SERVICES CATALOGUE

BGI

We Sequence. You Discover

Neo-Science

Neo Science Equipments & Chemicals Trading LLC

Bin Ghanim Tower, 10th floor, Office No 1002, Hamdan Street, Abu Dhabi, UAE

Contents

BGI Profile

Company Profile

BGI Genomics is one of the world's leading providers of genomic sequencing services and proteomic services, now serving customers in more than 66 countries. We provide academic institutions, pharmaceutical companies, health care providers and other organizations with integrated genomic sequencing and proteomic services and solutions across a broad range of applications spanning:

- Basic research covering human, plant, animal and microbial species
- Clinical research in human health
- Genetic testing and screening
- Agriculture and Biodiversity preservation and sustainability

We have more than 20 years of genomics experience helping our customers achieve their research goals by delivering rapid, high quality results using a broad array of cost-effective, cutting-edge technologies. Contact us to find out how we can assist with your project, we're confident we can help.

info@bgi.com
www.bgi.com

Platform Introduction

NGS Platforms

We provide services on a wide range of the industry's latest sequencing platforms to support the diverse research needs of our customers. Our global scale and 20 year experience enables quality services at market leading pricing:

- DNBSEQTM Platforms (MGI)
- Nanopore Platforms (Oxford Nanopore)
- Sequel Platforms (PacBio)
- HiSeq Platforms (Illumina)

DNBSEQ™-T7

MGISEQ-2000

Nanopore(PromethION)

Nanopore(MinION)

PacBio（Sequel Ⅰ,Ⅱ）

HiSeq 2500

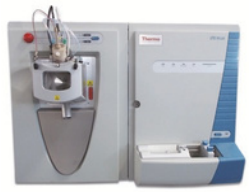
Mass Spectrometry Platforms

With more than 10 years of development, BGI has formed professional research platforms for proteomics, polypeptide omics and metabolomics based on high-throughput mass spectrometry detection technology. By integrating the advantages of high throughput sequencing research platforms with genomics and transcriptomics, BGI's mass spectrometry services provide partners with a complete research solution that enables multi-cluster correlation analysis. Our portfolio includes several high-precision and high-resolution mass spectrometers, which can carry out large-scale proteomics, metabolomics and detection for comprehensive research of target molecules.

• Orbital ion trap mass spectrometer

Orbitrap Fusion Lumos

Q-Exactive/ Q-Exactive HF/
Q-Exactive HF-X



LTQ Orbitrap velos

Eclipse

Exploris 480

• Triple quadruple pole mass spectrometer

QTRAP 6500+

TQS/TQD

TSQ Altis

TimsTOF

• Time of flight mass spectrometer (TOF)

TripleTOF 5600+

Xevo-G2-XS

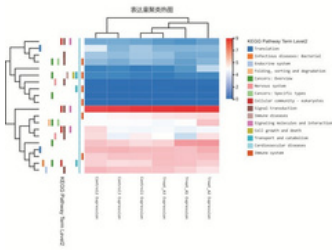
MicroTyper MS

Ultrafle Xtreme

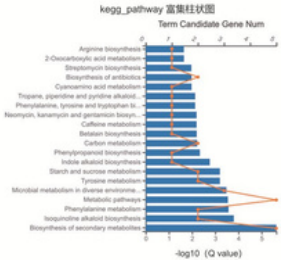
Dr.Tom System

Dr.Tom is a web-based solution for the convenient analysis, visualisation and interpretation of all types of RNA data, including small RNAseq, miRNA and lncRNA.
Designed by a team of expert RNA scientists and bioinformaticians at BGI with collective experience across thousands of RNA based research projects, Dr.Tom provides a wide range of intuitive and interactive data visualisation tools specifically designed to save users time in their differential expression or pathway analysis research.
In addition, powerful analysis tools and advanced algorithms allow users to mine their data to gain new insight and more value beyond standard available RNA analysis services.
Data from many of the world's leading databases have been integrated into the Dr.Tom system allowing users to reference and cross check all results and findings.
Dr.Tom is already relied upon by hundreds of scientists and researchers, and has shown itself to be a valuable and important tool in addition to any institution's own internal data curation and analysis efforts.

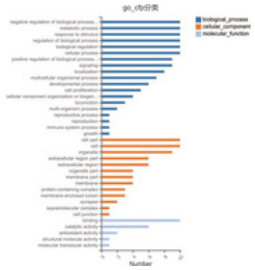
• Analysis Functionalities



Heat Map

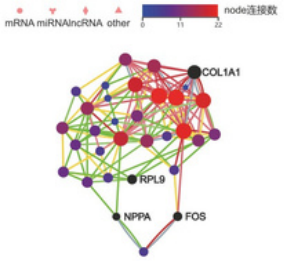


Bubble/Column Enrichment Charts

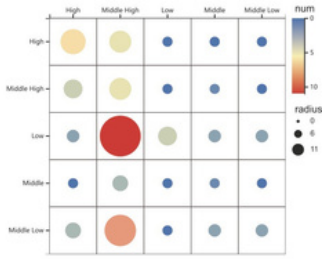


Classification Map

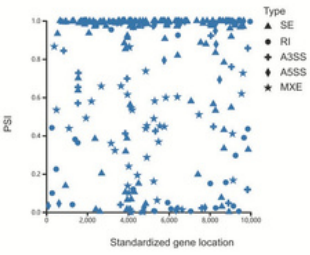
KDA



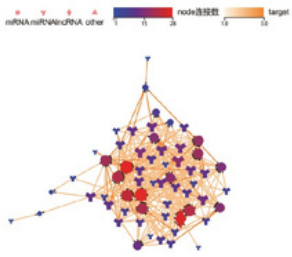
Interaction Charts



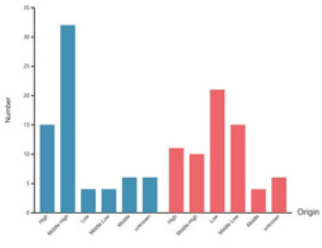
Chi-square Tests



Alternative Splicing Analysis



Relation Net



Multiomics Chart

Quality Management and Certifications

BGI's laboratories have a broad range of external quality and clinical certifications guaranteeing reliability and quality of service, including ISO9001, ISO14001, ISO45001, ISO/IEC 27001, ISO/IEC 17025 and CAP.

DNA Sequencing

01
02

Plant/Animal

Metagenomics

Human

Epigenetics

Plant and Animal *De novo* Sequencing

BGI Quick Facts

De novo sequencing refers to the sequencing of a novel genome without a reference sequence for alignment. The process of *de novo* genome sequencing involves the sequencing of small/large DNA fragments, assembling the reads into longer sequences (contigs) and finally ordering the contigs to obtain the entire genome sequence.
BGI is a recognized leader in *de novo* Whole Genome Sequencing and has extensive experience from the *de novo* sequencing and assembly of more than 400 species genomes.
We offer a complete suite of technologies to support your *de novo* sequencing projects, along with expert assistance in the planning of optimal sequencing and bioinformatics options, to ensure your project is a success.

Sequencing Service Specification

BGI plant and animal *de novo* sequencing services are executed utilizing multiple sequencing systems.

Sample Preparation and Services

- Library preparations (DNBSEQ™/Illumina, Nanopore PromethION, PacBio Sequel II etc)
- Various sequencing modes
- Raw data, standard and customized data analysis
- Available data storage and bioinformatics applications

Sequencing Quality Standard

- Guaranteed ≥90% of DNBSEQ™ clean bases with quality score of Q20
- Guaranteed ≥50Gb Nanopore pass data with Q>7
- Guaranteed ≥100Gb PacBio Sequel II CLR data Guaranteed ≥20Gb PacBio Sequel II CCS (HiFi library) data with accuracy greater than 99% except some complex species

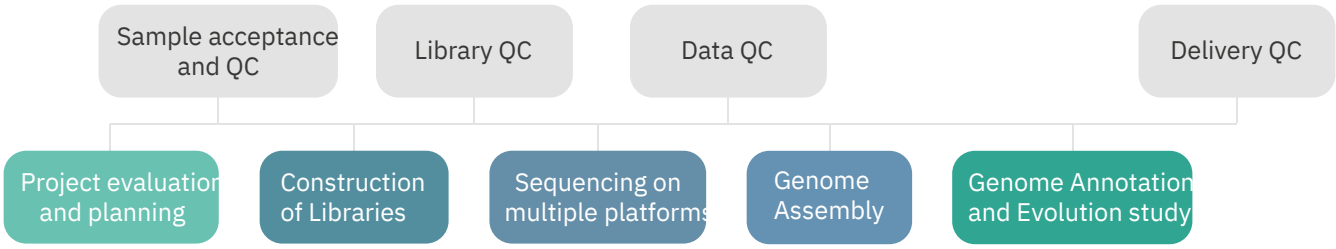
Turnaround Time For Species (genome size ≤ 5Gb)

- 70 working days from sample QC acceptance to filtered data availability
- 40/70 working days for the bioinformatics of common/complex genome assembly
- 30 working days for the bioinformatics of genome annotation
- 30 working days for the bioinformatics of genome annotation

Case by case for the species with genome size > 5Gb

Project Workflow

We care for your samples from the start through to the result reporting. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.



Sequencing Strategy

De novo sequencing usually requires a customized approach based on your subject species' genome size and complexity as well as the overall scientific objectives of the project.
Our plant and animal *de novo* sequencing services are usually performed using a combination of available platforms, including proprietary DNBSEQ™ NGS platforms augmented with Nanopore PromethION, PacBio Sequel II, Hi-C platforms for sequencing, library preparation and mapping. In addition, BGI offers extensive bioinformatics data analysis options for genome assembly, annotation and evolution studies.

Platform Tools	Library Type	Sequencing /USE	Recommend Sequencing Depth
DNBSEQ™	350bp Library	PE100/PE150	≥60X
	stLFR Library	PE150	≥100X
	Hi-C Library	PE150	≥100X
Nanopore PromethION	20K-40Kb Library	Read length ≥10Kb	≥100X
	Ultra long library (>40K)	Read length ≥20Kb	≥100X
PacBio Sequel II	15K-20Kb CCS (HIFI) Library	Read length ≥10Kb	≥50X
	20K-40K CLR library	Read length ≥20Kb	≥100X
	40K-60Kb CLR Library	Read length ≥20Kb	≥100X

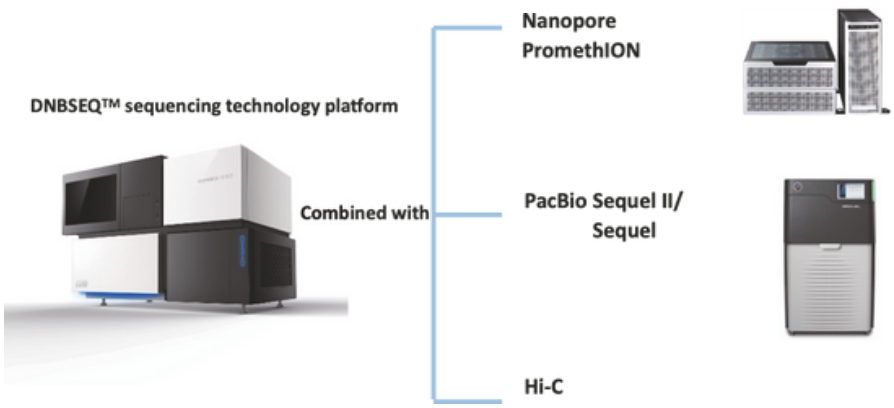
1. Packaging strategy for common genome:

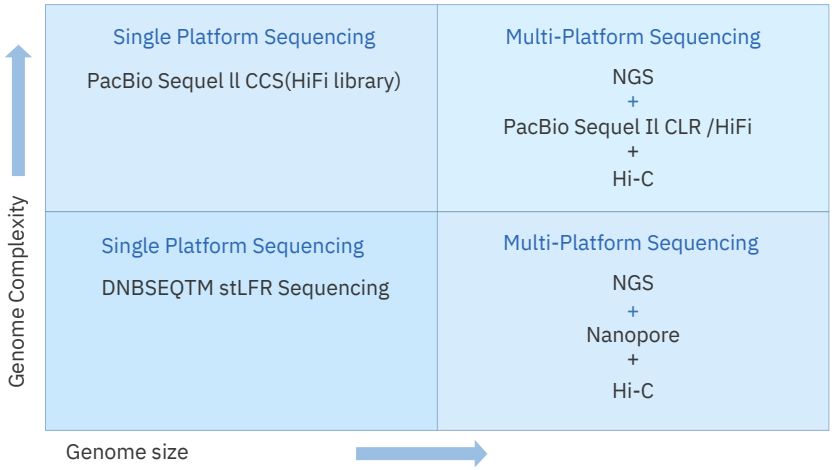
1) Nanopore PromethION 100X+ DNBSEQTM 60X;
2) PacBio Sequel II CLR 100X+ DNBSEQTM 60X;
3) stLFR 100X;3) PacBio Sequel II HiFi 50X;

2.Packaging strategy for highly heterozygous genome:

1) Nanopore PromethION 150X+ DNBSEQTM 60X;
2) PacBio Sequel II CLR 150X+ DNBSEQTM 60X;

Our sequencing specialists will work with you to design optimal strategies for your project, using platform combinations as appropriate for your project





Data Analysis

Besides clean data output, BGI offers a range of standard and customized bioinformatics pipelines for your plant and animal *de novo* sequencing projects.
Reports and output data files are delivered in industry standard file formats, of which BAM, .xls, .png. Raw FASTQ and FASTA data are available.

Genome Survey	1.Kmer estimation (Jellyfish + GenomeScope); 2. External pollution Analysis (BWA);
Genome Assembly (PacBio HiFi reads)	1. Assembly; 2.Assessment by short reads alignment; 3.BUSCO assessment;
Genome Assembly (PacBio CLR/Nanopore reads +NGS)	1. Reads correction; 2. Assembly; 3. Assembly result correction using long reads; 4. Assembly result correction using short reads; 5. BUSCO assessment;
Gene Annotation	1.Repeat annotation; 2.Gene prediction; 3.Gene function Annotation;
Evolution	1.Gene family identification (Animal TreeFam; Plant OrthoMCL; ≤10 species); 2.Phylogenetic tree construction; 3.Estimation of divergence time; 4.Genome synteny analysis; 5.Whole genome duplication analysis; 6.Gene family expansion and contraction analysis; 7. Positive selection analysis
Auxiliary Assembly	Hi-C data auxiliary assembly
stLFR Assembly	Genome assembly using stLFR data

Sample Requirements

We can process the DNA samples of plants and animals with the following general requirements (actual sample requirements for each specific project will depend on the number and type of libraries to be constructed). BGI also provides special sample extraction services to satisfy project requirements

Plant and Animal Genome ***De novo*** Sequencing (Genomic DNA)					
Platform	Sample type	Mass	OD	Integrity (AGE)	Sample Purity
Nanopore PromethION	20K-40Kb library	≥ 9μg	OD260/280: 1.6-2.2 OD260/230: 1.6-2.2	No degradation or little degradation with main band≥40Kb	No contamination with RNA, protein or salt ions; colorless and transparent; non-sticky
	Ultra long library (>40K)			No degradation or little degradation with main band≥50Kb	
	15K-20Kb HIFI library 20K-40Kb	≥ 15μg		No degradation or little degradation with main band≥30Kb	
PacBio Sequel II	CLR library 40K-60Kb	≥ 7μg	OD260/280: 1.6-2.2 OD260/230: 1.6-2.5	No degradation or little degradation with main band≥40Kb	
	CLR library stLFR library	≥ 9μg		No degradation or little degradation with main band≥50Kb	
DNBSEQ™	350bp library	≥500 ng		No degradation or little degradation with main band≥40Kb	
		≥1μg	-	No degradation or little degradation with main band≥20Kb	

Examples of recent ***de novo*** projects executed by BGI

Species	Heterozygosity	Genome Size	Sequencing Platform	Sequencing reads N50	ContigN50
Plant	0.89%	850M	Nanopore	22KB	7.3MB
Plant	0.20%	2.4G	Nanopore	30KB	23.7MB
Plant	0.80%	400M	Nanopore	30KB	17.3MB
Plant	1%	1.1G	Nanopore	23KB	6.5MB
Plant	1.10%	10G	Nanopore	30KB	1.6MB
Plant	0.38%	550M	Nanopore	21KB	10MB
Animal	0.30%	3G	PacBio CLR	22KB	27.8MB
Plant	0.30%	4G	PacBio CLR	20KB	6.6MB
Plant	0.10%	1.1G	PacBio CLR	20KB	17.0MB
Plant	0.40%	1.5G	PacBio CLR	25KB	1.5MB
Autopolyploid plant	3%	3G	PacBio HIFI	17KB	4.8MB
Plant	0.90%	650M	PacBio HIFI	15KB	8.6MB
Plant	0.80%	2.4G	PacBio HIFI	17KB	48.4MB
Animal	1.30%	1.4G	PacBio HIFI	18KB	7MB

Plant and Animal Whole Genome Re-Sequencing

Service Description

Plant and animal whole genome re-sequencing (WGRS) involves sequencing the entire genome of a plant or animal and comparing the sequence to that of a known reference genome. Re-sequencing of the plant and animal genome will identify genetic variations such as SNPs and Indels and elucidate other genetic changes of the sequenced species. WGRS is commonly used for the identification of functional genes and markers of important traits to facilitate molecular breeding and to improve agricultural production and conservation.

Highlights of Our Service Using DNBSEQ™ Technology

- Even coverage of reads
- Much less duplication
- True PCR-Free
- Index hopping free

Sequencing Service Specification

BGI Plant and Animal Whole Genome Re-Sequencing services are executed with DNBSEQ™ technology platforms.

Sample Preparation and Services

- PCR and PCR-Free library methods are available
- 100bp and 150bp paired-end sequencing available
- Raw data, standard and customized data analysis
- Available data storage and bioinformatics applications

Sequencing Quality Standard

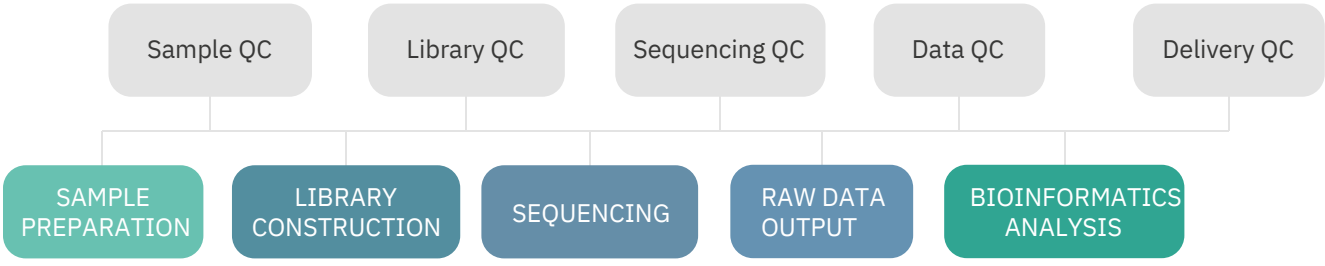
- Guaranteed ≥90% of clean bases with quality score of Q20
- Standard sequencing coverage of 10-30X is recommended for the study of individuals and 5-10X for population studies

Turnaround Time

- Typical 30 working days from sample QC acceptance to filtered raw data availability
- Expedited services are available; Contact your local BGI specialist for details

Project Workflow

We care for your samples from the start through to the result reporting. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.



Data Analysis

Besides clean data output, BGI offers a range of standard and customized bioinformatics pipelines for your whole genome re-sequencing project.
Reports and output data flies are delivered in industry standard file formats: FASTQ, BAM, VCF, .xls, .png

Standard Analysis	Library type
• Data Filtering • Alignment • Assembly of consensus sequences • SNP, InDel, SV, CNV calling, annotation and statistics	• Population evolution analysis • Point mutation detection (wild vs. mutant) • Linkage map construction and QTL mapping • GWAS analysis • BSA analysis

Customized Analysis
• Further customization of bioinformatics analysis to suit your unique project is available. • Please contact your BGI technical representative for details.

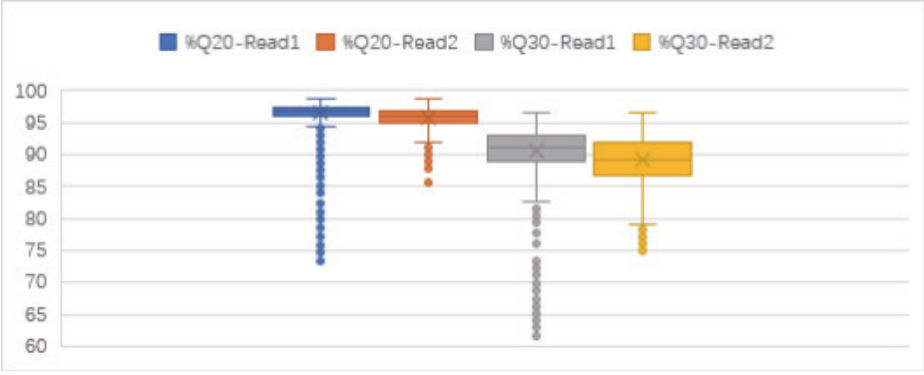
Sample Requirements

We can process your gDNA, whole blood, cell line, fresh frozen tissue and FFPE samples from a variety of species, with the following general requirements:

	DNA Amount and Concentration	Minimum Sample Volume
Regular Samples	Intact genomic DNA 1μg, Concentration12.5ng/μl	15μl
True PCR-Free	Intact genomic DNA 1.5μg, Concentration12.5ng/μl	15μl

High-Quality Data Performance

13,992 samples were delivered in the last two years. The data output and data quality remained very stable and high quality. The average Q20 and Q30 scores were 96.16% and 89.86% respectively for PE150 reads.



Data quality scores for plant and animal re-sequencing from DNBSEQ™ production line

Human Whole Genome Sequencing

Service Description

Whole Genome Sequencing (WGS) determines the complete human genome sequence all at once and provides the most comprehensive collection of an individual's genetic variation based on the human reference genome.
WGS can be applied to human genetics and evolution studies to detect genome-wide genetic variations, pathogenic and susceptibility genes, and to enable genetic diversity and evolution analysis. It can also be applied in translational research to provide information on cancer and other disease-associated mutations. It is one of the most important tools for precision medicine.
BGI offers a comprehensive range of WGS services for many sample types and a range of coverage levels.

Sequencing Service Specification

BGI Human Whole Genome Sequencing Services are performed with DNBSEQ™ sequencing technology, featuring DNA Nanoballs, linear Rolling Circle Replication, and cPAS technology for superior data quality.

Sample Preparation and Services

- PCR and PCR-free library preparation methods are available
- 100bp and 150bp paired end sequencing options
- Choice of sequencing depth: standard (~30x), deep (~60x) and low pass (less than 10x)
- Raw data, standard and customized bioinformatics analysis available
- Available data storage and bioinformatics applications
- CAP/CLIA laboratory services available

Sequencing Quality Standard

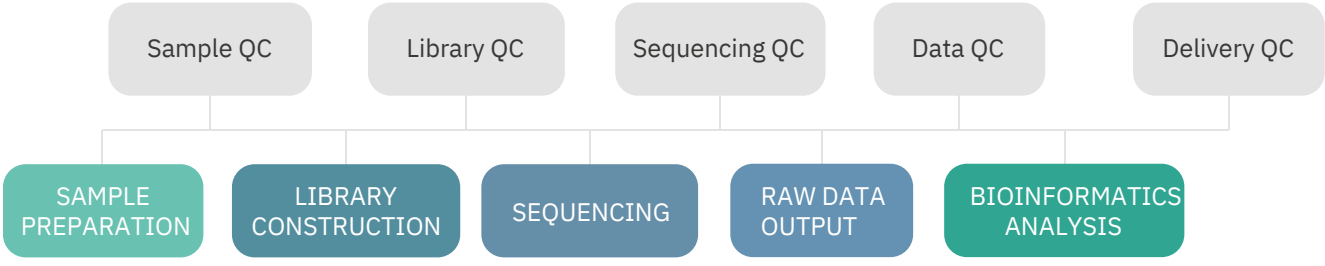
- Guaranteed ≥85% of bases with quality score of ≥Q30

Turnaround Time

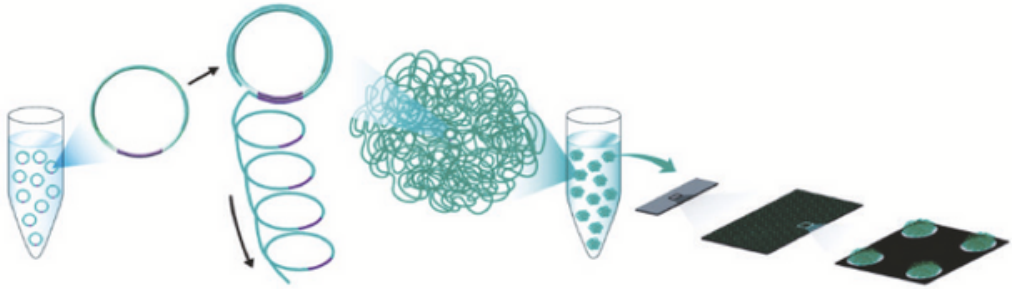
- Standard Turnaround: typical 18 working days from sample QC acceptance to filtered data availability
- Rapid Turnaround: 10 working days from sample QC acceptance to filtered data availability
- Expedited services are available for all WGS sequencing options, contact your local BGI specialist for details

Project Workflow

We care for your samples from the start through to the result reporting. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.



DNBSEQ™ is an innovative high-throughput sequencing technology, developed by BGI's Complete Genomics subsidiary in Silicon Valley. The system is powered by combinatorial Probe-Anchor Synthesis (cPAS), linear isothermal Rolling-Circle Replication and DNA Nanoballs (DNB™) technology, followed by high-resolution digital imaging.



The combination of linear amplification and DNB technology reduces the error rate while enhancing the signal. The size of the DNB is controlled in such a way that only one DNB is bound per active site on the flow cell. This densely patterned array technology provides optimal sequencing accuracy and increases flow cell utilization.

Data Analysis

In addition to clean data output, BGI offers a range of standard and customized bioinformatics pipelines for your human WGS project.
Reports and output data files are delivered in industry standard FASTQ, BAM. Excel formats with publication-ready tables and figures.

STANDARD BIOINFORMATICS ANALYSIS	CUSTOM ANALYSIS
• Filtering • Alignment • SNP calling and annotation • SNP validation and comparison • SNP functionality and conservation prediction • SNP Statistics per functional element • InDel calling and annotation • InDel validation and comparison • InDel statistics per functional element • CNV calling and annotation • SV calling and annotation	• Further customization of Bioinformatics analysis to suit your unique project is available; • Please contact your BGI technical representative.

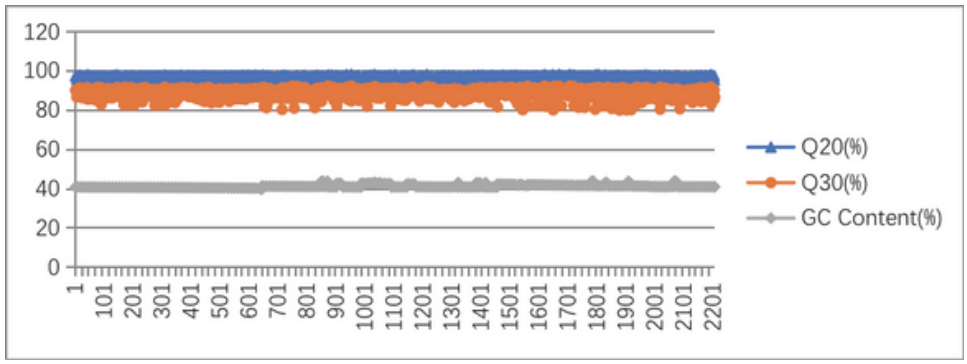
Sample Requirements

We can process your gDNA, saliva, blood, fresh frozen tissue, cell pellets and FFPE samples, with the following general requirements:

	Library type	DNA Amount and Concentration	Minimum Sample Volume
Regular Samples	PCR	Intact genomic DNA ≥1μg Concentration ≥12.5ng/μl	15 μl
	PCR-free	Intact genomic DNA ≥2μg Concentration ≥12.5ng/μl	15 μl
Low Input Samples	PCR	Intact genomic DNA ≥200ng Concentration ≥2.5ng/μl	15 μl

Stable and High-Quality Data Performance

2,206 Samples were randomly selected from over 20,000 samples that were sequenced at BGI's laboratories over a period of 6 months. The data output and data quality remained very stable over that period. The average Q20 and Q30 scores were 96.78% and 88.81% respectively. The average GC content was 41.25% without obvious base bias.



Stable Data quality scores and GC content from our DNBSEQ™ production line

Whole Exome Sequencing

Service Description

For many applications, Whole Exome Sequencing is a viable and cost-effective alternative for Whole Genome Sequencing. BGI has performed professional exome sequencing services for many years, to support human and mouse research and to benefit small and large-scale clinical trials and pharmaceutical drug development projects. Besides raw sequencing data output, BGI offers standard and custom bioinformatics services to suit your specific research needs.

Sequencing Specification

BGI Human Exome Sequencing Services are performed with DNBSEQ™ sequencing technology platforms, featuring cPAS and DNA Nanoballs(DNB™) for superior data quality.

Sample Preparation and Services

- Agilent SureSelect or IDT exome kit for library construction and enrichment, 100bp/150bp paired-end sequencing options available
- Clean data and advanced bioinformatics analysis are available in standard file formats
- Standard and custom bioinformatics data analysis
- Available data storage and bioinformatics applications

Sequencing Quality Standard

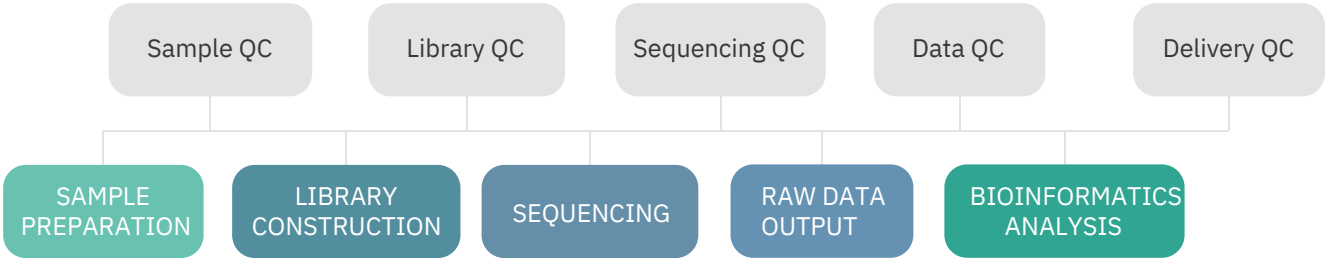
- Guaranteed ≥80% of bases with quality score of ≥Q30
- Standard sequencing coverage ≥100X; ≥200X is recommended for cancer samples

Turnaround Time

- Typical 25 days after sample acceptance for data delivery
- 10 days for rapid delivery service

Project Workflow

We care for your samples from the start through to the result reporting. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.



Data Analysis

Besides clean data output, BGI offers a range of standard and customized bioinformatics pipelines for your whole exome sequencing project.
Reports and output data files are delivered in industry standard file famats: BAM, xls, png.

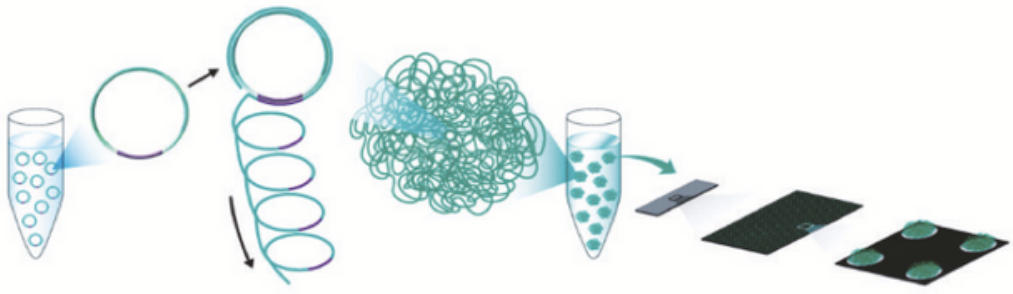
STANDARD CANCER MUTATION ANALYSIS	ADVANCED CANCER MUTATION ANALYSIS
• Data Filtering and QC • Align reads to the human reference genome • Germline SNP/InDel detection • Somatic SNV detection • Somatic InDel detection • Somatic CNV detection • Susceptibility gene screening • Verification of homology of paired samples • Tumor purity and ploidy analysis • Mutation spectrum and mutation signature analysis	• High frequency mutation gene statistics and pathway enrichment analysis • High frequency mutation gene correlation analysis (MRT) • Visual display of mutations • Driving gene prediction • Single-sample tumor cloning analysis • Tumor mutation burden analysis • Microsatellite instability

AVAILABLE ADVANCED ANALYSIS	
• Population genetics analysis • Complex disease analysis	• *De novo* mutation analysis for family sample • Tumor neoantigen prediction

Further customization of Bioinformatics analysis to suit your unique project is available.
Please contact your BGI technical representative.

DNBSEQ™ Sequencing Technology

DNBSEQ™ is an innovative high-throughput sequencing technology, developed by BGI's Complete Genomics subsidiary in Silicon Valley. The system is powered by combinatorial Probe-Anchor Synthesis (cPAS), linear isothermal Rolling-Circle Replication and DNA Nanoballs (DNB™) technology, followed by high-resolution digital imaging.



The combination of linear amplification and DNB technology reduces the error rate while enhancing the signal. The size of the DNB is controlled in such a way that only one DNB is bound per active site on the flow cell. This densely patterned array technology provides optimal sequencing accuracy and increases flow cell utilization.

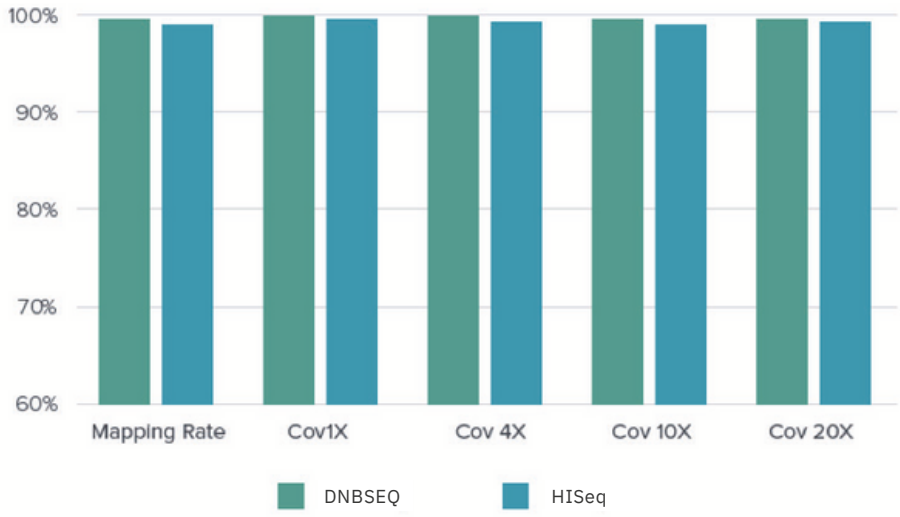
Sample Requirements

We can process your gDNA, Blood, Cell line, Fresh frozen tissue samples from a variety of species, with the following general requirements:

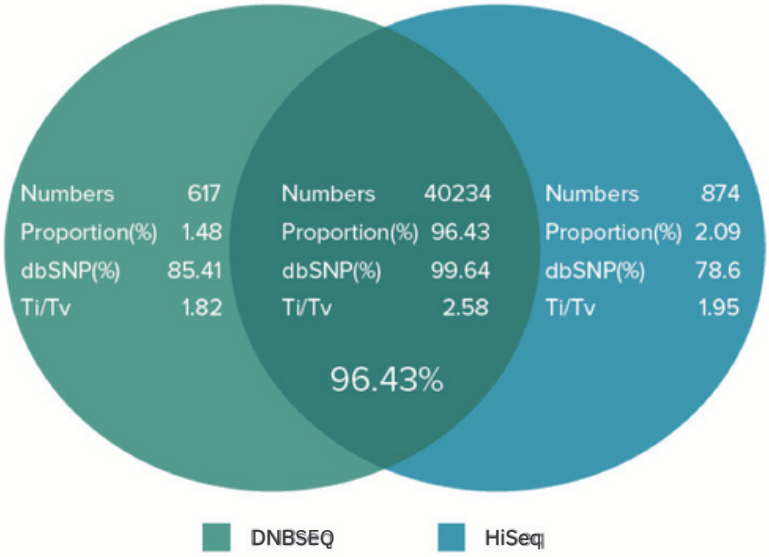
Exon Sequencing				
Sample type	Mass	Concentration	Integrity (AGE)	Sample Purity
Genomic DNA	≥1μg	≥2.5ng/μL	The band shown on gel electro-phoresis has little degradation, or of fragment size greater than 20kb.	No contamination with RNA, protein or salt ions; colorless and transparent; non-sticky.
FFPE DNA	≥100ng	≥2.5ng/μL	Fragment size greater than 500bp	No contamination with RNA, protein or salt ions; colorless and transparent; non-sticky.

Data Performance

Following is an example of typical DNBSEQ™ data output for a 100X WES project with standard sample NA12878, compared with data from the Illumina HiSeq 4000 system.

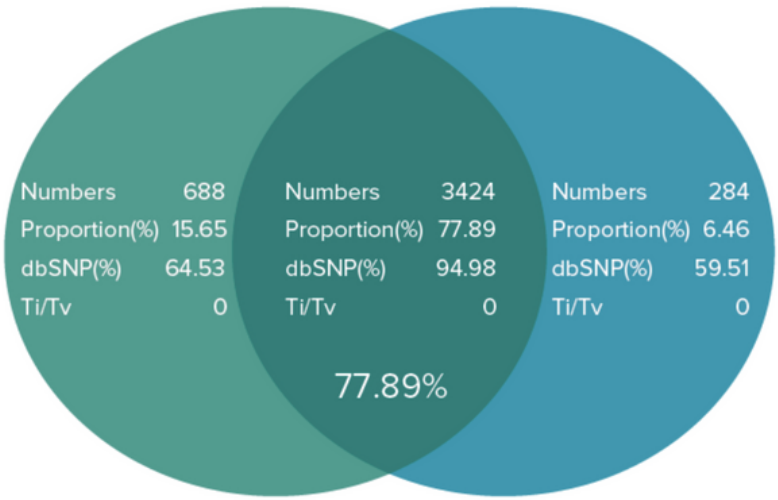


Bar-Graph showing the mapping rate and sequencing coverage of the samples using DNBSEQ™ and Illumina HiSeq 4000 platform of 100X WES.



NBSSEEQQHiSeq		
	PRECISION76.95%76.44%	
SENSITIVITY95.53%95.5%		

SNP calling performance from the NA12878 standard sample demonstrates good concordance between platforms



InDel ANALYSIS	DNBSEQ™	HiSeq
PRECISION	74.04%	79.77%
SENSITIVITY	87.60%	84.93%

InDel calling performance from the NA12878 standard sample demonstrates good concordance between platforms.
*Full demonstration data reports are available through your BGI account representative.

Bacterial *De Novo* Sequencing

Service Description

De novo sequencing refers to the sequencing of a novel genome where there is no reference sequence available for alignment.
The process of *de novo* genome sequencing involves the sequencing of small DNA fragments, and assembling the reads into longer sequences (contigs) and finally ordering the contigs to obtain the entire genome sequence.
With the advent of rapid, low-cost next-generation sequencing (NGS) technology and single molecule sequencing, researchers can now obtain complete genomes for bacteria at a lower cost. This has allowed large-scale genomic studies to be performed that were unimaginable just a few years ago.

Sequencing Service Specification

BGI bacterial *de novo* sequencing services are executed utilizing multiple sequencing systems.

Sample Preparation and Services

- Library preparations (DNBSEQTM, PacBio Sequel/Sequel II , Nanopore etc.)
- Various sequencing modes
- Standard and customized data analysis
- Available data storage and bioinformatics applications

Sequencing Quality Standard

- Guaranteed ≥90% of DNBSEQ™ clean bases with quality score of Q20.

Turnaround Time

- Typical 35 working days for bacterial primary assembly where only the DNBSEQTM sequencing technology platform is used.
- Typical 45 working days for bacterial complete map where multiple sequencing system is used.

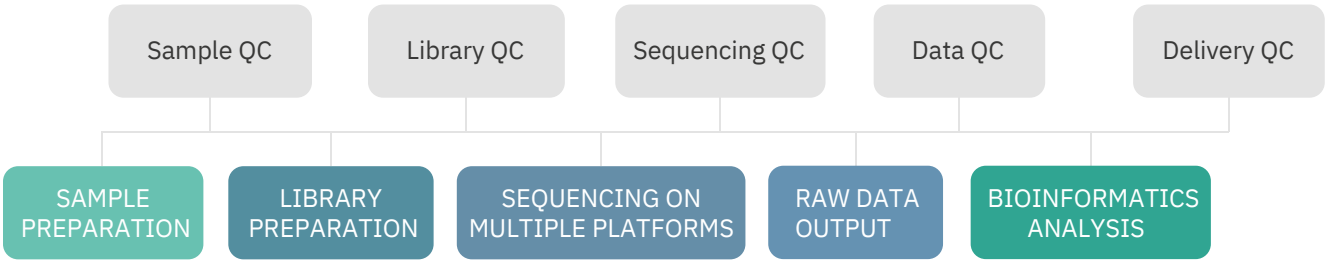
Sequencing Strategy

Our bacterial *de novo* sequencing service is usually performed on a combination of DNBSEQTM technology platforms and single molecule sequencing platforms (PacBio Sequel/Sequel II or Nanopore) for library preparation, sequencing, and mapping. In addition, BGI offers extensive bioinformatics data analysis options for genome assembly, annotation and evolution.

	Library Type	Read Length	Platform
Primary assembly	Short insert size library	PE150	DNBSEQTM
Complete map	Short insert size library + 10-15K library	PE150 />10K	DNBSEQTM + PacBio Sequel /Sequel II /Nanopore

Project Workflow

We care for your samples from the start through the result reporting. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.



Data Analysis

Besides clean data output, BGI offers a range of standard and customized bioinformatics pipelines for your sequencing project.
Reports and output data files are delivered in industry standard file formats: BAM, .xls, .png.

Standard Analysis	CUSTOM ANALYSIS
1. Data Filtering 2. Genome Assembly 3. Genome Component 1) Gene Component; 2) Repeat Sequence; 3) Non-coding RNA; 4) CRISPR; 5) Prophage; 6) Genomic Island. 4. Functional annotation based on KEGG, SwissProt, GO, VFDB, ARDB, CAZy et.al 5. Suggestions for data mining 6. Comparative genomic analysis	Further customization of Bioinformatics analysis to suit your unique project is available Please contact your BGI technical representative

Sample Requirements

We can process your extracted genomic DNA from bacterial samples with the following general requirements:

Platform	Sample type	Mass	OD	Integrity (AGE)	Sample Purity
PacBio Sequel /Sequel II	10-20Kb library	≥8μg	OD260/280: 1.6-2.2	The band on gel electrophoresis has little degradation, or has a fragment size greater than 40kb.	No contamination with RNA, protein or salt ions; colorless and transparent; non-sticky
PromethION	10-20Kb library	≥8μg	Nanopore OD260/230: 1.6-2.2		
DNBSEQ TM	350bp library	≥1μg		The band on gel electrophoresis has little degradation, or has a fragment size greater than 20kb.	

BGI also provides special sample extraction services upon demand.

16S/18S/ITS Amplicon Sequencing

Service Description

Over 99% of natural microorganisms cannot be isolated and cultured clonally, while traditional isolation and culture-dependent pose challenge for the study of microorganisms in their natural environment. Metagenomic studies of genetic material that is directly recovered from environmental samples have benefitted greatly from advanced NGS technology as a method for the exploration of microbial biodiversity. 16S and 18S rDNA are hypervariable regions in the 16S or 18S rRNA genes in bacteria and fungi, while ITS (Internal Transcribed Spacer) is the spacer DNA between the small-subunit and large-subunit rRNA genes in bacteria, fungi and archaea. Sequence comparison of 16S/18S/ITS regions is widely used in taxonomy and molecular phylogeny because of the easy amplification by PCR, even from low quantities of DNA, while they have a high degree of variation even between closely related species.

Sequencing Service Specification

BGI 16S/18S/ITS Amplicon Sequencing services are typically executed with the Illumina HiSeq 2500 or MiSeq sequencing system.

Sample Preparation and Services

- PCR will be used to isolate different 16S/18S/ITS regions
- PE250 sequencing by either MiSeq or HiSeq 2500, and PE300 by MiSeq
- Sequencing data are available in standard file formats
- Custom bioinformatic data analysis is available
- Data storage services are available

Sequencing Quality Standard

- Up to ≥ 80% of bases with a ≥Q30 quality score, depending on the chosen sequencing strategy.
- Recommended sequencing coverage is dependent on the complexity of the sample.

Turnaround Time

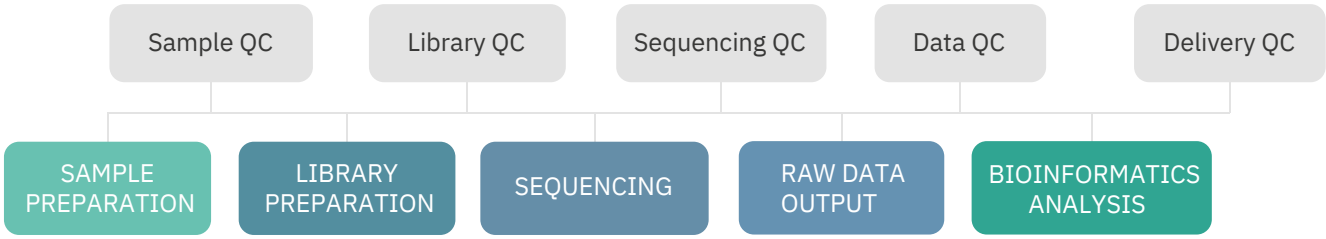
- Typical 30 working days from sample QC acceptance to data availability
- Expedited services are available, contact your local BGI specialist for details

Project Workflow

We care for your samples from the start through to the result reporting. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.

Data Analysis

Besides clean data output, BGI offers a range of standard and customized bioinformatics pipelines for your sequencing project.
Reports and output data files are delivered in industry standard file formats: FASTQ, .xls and .png.



STANDARD ANALYSIS

- Data filtration
- Overlap paired-end reads to form tags
- Tags are clustered into OTU. PCA, Venn diagram. A rank curve will be generated based on OTU abundance
- Species are classified by OTU annotation, based on which a species profiling histogram, heat map and a phylogeny tree will be provided
- Alpha diversity indexes are generated for in single sample
- Beta diversity and clustering analysis are performed for multiple samples
- Comparative analysis is used to screen significant differences in multiple samples

CUSTOMIZED ANALYSIS

Further customization of Bioinformatics analysis to suit your unique project is available: Please contact your BGI technical representative.

Sample Requirements

We can process intact genomic DNA or PCR fragment samples

Meta rDNA Amplicon Sequencing					
Sample type		MassConcentration	Integrity (AGE)		Sample Purity
Meta rDNA Amplicon	Genomic DNA	>0ng (above 50ng recommended)	>0ng/μL	Sample must be genomic DNA.	No contamination with RNA, protein or salt ions; colorless and transparent; non-sticky.
Meta 16S rDNA Amplicon PCR-free Library	PCR products	≥3μg	≥30ng/μL	The band on gel electrophoresis is clear without any smear.	

Meta rDNA Amplicon Sequencing includes Meta rDNA V3, V6, V4, V1-V3, V3-V4, V4-V5, V5-V6, ITS1, ITS2. regions amplicon library construction (18s and other regional or non-bgi standard method for building the library need to be communicated to the customization team). The key step of library construction is PCR amplification, which is affected by many factors such as purity, salt ions, pigment, and humic acid. Therefore, the quality of Meta rDNA Amplicon samples will be determined after library construction with PCR amplification.

Metagenomic Survey

Service Description

Over 99% of natural microorganisms cannot be isolated and cultured clonally. Traditional isolation and culture-dependent methods have limited the study of microorganisms in their natural environment. Metagenomic studies of genetic material that is directly recovered from environmental samples have benefitted greatly from advanced NGS technology as a method for the exploration of microbial biodiversity.
BGI's metagenomics survey service applies whole genome shotgun sequencing of DNA isolated from environmental samples, with the advantages of high throughput and high coverage. It can provide information not only on species composition and abundance, but also on functional genes, gene differences between samples, metabolic pathways and gene resource mining for bioactive products.

Sequencing Specification

BGI Metagenomic Survey Sequencing services are executed with DNBSEQ™ sequencing technology platforms.

Sample Preparation and Services

- 150bp Paired-End sequencing
- Data output, standard and custom data analysis
- Available data storage and bioinformatics applications

Sequencing Quality Standard

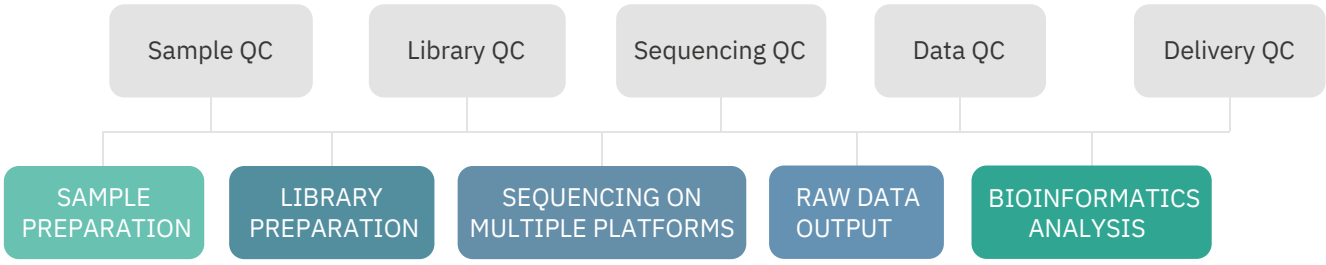
- Guaranteed ≥ 90% of bases with quality score ≥Q20

Turnaround Time

- Typical 40 working days after sample receipt for data delivery
- Expedited services are available, contact your local BGI specialist for details

Project Workflow

We care for your samples from the start through to the result reporting. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.



Data Analysis

Besides data output, BGI offers a range of standard and customized bioinformatics pipelines for your sequencing project.
Reports and output data files are delivered in industry standard file formats: BAM, .xls, .png.

Standard Analysis

- Data filtering
- Alignment
- Metagenomic *De novo* assembly
- Non-redundant gene catalogue
- Prophage transposable element prediction
- Functional annotation based on KEGG, CAZy, eggNOG, CARD
- Species composition and diversity analysis
- Quantitative and diferential analysis of gene abundance
- Principal component analysis

Customized Analysis

Further customization of Bioinformatics analysis to suit your unique project is available: Please contact your BGI technical representative.

Sample Requirements

We can process your exacted genomic DNA from a variety of environmental samples with the following general requirements:

Metagenomic Survey				
Sample type	Mass	Concentration	Integrity (AGE)	Sample Purity
Genomic DNA	≥1μg	≥12.5ng/μL	The band on gel electrophoresis has little degradation, or has a fragment size greater than 20kb.	No contamination with RNA, protein or salt ions; colorless and transparent; non-sticky.

Whole Genome Bisulfite Sequencing

Service Description

Methylation of DNA at the fifth position in cytosine (5-mC) is a stable epigenetic modification and plays an important role in many biological processes, including gene silencing, suppression of transposable elements, genomic imprinting and X chromosome inactivation. Detection and quantification of methylation are critical to understand gene expression and other processes subjected to epigenetic regulation.
Whole genome bisulfite sequencing (WGBS) is used to detect methylated cytosines by treating the DNA with sodium bisulfite before sequencing. WGBS has become the gold standard for studying genome-wide methylation at single base resolution.

Sequencing Service Specification

BGI Whole Genome Bisulfite Sequencing Services are executed with DNBSEQ™ sequencing technology platforms.

Sample Preparation and Services

- Library preparation, including bisulfite treatment
- PE100/PE150 sequencing
- Clean data, standard and customized data analysis
- Available data storage and bioinformatics applications

Sequencing Quality Standard

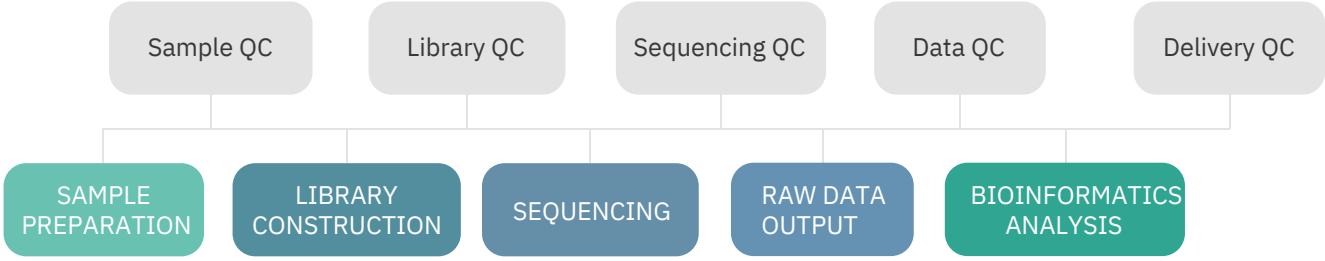
- Guaranteed ≥90% of clean bases with quality score of Q20
- Guaranteed ≥99% bisulfite conversion rate (For human/rat/mouse)
- Standard sequencing coverage ≥30X is recommended

Turnaround Time

- Typical 40 working days from sample QC acceptance to filtered raw data availability
- Expedited services are available; contact your local BGI specialist for details

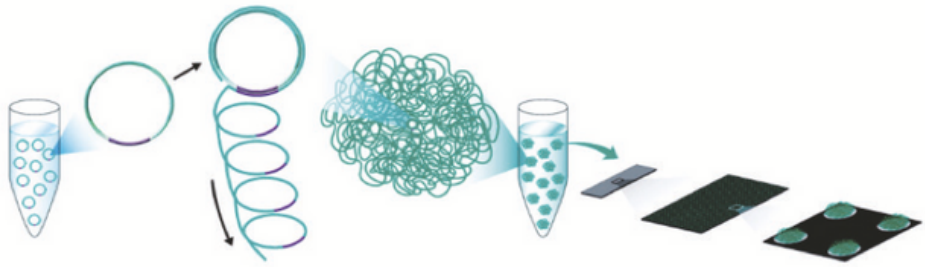
Project Workflow

We care for your samples from the start through to the result reporting. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.



DNBSEQ™ Sequencing Technology

DNBSEQ™ is an innovative high-throughput sequencing technology, developed by BGI's Complete Genomics subsidiary in Silicon Valley. The system is powered by combinatorial Probe-Anchor Synthesis (cPAS), linear isothermal Rolling-Circle Replication and DNA Nanoballs (DNB™) technology, followed by high-resolution digital imaging.



The combination of linear amplification and DNB technology reduces the error rate while enhancing the signal. The size of the DNB is controlled in such a way that only one DNB is bound per active site on the flow cell. This densely patterned array technology provides optimal sequencing accuracy and increase flow cell utilization.

Data Analysis

Besides clean data, BGI offers a range of standard and customized bioinformatics options for your whole genome bisulfite sequencing project.
Reports and output data flies are delivered in industry standard file formats: FASTQ, BAM, cout, .xls, .png.

Standard Analysis	Customized Analysis
• Data Filtering • Alignment • Data Quality Statistics • Methyl-cytosine identification • Differential Methylated Region (DMR) analysis • Statistics of methylation level • GO and Pathway analysis of DMRs-related gene	• Further customization of bioinformatics analysis to suit your unique project is available. • Please contact your BGI technical representative for details.

Sample Requirements

We can process your gDNA, whole blood, cell line, fresh frozen tissue samples, with the following general requirements:

	DNA Amount and Concentration	Minimum Sample Volume
Regular Samples	Intact genomic DNA 1μg,Concentration50ng/μl	15μl

DNBSEQ™ Performance

The human standard sample NA12878 was used to validate DNBSEQ™ WGBS. The PE100 sequencing data from DNBSEQ™ was compared to the PE150 sequencing data from the Illumina Novaseq 6000. Two datasets of each platforms are included in the comparison.
Both sequencing platforms show high mapping rates in the range of 82-86%, while the duplication rate of the DNBSEQ™ technology platforms are noticeable lower by around 10%. As a result, the average data-using rate of DNBSEQ™ platform is 14.3% greater than that of Illumina Novaseq 6000.
The GC-bias plot shows GC content has less impact on DNBSEQ™ technology platforms, whereas Illumina Novaseq 6000 prefers higher GC regions over the lower ones (Figure 2).

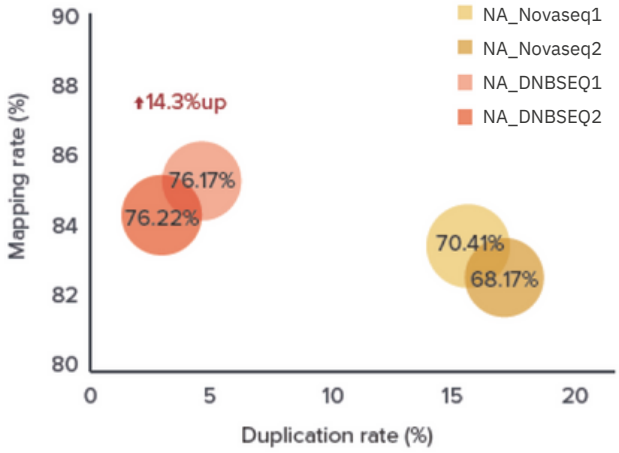


Figure 1 The mapping rate (Y axis), duplication rate (X axis), and data using rate (the size of the bubbles) of the four datasets. The data using rate in the bubbles refers to the proportion of valid reads (the duplicate removed mapped reads) data to the total filtered reads data.

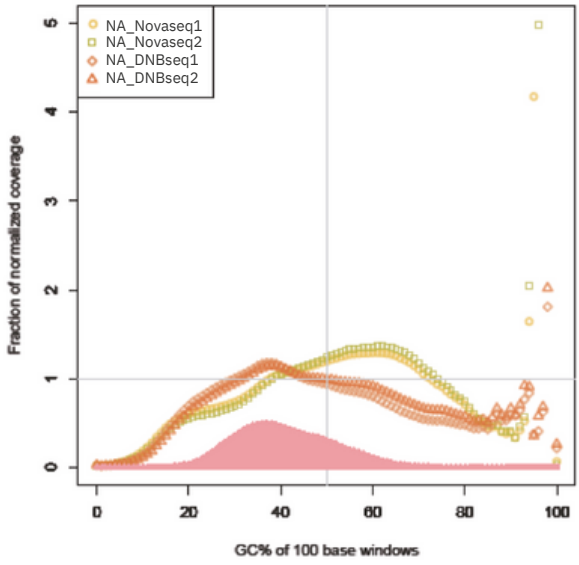


Figure 2. The GC-bias plot for the NA12878 genome. The yellow spots: GC composition distribution of the 2 datasets of Illumina Novaseq 6000. The orange spots: GC composition distribution of the 2 DNBSEQ™ datasets.

ChIP Sequencing

Service Description

ChIP-Seq, also known as ChIP-Sequencing, is widely used to analyze protein interactions with DNA. It combines chromatin immunoprecipitation (ChIP) with massively parallel DNA sequencing to identify binding sites of DNA as sociated proteins, and can be used to precisely map global binding sites for any protein of interest. ChIP sequencing offers higher resolution and more precise and abundant information in comparison with array-based ChIP-chip.
Besides clean sequencing data output, BGI offers standard, advanced and custom bioinformatics services to suit your specific research needs.

Confidence Great correlation with Hiseq data.

Low input As low as 5ng ChIP-ed DNA/sample for human sample.

Comprehensive analysis Correlation analysis between ChIP-Seq and RNA-Seq.

Sequencing Service Specification

BGI ChIP-Seq Services are performed with BGI's DNBSEQTM sequencing technology, featuring combinatorial probe-anchor synthesis (cPAS) and DNA Nanoballs (DNB) technology[1] for superior data quality.

Services

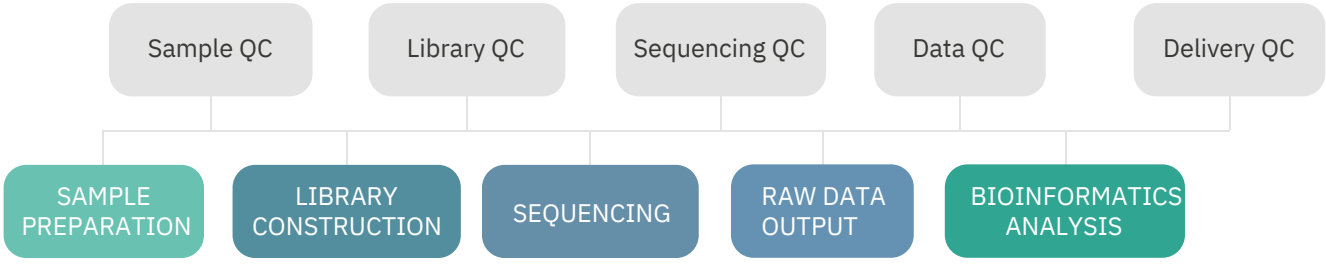
- 50bp Single-end sequencing reads
- Standard output 20 Million reads per sample
- Clean data and bioinformatics analysis are available in standard format
- Available data storage and bioinformatics applications
- Cloud-based data storage and delivery system

Turnaround Time

- Typical 25 working days from sample QC acceptance to final data delivery
- Expedited services are available, contact your local BGI specialist for details

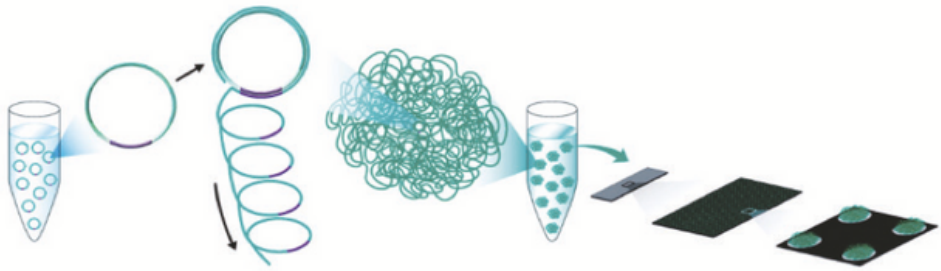
Project Workflow

We care for your samples from the start through to the result reporting. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.



DNBSEQ™ Sequencing Technology

DNBSEQTM is an innovative high-throughput sequencing technology, developed by BGI's Complete Genomics subsidiary in Silicon Valley. The system is powered by combinatorial Probe-Anchor Synthesis (cPAS), linear isothermal Rolling-Circle Replication and DNA Nanoballs (DNB™) technology, followed by high-resolution digital imaging.
The combination of linear amplification and DNB technology reduces the error rate while enhancing the signal. The size of the DNB is controlled in such a way that only one DNB is bound per active site on the flow cell. This densely patterned array technology provides optimal sequencing accuracy and virtually eliminates index-hopping when pooling samples.



Data Analysis

In addition to clean data output, BGI offers a range of standard, advanced and customized bioinformatics pipelines for your ChIP-Seq analysis project including the correlation analysis of differential expression genes and peak-related genes.
Reports and output data files are delivered in industry standard FASTQ, Excel file formats with publication-ready tables and figures.

STANDARD BIOINFORMATICS ANALYSIS	ADVANCED BIOINFORMATICS ANALYSIS
Data Filtering Alignment to the reference genome Peak scanning and annotation Identify Differential peaks between samples Identify Differential peaks between groups Differential peaks annotation	Motif analysis

CUSTOMIZED BIOINFORMATICS ANALYSIS
Statistics of epigenomics modification value and mRNA expression value of related genes; The distribution of epigenomics modification values in different gene categories; Overall connection of epigenomics modification level and the mRNA expression level in different gene categories; The relationship of the ratio of epigenomics modification and the ratio of mRNA expression in a pair of samples; Clustering analysis based on epigenomics modification level and mRNA expression level; Calculate different level of epigenomics modification when mRNA expression value is different; GO, Pathway analysis, related functional excavation and verification of the genes that differences exist in both epigenomics modification and mRNA expression; Further customization of bioinformatics analysis to suit your unique project is available: Please contact your BGI technical representative

Sample Requirements

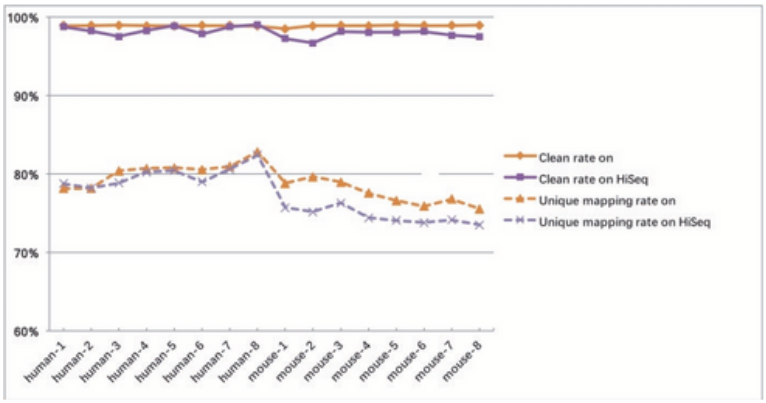
We can process DNA samples of human, plant, animal and microbial species, with the following general requirements:

	RNA Amount and Concentration	Minimum Volume
ChIP-ed DNA	amount ≥ 10ng, concentration ≥ 1 ng/μl	15μl

Sequencing Service Specification

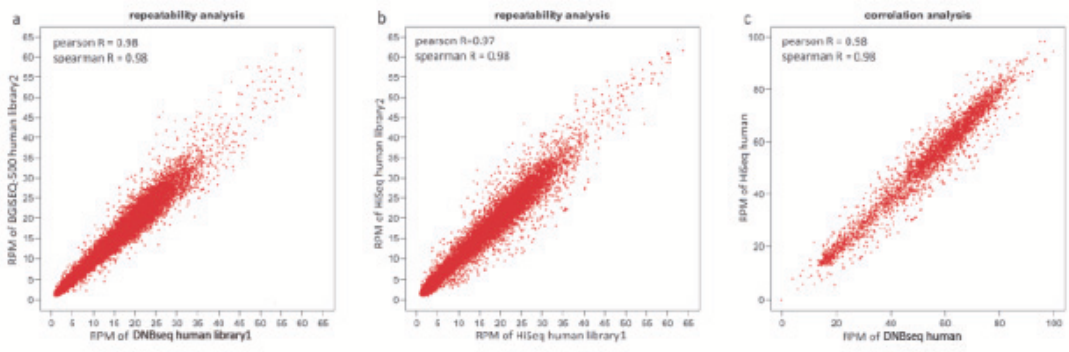
1) Excellent agreement of data quality between HiSeq and DNBSEQ™

The plot below shows the high sequencing data quality of human HeLa cell libraries and mouse libraries on DNBSEQTM as compared to the HiSeq platform. By applying the same filtering criteria, the clean rate and unique mapping rate from the same samples are consistent on both platforms.



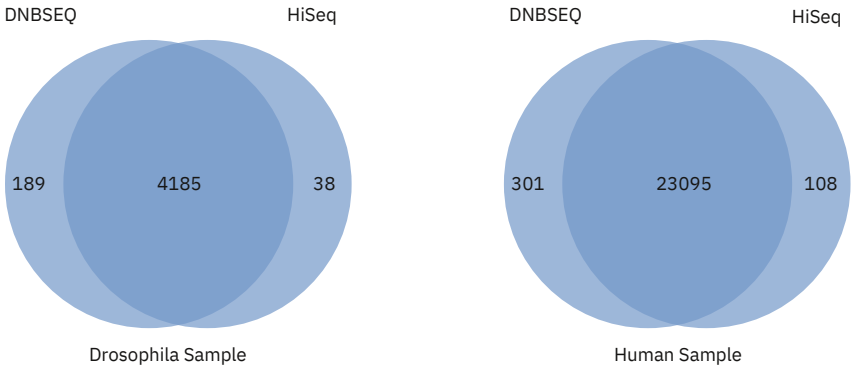
2) Excellent agreement of reads per million (RPM) between HiSeq and DNBSEQ™

The scatter plots below compare the reads per million (RPM) values of human HeLa cell sample sequenced on the DNBSEQTM and HiSeq platforms. The coefficients of sequencing repeatability on both platforms are high (Figures a, b). There is a high correlation between DNBSEQTM and HiSeq platforms as well (Figure c).



3) Excellent agreement of peak number detection between HiSeq and DNBSEQ™

The figures below show the comparison of detected peak numbers on both HiSeq and DNBSEQTM platforms from the same samples[2]. The common peak detection rate of both drosophila and human HeLa cell sample are higher than 99%, while both platforms have a small amount of their own uniquely detected peaks.



References

[1] Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, Carnevali P, Nazarenko I, Nilsen GB, Yeung G, et al. *Science* (2010) 327(5961):78–81.
http://science.sciencemag.org/content/327/5961/78.full
[2] ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Landt S G, Marinov G K, Kundaje A, et al. *Genome research*, 2012, 22(9): 1813-1831.
https://genome.cshlp.org/content/22/9/1813.short

RNA Sequencing

- DNBSEQ
- DNBSEQ UMI Small RNA Sequencing
- DNBSEQ
- Iso-Seq/PacBio Transcriptome Sequencing

DNBSEQ™ RNA Sequencing (Transcriptome)

Service Description

Transcriptome sequencing is used to reveal the presence, quantity and structure of RNA in a biological sample under specific conditions. Compared to hybridization, based RNA quantification methods such as microarray analysis, sequencing-based transcriptome detection can quantify gene expression with low background, high accuracy and high levels of reproducibility within a large dynamic range.
In addition, transcriptome sequencing does not require an existing genome sequence and can detect mutations, splice variants and fusion genes that cannot be detected by microarrays.

Sequencing Service Specification

BGI transcriptome sequencing services are executed with DNBSEQ™sequencing technology, featuring cPAS and DNA Nanoballs (DNB™) technology for superior data quality.

Sample Preparation and Services

- Multiple choices for mRNA enrichment and rRNA removal
- Stranded library
- 100bp and 150bp paired-end sequencing options available
- ≥30 Million reads per sample recommended
- Sequencing Data and bioinformatics analysis are available in standard file formats
- Advanced RNA data visualization and data mining with Dr.Tom system

Sequencing Quality Standard

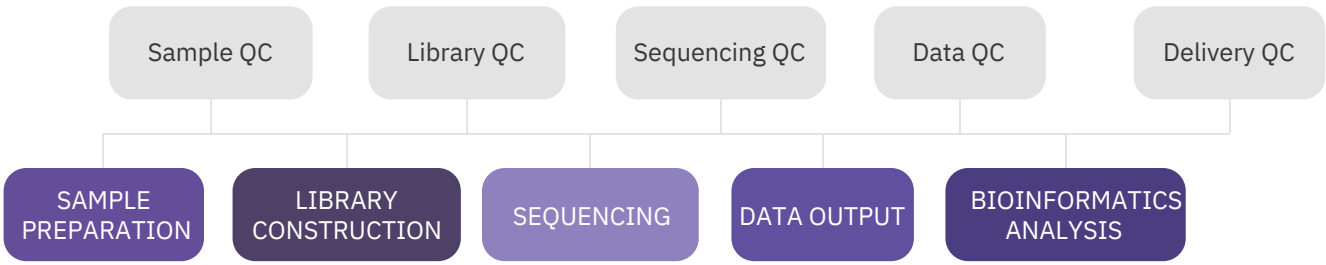
- Guaranteed ≥80% of bases with quality score of ≥Q30

Turnaround Time

- Typical 18 working days from sample QC acceptance to data availability.
- Expedited services are available.

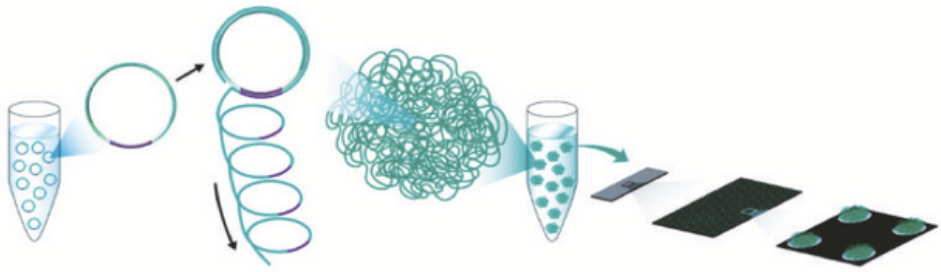
Project Workflow

We care for your samples from the start through to the result reporting. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.



DNBSEQ™ Sequencing Technology

DNBSEQTM is an innovative high-throughput sequencing technologysolution, developed by BGI's Complete Genomics subsidiary in Silicon Valley. The system is powered by combinatorial Probe-Anchor. Synthesis (cPAS), linear isothermal Rolling-Circle Replication and DNA Nanoballs (DNB™) technology, followed by high-resolution digital imaging.



The combination of linear amplification and DNB™ technology reduces the error rate while enhancing the signal. The size of the DNB is controlled in such a way that only one DNB is bound per active site on the DNBSEQ™ flow cell. This densely patterned array technology provides optimal sequencing accuracy and increases flow cell utilization.

Data Analysis

In addition to raw data output, BGI offers a range of standard and customized bioinformatics pipelines for your transcriptome sequencing project. Reports and output data files are delivered in industry standard file formats: FASTQ, BAM and Excel.

Standard Bioinformatics Analysis	Data Visualization and Customized Analysis with the Dr. Tom System
• Quantitative expression profiles • Alternative splicing analysis • Fusion gene analysis (human only) • Time series analysis • Pathway enrichment analysis • Hierarchical clustering analysis • Protein-Protein Interaction (PPI) analysis • Gene ontology analysis	• Interactive data visualization tools for Expression Analysis, Gene Set Enrichment Analysis, Association Analysis and More. • Access world-leading Databases for Powerful Data Mining • AI-based Literature Retrieval for Easy Referencing

Sample Requirements

We can process your total RNA, blood, cell line, FFPE, fresh frozen tissues and single cell samples from a variety of species, with the following general requirements:

Sample	Species	Amount	Concentration (ng/μL)	RIN/RQN value	28S/18S	DV200
Total RNA	Human/Mouse/Rat (non-whole blood)	≥200ng	≥10	≥7	≥1.0	N/A
	Human (whole blood)	≥500ng	≥40	≥7	≥1.0	N/A
	Human (FFPE)*	≥200ng	≥70	≥2	N/A	≥30%
	Insect	≥ 1 μg	≥40	N/A	N/A	N/A
	Other Animals	≥ 1 μg	≥40	≥7	≥1.0	N/A
	Plant	≥ 1 μg	≥40	≥6	≥1.0	N/A
	Fungi	≥ 1 μg	≥40	≥6.5	≥1.0	N/A

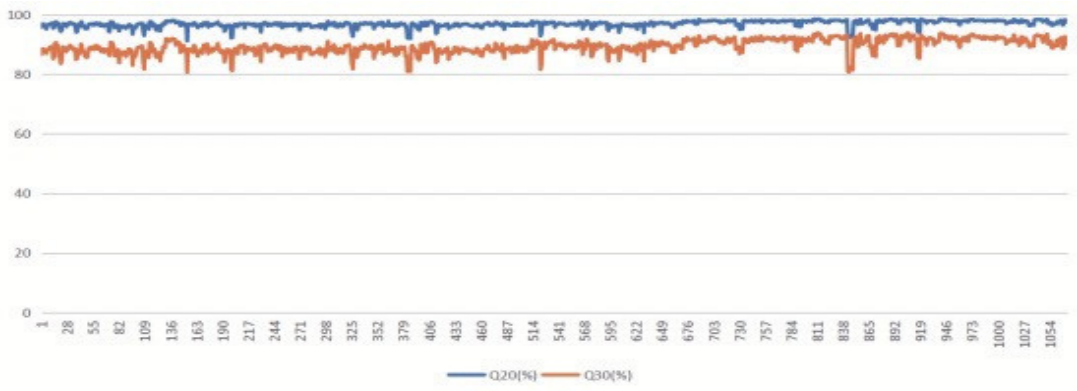
Sample Type (For Human Samples)	FFPE	Whole Blood	Tissue	Cell Line
Requirement	≥ 5 microns thick per slides	≥1 mL	≥30mg	≥2*105 cells

≥ 5 Slides

Low-input transcriptome sequencing is available.

Stable and High-Quality Data Performance

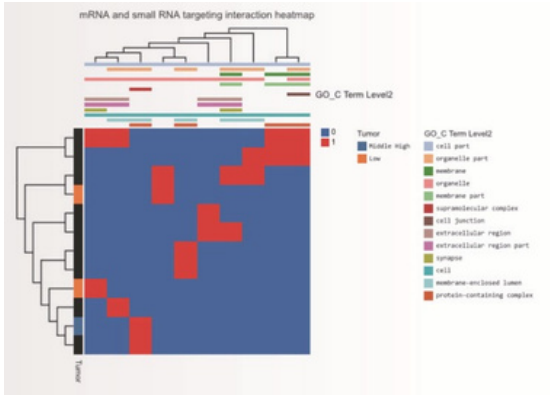
1,072 samples were randomly selected from over 10,000 samples that were sequenced at BGI's laboratory over a period of 6 months.
The data output and data quality remained stable over that period. The average Q20 and Q30 scores were 97%and 89.5% respectively.



Core Capabilities of Dr.Tom system

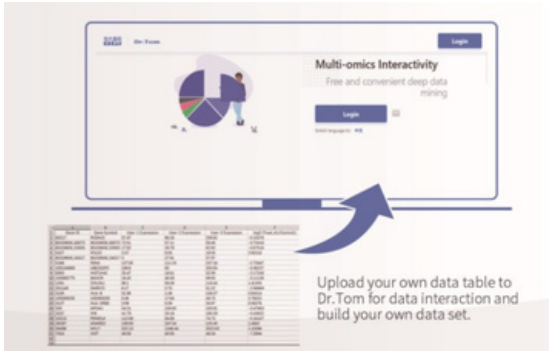
Association Analysis

With a simple click Dr.Tom lets users detect RNA association with target genes, based on their interaction relationship (such as PPI, Target, Co-expression, ceRNA, GGI and RNAplex), or based on the position relationship (such as upstream and downstream position).



Custom Datasets

Customers can upload their own gene expression data, using tool boxes for graphing and visualization, and construct their own gene annotation database for enrichment, clustering and multi-omics association analysis.



DNBSEQTM UMI Small RNA Sequencing

Service Description

Small RNAs are a type of non-coding RNA (ncRNA) molecule that are less than 200nt in length. They are often involved in gene silencing and post-transcriptional regulation of gene expression. Small RNA sequencing is used to discover novel small RNAs, examine the differential expression of all small RNAs and to characterize variations with single-base resolution.

Unique Molecular Identifiers (UMIs) can be utilized to eliminate undesirable PCR duplicates derived from a single molecule. After PCR, molecules sharing a UMI are assumed to be derived from the same input molecule. As such, UMI counts offer superior results to counting reads, leading to more accurate estimates of quantitative small RNA expression[1]. UMI technology is especially beneficial to customers doing research on rare and precious samples or samples containing less RNAs, such as exosomes.

Our DNBSEQTM Small RNA sequencing service with UMI technology delivers accurate, affordable and high-quality sequencing data to support your academic and clinical research applications.

Sequencing Service Specification

DNBSEQTM Small RNA Sequencing Services are performed with the DNBSEQ™ technology platforms. Featuring cPAS and DNA Nanoballs (DNBTM) technology for superior data quality[2].

Sample Preparation and Services

- 50bp single-end sequencing reads
- Standard output 20 Million reads per sample
- UMI technology to enhance the quantification accuracy
- Sequencing data and bioinformatics analysis are available in standard file formats
- Advanced RNA data visualization and data mining with Dr.Tom's system

Sequencing Quality Standard

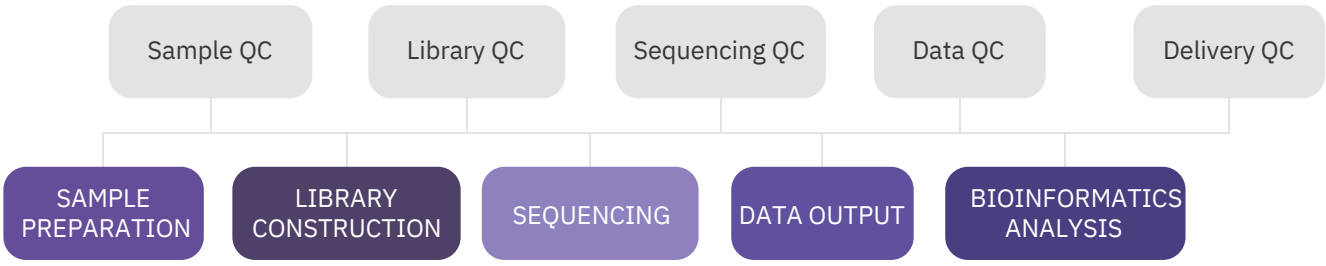
- Guaranteed ≥80% of bases with quality score of ≥Q30

Turnaround Time

- Typical 27 working days from sample QC acceptance to data availability.
- Expedited services are available.

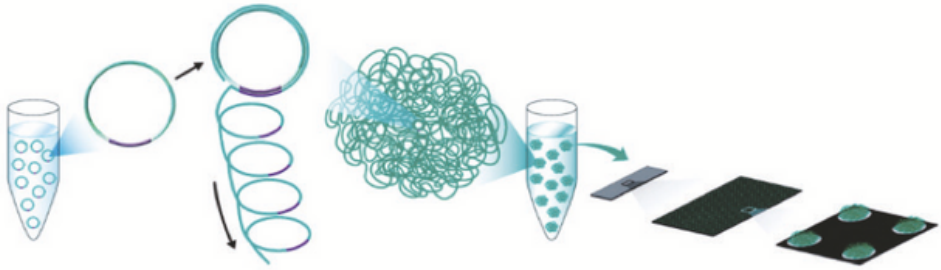
Project Workflow

We care for your project from the receipt of samples through to the reporting of results. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.



DNBSEQ™ Sequencing Technology

DNBSEQTM is an innovative high-throughput sequencing technology solution, developed by BGI's Complete Genomics subsidiary in Silicon Valley. The system is powered by combinatorial Probe-Anchor. Synthesis (cPAS), linear isothermal Rolling-Circle Replication and DNA Nanoballs (DNB™) technology, followed by high-resolution digital imaging.



The combination of linear amplification and DNB™ technology reduces the error rate while enhancing the signal. The size of the DNB is controlled in such a way that only one DNB is bound per active site on the DNBSEQ™ flow cell. This densely patterned array technology provides optimal sequencing accuracy and increases flow cell utilization.

Data Analysis

In addition to data output, BGI offers a range of standard and customized bioinformatics pipelines for your small RNA sequencing project.
Reports and output data files are delivered in industry standard FASTQ, and Excel file formats with publication-ready tables and figures.

STANDARD BIOINFORMATICS ANALYSIS

- Data filtering
- Length distribution of small RNAs
- Analysis of common and specific sequences between two samples
- Small RNAs distribution across selected genome
- Identification of rRNAs, tRNAs, snRNAs, snoRNAs, etc.
- Identification of repeat associated small RNAs
- Identification of small RNA sequences which could align to exon/intron
- Identification of known miRNAs by aligning to designated part of miRBase
- Analysis of the expression pattern of known miRNAs
- Classification of small RNAs into several categories based on customer preference
- Prediction of novel miRNAs and their secondary structures by Mireap and miRDeep from unannotated small RNAs
- Family analysis of known miRNAs

DR.TOM SYSTEM ANALYSIS

- miRNA target gene analysis
- GO and Pathway annotation for target genes
- miRNA-mRNA interaction, lncRNA-mRNA interaction analysis
- Co-expression Interaction Nnetwork Aanalysis

Sample Requirements

We can process your human, plant or animal samples, with the following general requirements:

Sample type	RNA Amount and Concentration	Minimum Sample Volume	Quantitative Result
Regular sample	Mass ≥1μg Concentration ≥50ng/μl	15 μl	RIN≥6.5 (plant) RIN≥8.0 (human/animal)
FFPE RNA	Mass ≥1μg Concentration ≥50ng/μl	15 μl	RIN≥2.0 DV200≥30%
Small RNA of plasma/serum/exosome	Mass≥20ng Concentration ≥1ng/μl	15 μl	N/A

Stable and High-Quality Data Performance

Technical Reproducibility[3]

To demonstrate the high technical reproducibility of DNBSEQTM technology platforms, six human brain samples, two heart samples and two blood samples were sequenced. Reproducibility was assessed by using six technical replicates of human brain sample (see Fig. 1). The median correlation between the six replicates was 0.98, and the 25% and 75% quantile were 0.98 and 0.99, respectively.

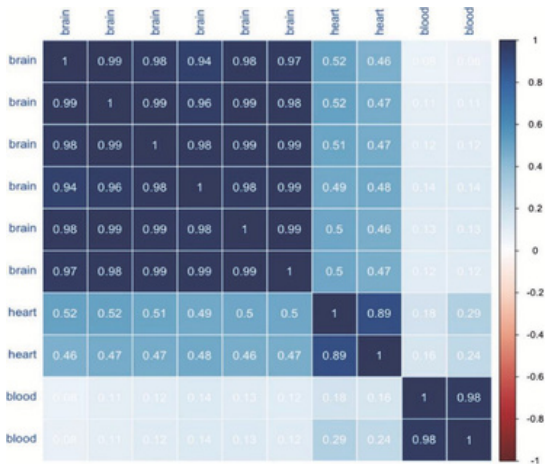


Fig. 1 Correlation matrix of brain (six technical replicates), heart (two technical replicates), and blood (two biological replicates) samples, sequenced by the DNBSEQTM technology platforms.

References

[1] Fu Y, Wu PH, Beane T, Zamore PD, Weng Z: . Elimination of PCR duplicates in RNA-seq and small RNA-Seq using unique molecular identifiers. *BMC Genomics*, 2018;,19:531.

[2] Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, Carnevali P, Nazarenko I, Nilsen GB, Yeung G, Dahl F, Fernandez A, Staker B, Pant KP, Baccash J, Borcherding AP, Brownley A, Cedeno R, Chen L, Chernikoff D, Cheung A, Chirita R, Curson B, Ebert JC, Hacker CR, Hartlage R, Hauser B, Huang S, Jiang Y, Karpinchyk V, Koenig M, Kong C, Landers T, Le C, Liu J, McBride CE, Morenzoni M, Morey RE, Mutch K, Perazich H, Perry K, Peters BA, Peterson J, Pethiyagoda CL, Pothuraju K, Richter C, Rosenbaum AM, Roy S, Shafto J, Sharanhovich U, Shannon KW, Sheppy CG, Sun M, Thakuria JV, Tran A, Vu D, Zaranek AW, Wu X, Drmanac S, Oliphant AR, Banyai WC, Martin B, Ballinger DG, Church GM, Reid CA: . Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*, 2010;,327:78-81.

[3] Fehlmann T, Reinheimer S, Geng C, Su X, Drmanac S, Alexeev A, Zhang C, Backes C, Ludwig N, Hart M, An D, Zhu Z, Xu C, Chen A, Ni M, Liu J, Li Y, Poulter M, Li Y, Stahler C, Drmanac R, Xu X, Meese E, Keller A: cPAS-based sequencing on the BGISEQ-500 to explore small non-coding RNAs. *Clin Epigenetics* 2016;8:123.

DNBSEQ™ Long Non-Coding RNA Sequencing

Service Description

Long non-coding RNAs (lncRNAs) are a large class of transcribed RNA molecules with a length of more than 200 nucleotides that do not encode proteins. lncRNAs are thought to encompass nearly 30,000 different transcripts in humans, hence lncRNA transcripts account for the major part of the non-coding transcriptome.
lncRNA discovery is still at an early stage and only a small proportion of lncRNAs have so far been investigated. Although we can start to classify different types of lncRNA functions, we are still far from being able to predict the function of new lncRNAs.
BGI offers expression profiling as one way to uncover the function of lncRNA. Identifying lncRNAs that are differentially expressed during development or in particular situations can shed light on their potential functions. Alternatively, looking for lncRNAs and protein-coding genes whose expression is correlated, can indicate co-regulation or related functions.

Sequencing Service Specification

DNBSEQTM Long Non-coding RNA Sequencing Services are performed with DNBSEQ™ technology platforms, featuring cPAS and DNA Nanoballs (DNBTM) technology for superior data quality.

Sample Preparation and Services

- Paired-end 100 bp
- Strand specific library with rRNA removal
- Data and bioinformatics analysis are available in standard file formats
- Advanced RNA data visualization and data mining with Dr.Tom system

Sequencing Quality Standard

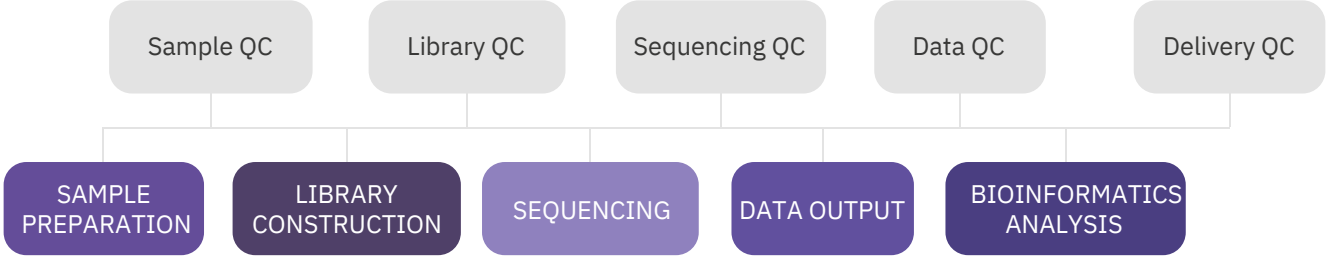
- Guaranteed ≥80% of bases with quality score of ≥Q30

Turnaround Time

- Typical 27 working days from sample QC acceptance to data availability.
- Expedited services are available.

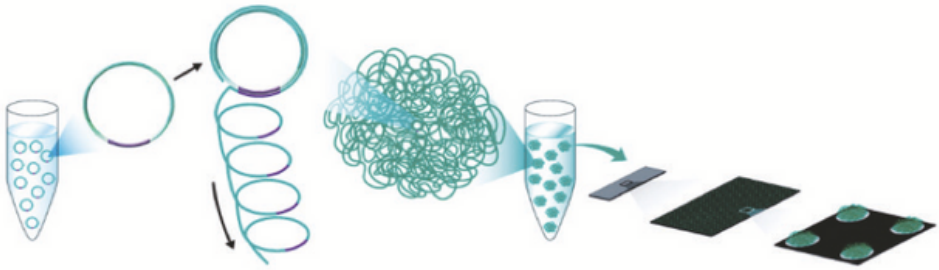
Project Workflow

We care for your samples from the start through to the result reporting. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.



DNBSEQ™ Sequencing Technology

DNBSEQ™ is an innovative high-throughput sequencing technology, developed by BGI's Complete Genomics subsidiary in Silicon Valley. The system is powered by combinatorial Probe-Anchor. Synthesis (cPAS), linear isothermal Rolling-Circle Replication and DNA Nanoballs (DNB™) technology, followed by high-resolution digital imaging.



The combination of linear amplification and DNB™ technology reduces the error rate while enhancing the signal. The size of the DNB is controlled in such a way that only one DNB is bound per active site on the DNBSEQ™ flow cell. This densely patterned array technology provides optimal sequencing accuracy and increases flow cell utilization.

Data Analysis

In addition to data output, BGI offers a range of standard and customized bioinformatics pipelines for your Long non-coding RNA sequencing project. Reports and output data files are delivered in industry standard file formats: FASTQ, BAM and Excel.

STANDARD BIOINFORMATICS ANALYSIS
• Identification of mRNA and LncRNA • Quantification and differential expression analysis • Gene ontology analysis and Pathway enrichment analysis

DR.TOM SYSTEM ANALYSIS
• Transcription factor prediction (For AnimalTFDB/PlantTFDB) • GSEA Analysis • Rfam、Pfam、Reactome、COG、EggNOG and InterPro annotation; • miRNA-mRNA interaction, lncRNA-mRNA interaction analysis • Protein-Protein Interaction (PPI) analysis • Co-expression Interaction Network Analysis • Custom dataset table upload

Sample Requirements

We can process your total RNA, blood, cell line, FFPE, fresh frozen tissues samples from a variety of species, with the following general requirements.
Purity: OD260/280 ≥ 2.0, OD260/230 ≥ 2.0, without degradation and DNA contamination

Species	RNA Amount and Concentration	Quantitative Result
Human, mouse and rat	Total RNA ≥200ng Concentration ≥20ng/μl	RIN≥7.0 28S/18S≥1.0

Stable and High-Quality Data Performance

High Reproducibility

The reproducibility among 3 technical replicates is very high. (Pearson value>0.99)

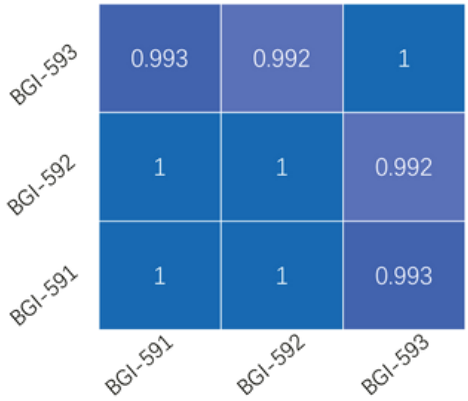


Fig 1. Technical Replicates performance

Iso-Seq/PacBio Transcriptome Sequencing

Service Description

RNA Sequencing (RNA-Seq) has become the most frequently used method for the majority of researchers conducting gene expression profiling. However, it is difficult to obtain a complete picture of the transcriptome because short reads cannot accurately assemble complex transcripts.
“Isoform Sequencing” (Iso-seq) developed by Pacific Biosciences (PacBio), is based on long-read sequencing technology. The unique long-read sequencing feature allows this method to identify new isoforms with extraordinary precision. The Iso-Seq application generates full-length cDNA sequences — from the 5’ end of transcripts to the poly-A tail—eliminating the need for transcriptome reconstruction using isoform-inference algorithms. The Iso-Seq method provides accurate information about alternatively spliced exons and transcriptional start sites. It also reveals information about poly-adenylation sites for transcripts across the full complement of isoforms within targeted genes or the entire transcriptome.

Sequencing Service Specification

BGI Iso-Seq/PacBio transcriptome sequencing services are executed on the Sequel I/II platform.

Sample Preparation

• Library preparation - Standard Iso-Seq library/Multi-throughput Iso-Seq library/polyA Iso-Seq library

Sequencing Quality Standard

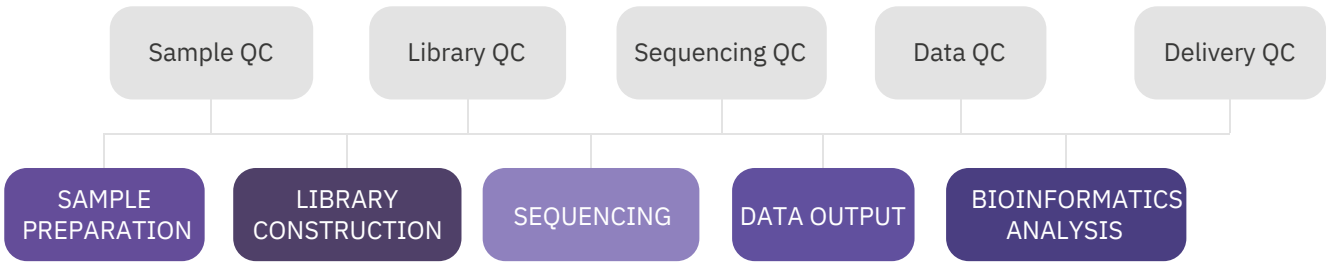
• 20Gb sequencing data per sample is recommended

Turnaround Time

• Typical 40 working days from sample QC acceptance to sequencing data availability

Project Workflow

We care for your project from the start through to reporting of results. Highly experienced laboratory professionals follow strict quality procedures to ensure the integrity of your results.



Data Analysis

Besides clean data, BGI offers a range of standard and customized bioinformatics options for your Iso-seq/PacBio transcriptome sequencing.
Reports and output data flies are delivered in industry standard file formats: FASTQ, BAM, cout, .xls, .png

Standard Analysis (without reference)

- Remove the low-quality reads and short reads
- Identify the full-length, non-chimeric transcripts and non-full-length, non-chimeric transcripts
- Build similarity graph using BLASR, get cluster consensus
- Polish the consensus sequences and get high quality full-length, non-chimeric transcripts
- Merge consensus sequences of all libraries and remove redundancy
- Annotation of the full-length non-chimeric transcripts (Nr、Nt、Swissprot、KEGG、GO、COG and Interpro)
- CDS prediction
- SSR prediction

Standard Analysis (with reference)

- Remove the low-quality reads and short reads
- Identify the full-length, non-chimeric transcripts and non-full-length, non-chimeric transcripts
- Build similarity graph using BLASR, get cluster consensus
- Polish the consensus sequences and get high quality full-length, non-chimeric transcripts
- Align to the reference genome with GMAP
- Merge consensus sequences of all libraries and remove redundancy
- Transcripts classification
- Novel transcripts analysis
- Long noncoding RNA prediction
- Splicing site detection and annotation
- Gene fusion detection and annotation

Multi-throughput Iso-Seq

- Analysis of Standard Iso-Seq + Gene/Transcripts quantification; differentially expressed gene detection and annotation

PolyA Iso-Seq

- Analysis of Standard Iso-Seq + Gene/Transcript quantification; differentially expressed gene detection and annotation; polyA length analysis

Customized Analysis

- Further customization of bioinformatics analysis to suit your needs is available.
- Please contact your BGI technical representative for details.

Sample Requirements

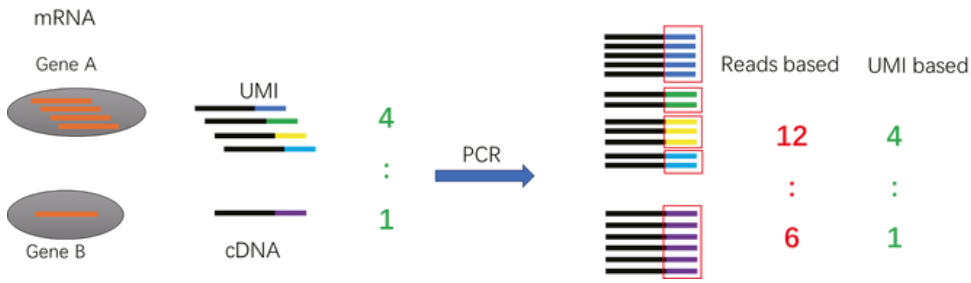
We can process your total RNA, whole blood, cell line, fresh frozen tissue samples, with the following general requirements:

	Total RNA Amount	Concentration, RIN, 28S/18S	Minimum Sample Volume
Recommended	m≥3μg	c≥300ng/μL RIN≥8 28S/18S(23S/16S)≥1.4	15 μl
Required	1μg≤m<3μg		15 μl

Highlights of BGI Iso-Seq Service

1) Absolute Quantification with UMI Technology

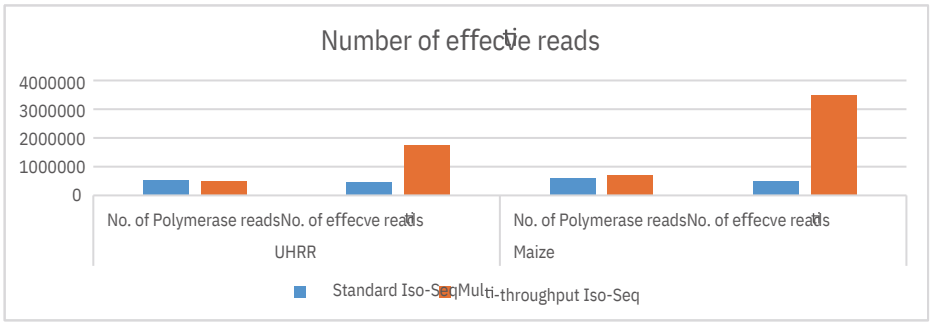
Each original transcript is marked by a unique molecular identifier (UMI), which includes 6-8 random bases. Counting the copy number of transcripts with a UMI based approach enables accurate quantification of the isoform without the interference of sequencing duplication. The UMI technology is built-in multi-throughput Iso-Seq and polyA ISO-Seq workflows. Both the UMI technology and the multi-insert sequential ligation technique are BGI patent-protected.



UMI enables more accurate gene quantification.

2) Greater Transcripts Detection with Multi-throughput Iso-Seq

Compared with standard Iso-Seq, Multi-throughput Iso-Seq can obtain 3-5 times more effective sub-reads and allow users to detect double amount of transcripts with the same volume of sequencing data.



More effective sub-reads have been detected by Multi-throughput Iso-Seq

Notes:

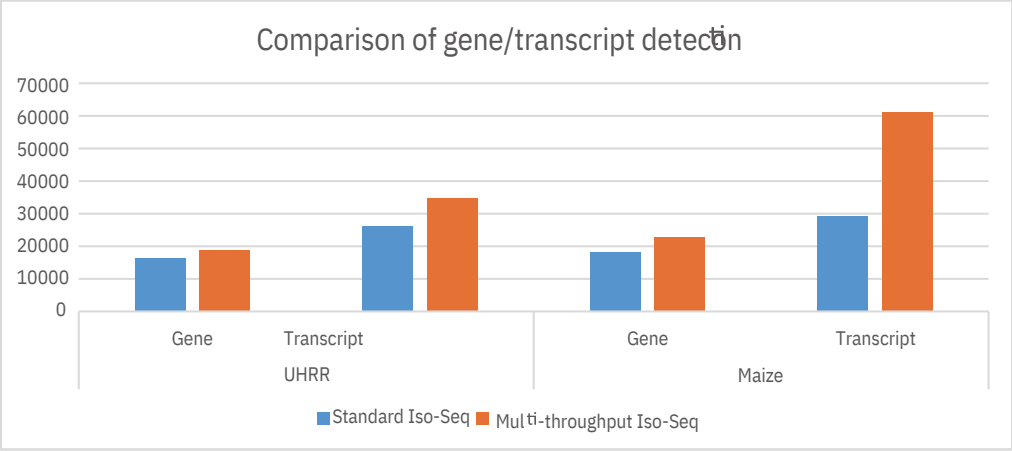
Total sequencing data amount as follows:
UHRR-9.47 Gb (Standard Iso-Seq); 9.57 Gb (Multi-throughput Iso-Seq);
Maize-11.04 Gb (Standard Iso-Seq); 19.3 Gb (Multi-throughput Iso-Seq).

Polymerase-reads :

The number of polymerases generated high quality reads. Polymerase reads will be then trimmed to preserve only the high-quality region, which includes bases from adaptors and single or multiple passes around a circular template.

Effective reads:

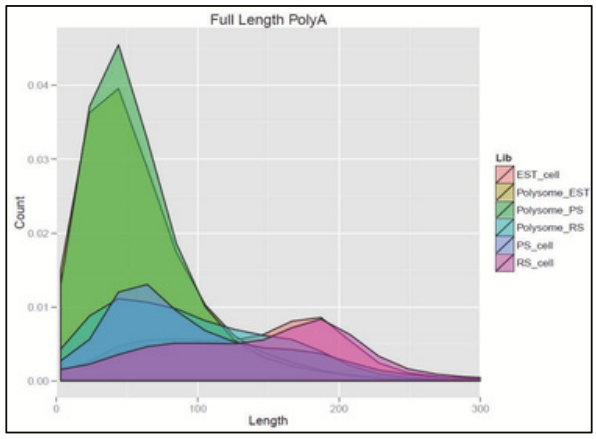
Each cDNA template molecule is considered as an “insert” and each pass through the insert is called a effective read. A polymerase read made by multi-throughput Iso-Seq can contain more than one unique inserts.



More genes/transcripts have been detected by Multi-throughput Iso-Seq.

3）polyA length analysis with polyA Iso-Seq

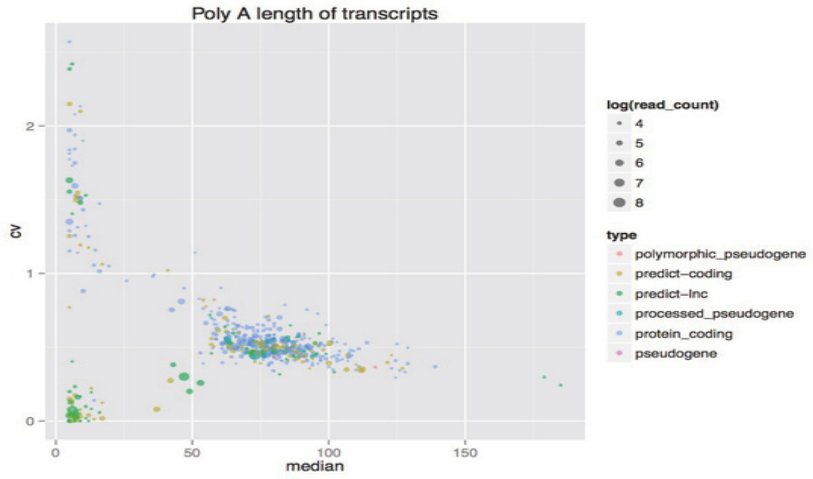
A variety of studies have reported the importance of polyA tail length for gene expression activity[1-3]. The polyA Iso-Seq service provides additional information of polyA sequence, such as the length distribution of polyA, ; the occurrence frequency of other bases in polyA and the correlation between polyA length and gene expression.



The length distribution of polyA

AAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA (1);
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA(35);
ACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA(1);
AAAAAAAAAAAAAAAAAAAAAAAAA (5);
ATAAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAA(1);
AAAAAAAAAACAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAA(1);
AAAAAAAAAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA (1);

The occurrence frequency of other bases in polyA



The correlation between polyA length and gene expression

References

[1] Subtelny A O, Eichhorn S W, Chen G R, et al. Poly (A)-tail profiling reveals an embryonic switch in translational control[J]. *Nature*, 2014, 508(7494): 66-71.
[2] Lim J, Lee M, Son A, et al. mTAIL-seq reveals dynamic poly (A) tail regulation in oocyte-to-embryo development[J]. *Genes & development*, 2016, 30(14): 1671-1682.
[3] Eichhorn S W, Subtelny A O, Kronja I, et al. mRNA poly (A)-tail changes specified by deadenylation broadly reshape translation in Drosophila oocytes and early embryos[J]. *Elife*, 2016, 5: e16955.

Mass Spectrometry

Protein Identification and Profiling

Mass Spectrometry (MS)–based proteomics has matured into an attractive technology for the analysis of protein composition, modifications, and dynamics in cell biology [1]. BGI has developed a comprehensive protein identification pipeline, including sample preparation, protein separation and purification, followed by mass spec analysis, and finally a systematic bioinformatics analysis.

Protein Identification

Proteome Profiling

Service Description

Modern liquid chromatography-mass spectrometry provides a means to routinely conduct non-targeted protein identification. The challenges this approach poses can be simplified with BGI's assistance.

Proteome Profiling

Proteome Profiling is a cost effective solution for the simultaneous monitoring of hundreds to thousands of proteins. We perform sample digestion using sequencing-grade trypsin or alternative proteolytic methods, and trypsin digested peptide samples are then analyzed using nano-flow LC-MS/MS. The label free area under the curve (AUC) method provides information on protein relative abundances[2]. Besides, our HPLC sample fractionation systems are highly recommended for complex sample types to reduce sample complexity, as well as to maximize protein sequence coverage and detection dynamic range [3].

Gel Spot Identification

We can identify proteins present in gel spots. Our service includes destaining, in-gel digestion, nano-flow LC-MS/MS, and peptide spectral match data analysis.

Gel Band Identification

We can identify proteins present in gel bands or protein solution.

Technical Specifications

Gel Spot/Gel Band Identification Sample Preparation and Services

- Digestion performed using sequencing-grade trypsin
- Off-line sample fractionation using HPLC technology
- Each fraction analyzed using nano-flow LC-MS/MS Peptide Spectral Match (PSM) data utilized for constructing proteomic profile
- Peptide Spectral Match (PSM) data utilized for constructing proteomic profile

Gel Spot/Gel Band Identification Sample Preparation and Services

- Digestion performed using sequencing-grade trypsin
- Nano-flow LC-MS/MS

Data Analysis

- Data analysis and validation performed with Mascot or SEQUEST
- GO (Gene ontology) category analysis
- COG (Cluster of Orthologous Groups of proteins) category analysis
- Pathway analysis

Quality Standard

Proteome Profiling, Gel Spot and Gel Band Identification summary includes all methods and data analysis. Reports provided in Excel or PDF format; RAW files available.

Turnaround Time

Typical 4-5 weeks from sample QC acceptance to data report delivery for Proteome Profiling and typical 2 weeks after sample receipt for raw data delivery for Gel Spot/Gel Band Identification.

Sample Requirements

Gel Spot/Gel Band Identification:

Sample \ Product	Gel Spot Identification	Gel Band Identification	Modification Site Identification
Comassive or silver stained protein spot/band	Visible spot	Protein≥1μg	Protein ≥ 1 μg, Projects that need enrichment is not accepted
Single protein (protein solution purified after IP or CO-IP)	/	≥ 5 μg, ≥ 0.5 μg/μL	≥ 10 μg, ≥ 0.5 μg/μL. If enrichment is required, the amount of sample sent needs to be increased by 3 times.
Note	/	/	Projects that need enrichment for acetylation and ubiquitination studies are not accepted

Proteome Profiling:

	Sample Type	Amount	
		Recommend	Minimum
Animal	Animal internal organs (heart, liver, spleen, lung, kidney), skin, muscle, brain, etc.	≥ 10 mg	≥ 5 mg
	Mollusks (Toxoplasma, Schistosomiasis, Drosophila, Acarid, Plutella xylostella, Laodelphax, Cestode, Cicada, Hematodinium, etc.)	≥ 10 mg	≥ 5 mg
Cell	Suspended cells, adherent cells	≥ 1×107	≥ 1×106
	Cell culture supernatant	≥ 20 mL	
Exosome	Exosome isolated by customer	≥ 100 μg, ≥ 0.5 μg/μL	
Fluid	Plasma, serum (remove highly-abundant protein)	≥ 200 μL	≥ 200 μL
	Amniotic fluid, cerebrospinal fluid, semen, etc. (remove highly-abundant protein)	≥ 1 mL	≥ 500 μL
	Amniotic fluid, cerebrospinal fluid, semen, etc. (with highly-abundant protein) Saliva, milk	≥ 200 μL	≥ 100 μL
	Urine	≥ 200 μL	≥ 100 μL
	Tears	≥ 50 mL	≥ 20 mL
	Twigs of plants (leaf buds, tender leaves), algae	≥ 25 μL	≥ 20 μL
Plant	Old leaves, roots, stems, bark of plants	≥ 500 mg	≥ 250 mg
	Plant buds, pollen	≥ 1 g	≥ 500 mg
	Plant seeds (rice/wheat seeds, etc.), fruits (apples, peaches, pears)	≥ 100 mg	≥ 50 mg
	Prokaryotic bacteria (E. coli, Staphylococcus aureus, etc.), fungi (yeast, etc.)	≥ 1 g	≥ 500 mg
Microorganism	Complex protein solution, protein powder	Thallus ≥ 50 mg, cells ≥ 5×106	
Protein solution	Feces	≥ 200 μg, ≥ 100 μg, ≥ 0.5 μg/μL	≥ 0.5 μg/μL
Others		≥ 500 mg	≥ 100 mg

References

[1]Gillet, L.C., A. Leitner, and R. Aebersold, Mass Spectrometry Applied to Bottom-Up Proteomics: *Entering the High-Throughput Era for Hypothesis Testing*. Annu Rev Anal Chem (Palo Alto Calif), 2016. 9(1): p. 449-72.
[2]Zhang, Y., et al., Tissue-Based Proteogenomics Reveals that Human Testis Endows Plentiful Missing Proteins. *Journal of Proteome Research*, 2015. 14(9): p. 3583-3594.
[3]Chandramouli, K. and P.Y. Qian, Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. *Hum Genomics Proteomics*, 2009.

Quantitative Proteomics

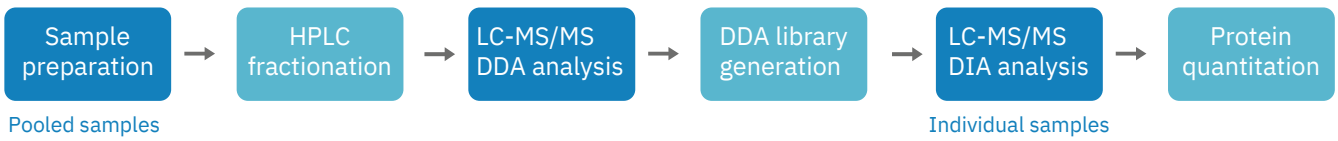
BGI supports multiple workflows for proteomic quantitation. We have developed innovative solutions for optimizing the process of accurately quantifying proteins in complex biological matrices and other complex mixtures. Our quantitative services mainly include Label-Free Data-Independent Acquisition (DIA), Isobaric Label (iTRAQ, TMT and IBT tag), Rapid DDA Proteomics, Targeted Proteomics (MRM/PRM), Nanoproteomics, FFPE Proteomics and Metaproteomics. We are happy to work with you to determine the best quantitative proteomics strategy for your project [1].

Service Description

Our quantitative proteomics services are performed using nano-flow liquid chromatography and high resolution Orbitrap mass spectrometry. Sample digestion is performed using sequencing-grade trypsin or alternative proteolytic method. We can provide Targeted MRM/PRM Peptide Quantitation service using micro-flow or analytical flow rates for high sample throughput.

Label-Free DIA Proteomics

Data-Independent Acquisition (DIA) is a label-free quantitative technique which provides highly consistent quantitation and broad proteome coverage. We create a customized spectral library with your specific sample and then analyze individual samples using nano-flow LC-MS/MS with DIA scanning mode [2]. Label-Free DIA Quantitative Proteomics service is ideal for long-term projects or projects with large sample sets which require accurate and reproducible quantitation.



Isobaric Label Proteomics

Isobaric Label services incorporate sample multiplexing to provide deep proteome coverage and highly precise quantitation of small or medium-sized sample sets [3]. BGI provides Isobaric Label Quantitative Proteomics service using iTRAQ (Isobaric Tags for Relative and Absolute Quantitation), IBT (Isobaric Tags for Quantitation) and TMT (Tandem Mass Tags for Quantitation) technologies developed by AB SCIEX, BGI and Thermo Scientific, respectively.

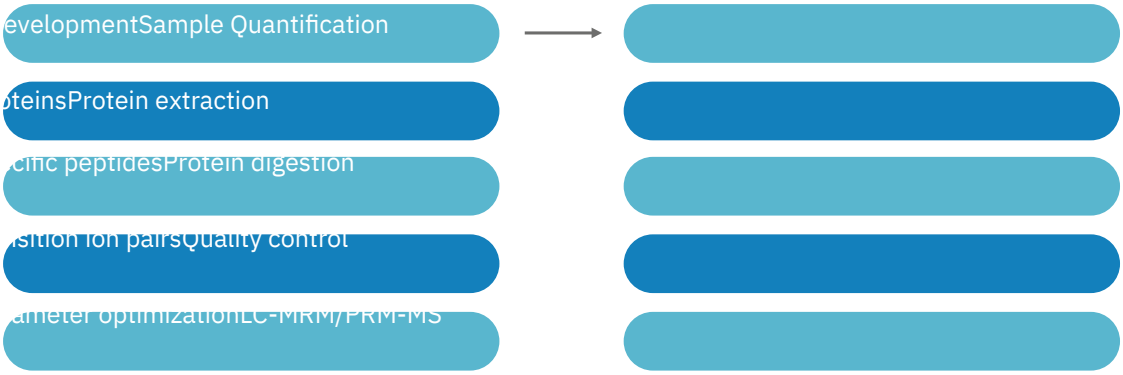
Rapid DDA Proteomics

In contrast to label-dependent quantitative proteomics such as the iTRAQ method, label-free data-dependent acquisition (DDA) proteomics doesn't need to use expensive stable isotopes for labeling of proteins. This allows accurate protein quantification of samples without additional error-prone in vitro labelling reactions and enables fast proteome identification and quantification [4].

Targeted Proteomics

Traditional target protein analysis is based on antibody techniques, which have limitations in throughput and specificity. However, multiple reaction monitoring (MRM) and parallel reaction monitoring (PRM) approaches [5] are able to resolve these obstacles.

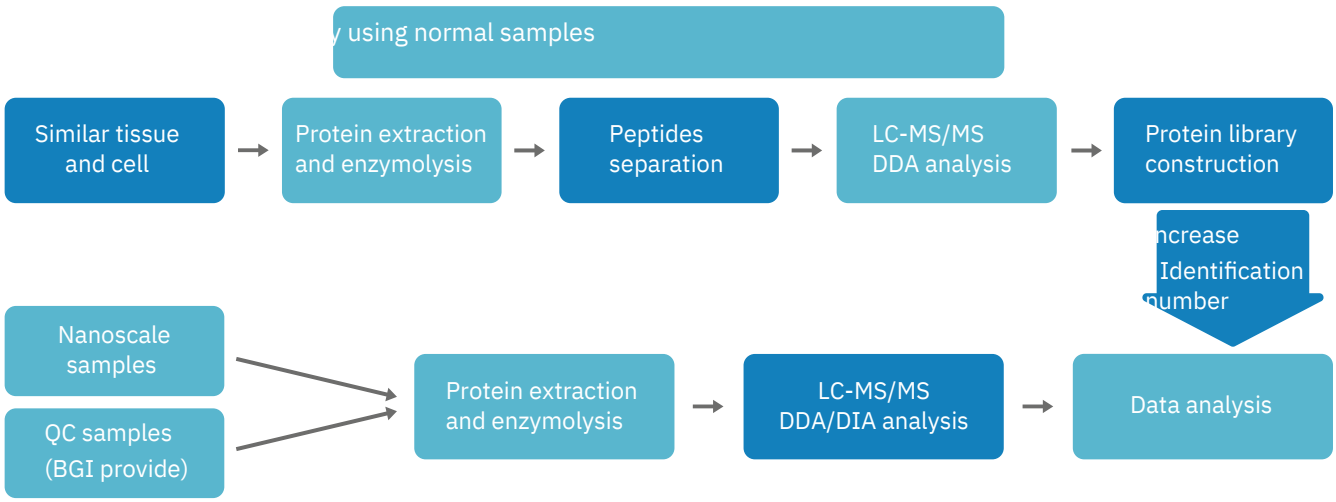
As the gold standard for targeted proteomics research based on MS platforms, MRM/PRM technology can quickly and accurately determine dozens of proteins simultaneously in complex biological samples and solve slow verification problems.



Nanoproteomics

Nanoproteomics refers to quantitative proteomics analysis of small cell populations (typically < 5,000 cells) by the combination of in-situ cleavage and DDA label free technology, and offers possibilities for analysis unavailable with conventional protein extraction and mass spectrometry [6].

BGI has extensive experience in the field of nanoproteomics spanning cellular heterogeneity research across rare cell populations, hard-to-obtain clinical specimens, and pathological tissues.



FFPE Proteomics

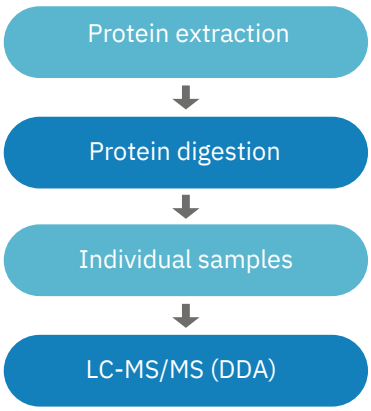
'FFPE' samples refer to Formalin-Fixed Paraffin-Embedded samples, and is a form of sample preservation and preparation for biopsy specimens. FFPE samples can be stably stored for decades at room temperature and is the gold standard for histoppathological analysis [7]. BGI has extensive experience in the field of FFPE proteomics of tumor heterogeneity research.

FFPE sample analysis when there are suitable library construction samples:

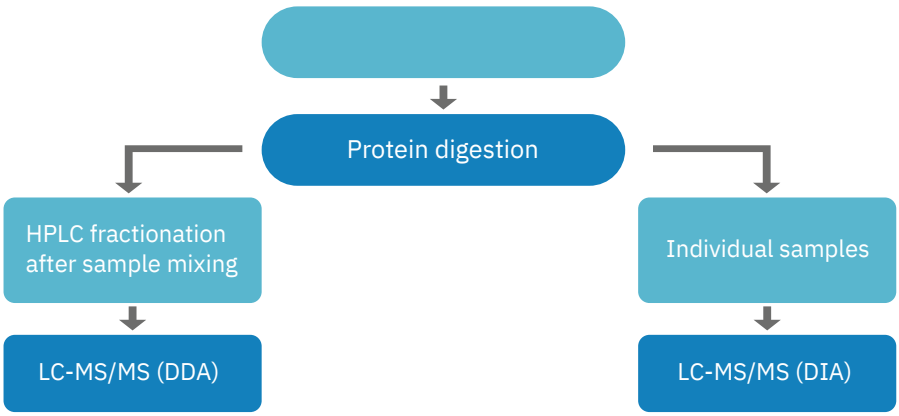
FFPE sample analysis when there are no suitable library construction samples:

Metaproteomics

Since its inception in 2004, metaproteomics aims at the large-scale characterization of all proteins expressed by environmental microbiota at a given point in time, which can better promote the in-depth study of the diversity, structure and potential gene function of the microbial community [8].
BGI has extensive experience in the field of metaproteomics spanning microorganisms research across human fecal samples, water samples, and fermentation liquor samples.
DDA analysis can be carried out if there are no suitable library construction samples:



DIA analysis can be carried out if there are suitable library construction samples:



Technical Specifications

Our quantitative proteomics services are performed using nano-flow liquid chromatography and cutting-edge mass spectrometer (Q Exactive HF/HF-X, Orbitrap Fusion Lumos , QTRAP 6500, Eclipse，Exploris 480 and TimsTOF). Sample digestion is performed using sequencing-grade trypsin or alternative proteolytic method. We can provide Targeted Peptide Quantitation services using micro-flow or analytical flow rates for high sample throughput.

Label-Free DIA Proteomics Sample Preparation and Services

- The QC analysis of enzymatic hydrolysis is strictly controlled by LC-MS
- DDA-Spectral Library generated using HPLC technology and sample fractionation

Isobaric Label Proteomics Sample Preparation and Services

- The QC of enzymatic hydrolysis and labeling steps are strictly controlled by LC-MS DDA analysis
- Increase proteome depth of coverage using HPLC sample fractionation
- Each fraction analyzed using nano-flow LC-MS/MS

Rapid DDA Proteomics Sample Preparation and Services

- The QC analysis of enzymatic hydrolysis is strictly controlled by LC-MS

Targeted Proteomics Sample Preparation and Services

- Customized LC method and MRM/PRM-acquisition MS method

Nanoproteomics Sample Preparation and Services

- The QC analysis of enzymatic hydrolysis is strictly controlled by LC-MS

FFPE Proteomics Sample Preparation and Services

- The QC analysis of enzymatic hydrolysis is strictly controlled by LC-MS

Metaproteomics Sample Preparation and Services

- The QC analysis of enzymatic hydrolysis is strictly controlled by LC-MS

Data Analysis

- Data analysis and validation performed with Mascot and SEQUEST
- GO (Gene Ontology) category analysis
- Pathway analysis
- DEPs (deferentially expressed proteins) cluster analysis
- DEPs GO enrichment analysis
- DEPs pathway enrichment analysis

Quality Standard

Label-Free DIA Proteomics, Isobaric Label Proteomics, Rapid DDA Proteomics, Targeted Proteomics, Nanoproteomics, FFPE Proteomics, and Metaproteomics include all methods and data analysis. Reports provided in Excel or PDF format; RAW files available.

Turnaround Time

Typical 3-8 weeks from sample QC acceptance to data report delivery for Quantitative Proteomics.

Sample Requirements

We accept protein samples in a variety of formats. For attaining maximum proteome coverage, we recommend utilizing our sample fractionation services, performed using our off-line HPLC platform.

Label-Free DIA Proteomics:

	Sample Type	Amount	
		Recommend	Minimum
Animal	Common animal tissues: animal internal organs (heart, liver, spleen, lung, kidney), skin, muscle, brain, etc.	≥ 5 mg	≥ 1 mg
	Mollusks (Toxoplasma, Schistosomiasis, Drosophila, Acarid, Plutella xylostella, Laodelphax, Cestode, Cicada, Hematodinium, etc.)	≥ 5 mg	≥ 2 mg
Cell	Suspended cells, adherent cells	≥ 1×107	≥ 1×106
	Cell culture supernatant	≥ 5 mL	
Exosome	Exosome isolated by customer	≥ 20 μg, ≥ 0.5 μg/μL	
Fluid	Plasma, serum (remove highly-abundant protein)	≥ 200 μL	≥ 100 μL
	Plasma, serum (with highly-abundant protein)	/	
	Amniotic fluid, cerebrospinal fluid, semen, etc. (remove highly-abundant protein)	≥ 1 mL	≥ 500 μL
	Amniotic fluid, cerebrospinal fluid, semen, etc. (with highly-abundant protein) Saliva, milk	≥ 200 μL	≥ 100 μL
	Urine	≥ 200 μL	≥ 100 μL
	Tears	≥ 30 mL	≥ 15 mL
	Twigs of plants (leaf buds, tender leaves), algae	≥ 15 μL	≥ 10 μL
Plant	Old leaves, roots, stems, bark of plants	≥ 300 mg	≥ 200 mg
	Plant buds, pollen	≥ 1 g	≥ 500 mg
	Plant seeds (rice/wheat seeds, etc.), fruits (apples, peaches, pears)	≥ 100 mg	≥ 50 mg
	Prokaryotic bacteria (E. coli, Staphylococcus aureus, etc.), fungi (yeast, etc.)	≥ 1 g	≥ 500 mg
Microorganism	Thallus ≥ 50 mg，cells ≥ 5×106		
Protein solution	Complex protein solution, protein powder Feces	≥ 40 μg, ≥ 0.5 μg/μL	≥ 20 μg, ≥ 0.5 μg/μL
Others		≥ 300 mg	≥ 100 mg

Isobaric Label Proteomics:

	Sample Type	Amount	
		Recommend	Minimum
Animal	Common animal tissues: animal internal organs (heart, liver, spleen, lung, kidney), skin, muscle, brain, etc.	≥ 10 mg	≥ 1 mg
	Mollusks (Toxoplasma, Schistosomiasis, Drosophila, Acarid, Plutella xylostella, Laodelphax, Cestode, Cicada, Hematodinium, etc.)	≥ 10 mg	≥ 2 mg
Cell	Suspended cells, adherent cells	≥ 1×107	≥ 1×106
	Cell culture supernatant	≥ 10 mL	
Exosome	Exosome isolated by customer	≥ 50 μg, ≥ 0.5 μg/μL	
Fluid	Plasma, serum (remove highly-abundant protein)	≥ 200 μL	≥ 200 μL
	Plasma, serum (with highly-abundant protein)	≥ 10 μL	≥ 3 μL
	Amniotic fluid, cerebrospinal fluid, semen, etc. (remove highly-abundant protein)	≥ 1 mL	≥ 500 μL
	Amniotic fluid, cerebrospinal fluid, semen, etc. (with highly-abundant protein) Saliva, milk	≥ 200 μL	≥ 100 μL
	Urine	≥ 200 μL	≥ 100 μL
	Tears	≥ 30 mL	≥ 15 mL
	Twigs of plants (leaf buds, tender leaves), algae	≥ 15 μL	≥ 10 μL
Plant	Old leaves, roots, stems, bark of plants	≥ 300 mg	≥ 200 mg
	Plant buds, pollen	≥ 1 g	≥ 500 mg
	Plant seeds (rice/wheat seeds, etc.),	≥ 100 mg	≥ 50 mg
	fruits (apples, peaches, pears)	≥ 1 g	≥ 500 mg
Microorganism	Prokaryotic bacteria (E. coli, Staphylococcus aureus, etc.), fungi (yeast, etc.)	Thallus ≥ 50 mg,	cells ≥ 5×106
Protein solution	Complex protein solution, protein powder	≥ 100 μg, ≥ 0.5 μg/μL	≥ 50 μg, ≥ 0.5 μg/μL
Others	Feces	≥ 300 mg	≥ 100 mg

Rapid DDA Proteomics:

	Sample Type	Amount	
		Recommend	Minimum
Animal	Common animal tissues: animal internal organs (heart, liver, spleen, lung, kidney), skin, muscle, brain, etc.	≥ 5 mg	≥ 1 mg
	Mollusks (Toxoplasma, Schistosomiasis, Drosophila, Acarid, Plutella xylostella, Laodelphax, Cestode, Cicada, Hematodinium, etc.)	≥ 5 mg	≥ 1 mg
Cell	Suspended cells, adherent cells	≥ 1×105	≥ 1×105
	Cell culture supernatant	≥ 5 mL	
Exosome	Exosome isolated by customer	≥ 10 μg, ≥ 0.5 μg/μL	
Fluid	Plasma, serum (remove highly-abundant protein)	≥ 200 μL	≥ 100 μL
	Plasma, serum (with highly-abundant protein)	/	
	Amniotic fluid, cerebrospinal fluid, semen, etc. (remove highly-abundant protein)	≥ 1 mL	≥ 500 μL
	Amniotic fluid, cerebrospinal fluid, semen, etc. (with highly-abundant protein) Saliva, milk	≥ 100 μL	≥ 50 μL
	Urine	≥ 50 μL	≥ 20 μL
	Tears	≥ 20 mL	≥ 10 mL
	Twigs of plants (leaf buds, tender leaves), algae	≥ 5 μL	≥ 3 μL
Plant	Old leaves, roots, stems, bark of plants	≥ 200 mg	≥ 100 mg
	Plant buds, pollen	≥ 500 mg	≥ 500 mg
	Plant seeds (rice/wheat seeds, etc.), fruits (apples, peaches, pears)	≥ 20 mg	≥ 10 mg
	Prokaryotic bacteria (E. coli, Staphylococcus aureus, etc.), fungi (yeast, etc.)	≥ 500mg	≥ 200 mg
Microorganism	Complex protein solution, protein powder	Thallus ≥ 50 mg，cells ≥ 5×106	
Protein solution	Feces	≥ 20 μg, ≥ 0.5 μg/μL	≥ 5 μg, ≥ 0.5 μg/μL
Others		≥ 200 mg	≥ 100 mg

Targeted Proteomics:

	Sample Type	Amount	
		Recommend	Minimum
Animal	Animal internal organs (heart, liver, spleen, lung, kidney), skin, muscle, brain, etc.	≥ 10 mg	≥ 1 mg
	Mollusks (Toxoplasma, Schistosomiasis, Drosophila, Acarid, Plutella xylostella, Laodelphax, Cestode, Cicada, Hematodinium, etc.)	≥ 10 mg	≥ 2 mg
Cell	Suspended cells, adherent cells	≥ 1×107	≥ 1×106
	Cell culture supernatant	≥ 10 mL	
Exosome	Exosome isolated by customer	≥ 50 μg, ≥ 0.5 μg/μL	
Fluid	Plasma, serum (remove highly-abundant protein)	≥ 200 μL	≥ 200 μL
	Plasma, serum (with highly-abundant protein)	≥ 10 μL	≥ 3 μL
	Amniotic fluid, cerebrospinal fluid, semen, etc. (remove highly-abundant protein)	≥ 1 mL	≥ 500 μL
	Amniotic fluid, cerebrospinal fluid, semen, etc. (with highly-abundant protein) Saliva, milk	≥ 200 μL	≥ 100 μL
	Urine	≥ 200 μL	≥ 100 μL
	Tears	≥ 30 mL	≥ 15 mL
	Twigs of plants (leaf buds, tender leaves), algae	≥ 15 μL	≥ 10 μL
Plant	Old leaves, roots, stems, bark of plants	≥ 300 mg	≥ 200 mg
	Plant buds, pollen	≥ 1 g	≥ 500 mg
	Plant seeds (rice/wheat seeds, etc.),	≥ 100 mg	≥ 50 mg
	fruits (apples, peaches, pears)	≥ 1g	≥ 500 mg
Microorganism	Prokaryotic bacteria (E. coli, Staphylococcus aureus, etc.), fungi (yeast, etc.)	Thallus ≥ 50 mg，	cells ≥ 5×106
Protein solution	Complex protein solution, protein powder	≥ 100 μg, ≥ 0.5 μg/μL	≥ 50 μg, ≥ 0.5 μg/μL
Others	Feces	≥ 300 mg	≥ 100 mg

Nanoproteomics:

	Sample Type	Amount	Note*
Library construction samples	Mammalian cells from similar or adjacent parts	5*105	Mammalian cells fromsimilar or adjacent parts
	Tissue from similar or adjacent parts	10 mg	
Mammalian cells		> 100	Collected cells or tissues (not LCM) with 200 μL low adsorption PCR tube and stored in 2-5 μL PBS solution
Nanoscale samples	Egg cell or early embryo cell	> 1, and 3-5 is better	
	Tissue	1mg	
	LCM (Laser capture microdissection)	Thickness of 10 μm, > 10mm2	

FFPE Proteomics:

	Sample Type	Amount	Note*
Library construction samples	Fresh samples of relevant tissue	> 5 mg	Mammalian cells fromsimilar or adjacent parts
	FFPE tissue block	> 1 mg	
	FFPE slice	Minimum: 150 mm2;3 slices, thickness of 10 μm, > 50 mm2 Recommend: 500 mm2;10 slices, thickness of 10 μm, > 50 mm2	
DIA analysis samples	FFPE slice	≥ 1 slice, thickness of 10 μm, > 50 mm2; non-staining or hematoxylin-eosin staining > 1 mg; non-staining	FFPE samples t ransport at room temperature; Fresh tissue samples transport with dry ice
	FFPE tissue block	Cells > 5000; non-staining or hematoxylin-eosin staining	
	LCM FFPE (Laser capture microdissection)		

Metaproteomics:

Please be sure to send the sample after enrichment.

References

[1] Cheung, C.H.Y. and H.F. Juan, Quantitative proteomics in lung cancer. *J Biomed Sci*, 2017. 24(1): p. 37.

[2] Searle, B.C., et al., Chromatogram libraries improve peptide detection and quantification by data independent acquisition mass spectrometry. *Nat Commun*, 2018. 9(1): p. 5128.

[3] Rauniyar, N. and J.R. Yates, 3rd, Isobaric labeling-based relative quantification in shotgun proteomics. J Proteome Res, 2014. 13(12): p. 5293-309.

[4] Krasny L, et al., Data-independent acquisition mass spectrometry (DIA-MS) for proteomic applications in oncology. *Mol Omics*, 2020. 9(4).

[5] Method of the Year 2012. *Nat Methods*, 2013. 10(1).

[6] L. Yi, et al., Advances in microscale separations towards nanoproteomics *applications, 2017*.1523: 40–48.

[7] Marchione D M, et al., HYPERsol: High-Quality Data from Archival FFPE Tissue for Clinical Proteomics. *J Proteome Res*, 2020 Feb 7; 19(2): 973–983.

[8] Kleiner M. Metaproteomics: Much More than Measuring Gene Expression in Microbial Communities. *MSystems*, 2019 May-Jun; 4(3): e00115-19.

PTM Proteomics

Service Description

Post-translational modification (PTM) of proteins refers to the chemical changes proteins may undergo after translation. Common PTMs include Phosphorylation, Glycosylation, Acetylation and Ubiquitination. As a result, identifying and understanding PTMs is key for the study of cell biology and the treatment and prevention of disease.
BGI has extensive experience in the field of PTM Proteomics [1,2] and has developed reliable workflows using market leading technologies and a bioinformatics infrastructure that is second to none.

Protein Phosphorylation Analysis

Protein phosphorylation analysis mainly includes phosphorylated protein identification (single protein phosphorylation identification/phosphoproteome profiling) and phosphorylated protein quantification.
Single phosphoprotein identification aims to identify the phosphorylation-site on purified protein Phosphoproteome profiling is the method of detecting all of the phosphorylated proteins and their phosphorylation sites in the organisms based on phosphopeptides enrichment and high-throughput MS technologies, including the processes of sample preparation, phosphopeptides enrichment, LC-MS/MS analysis and bioinformatics analysis.

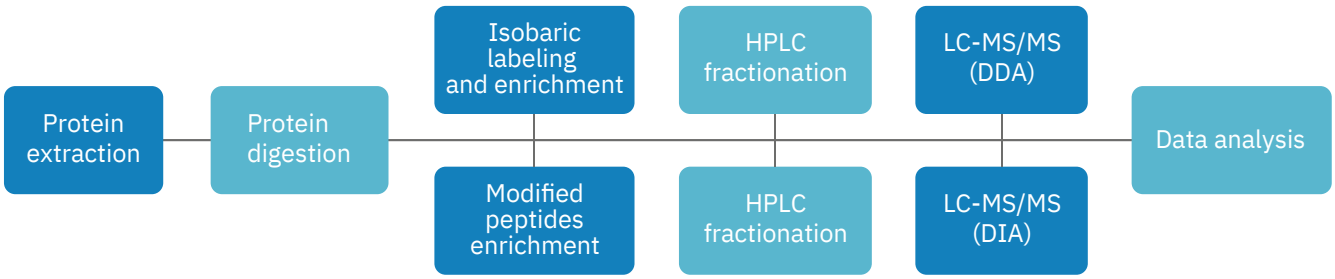
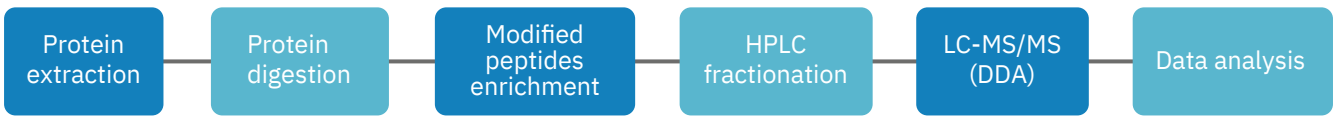


Figure 2. Phosphoproteome quantification workflow

Other PTM Proteomics Services

Protein Acetylation Analysis
Protein N glycosylation Analysis



Service Specifications

PTM Proteomics services are performed using nano-flow liquid chromatography and cutting-edge mass spectrometer (Q Exactive HF/HF-X and Orbitrap Fusion Lumos).

Sample Preparation and Services

- Digestion performed using sequencing-grade trypsin
- Off-line sample fractionation using HPLC technology
- Each fraction analyzed using nano-flow LC-MS/MS
- TiO2 enrichment for phosphoproteomics; PTM-104 antibody enrichment for acetylated proteins; HILIC column enrichment for glycosylated proteins

Protein Phosphorylation Analysis

- Purified protein phosphorylation identification
- Phosphoproteome profiling
- Protein Phosphorylation iTRAQ/IBT/TMT Quanitification Analysis
- Protein Phosphorylation label-free DIA Quantification Analysis
- Protein phosphorylation PRM

1. Full spectrum identification
2. Sample quantification

Protein Acetylation Analysis

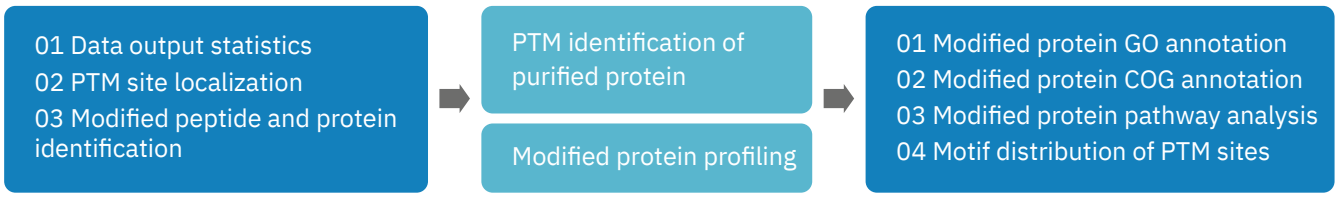
- Purified protein acetylation identification
- Acetylated protein full spectrum identification
- Acetylated proteome quantification

Protein N glycosylation Analysis

- Purified protein N glycosylation identification
- N glycosylated protein full spectrum identification
- N glycosylated proteome quantification

Data Analysis

Standard Protein PTM Identification:



Standard Protein PTM Quantification:

Identification

01 Statistics of modified protein identification
02 Statistics of PTM sites
03 Quality evaluation of modified protein identification

Quantification

01 Statistics of differential modified peptides
02 Quantitative repeatability assessment
03 Cluster analysis of differential modified peptides

Protein

PTM

Protein Function Annotation

01 GO annotation
02 COG/KOG annotation
03 Pathway annotation

Differential Protein Function Enrichment

01 GO enrichment analysis
02 Pathway enrichment analysis
03 COG/KOG annotation
04 Protein interaction analysis
05 Protein subcellular localization analysis

Customized Services:

Kinase prediction analysis
Quantitative proteomics and phosphoproteomics correlation analysis

Quality Standard

PTM proteomics services include all methods and data analysis. Reports provided in Excel or PDF format; RAW files available.

Turnaround Time

PTM Type	Turn Around Time from sample QC acceptance to data report delivery
Phosphorylation	3-5 weeks
Glycosylation	4-5 weeks
Acetylation	3-5 weeks

Sample Requirements (eg: Phosphoproteome Analysis)

We accept protein samples in a variety of formats. For attaining maximum proteome coverage, we recommend utilizing our sample fractionation services, performed using our off-line HPLC platform.

	Sample Type	Amount	Phosphorylation	Glycosylation	Acetylation
Animal	Animal internal organs (heart, liver, spleen, lung, kidney), skin, muscle, brain.	Recommended	≥ 20-30 mg	≥ 20 mg	≥ 200 mg
		Minimum	≥ 10 mg	≥ 10 mg	≥ 100 mg
	Mollusks(including Toxoplasma, Schistosomiasis, Drosophila, Acarid, Plutella xylostella, Laodelphax, Cestode, Cicada, Hematodinium)	Recommended	≥ 20-30 mg	≥ 20 mg	≥ 200 mg
		Minimum	≥ 10 mg	≥ 10 mg	≥ 100 mg
Cell	Suspended cells, adherent cells	Recommended	≥ (2-5) ×107	≥ 2×107	≥ 1×108
		Minimum	≥ 2×107	≥ 2×107	≥ 1×108
Plant	Plants (leaf buds, tender leaves), algae	Recommended	≥ 2 g	≥ 2 g	≥ 5 g
		Minimum	≥ 1 g	≥ 1 g	≥ 3 g
	Old leaves, roots, stems, bark of plants	Recommended	≥ 4 g	≥ 4 g	≥ 15 g
		Minimum	≥ 2 g	≥ 2 g	≥ 5 g
	Plant buds, pollen	Recommended	≥ 200 mg	≥ 200 mg	≥ 1 g
		Minimum	≥ 100 mg	≥ 100 mg	≥ 500 g
	Plant seeds (including rice/wheat seeds),fruits (including apples, peaches, pears)	Recommended	≥ 500 mg	≥ 500 mg	≥ 5 g
		Minimum	≥ 200 mg	≥ 200 mg	≥ 2 g
Microorganism	Prokaryotic bacteria (including E. coli, Staphylococcus aureus), fungi (including yeast)	Recommended	Thallus ≥ 200 mg cells ≥ 2×107	Thallus ≥ 200 mg cells ≥ 2×107	Thallus ≥ 1 g cells ≥ 1×108
Protein solution	Complex protein solution, protein powder	Recommended	≥ 1-3 mg, ≥ 0.5 μg/μL	≥ 1 mg, ≥ 0.5 μg/μL	≥ 10 mg, ≥ 0.5 μg/μL
		Minimum	≥ 0.5 mg, ≥ 0.5 μg/μL	≥ 1 mg, ≥ 0.5 μg/μL	≥ 5 mg, ≥ 0.5 μg/μL

References

[1] Zhang, Y., et al., Evidence for Differential Glycosylation of Trophoblast Cell Types. *Mol Cell Proteomics*, 2016 Jun;15(6):1857-66. doi: 10.1074/mcp.M115.055798.

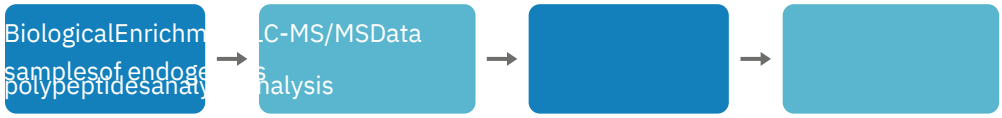
[2] Hao P L, et al., Correction of Errors in Tandem Mass Spectrum Extraction Enhances Phosphopeptide Identification. *J. Proteome Res.,* 2013, 12, 12, 5548–5557. doi: 10.1021/pr4004486.

Peptidomics

We have developed innovative solutions for optimizing the process of comprehensive qualitative and quantitative analysis of all endogenous peptides in complex biological matrices and other complex mixtures.

Service Description

Peptidomics is an emerging field filling the research gap between proteomics and metabolomics and can be defined as the comprehensive qualitative and quantitative analysis of all endogenous peptides from biological samples at a specific time and location.
BGI has extensive experience in the field of peptidomics spanning disease biomarker study, off target effects of drugs, food processing and fermentation research, the molecular mechanisms of animals and plants, and cell antigen epitopes prediction [1,2].



Technical Specification

Peptidomics services are performed using nano-flow liquid chromatography and cutting-edge mass spectrometer (Q Exactive HF and Orbitrap Fusion Lumos).

Sample Preparation and Services

- Enrichment of endogenous polypeptides
- 1h for LC-MS/MS analysis

Data Analysis

- Data analysis and validation performed with Mascot or SEQUEST
- Identification of peptides and proteins
- Peptide quantification
- Protein GO/COG/Pathway annotation
- Identified peptide matching analysis of specific protein

Quality Standard

- Summary includes all methods data analysis
- Reports provided in Excel or PDF format, RAW files available

Turnaround Time

Typical 4-5 weeks from sample QC acceptance to data report delivery

Sample Requirements

We accept samples in a variety of formats.

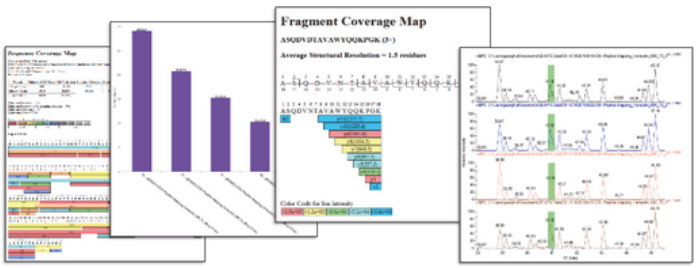
Sample Type		Amount	Comment
Body fluids	Serum/Plasma	200 μL	100 μL/one experiment
	Saliva	5 mL	2 mL/one experiment
	Urine	50 mL	20 mL/one experiment
Enriched polypeptides		2 μg	For complex polypeptide samples, the sample amount of molecular weight less than 10kDa is > 2 μg (refer to A280 detection value), and the purity is > 90%. And it does not contain special and hard-to-remove substances, such as SDS, Triton x, chaps, etc

References

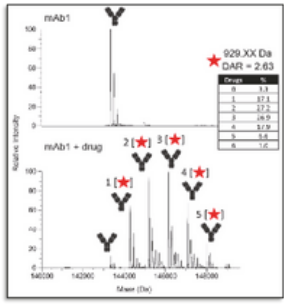
[1] Chen X Y, et al., High-Throughput Identification of Putative Antimicrobial Peptides from Multi-Omics Data of the Lined Seahorse (Hippocampus erectus). *Mar. Drugs*, 2020, 18(1), 30. doi: 10.3390/md18010030.
[2] Yi Y H, et al., High Throughput Identification of Antihypertensive Peptides from Fish Proteome Datasets. *Mar Drugs,* 2018 Oct 2;16(10):365. doi: 10.3390/md16100365.

Biologics Characterization

BGI provides LC-MS/MS peptide mapping services to support biopharmaceutical and biotechnology applications. Our Peptide Mapping services are designed to characterize and monitor the molecular details of a therapeutic protein drug at each position in the amino acid sequence [1,2]. Meanwhile, Disulfide Mapping and Epitope Mapping services are provided as well. BGI has also developed cutting edge Intact Mass services including Native MS utilizing multiple HPLC separation strategies combined with high resolution Orbitrap mass spectrometry [3,4,5].



Peptide Mapping



Intact Mass

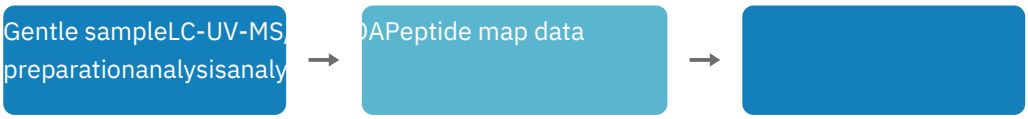
Service Description

Peptide Mapping

Sequence Verification

We provide amino acid sequence verification services and PTM mapping for purified proteins. Our service is useful for accurately validating expression of recombinant proteins or identifying process related PTMs on biologic drugs.

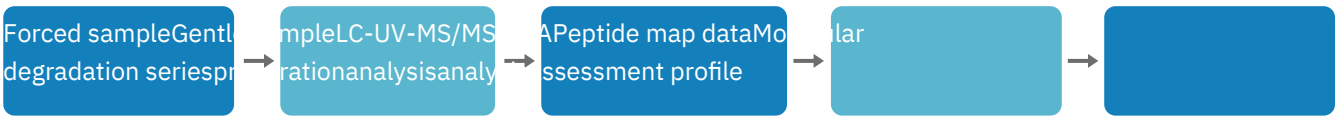
Individual samples



Molecular Assessment

We can help you characterize amino acid sequence liability. We provide a Molecular Assessment service to monitor specific amino acid modifications in response to a forced degradation treatment customized to meet your project needs. Our service is optimized for high dynamic range peptide sequence coverage.

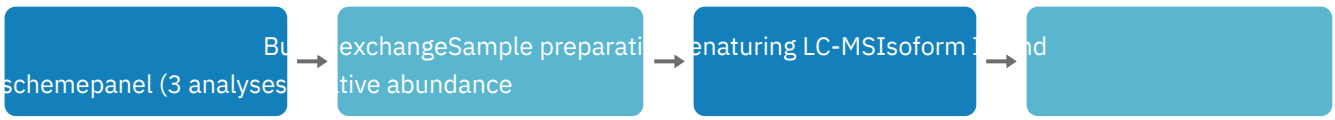
Sample batchIndividual samples



Intact Mass

Our service incorporates a three way analysis of the untreated, deglycosylated, and reduced preparations of a biologic drug using reverse phase LC-MS. Our Intact Mass service is designed to accommodate screening and confirmation projects with large sample numbers [3].
Our scientific team have developed unique, cutting edge methodologies which can support a wide range of biologic intact mass applications. We can help you measure chain pairing of bispecific antibody samples. We also provide a complete solution for measuring drug-to-antibody ratio (DAR) on antibody-drug conjugate (ADC) samples or any other protein-small molecule conjugates.

Denaturing conditions

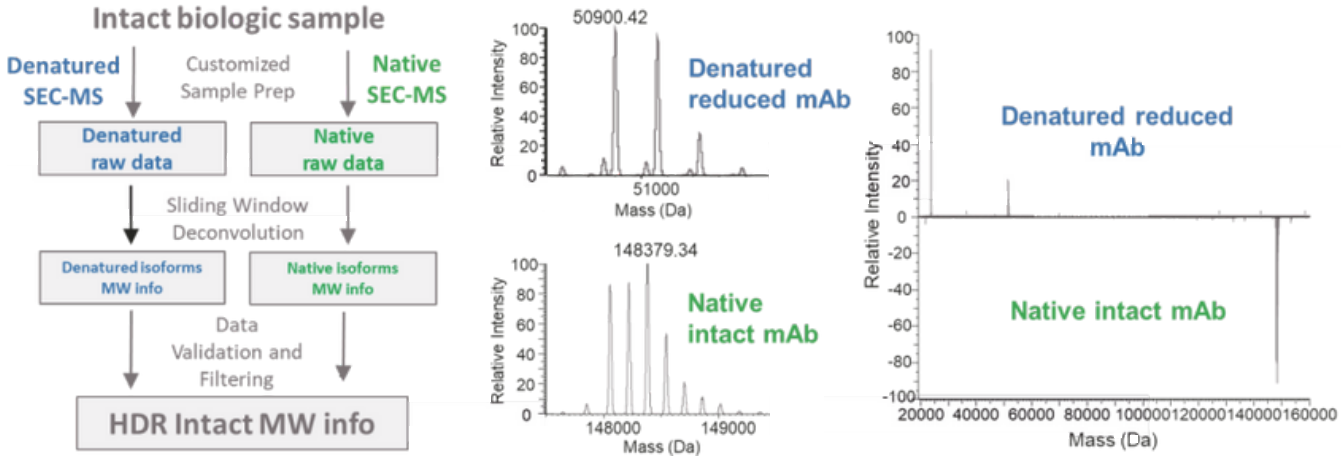


High Dynamic Range (HDR) Intact Mass

BGI provides a service for intact mass analysis of complex mixtures of biologic isoforms, including moderate-to-high glycosylation or small molecule conjugation.
High Dynamic Range (HDR) Intact Mass service combines data from both denaturing and native conditions. This service is based on size exclusion (SEC) LC-MS technology to directly measure masses using both denaturing and native conditions. This service can provide a highly confident DAR measurement of intact cysteine-linked ADCs.

Native and denaturing conditions

HDR Intact Mass service incorporates 5 key analyses, including native masses for (1) intact and (2) Ides-digested subunit (for antibody-related biologics), and denatured masses for (3) intact, (4) deglycosylated, and (5) disulfide bond reduced form.
We are happy to work with you to determine the best sample preparation strategy for comprehensive intact mass data of your biologic sample.
Please ask us about our specialized services for measuring charge variant profile or protein-ligand binding using native MS technology [4,5].



Technical Specifications

Peptide Mapping services are performed using analytical flow liquid chromatography, UV detection, and high resolution Orbitrap mass spectrometry.

Intact Mass services are performed using analytical flow liquid chromatography, UV detection, and high resolution Orbitrap mass spectrometry.

Peptide Mapping Sample Preparation Services

- Gentle digestion performed using sequencing-grade trypsin or alternative proteolysis designed to minimize sample preparation artifacts
- In-line UV detection also possible. Optimized for minimal sample oxidation
- Each fraction analyzed using 150 min analytical flow LC-MS/MS using a Q Exactive HF-X Orbitrap mass spectrometer
- Data analysis and validation performed with BioPharma Finder software
- We can work with you to develop customized Multi-Attribute Method (MAM) services using Chromeleon 7 software (21 CFR Part 11 compliant)

Intact Mass Sample Preparation Services

- Includes buffer exchange and denaturing sample preparation panel including denatured intact, de-glycosylation, and disulfide bond reduction
- Intact Mass service utilizes 20 min analytical flow reverse phase LC-MS/MS using Q Exactive HF-X BioPharma mass spectrometer
- High Dynamic Range (HDR) Intact Mass service utilizes 30 min SEC-MS performed using both denaturing and native MS-friendly mobiles phase.
- HDR Intact Mass service includes native LC-MS measurement of intact and subunit masses

Data analysis and validation performed with BioPharma Finder software

Data Analysis

Peptide Mapping:

- Data analysis and validation performed with BioPharma Finder software
- Peptide mapping method for identifying sequence variants, unknown modification, and PTM co-occupancy
- Amino acid site PTM status and relative abundance

Intact Mass:

1. Data analysis and validation performed with BioPharma Finder software.
2. Service allows protein isoforms to be identified by exact intact mass and quantified relative to other isoforms present

Quality Standard

- Summary includes all methods and data analysis
- Reports provided in Excel or PDF format, RAW files available

Turnaround Time

Typical 3 weeks from sample QC acceptance to data report delivery

Sample Requirements

Sample Type	Amount and Concentration		Minimum Sample Volume
Purified biologic sample in liquid solution	Recommended	1000 μg; 10 μg/μL	100 μL
	Required	10 μg; 1 μg/μL	10 μL

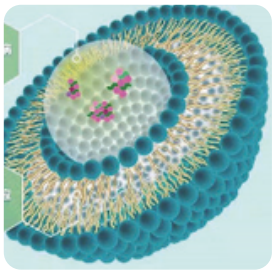
References

[1]. Lauber MA, Koza SM, McCall SA, Alden BA, Iraneta PC, Fountain KJ. High-resolution peptide mapping separations with MS-friendly mobile phases and charge-surface-modified C18. *Anal Chem*. 2013 Jul 16;85(14):6936-44. doi: 10.1021/ac401481z.
[2]. Mouchahoir T, Schiel JE. Development of an LC-MS/MS peptide mapping protocol for the NISTmAb. *Anal Bioanal Chem*. 2018 Mar;410(8):2111-2126. doi: 10.1007/s00216-018-0848-6.
[3]. Schachner L, Han G, Dillon M, Zhou J, McCarty L, Ellerman D, Yin Y, Spiess C, Lill JR, Carter PJ, Sandoval W. Characterization of Chain Pairing Variants of Bispecific IgG Expressed in a Single Host Cell by High-Resolution Native and Denaturing Mass Spectrometry. *Anal Chem*. 2016 Dec 20;88(24):12122-12127. doi:10.1021/acs.analchem.6b02866.
[4]. Bailey AO, Han G, Phung W, Gazis P, Sutton J, Josephs JL, Sandoval W. Charge variant native mass spectrometry benefits mass precision and dynamic range of monoclonal antibody intact mass analysis. *MAbs*. 2018 Nov-Dec;10(8):1214-1225. doi: 10.1080/19420862.2018.152113
[5]. Ren C, Bailey AO, VanderPorten E, Oh A, Phung W, Mulvihill MM, Harris SF, Liu Y, Han G, Sandoval W. Quantitative Determination of Protein-Ligand Affinity by Size Exclusion Chromatography Directly Coupled to High-Resolution Native Mass Spectrometry. *Anal Chem*. 2019 Jan 2;91(1):903-911. doi:10.1021/acs.analchem.8b03829.

Metabolomics Services

Metabolomics is an emerging and rapidly evolving “omic-level” research technology tool, which involves quantitative and qualitative metabolite assessments in biological systems. Combining high-throughput analytical chemistry and multivariate data analysis, metabolomics offers a brand-new perspective for studies on metabolic mechanisms. It offers tremendous promise for diverse applications in various fields such as medical, nutrition, environmental, and agricultural sciences. BGI provides scalable, customizable metabolomics services across the research continuum from discovery to targeted analysis.

Untargeted Metabolomics



Lipidomics Profiling

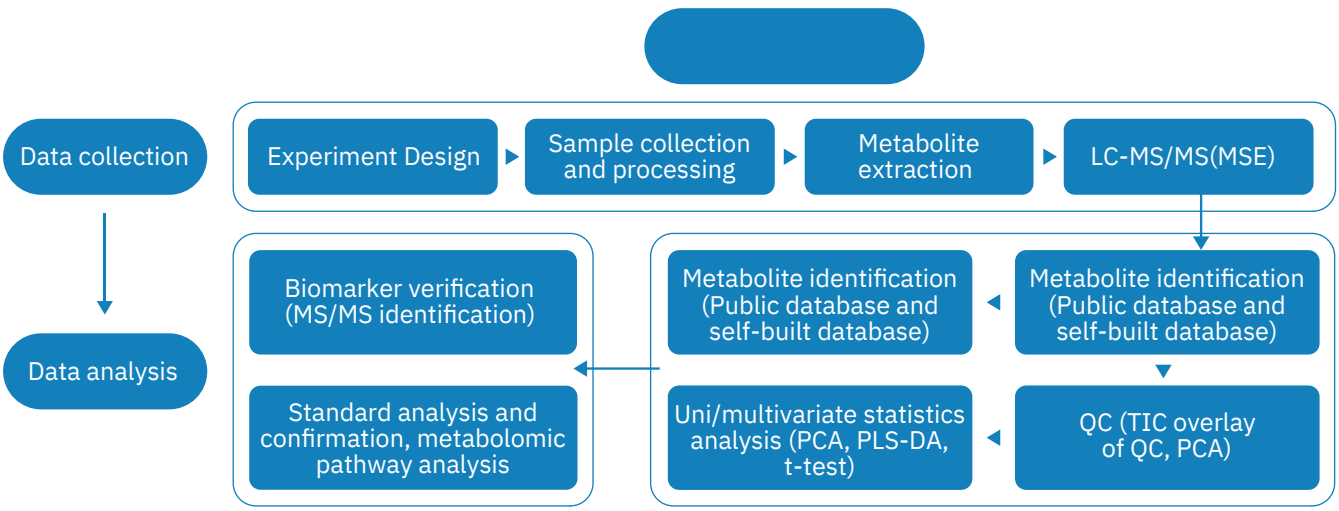
Targeted Metabolomics

Service Description

BGI has extensive experience in the field of metabolomics and offers a wide range of services including overall metabolite profiles as well as targeted small molecule analysis services [1]. By applying state-of-the-art LC-MS/MS systems and techniques, BGI offers validated metabolomics workflows including: Untargeted metabolomics (water/lipid-soluble), Lipidomics profiling (FA, GL, GP, SP, ST, PR, SL, and PK) and Targeted metabolomics (vitamins, amino acids, hormones and 700+ kinds of small molecules).

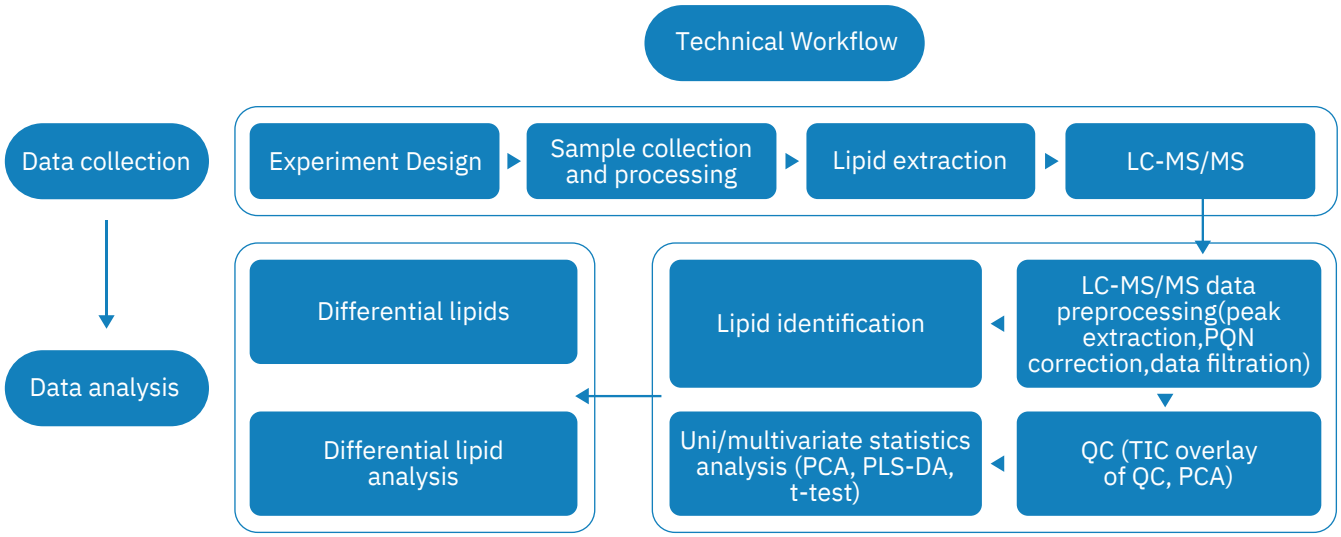
Untargeted Metabolomics

BGI provides LC-MS/MS untargeted metabolomics services to support biomedical and biotechnology applications. Our Untargeted Metabolomics services are designed to obtain a metabolite profile and screen for differentially-expressed molecules in the sample.
Metabolites are identified using self-built metabolite standard database (3000+) and Thermo mzCloud standard database. MzCloud is an online standard database of over 18,000 compounds containing both MS1 and MS2 spectra information.



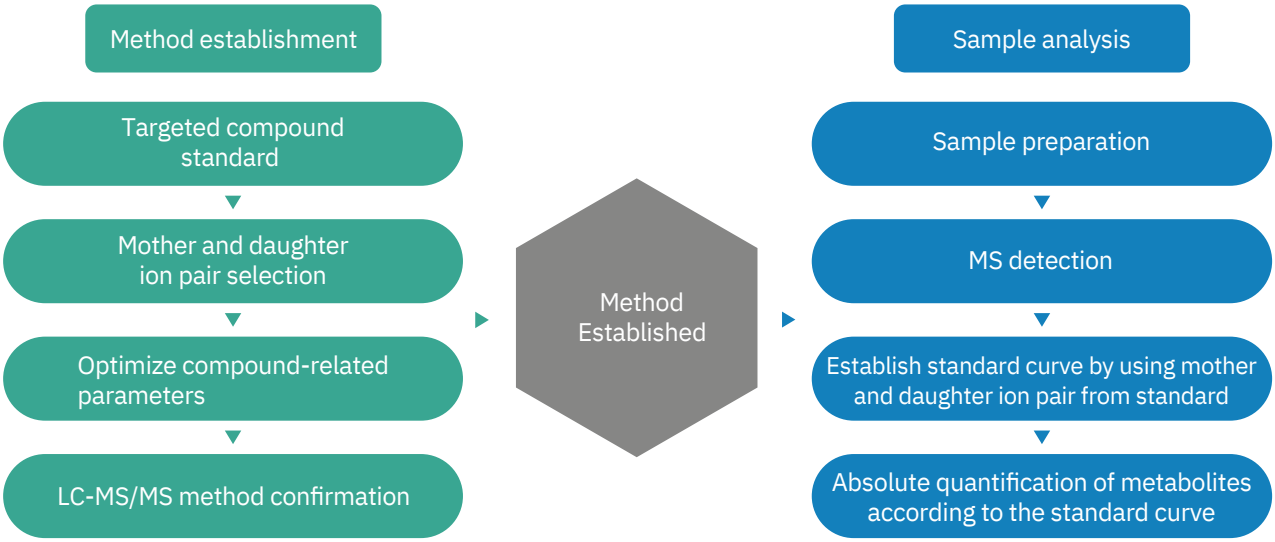
Lipidomics Profiling

BGI lipidomics services are designed to obtain a lipid profile and screen for deferentially-expressed lipids in the sample [2]. BGI's lipidomics service enables confirmation of lipid structure characteristics such as lipid molecule head structure, fatty acid branch chain length and unsaturated double bond number by comparing the characteristic signals of precursor ions, MS2 fragments and neutral loss.
LipidSearch is equipped with the largest lipids database, covering 8 categories of lipids, 300 lipid subclasses, and containing more than 1.7 million lipid ions and their predictive fragment ions information.



Targeted Metabolomics

BGI's targeted metabolomics service involves the detection and analysis for specific metabolites. By quantifying the small molecules with reference to established standards, BGI' s targeted metabolomics service can accurately monitor the dynamic metabolic process, uncover relevant metabolic mechanisms and verify potential metabolic biomarkers.
By applying state-of-the-art LC-MS/MS systems and techniques, BGI offers validated workflows of targeted metabolomics (vitamins, amino acids, hormones and 700+ kinds of small molecules).



Technical Specifications

BGI has extensive experience in the field of metabolomics and offers a wide range of services including overall metabolite profiles as well as targeted small molecule analysis services [1]. By applying state-of-the-art LC-MS/MS systems and techniques, BGI offers validated metabolomics workflows including: Untargeted metabolomics (water/lipid-soluble), Lipidomics profiling (FA, GL, GP, SP, ST, PR, SL, and PK) and Targeted metabolomics (vitamins, amino acids, hormones and 700+ kinds of small molecules).

Untargeted Metabolomics Sample Preparation Services

- Increase Metabolome depth of coverage using both ACQUITY UPLC BEH C18 and Amide column
- (+) and (-) ESI detection by using Waters ACQUITY UPLC and cutting-edge mass spectrometer (Q Exactive HF/HF-X and Q-TOF 5600)

Lipidomics Profiling Sample Preparation Services

- Increase metabolome depth of coverage using ACQUITY UPLC CSH C18 column
- (+) and (-) ESI detection by a Q Exactive HF/HF-X mass spectrometer

Targeted Metabolomics Sample Preparation Services

- Custom-developed UPLC-MRM/MS methods combine Isotope internal standard technology
- Using LC-MS (SCIEX QTRAP 4500/5500/6500 and Waters Xevo TQ-S), GC-MS (Thermo TSQ 9000) and ICP-MS (Agilent 7700) to increase the scope of detection

Data Analysis

Untargeted Metabolomics and Lipidomics Profiling:

- Data analysis performed with metaX3 and BGI library
- T-test, PCA, PLS-DA for differential metabolite identification
- Cluster analysis
- Metabolite pathway annotation
- ROC analysis
- Metabolite correlation analysis

Targeted Metabolomics:

- Data analysis performed with metaX3 and BGI library
- Data quality control
- Standard curves
- Quantification results

Quality Standard

- Summary includes all methods and data analysis
- Reports provided in Excel or PDF format; RAW files available

Turnaround Time

Typical 3-5 weeks after sample receipt for raw data delivery

Sample Requirements

We accept metabolite samples in a variety of formats. Contact your BGI account representative for details

Untargeted Metabolomics and Lipidomics Profiling:

Sample Type	Recommended Sample Amount	Minimum Sample Amount
Serum, plasma	≥ 250 μL	≥ 50 μL
Urine	≥ 500 μL	≥ 50 μL
Animal and clinical tissues	≥ 200 mg	≥ 25 mg
Feces and intestinal contents	≥ 200 mg	≥ 50 mg
Cell	≥ 1×107	≥ 5×106
Microorganism	≥ 1×107 or ≥ 100 mg	≥ 5×106 or ≥ 25 mg
Culture medium, fermentation medium	≥ 1 mL	≥ 100 μL
Plant tissue	≥ 1 g	≥ 100 mg
Milk	≥ 1 mL	≥ 1 mL
Other body fluids (amniotic fluid, saliva, hemolymph, cerebrospinal fluid, etc.)	≥ 250 μL	≥ 50 μL

References

[1] Liu, R., et al., Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nat Med*, 2017. 23(7): p. 859-868.
[2] Zhong, H., et al., Lipidomic profiling reveals distinct differences in plasma lipid composition in healthy, prediabetic, and type 2 diabetic individuals. *Gigascience*, 2017. 6(7): p. 1-12.
[3] Wen, B., et al., metaX: a flexible and comprehensive software for processing metabolomics data. *BMC Bioinformatics*, 2017. 18(1): p. 183.

Multi-Omics

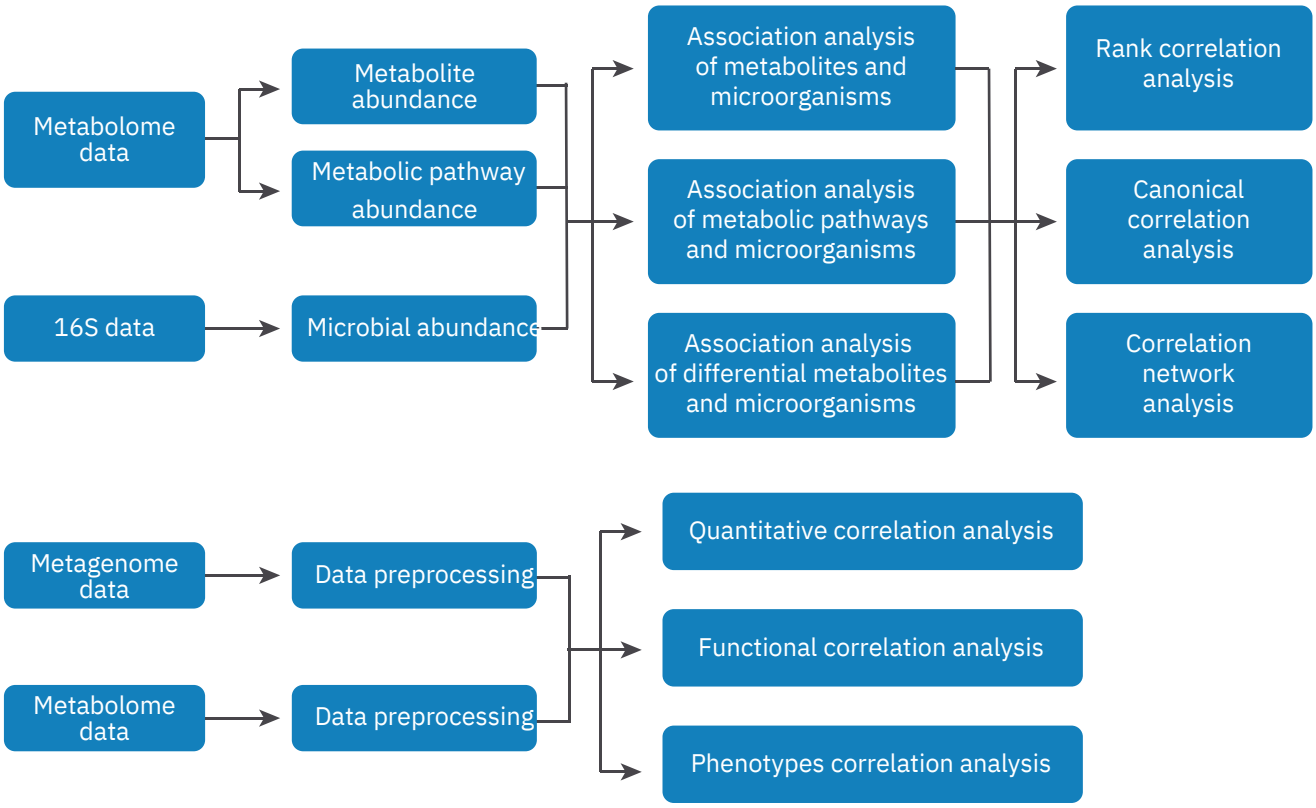
Multi-omics is the integrative biological analysis of different data sets from single omics areas for new insight. An integrated multi-omics approach to research enables a more comprehensive understanding of genotypic, phenotypic and environmental relationships and their association to disease and health of an organism.

BGI offers multi-omics services to look across genomics, transcriptomics, epigenomics, proteomics and metabolomics, with the flexibility to customize solutions that meet your specific needs [1,2,3]. All projects are supported by a bioinformatics infrastructure that is second to none.

Service Description

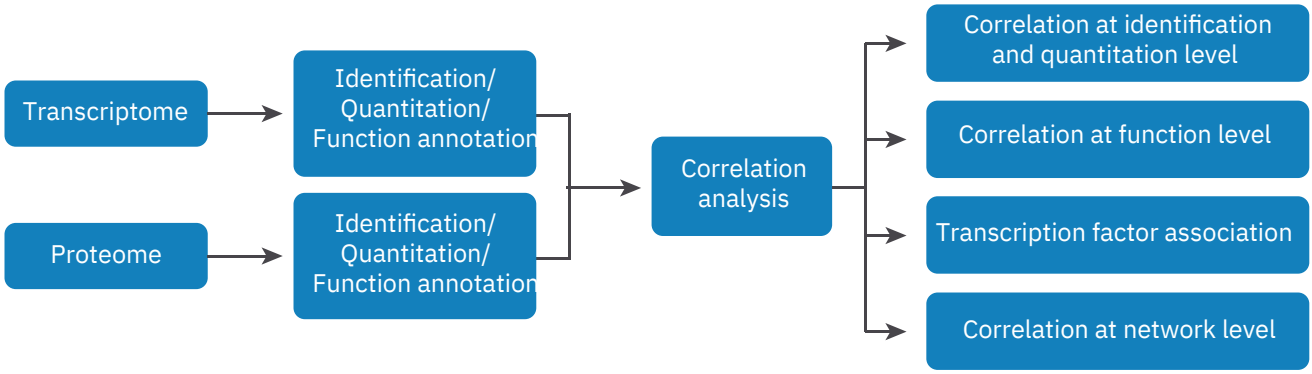
16S/Metagenome + Metabolome Correlation Analysis

In recent years, research on gut related diseases has developed rapidly and it is now thought that nearly 90% of diseases may be related to gut microbiota health. 16S/Metagenome + metabolome correlation analysis enables researchers to establish an association model between host metabolism and gut microbiota and explore the causal relationships between microbes and disease [1,4].



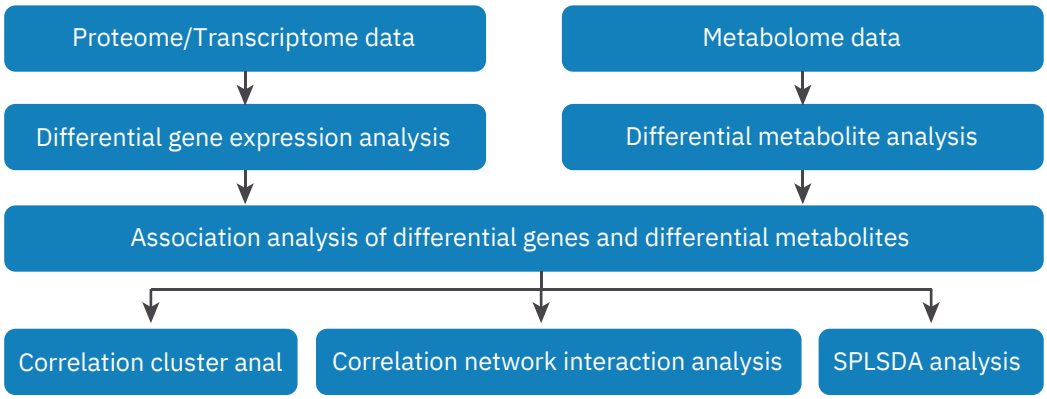
Proteome + Transcriptome Correlation Analysis

Using a multi-omics approach to correlate transcriptomics with proteomics data provides a more comprehensive overview of expression patterns and enables researchers to interpret deeper biological implications [2,5].



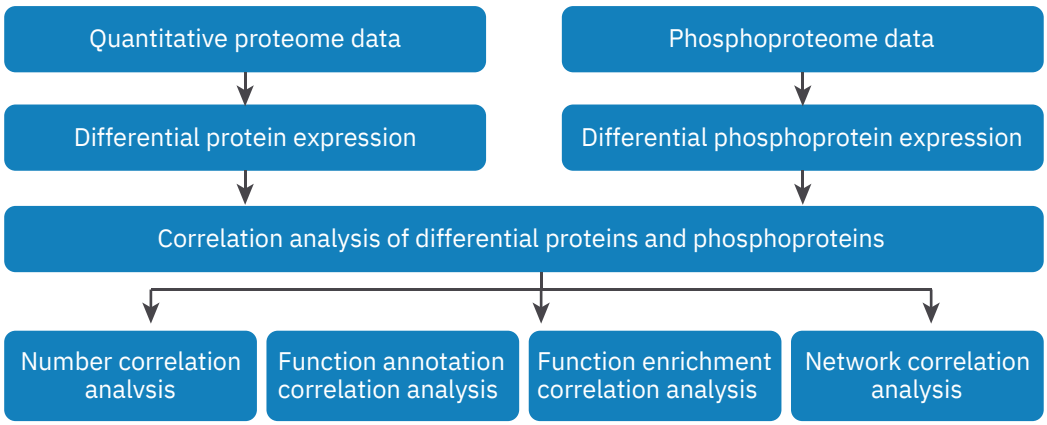
Transcriptome/Proteome + Metabolome Correlation Analysis

Transcriptome/Proteome + metabolome correlation analysis enables researchers to explore the biological problems from the two directions of "cause" and "result", This lets researchers better understand the regulation mechanism of biological systems and helps explain the influence of gene function and environmental factors on biological phenotype [3,6].



Quantitative Proteome + Phosphoproteome Correlation Analysis

Protein phosphorylation is one of the most basic modification types. A protein function panorama and precise positioning of the leading regulatory role for interpretation of molecular mechanisms can be obtained with BGI's expert quantitative proteomics and phosphoproteomics correlation analysis services.
BGI provides quantitative proteomics and phosphoproteomics correlation analysis spanning across diseases biomarkers research, growth and development research, regulation mechanism research, and target drug research.



Technical Specification

16S/Metagenome + Metabolome Correlation Analysis

- 16S/Metagenome sequencing platform: 16S/Metagenome services
- Metabolome platform: Untargeted/Targeted Metabolomics and Lipidomics services

Proteome + Transcriptome Correlation Analysis

- Transcriptome sequencing platform: RNA-Seq services
- Proteome platform: Untargeted and Targeted Quantitative Proteomics services

Transcriptome/Proteome + Metabolome Correlation Analysis

- Transcriptome sequencing platform: 16S/Metagenome services
- Proteome platform: Untargeted and Targeted Quantitative Proteomics services
- Metabolome platform: Untargeted/Targeted Metabolomics and Lipidomics services

Quantitative Proteome + Phosphoproteome Correlation Analysis

- Quantitative Proteome platform: Untargeted and Targeted Quantitative Proteomics services
- Phosphoproteome platform: Phosphoproteomics services

Quality Standard

- Summary includes all methods and data analysis

Reports provided in Excel or PDF format; RAW files available

Turnaround Time

Typical 2 weeks from LC-MS/MS raw data acceptance to data report delivery

Sample Requirements

Sample selection for 16S/Metagenome + Metabolome Correlation Analysis should be as consistent as possible.

Integrative Analysis	Biological duplicates when 2 groups	Biological duplicates when ≥ 3 groups
Metagenome + metabolome	≥ 8	≥ 6
16S+ metabolome	≥ 8	≥ 4
Clinical samples	≥ 30	

Sample selection for (1) Proteome + Transcriptome Correlation Analysis, (2) Transcriptome/Proteome + Metabolome Correlation Analysis and (3) Quantitative Proteome + Phosphoproteome Correlation Analysis should be as consistent as possible. At least 2 groups. At least 3 biological replicates in each group.

References

[1] Liu R., et al., Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nat Med*, 2017. 23(7): p. 859-868. doi: 10.1038/nm.4358.
[2] Dai F, Wang Z. et al. Transcriptomic and proteomic analyses of mulberry (Motus atropurpurea) fruit response to Ciboria carunculoides. *J Proteomics*. 2019 Feb 20; 193: 142-153. doi: 10.1016/j.jprot.2018.10.004
[3] Gao H Y. et al. Transcriptomics and metabolomics analyses reveal the differential accumulation of phenylpropanoids between Cinnamomum cassia Presl and Cinnamomum cassia Presl var. macrophyllum Chu. *Industrial Crops and Products,* Volume 148, 2020, 112282, ISSN 0926-6690, doi: 10.1016/j.indcrop.2020.112282.
[4] Liu H, Chen X. et al. Alterations in the gut microbiome and metabolism with coronary artery disease severity. *Microbiome*. 2019 Apr 26;7(1):68.
[5] Becker K, Bluhm A, Casas-Vila N. et al. Quantifying post-transcriptional regulation in the development of Drosophila melanogaster. *Nat Commun*. 2018 Nov 26;9(1):4970.
[6] Li Y N, Chen X. et al. Hepatotoxicity study of combined exposure of DEHP and ethanol: A comprehensive analysis of transcriptomics and metabolomics. *Food Chem Toxicol.* 2020 Jul; 141: 111370. doi: 10.1016/j.fct.2020.111370.

Customized Solutions

BGI 10X Genomics High Throughput Single Cell RNA-Seq Service Description

BGI 10X Genomics High Throughput Single Cell RNA-Seq Service Description

Introduction

Service Description

Thisservice is based on 10X Genomics high throughput single cell processing platform, providing an integrated service for sample preparation, cell sorting, droplet formation, library construction, sequencing and bioinformatics. By analyzing the difference between gene expression levels among single cells, 10X Genomics high throughput single cell RNA-seq describes the heterogeneity between cells and facilitates the study of biological interactions.

Explanation of Key Terms

GEM (10X Gel bead in Emulsion), a droplet reaction system composed of mRNA, bead and emulsion where the subsequent cell sorting, amplification, library construction reaction take place. More than 90% gel beads can be encapsulated in GEM. Literally 100% reaction product can be marked with the unique barcode.

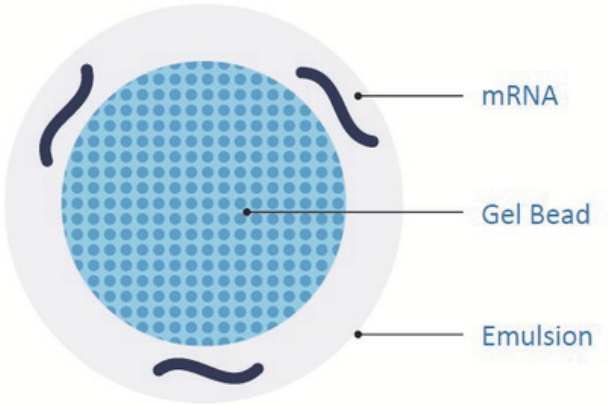


Fig 1. GEM (Gel in Emulsion) structure

The Advantages of the BGI 10X Single Cell Platform

• High throughput: 8 channels on one chip. Each time up to 80,000 cells are sorted and amplified;
• Integrated automatic processing platform: highly automatic reaction, able to completes thousands of cell sorting within 10 minutes. Amplification and library construction are done in the reaction system. Cell Ranger pipelines provide direct analysis results;
• Lower standard cost for each individual cell;
• Dr.Tom Single Cell RNA-Seq Data Visualisation System for Effcient Analysis and Discovery;

Service Workflow

Service Flow Chart

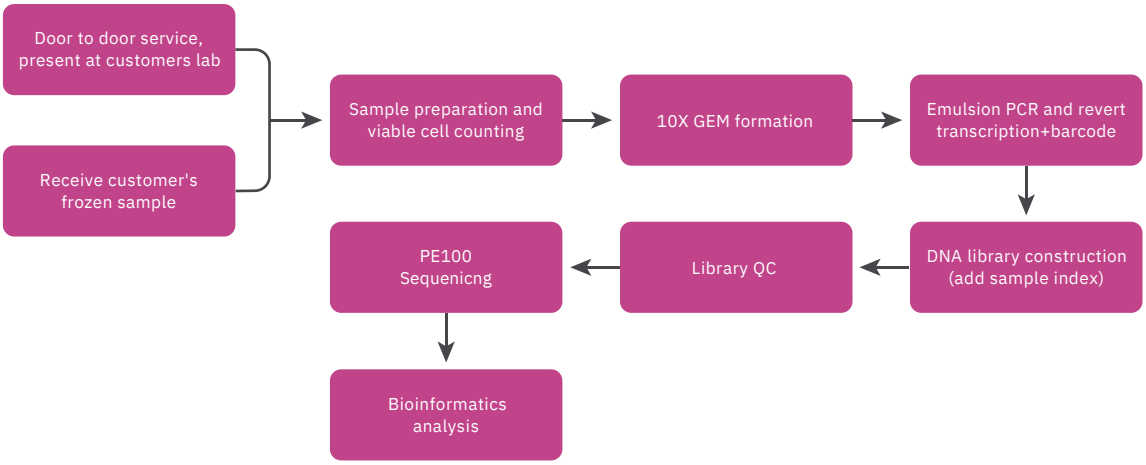


Fig 2. Service Flow Chart

10X Genomics Library Construction Mechanism

By using microfluidic chip, beads with unique barcode and cell are encapsulated in droplet and collected. In the droplet, the single cell is lysed to release mRNA, which is then ligated with barcode on the beads to form Single Cell GEMs. After that, the revert transcription is taken place in droplet to generate cDNA. The emulsion is subsequently broken down to release cDNA for library construction. By identifying the cell barcode on each library, it can be recognized which cell the target region from. Each sequence has its own index to identify which sample it comes from.

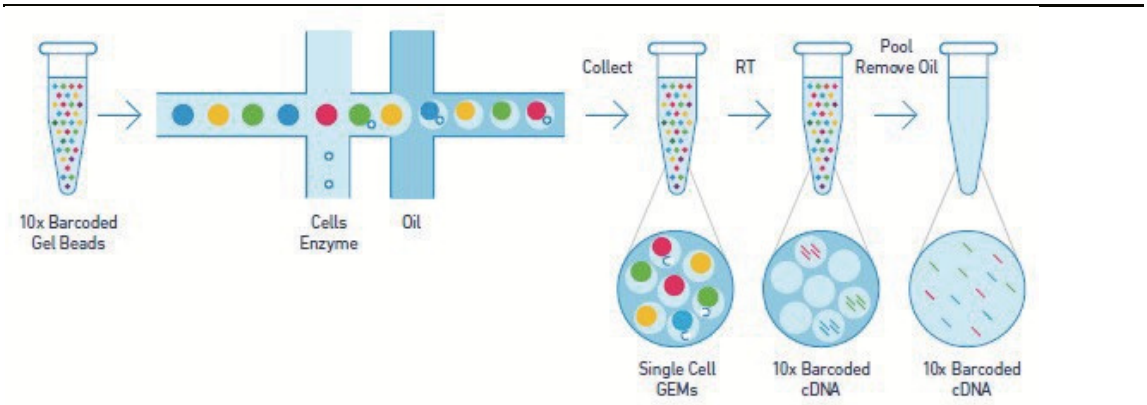


Fig 3. 10X Genomics Library Construction Mechanism

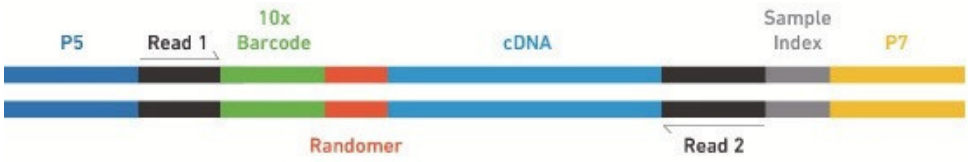


Fig 4. 10X Genomics Library Sequence Paradigm

Sequencing Strategy

Sequencing Platform

V3 kit of 10X Genomics single cell RNA-Seq applies to MGISEQ-2000 or HiSeq X ten/NovaSeq platform.

Sequencing Strategy

PE100.

Recommended Data Volume

PBMC sample: 50,000 raw reads for each single cell are recommended.
Tumor tissue sample: 100-200k raw reads for each single cell.

Bioinformatics Workflow

Customers can choose two kinds of bioinformatics analysis service: standardized Cell Ranger analysis or customized bioinformatics analysis performed according to customers' needs. (See Fig 5) Customers can also use the web-based Dr. Tom data visualization system to further interpret and analyze the data.

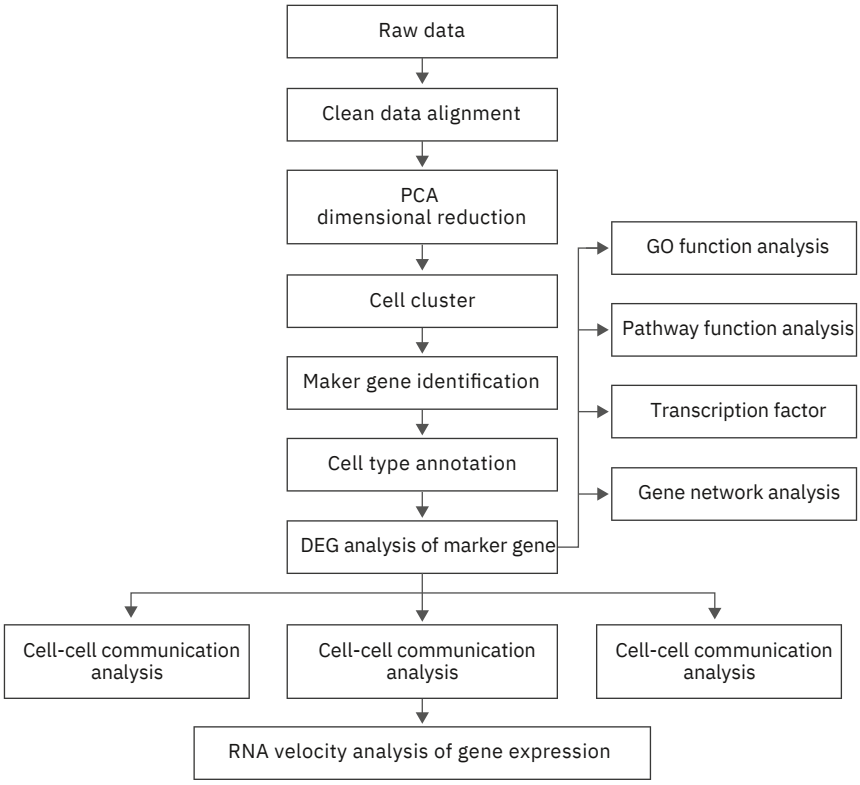


Fig 5. Bioinformatics Workflow

Bioinformatics Analysis	Bioinformatics Analysis Contents
Standard Bioinformatics Analysis	1) Sequencing Result; 2) Comparison Result; 3) Quantitative Analysis; 4) Data QC; 5) Cell Cluster; 6) Differential Gene Identification Between Samples; 7) Cell Cluster Annotation; 8) Marker Gene/Differentially Expressed Genes GO Pathway; 9) Marker Gene/Differentially Expressed Genes KEGG Pathway Enrichment; 10) RNA Velocity; 11) CellPhone DB; 12) Pseudotemporal Ordering;
Dr. Tom Bioinformatics Analysis (ref)	Dr. Tom Single Cell Data Cross Mining System (12 months free access) 1, Customized Cell Cluster Analysis 2, Customized Cell Cluster Annotation By Database of your choice 3, Gene-Protein Interaction Analysis 4, Gene Function Enrichment Analysis 5, Dynamic Interactive Analysis of Cell Type And Cell Trajectory Analysis; 6, Different Dimension Analysis of Differential Genes Between Cell Types/Samples
Customized Bioinformatics Analysis	Customized according to client's needs

The Demo Report Of 10X Genomics Single Cell 3'RNA-Seq:

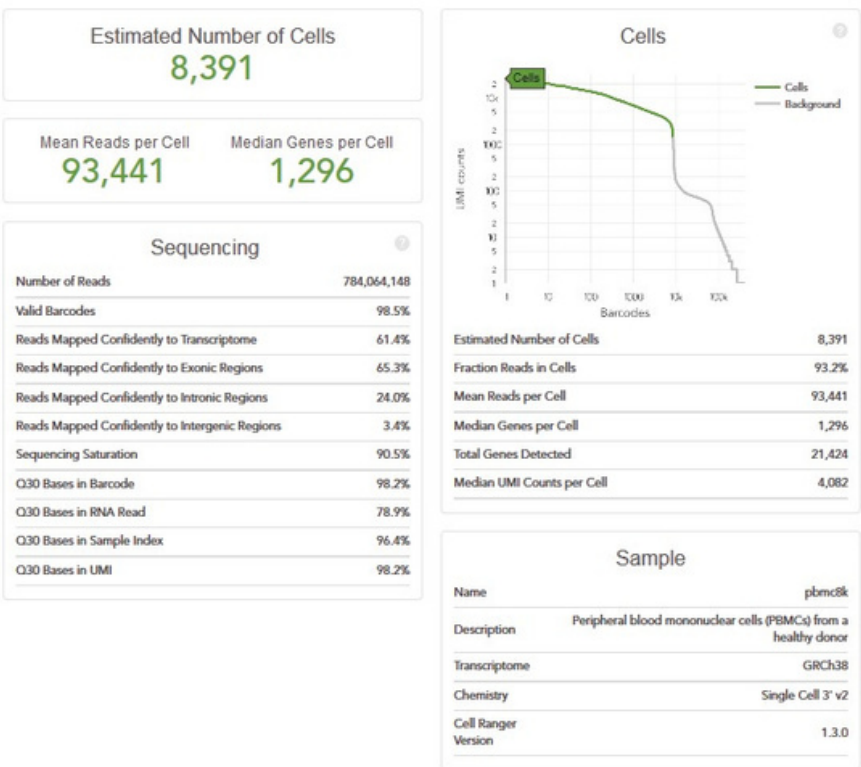


Fig 6A. Cell Split, Counting And Quality Control

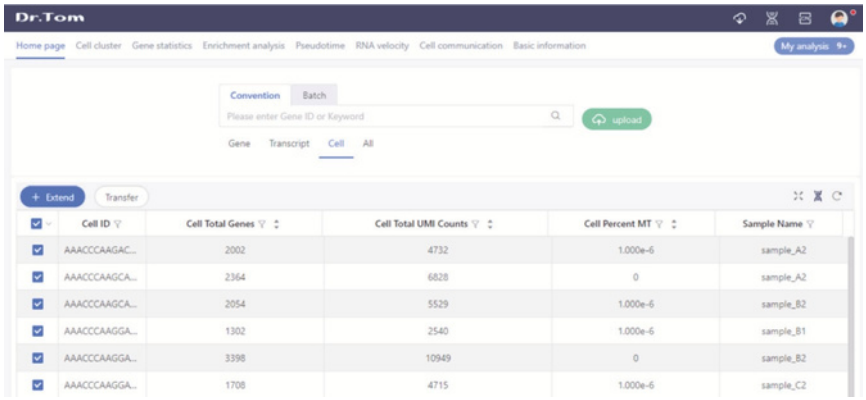


Fig 6B. Sequencing Data Bioinformatics Analysis Result in Dr.Tom

Fig 6C. Cluster Specific Marker Gene Expression Heatmap in Dr.Tom.

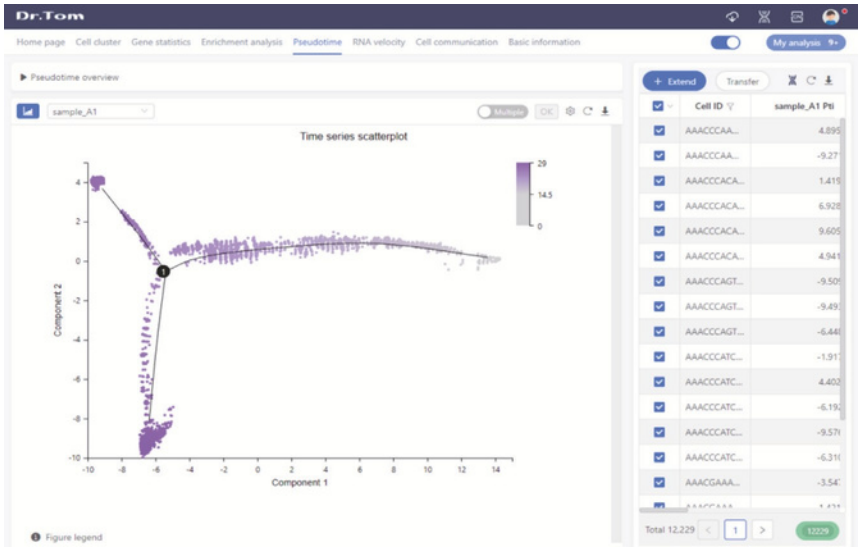


Fig 6D. Pseudotime overview in Dr.Tom

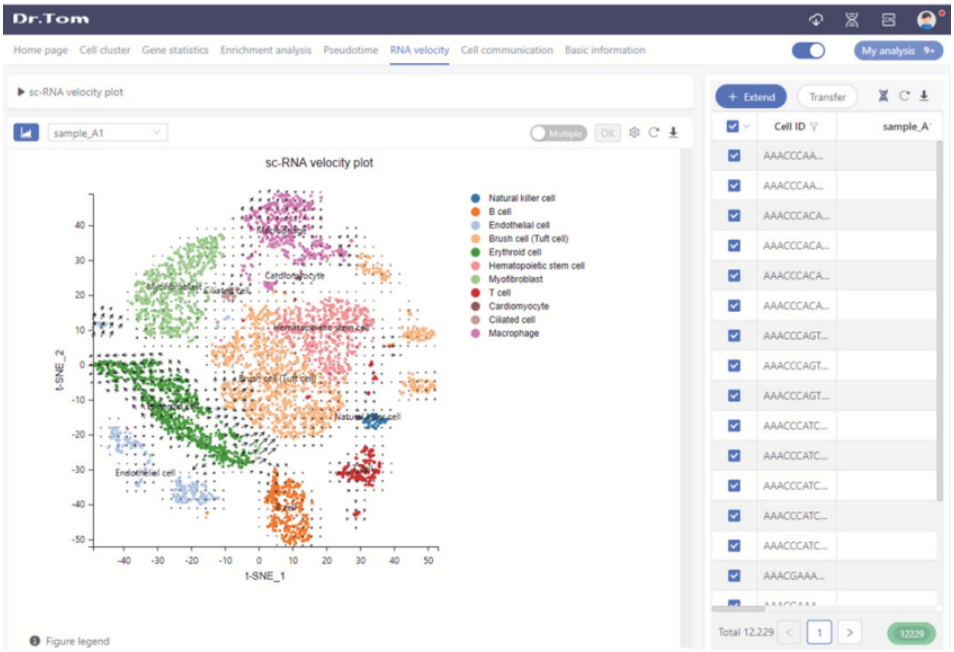


Fig 6E. SC-RNA velocity plot in Dr.Tom

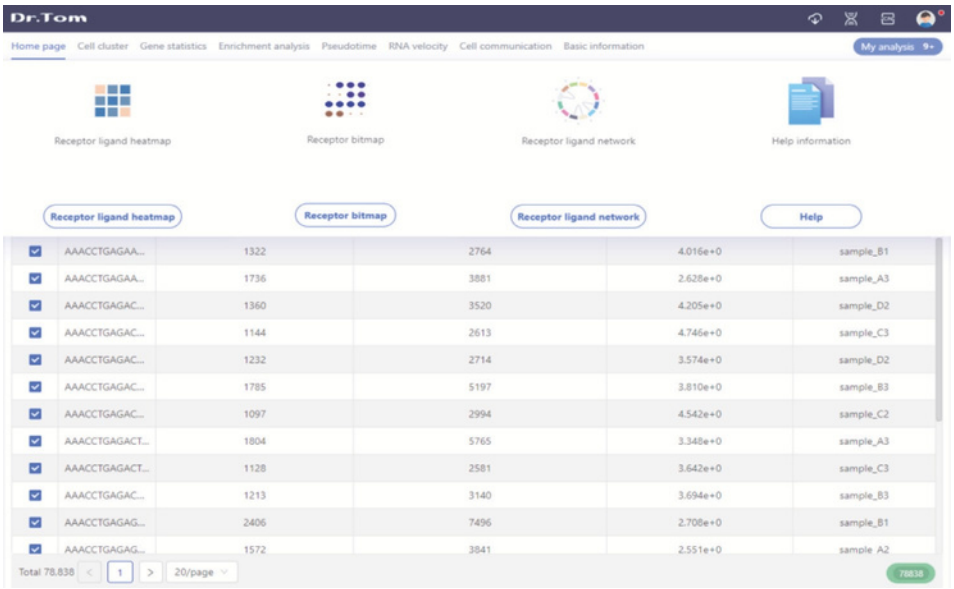


Fig 6F. Cell communication in Dr.Tom

Turnaround Time

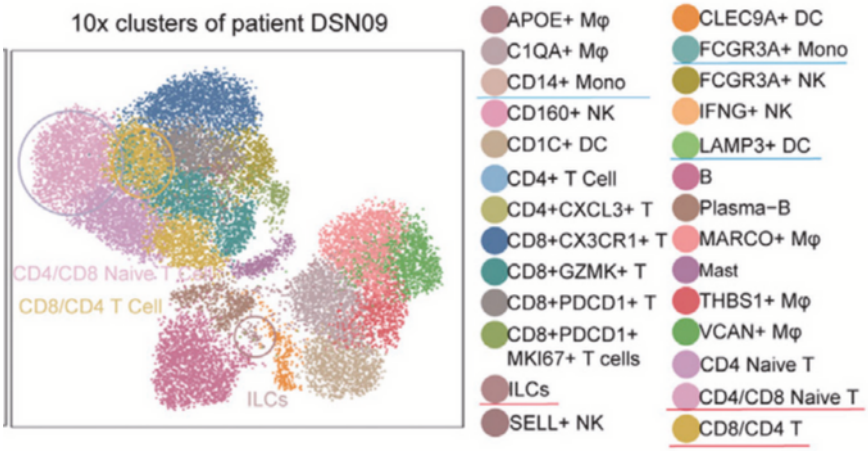
The project is launched after cell recovery, microscopic examination and advance payment be made, not including the delay caused by low quality sample.

Platform	Without Bioinformatics Analysis (Working Days)	With bioinformatics analysis (working days)
DNBSEQ™	25	35

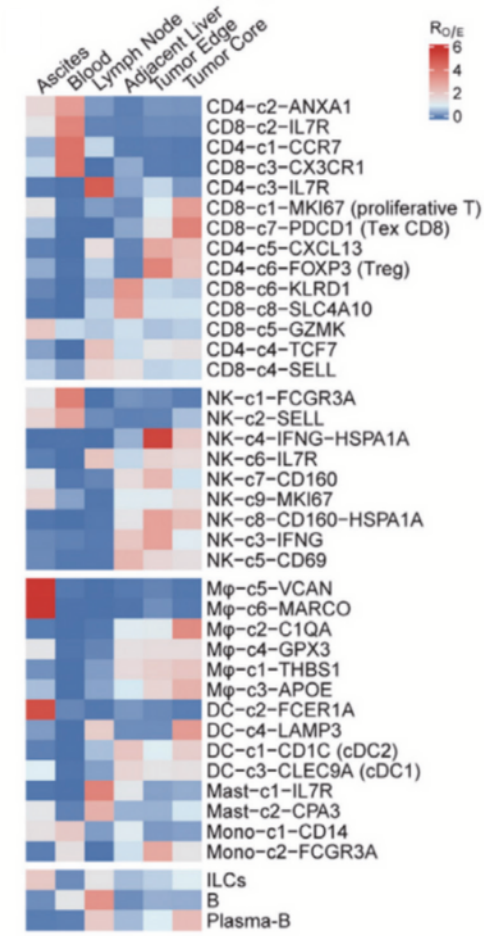
Case Study

Case Study 1

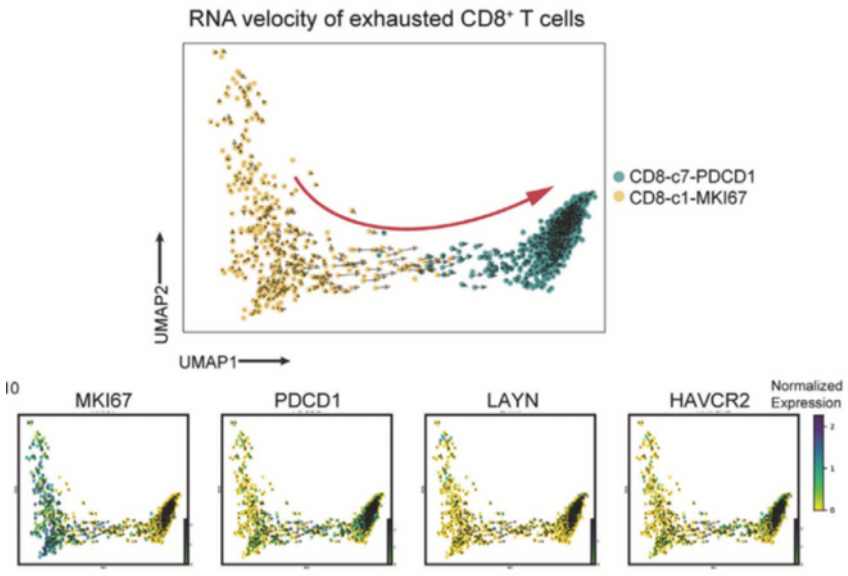
Landscape and Dynamics of Single Immune Cells in Hepatocellular Carcinoma
Qiming Zhang et al. (2019) *Cell* 179 829-845
To study the immune microenvironment of hepatocellular carcinoma(HCC) the single cell transcriptomes of CD45+ immune cells from five immune-relevant sites: tumor, adjacent liver, hepatic lymph node(LN), blood, and ascites of 16 treatment-naïve liver cancer patients are produced.20 clusters of cells are identified using 10X data.



Using Uniform Manifold Approximation and Projection (UMAP) to identify the tissue preference of each cluster estimated by Ro/e based on 10X data. The tissue enrichment is quantified based on the ratio of observed to expected cell numbers in each cluster(Ro/e) using the 10X data. In T cells, CD4-c6-FOXP3 and CD8-c7-PDCD1, corresponding to regulatory T(Treg) cells and exhausted CD8+ T(Tex) cells, respectively, were enriched in tumors. In NK cell clusters, NK-c3-IFNG, NK-c4-IFNG-HSPA1A, NK-c7-CD160, and NK-c8-CD160-HSPA1A were enriched in tumor compared with other tissues. M4-c5-VCAN, M4-c6-MARCO, and DCc2-FCER1A were dominantly enriched in ascites.

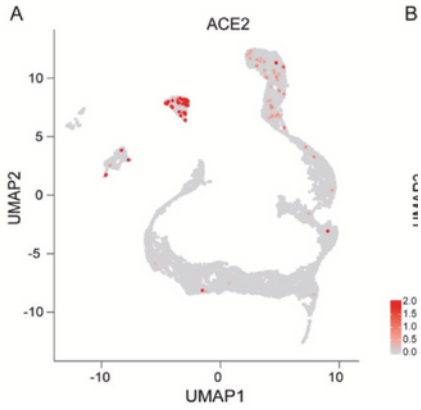


Using RNA velocity, a method inferring precursor-progeny cell dynamics, a clear directional flow from proliferative T cells to Tex cells is observed.

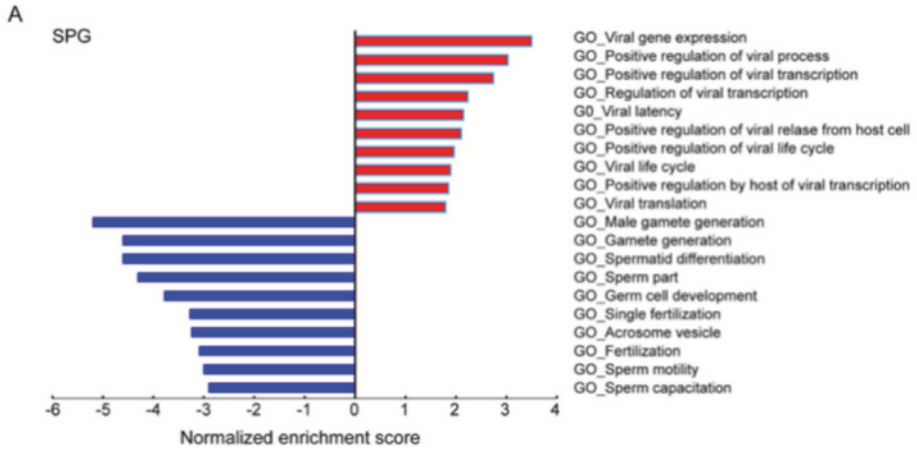


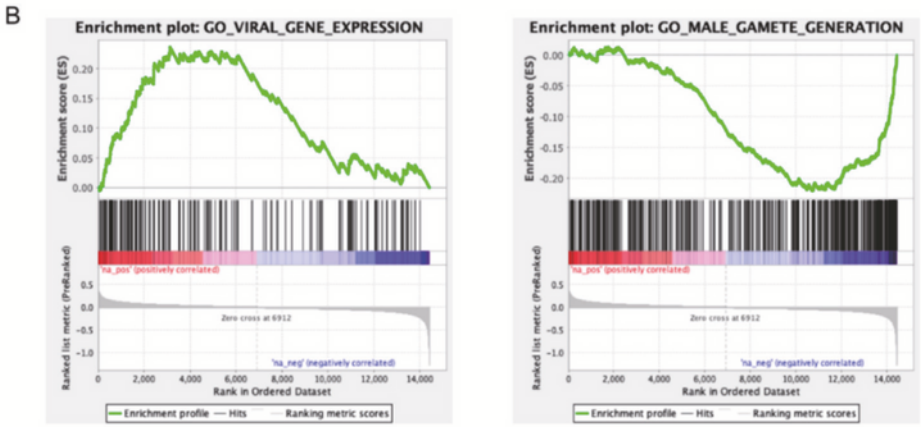
Case Study 2

cRNA-seq Profiling of HumanTestes Reveals the Presence of the ACE2 Receptor, A Target for SARS-CoV-2 Infection in Spermatogonia, Leydig and Sertoli Cells
Zhenpin Wang, Xiaojiang Xu, (2020), Cell, 9, 920
ACE2 is shown to be the cell entry receptor of SARS-Cov-2.
To investigate the expression pattern of ACE2 in adult human testes at the single cell level, UMAP is used to reveal that ACE2 was primarily enriched in spermatogonia and Leydig and Sertoli cells.



Gene ontology(GO) enrichment analysis(7C) and GSEA(7D) was performed by comparing ACE2-positive cells with ACE2-negative cells. 24 GO terms associated with viral reproduction and transmission were positively enriched in ACE2-positive spermatogonia. Multiple GO terms related to male reproduction were significantly decreased in ACE2-positive spermatogonia. GSEA results suggest that male germ cell specific genes and genes that are collectively involved in spermatogenesis are compromised.





Sample Submission Guidelines

Due to sample preservation limitations and chip order turnaround time, please contact production department one week in advance before sample shipment.

Sample Type

Cell culture (cell line and primary cell culture etc): Complete and active cell no less than 10^6/ml and the survival rate of cell should be greater than 80%.

Sample Processing And Delivery Requirement

A.For cell line:

1) Make sure the cells are in good condition under microscope, and the cell viability is above 80%.
2) Resuspend cells with cell freezing medium to a final cell density of 2×106-1×107/ml and total volume is no less than 1ml/tube. (Cell freezing medium: 90% fetal bovine serum (FBS + 10% DMSO).
3) Transfer 250ml of isopropyl alcohol (it can be re-used for 3-5 times) to gradient freezing container at room temperature. Put the cryopreservation tube into the gradient freezing container, and put it at -80℃overnight.
4) Keep samples at -80℃ for short-term storage, or keep them in liquid nitrogen for long-term preservation. Transport samples with enough dry ice.

B.For PBMC sample:

1) Extract 5ml human peripheral blood into an EDTA vacuum blood collection tube. Incubate lymphocyte separation medium(LSM) at 20℃.
2) Dilute human blood with equal volume of HBSS.
3) Aseptically transfer 2ml LSM into a 15ml centrifuge tube. Then carefully pipette 4ml diluted blood into 15ml centrifuge tube along tube wall, ensuring a distinct blood-LSM interface.
Note: DO NOT DISTURB BLOOD-LSM INTERFACE or MIX TOGETHER.
4) Centrifuge the tube at 500g at 20℃for 20min with no brake or lowest brake. There would be three layers in the centrifuge tube: the upper layer is plasma (including cell debris); the lower layer is erythrocytes; the middle area includes a white layer which is consisted of mononuclear cell sand lymphocytes.
5) Discard the upper plasma, carefully aspirate mononuclear cells and lymphocytes and transfer them to a new 15ml centrifuge tube.
6) Add 5ml HBSS to the centrifuge tube and centrifuge for 10min at 300g. Discard the supernatant.
7) Resuspend cells with 1ml HBSS. Count cells under microscope. Freeze cells (final concentration 2×10^6-1x10^7/ml/tube).
8) The frozen storage and transport methods are consistent with cultured cells.

C.For tissue:

1) Once the live tissue is obtained (from surgery for example), they should be put in a 50ml centrifuge tube, petri dish or culture dish on ice box, and covered with sterilized ice-cold PBS immediately.
2) Use a sterilized ophthalmic scissor or a clean scalpel to cut samples into small pellets about 3 to 4 mm in diameter. This step should be done in ice-cold PBS solution. The ophthalmic scissors or scalpel should be sharp, so the tissue sample is cut into pellets, not squeezed into pieces.
3) Transport tissue pellets to a 50ml centrifuge tube, then centrifuge it at 4℃, 500g for 5 minutes.
4) Check the suspension in the tube, if the suspension appears to be cloudy, then repeat the centrifuge.
5) Carefully draw and discard the supernatant. Do not disturb tissue pellets at the bottom of the tube.
6) Prepare enough cryopreservation solution, we recommend 10% to 15% DMSO in FBS. Usually, the volume ratio of tissue to cryopreservation solution is 1:9. Mix cryopreservation solution well and put on ice.
7) Slowly resuspend tissue pellets in cryopreservation solution by wide bore pipette tips.
8) Transport 1ml to 1.5ml tissue pellets suspension to a 2ml cryotube.
9) Put the cryotube in cell freezing container, and put in -80℃ fridge overnight.

06

Data Analysis Solution

Dr. Tom

Dr. Tom

An Advanced and Intuitive RNA Data Visualisation System for Efficient Analysis and Discovery

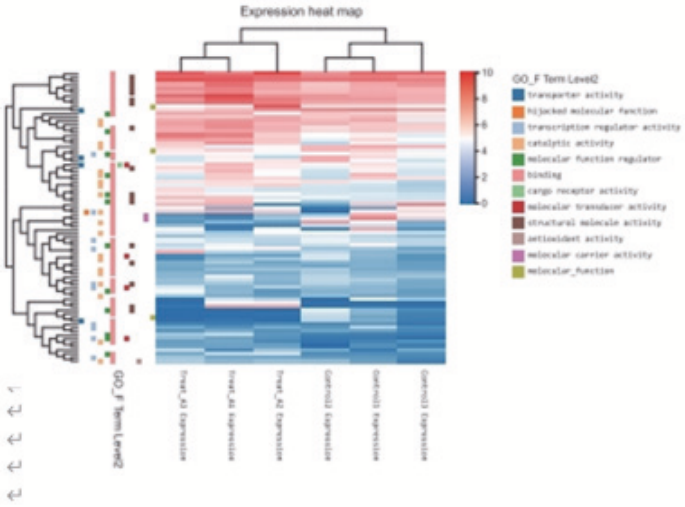
Introduction

Dr. Tom is a web-based solution for the convenient analysis, visualisation and interpretation of all types of RNA data, including small RNAseq, miRNA and lncRNA. Designed by a team of expert RNA scientists and bioinformaticians at BGI with collective experience across thousands of RNA based research projects, Dr. Tom provides a wide range of intuitive and interactive data visualisation tools specifically designed to save you time in your differential expression or pathway analysis research.
In addition, powerful analysis tools and advanced algorithms allow you to mine your data to gain new insight and more value beyond standard available RNA analysis services.
Data from many of the world's leading databases have been integrated into the Dr. Tom system allowing users to reference and cross check all results and findings.
Dr. Tom is already relied upon by hundreds of scientists and researchers, and has shown itself to be a valuable and important tool in addition to any institution's own internal data curation and analysis efforts.

Core Capabilities

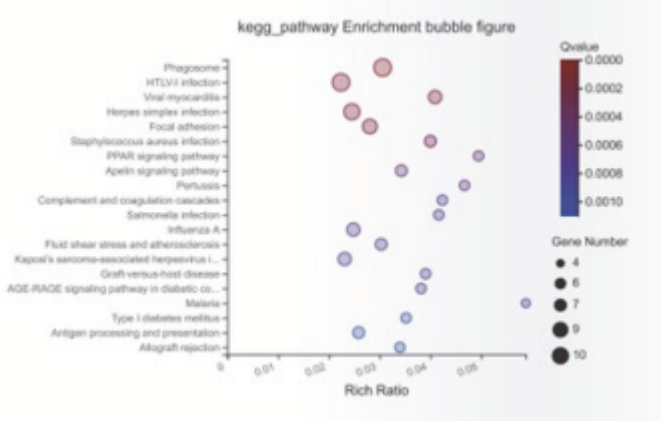
Expression Analysis

Dr-Tom's detailed, interactive heatmap functionality can be used to quickly identify genes that are commonly regulated. With simple point-and click action, data can be selected and manipulated to show clusters under different pathways.



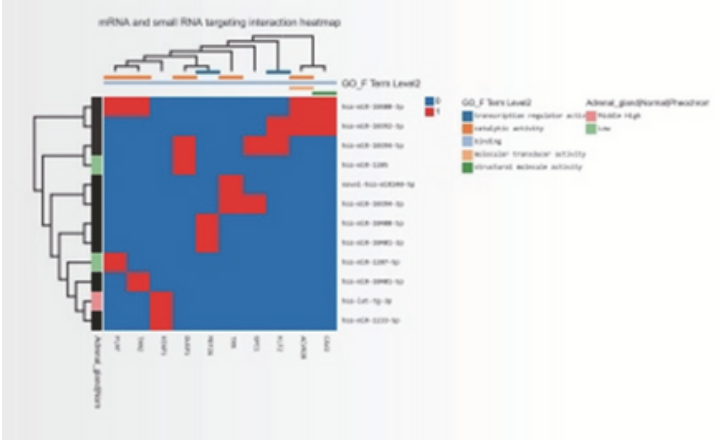
Gene Set Enrichment Analysis

Dr-Tom accesses both free and licensed KEGG databases to allow users to conveniently and quickly find statistically significant trends in the large lists of genes generated by many functional genomics techniques and bioinformatics analyses approaches.



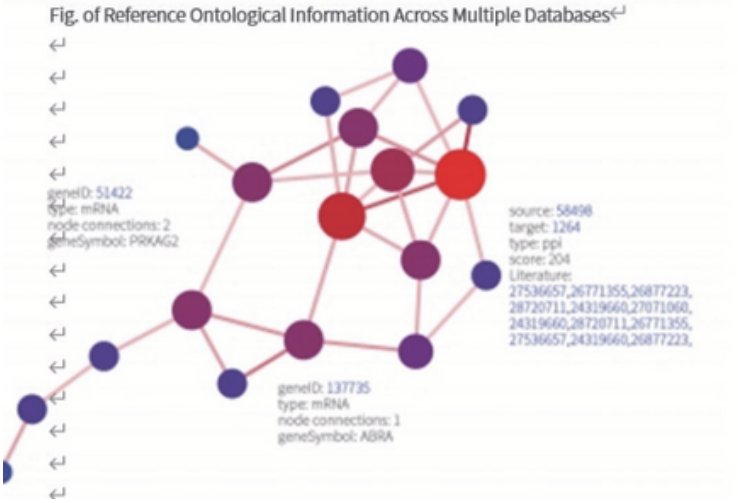
Association Analysis

With a simple click Dr. Tom lets users detect RNA association with target genes, based on their interaction relationship (such as PPI, Target, Co-expression, ceRNA, GGI and RNAplex), or based on the position relationship (such as upstream and downstream position).



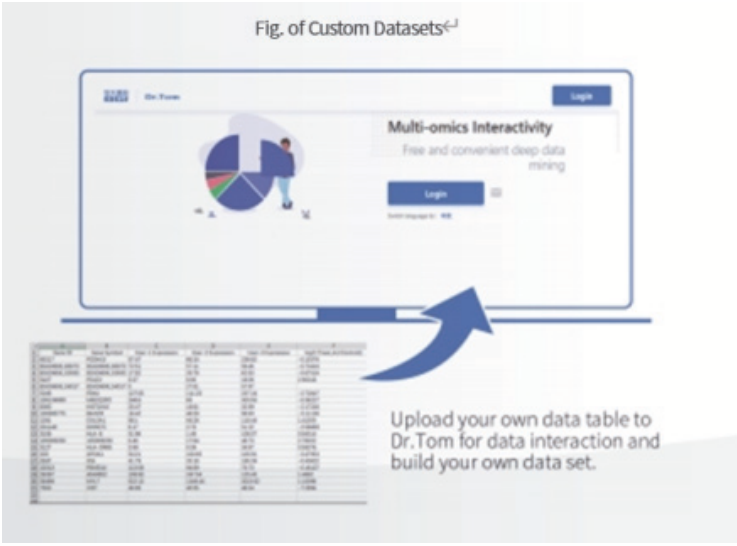
Reference Ontological Information Across Multiple Databases

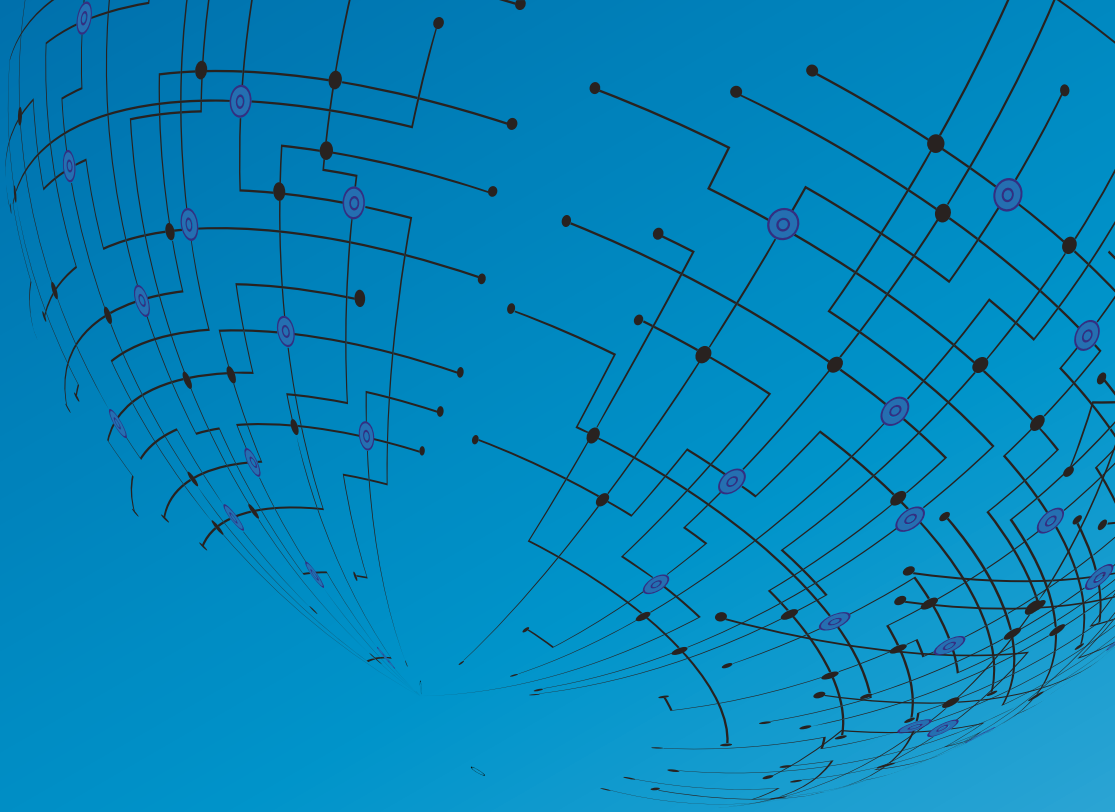
Dr. Tom is able to reference multiple-databases for association analysis，including TCGA, NCBI and many more. This allows a user to quickly and conveniently view comprehensive ontological information for any gene of interest, including annotation, sequences, expression level, and a list of relevant published papers.



Custom Datasets

Customers can upload their own gene expression data, using tool boxes for graphing and visualisation, and construct their own gene annotation database for enrichment, clustering and multi-omics association analysis.





www.bgi.com
info@bgi.com

BGI Americas

One Broadway, 3rd Floor
Cambridge, MA 02142
U.S.A.

BGI Japan

Kobe KIMEC Center BLDG. 8F 1-5-2
Minatojima-minamimachi,
Chuo-ku, Kobe 650-0047 Japan

BGI Europe

Ole Maaløes Vej 3,
DK-2200 Copenhagen N,
Denmark

BGI Australia

L6,CBCRC, 300 Heston Road,
erston, Brisbane,
Queensland 4006, Australia

BGI Asia

Building NO.7, BGI Park,
No.21 Hongan 3rd Street,
Yantian District, Shenzhen
518083,China

For Research Use Only. Not for use in diagnostic procedures (except as specifically noted).

Published: October 2021
Version: 1.0