (%) [7]	Percentage of Genome with ⩾ 10X coverage (%) [8]	Percentage of Genome with ⩾ 20X coverage (%) [9]
Novo 1	90.4	99.85	89.57	99.72	30.16	99.04	98.72	98.11	90.27
Novo 2	109.9	99.84	90.41	99.78	36.57	99.78	99.54	98.87	93.82
Novo 3	147.9	99.75	90.42	99.73	50.4	99.05	98.82	98.51	97.71
Novo 4	166.5	99.90	90.49	99.73	55.84	99.74	99.44	98.60	96.34
Novo 5	182.9	99.69	90.81	99.76	60.32	99.81	99.62	99.33	98.50

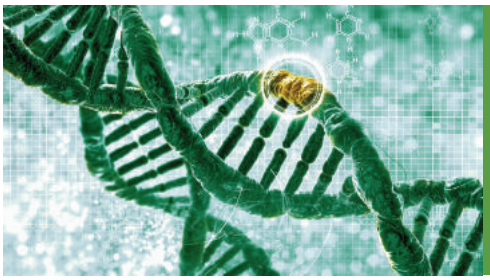
1 Original sequencing data (in gigabases).
2 Percentage of clean reads from all raw reads.
3 Percentage of reads with an average quality greater than Q30.
4 Percentage of total reads that mapped to the reference genome (UCSC hg38).
5 Average sequencing depth.
6 Percentage of genome covered by sequencing.
7 Percentage of bases in genome with a sequencing depth 4×.
8 Percentage of bases in genome with a sequencing depth 10×.
9 Percentage of bases in genome with a sequencing depth 20×.

Publications using Novogene's expertise

Year	Journal	Article
2019	Cell	Large-scale whole-genome sequencing of three diverse Asian populations in Singapore
2019	Lung Cancer	A study of ALK-positive pulmonary squamous-cell carcinoma: From diagnostic methodologies to clinical efficacy.
2019	BMC Genet	Characterization of the extra copy of TPOX locus with tri-allelic pattern.
2019	Oncogene	Hybrid sequencing-based personal full-length transcriptomic analysis implicates proteostatic stress in metastatic ovarian cancer.
2019	Anat Rec	An improved NGS library construction approach using DNA isolated from human cancer formalin-fixed paraffin-embedded samples
2018	Int J Cancer	Characteristics of genomic alterations of lung adenocarcinoma in young never-smokers.
2018	Cell Physiol Biochem	Genome evolution analysis of recurrent testicular malignant mesothelioma by whole-genome sequencing

Novogene

Human Whole Exome Sequencing



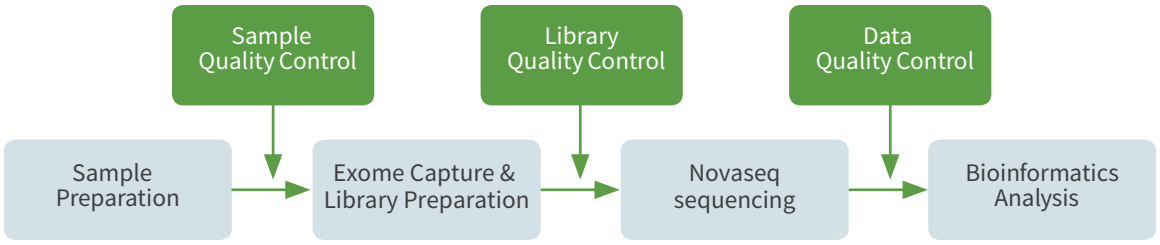
Exome sequencing provides a cost-effective alternative to whole genome sequencing, as it targets only the protein coding region of the human genome responsible for a majority of known disease-related variants. Whether you are conducting studies in rare mendelian disorders, complex disease, cancer research, or human population studies. Novogene' s comprehensive human whole exome sequencing (hWES) service provides a high-quality, affordable, and convenient solution.

Extensive experience with >3000 projects

Data quality exceeds Illumina's official guarantee

In house pipeline to meet different analysis requirement

Project workflow



Sequencing parameter

Platform	Illumina NovaSeq 6000
Read length	Pair-end 150
Recommended Sequencing Depth	For Mendelian disorder/rare disease: effective sequencing depth above 100× (12G) For tumor samples: effective sequencing depth above 200× (24G)
Data quality	Guarantee Q30 ⩾80%
Turnaround time	23 working days from verification of sample quality to data delivery （<24 samples）

Samples requirement

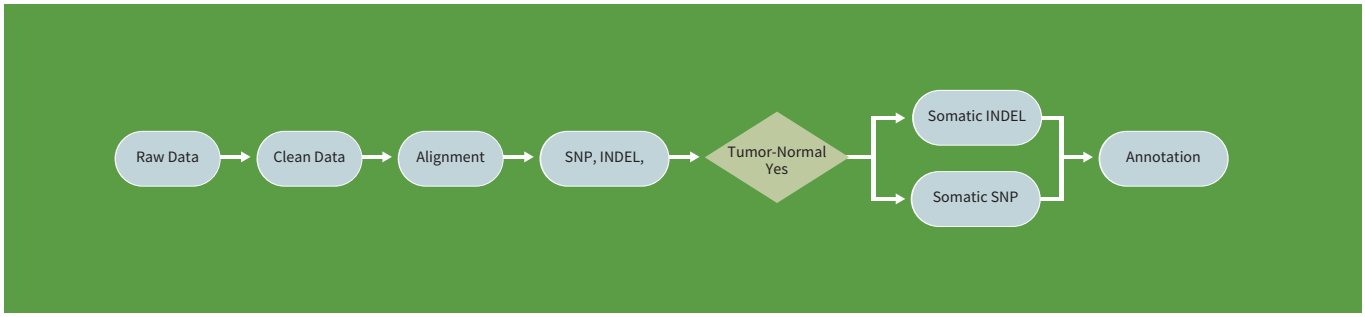
Library Type	Sample Type	Amount Required	Volume	Concentration	Purity
Short insert library (180-280 bp)	Genomic DNA	⩾400 ng	⩾ 20 μL	⩾ 20 ng/μL	OD260/280=1.8~2.0 No degradation, no contamination
	FFPE Genomic DNA	⩾800 ng	-	-	Fragments should be longer than 1000 bp

NOVOGENE (UK) COMPANY LTD.
25 Cambridge Science Park
Milton Road
Cambridge, CB4 0FW
United Kingdom

Tel: +44(0)1223 628750
Eml: europe@novogene.com
Web: www.novogene.com
China · China Hong Kong · Singapore · UK · USA

Follow us on LinkedIn

Analysis Pipeline



Standard Analysis

- Data quality control
- Alignment with reference genome, statistics of sequencing depth and coverage
- SNP/InDel/CNV calling, annotation and statistics
- Somatic SNP/InDel/CNV calling, annotation and statistics (paired tumor samples)

Advanced Analysis

- Tumor evolution analysis (Cancer)
- Tumor neoantigen identification (Cancer)
- Candidate variant identification (Disease)
- Linkage analysis (Disease)
- Xenograft tumor analysis (PDX)

…….

Novogene Data

Below are Novogene data from human genome sequencing projects.

Sample name	Raw Data[1]	Effective (%) [2]	Q30 (%)[3]	Mapped (%) [4]	Average sequencing depth[5]	Genome coverage (%) [6]	Percentage of Genome with ⩾ 4X coverage (%) [7]	Percentage of Genome with ⩾ 10X coverage (%) [8]	Percentage of Genome with ⩾ 20X coverage (%) [9]
Novo 1	6.5	98.17	88.74	99.72	67.3	99.7	99.4	98.1	92.2
Novo 2	9	98.99	90.20	99.86	91.54	99.9	99.7	99.2	97.1
Novo 3	12.3	98.57	93.19	99.88	117.46	99.9	98.5	98.2	94.8
Novo 4	15.2	98.71	93.22	99.81	179.56	99.6	99.3	98.5	96.6
Novo 5	18.7	98.96	93.53	99.85	188.92	99.8	99.5	98.6	96.4
Novo 6	19.8	98.81	92.41	99.84	215.3	99.6	99.4	98.8	97.5

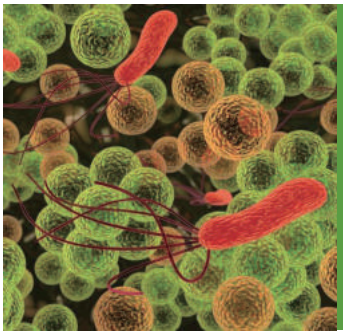
1 Original sequencing data (in gigabases).
2 Percentage of clean reads from all raw reads.
3 Percentage of reads with an average quality greater than Q30.
4 Percentage of total reads that mapped to the reference genome (UCSC hg38).
5 Average sequencing depth.
6 Percentage of genome covered by sequencing.
7 Percentage of bases in genome with a sequencing depth 4x.
8 Percentage of bases in genome with a sequencing depth 10x.
9 Percentage of bases in genome with a sequencing depth 20x.

Publications using Novogene's expertise

Year	Journal	Article
2019	Cancer Res	Multiregion sequencing reveals the genetic heterogeneity and evolutionary history of osteosarcoma and matched pulmonary metastases
2019	Neural Plasticity	New genotypes and phenotypes in patients with 3 subtypes of Waardenburg Syndrome identified by diagnostic next-generation sequencing.
2019	International Journal of Cancer	Preliminary exploration of potential molecular therapeutic targets in recurrent and metastatic parathyroid carcinomas.
2019	Eur Respir J	Germline BMP9 mutation causes idiopathic pulmonary arterial hypertension
2019	Nature Communications	Recurrent GNAQ mutation encoding T96S in natural killer/T cell lymphoma
2018	J Cell Mol Med	Whole-exome sequencing identifies a novel mutation of GPD1L
2018	Cell	Mutational landscape of secondary Glioblastoma guides MET-Targeted trial in brain tumor

Novogene

Amplicon-based metagenomic sequencing



Amplicon-based metagenomic sequencing is frequently used to identify and differentiate microbial species. Short (< 500 bp) hypervariable regions of conserved genes or intergenic regions are amplified by PCR, sequenced using NGS technology, and compared against microbial databases for taxonomic identification. Most commonly used targets are 16S rRNA gene of bacteria and archaea, 18S rRNA gene, or two internal transcribed spacers (ITS) of fungi.

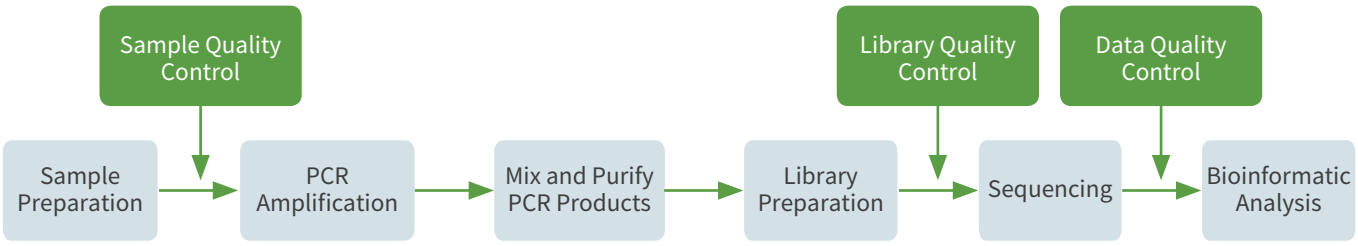
At Novogene, we have sequenced over 200,000 microbial samples for our customers. Our standard bioinformatics analyses include OTU analysis, species annotation, alpha-diversity analysis, beta-diversity analysis, and multi-variate statistical analysis. Applications range from identifying a single species in pure culture to characterizing the microbiota of animals or plants to comparing species diversity and population structure in various environmental sources or geographic regions. Our specialists can advise you on the appropriate analyses for your project.

Extensive experience with over 30,000 samples

Data quality exceeds Illumina's official guarantee

In house pipeline to meet different analysis requirement

Project workflow



Sequencing parameter

Platform	Illumina NovaSeq 6000 SP
Read length	Pair-end 250
Recommended Sequencing Depth	Recommended: 30K/50K/100K raw tags
Data quality	Guaranteed Q30 ⩾ 75%
Turnaround time	Within 3 weeks from project verification to data releasing without bioinformatic analysis (<24 samples)

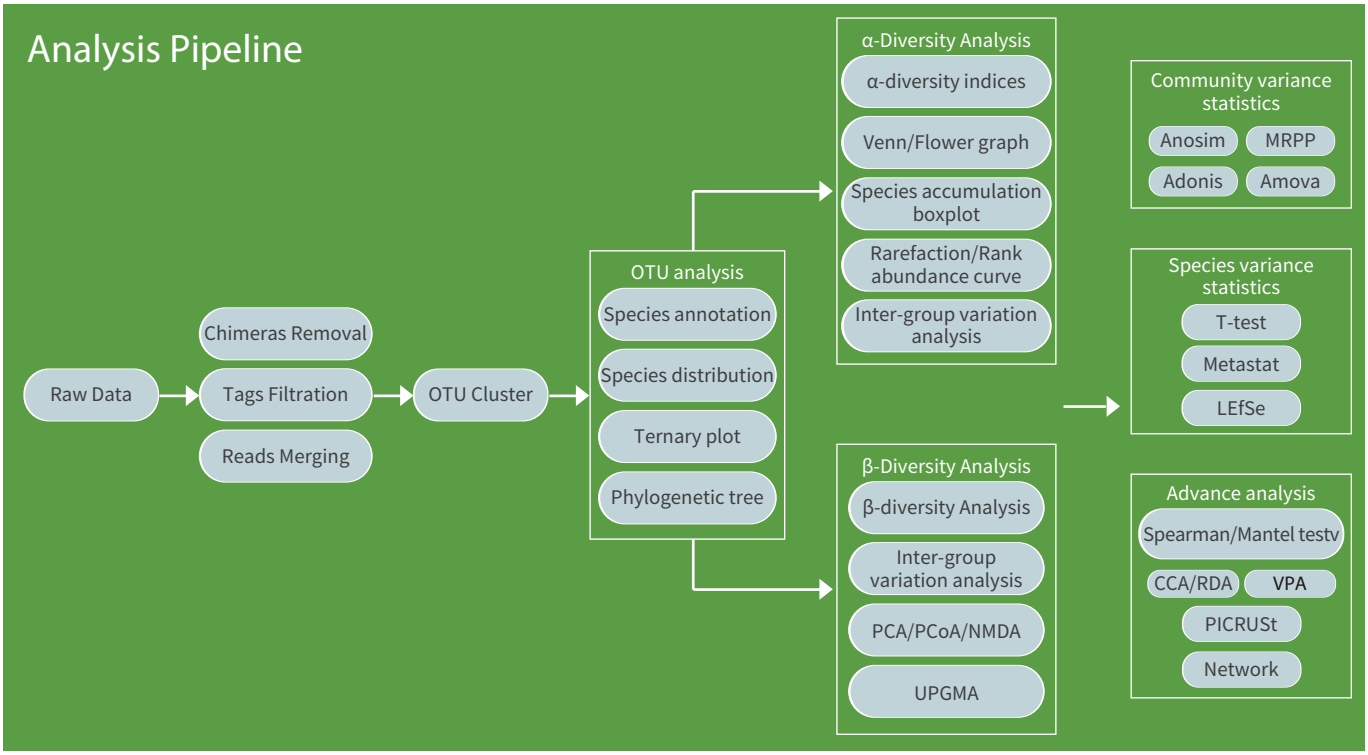
Samples requirement

Library Type	Sample Type	Required	Volume	Concentration	Purity (NanoDrop™)
Amplicon library	Genomic DNA	⩾ 200 ng	⩾ 12 μL	⩾ 20 ng/μL	OD260/280 = 1.8~2.0, no degradation, no contamination
	PCR Products (Fragment size: ⩽ 470 bp)	⩾ 200 ng (pooled samples/library)			
		⩾ 1.5 μg (one sample/library)			

NOVOGENE (UK) COMPANY LTD.
25 Cambridge Science Park
Milton Road
Cambridge, CB4 0FW
United Kingdom

Tel: +44(0)1223 628750
Eml: europe@novogene.com
Web: www.novogene.com
China · China Hong Kong · Singapore · UK · USA · Netherlands · Japan

Follow us on LinkedIn



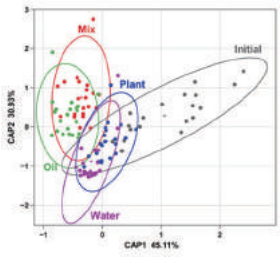
Project Example

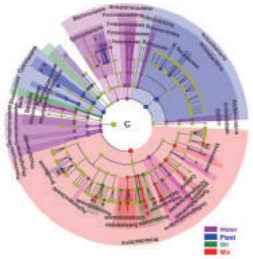
The following study utilized Novogene' s amplicon-based metagenomic sequencing services.

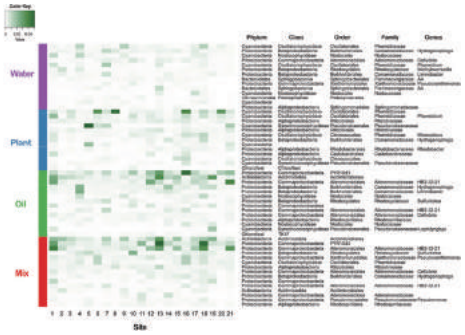
Plant growth and oil contamination alter the diversity and composition of bacterial communities in agricultural soils across China

Land Degrad Dev, 2018, 29:1660–1671.

The dynamics of microbial diversity in response to biotic and abiotic disturbances provide a sensitive indicator for evaluating the potential stability and degradation of soils in agro - ecosystems. To determine the effect on soil bacterial communities of disturbances by plant (Robinia pseudoacacia) growth and oil contamination, 16S rRNA genes were sequenced. Abiotic and biotic disturbances, including treating the soils with sterile water, crude oil, and/or an invasive plant, altered the bacterial community structure in the soils, increased bacterial richness, and reduced bacterial dispersion.



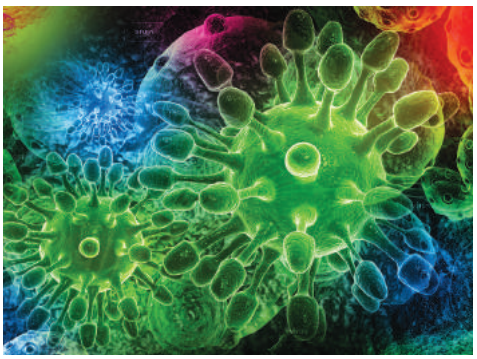




Publications using Novogene's expertise

Year	Journal	Article
2019	Front. Microbiol	Effect of fermented corn-soybean meal on serum immunity, the expression of genes related to gut immunity, gut microbiota, and bacterial metabolites in grower-finisher pigs
2019	The ISME Journal	Wastewater treatment plant resistomes are shaped by bacterial composition, genetic exchange, and upregulated expression in the effluent microbiomes
2019	Nature communications	Pathogen-targeting glycovesicles as a therapy for salmonellosis
2018	Front Microbiol	Metagenomics investigation of agarlytic genes and genomes in mangrove sediments in China: A potential repertory for carbohydrate-active Enzymes
2018	Microbiome	Gut-dependent microbial translocation induces inflammation and cardiovascular events after ST-elevation myocardial infarction
2017	Tumori	Performance comparison of NextSeq and Ion Proton platforms for molecular diagnosis of clinical oncology
2016	J Dairy Sci	Characterization of the indigenous microflora in raw and pasteurized buffalo milk during storage at refrigeration temperature by high-throughput sequencing

Shotgun Metagenomics



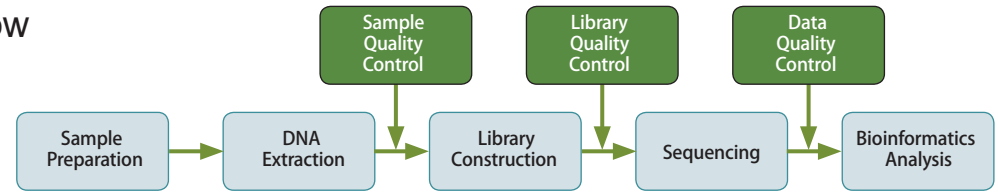
Shotgun metagenomics is a powerful technique for studying microbial communities in their natural habitat, with a broad range of applications. In shotgun metagenomics, genomes from environmental samples are analyzed without the prior isolation and cultivation of individual species.

Novogene helps you with expertise in NGS sequencing and bioinformatics analysis to explore the abundant genetic repertoire of microbial communities and to identify the species, genes, and pathways present in their samples. With shotgun metagenomics sequencing on Illumina platforms, assembly-first strategy and bioinformatics analyses, we could provide you with high-quality and publication-ready data including gene predictions, function annotations, and taxonomic annotations.

Novogene's Advantages

- Highly experienced: We have completed thousands of shotgun metagenomic sequencing and have published dozens of metagenomic studies together with our clients.
- Outstanding service: We provide high-quality sequencing, an efficient standard workflow, and bioinformatics analyses at a cost-effective price.
- Effective methodology: Our techniques enhance the generation of data from low-abundance species.
- Comprehensive analysis: Expert bioinformatics analysis with three databases (KEGG, eggNOG, CAZy) provides comprehensive data on annotated genes and metabolic pathways.

Project Workflow



SEQUENCING STRATEGY

- 350-bp insert DNA library
- Illumina platform, paired-end 150 bp.

TURNAROUND TIME

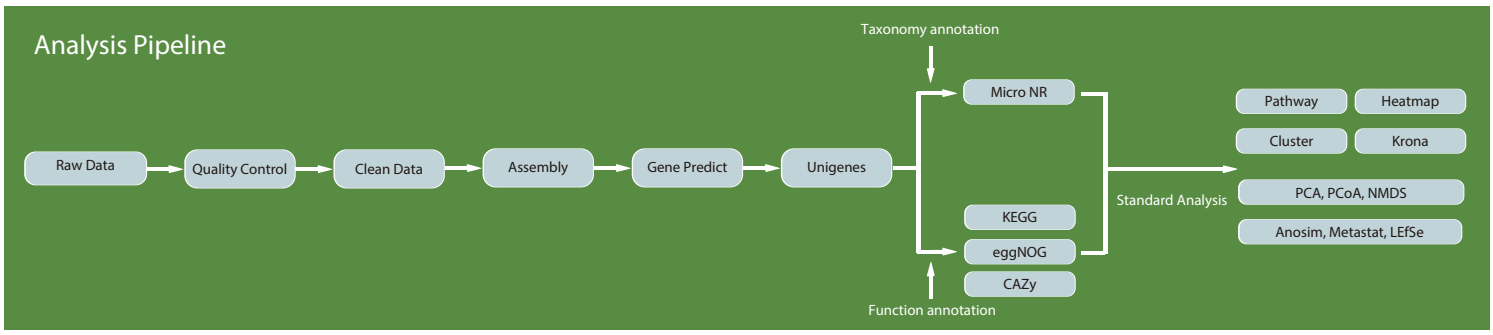
- Within 15 working days from verification of sample quality (without data analysis)
- Additional 10 working days for data analysis.

SAMPLE REQUIREMENTS

- DNA amount: ⩾ 200 ng
- DNA concentration: ⩾ 10 ng/μl
- Purity: OD260/280 = 1.8 - 2.0 without degradation or RNA contamination.

DATA QUALITY GUARANTEE

- Q30 ⩾ 80%.



NOVOGENE (UK) COMPANY LTD.
25 Cambridge Science Park
Milton Road
Cambridge, CB4 0FW
United Kingdom

Tel: +44(0)1223 628750
Eml: europe@novogene.com
Web: www.novogene.com
China · China Hong Kong · Singapore · UK · USA

Follow us on LinkedIn

List of Analyses

- Our standard analysis package includes gene prediction, function a nnotation, and taxonomic annotation, mPATH, heatmaps, PCA, Kr ona, cluster analysis, MetaStats, and OG-Taxa.
- Our advanced analysis package includes MRPP, ANOSIM, NMDS (No n-metric Multidimensional Scaling), CCA/RAD, and LEFSe (LDA Effe ct Size).
- For other analysis, please contact your local Novogene represen- tative or europe@novogene.com.

Demo Data

The following table shows sample data from sequencing projects conducted by Novogene. The effective rate, comparing clean data to raw data, was very high, with an average of over 94%, indicating that the base cal ling was highly accurate.

Sample	Insert Size (bp)	Raw Data	Clean Data	Clean Q20	Clean Q30	GC (%)	Effective Rate (%)
Test 1	300	5,491.78	5,273.46	94.06	88.54	52.17	96.03
Test 2	300	5,263.63	5,004.54	94.20	88.82	51.13	95.08
Test 3	300	5,471.88	5,090.86	93.74	87.90	54.05	93.04
Test 4	300	5,337.27	5,142.07	93.81	88.21	50.45	96.34
Test 5	300	5,781.12	5,700.68	95.97	91.90	42.47	98.61
Test 6	300	4,325.69	4,259.49	94.20	88.62	50.23	98.47

Project Example

The following study utilized Novogene's expert metagenomics services.

Impacts of the Three Gorges Dam on Microbial Structure and Potential Function.

Scientific Reports, 5:8605 (2015).

The Three Gorges Dam has significantly altered ecological and environmental conditions within the reservoir region, but how these changes affect bacterioplankton structure and function is unknown. Here, three widely accepted shotgun metagenomic tools were employed to study the impact of damming on the bacterioplankton community in the Xiangxi River. Our results indicated that bacterioplankton communities were both taxonomically and functionally different between backwater and riverine sites, which represent communities with and without direct dam effects, respectively. There were many more nitrogen cycling Betaproteobacteria (e.g., Limnohabitans), and a higher abundance of functional genes and KEGG orthology (KO) group sinvolved in nitrogen cycling in the riverine sites, suggesting a higher level of bacterial activity involved in generating more nitrogenous nutrients for the growth of phytoplankton. Additionally, the KO categories involved in carbon and sulfur metabolism, as well as most of the detected functional genes, also showed clear backwater and riverine patterns. As expected, these diversity patterns all significantly correlated with environmental characteristics, confirming that the bacterioplankton communities in the Xiangxi River were markedly affected by environmental changes caused by the Three Gorges Dam. This study provides the first comparative shotgun metagenomic insight into evaluating the impact of the large dam on microbial function.

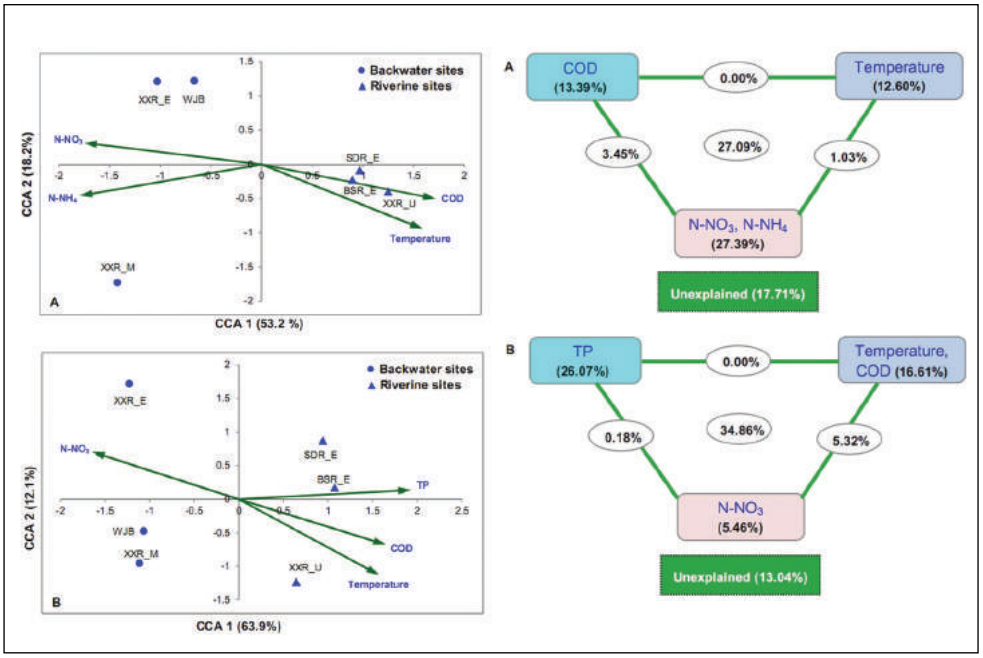


Figure 1. Canonical correspondence analysis (CCA) shows the relationships between environmental variables variables and the bacterial OTUs (A) and functional genes (B). Only variables that were significantly correlated with the community (forward selection with Monte Carlo test, P = 0.05) are shown. Abbreviations: TP, total phosphorus; N-NH4, ammonium nitrogen; N-NO3, nitrate nitrogen; COD, chemical oxygen demand.

Figure 2. Variance partitioning canonical correspondence analysis (CCA) shows the relative effects of multiple variables on the composition of bacterial taxa (A) and functional genes (B). The squares represent the effect of individual variables by partitioning out the effects of the other variables.The ellipses between the squares represent the combined effects from the variables on either side of the ellipse. The combined effects of all variables are shown by the ellipse in the center. The square at the bottom of each figure represents the effect that could not be explained by any of the variables tested. Abbreviations: TP, total phosphorus; N-NH4, ammonium nitrogen; N-NO3, nitrate nitrogen; COD, chemical oxygen demand.

EXAMPLES OF PUBLICATIONS USING NOVOGENE'S EXPERTISE

Year	Journal	Article
2018	Nature Communications	Hypoxia induces senescence of bone marrow mesenchymal stem cells via altered gut microbiota
2017	Microbiome	Gut microbiota dysbiosis contributes to the development of hypertension
2016	Water Research	Changes of resistome, mobilome and potential hosts of antibiotic resistance genes during the transformation of anaerobic digestion from mesophilic to thermophilic
2015	Water Research	Metagenomic insights into Cr(VI) effect on microbial communities and functional genes of an expanded granular sludge bed reactor treating high-nitrate wastewater
2014	Environment Science and Technology	Prevalence of antibiotic resistance genes and bacterial pathogens in long-term manured greenhouse soils as revealed by metagenomic survey