SERVICE OVERVIEW

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

Data Analysis

- Clean data, standard and customized data analysis options available
- Data storage services and bioinformatics applications available upon request

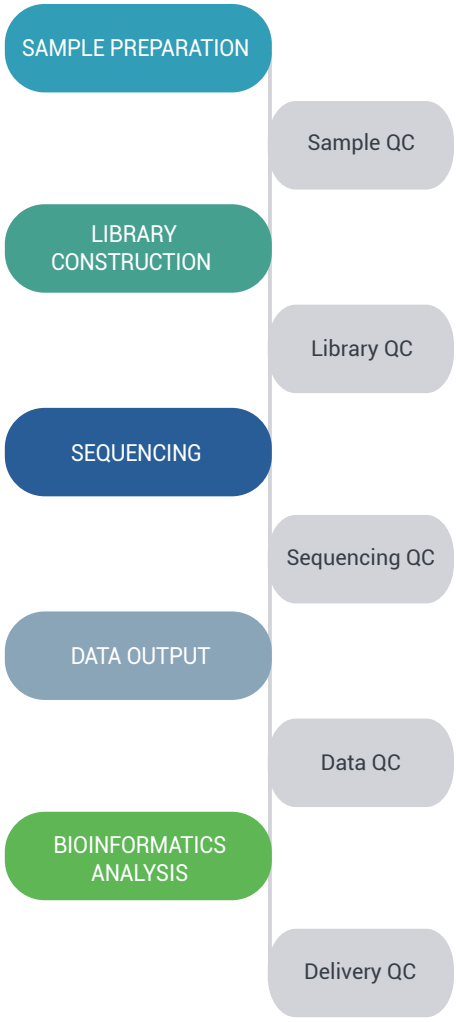
Turnaround Time

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

Quality Data | Fast TAT | Cost Effective

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 sequencing 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

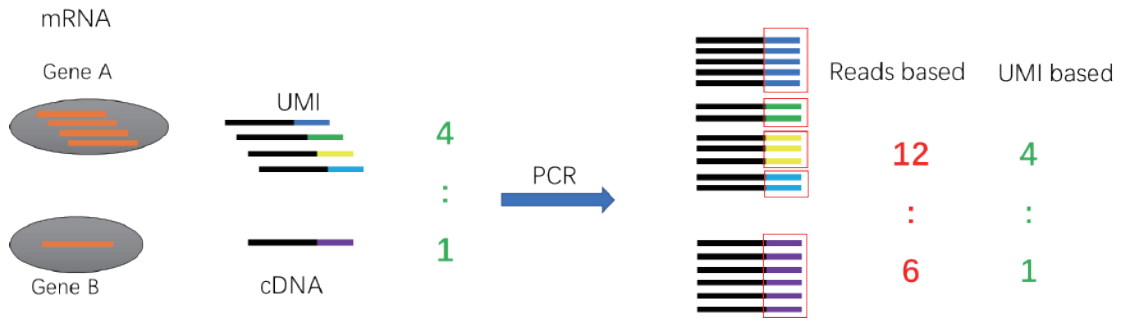
We can process your total RNA with the following general requirements:

	Total RNA Amount	Concentration, RIN, 28S/18S	Minimum Sample Volume
Recommended	m≥3 μg	c≥285ng/μL RIN≥7.5 28S/18S(23S/16S)≥1.2	15 μl
Required	1.5 μg≤m<3 μg		15 μl

Highlights of BGI Iso-Seq Service

1) Absolute Quantification with UMI Technology

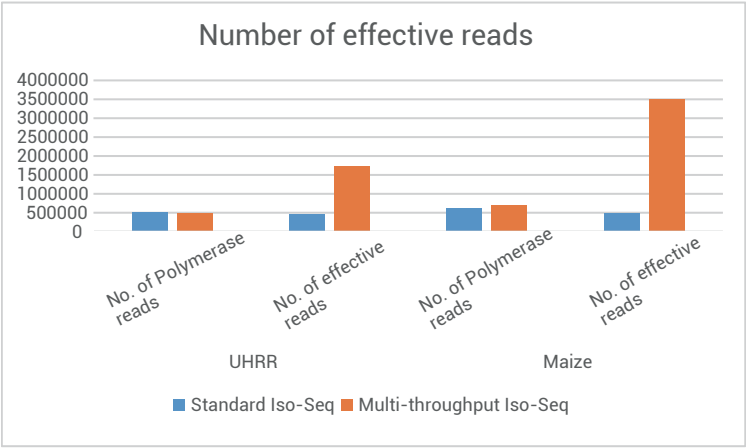
Each original transcript is marked by a unique molecular identifier (UMI), which includes 6-8 random bases. Counting the copy number of transcripts with a UMI based approach enables accurate quantification of the isoform without the interference of sequencing duplication. The UMI technology is built-in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



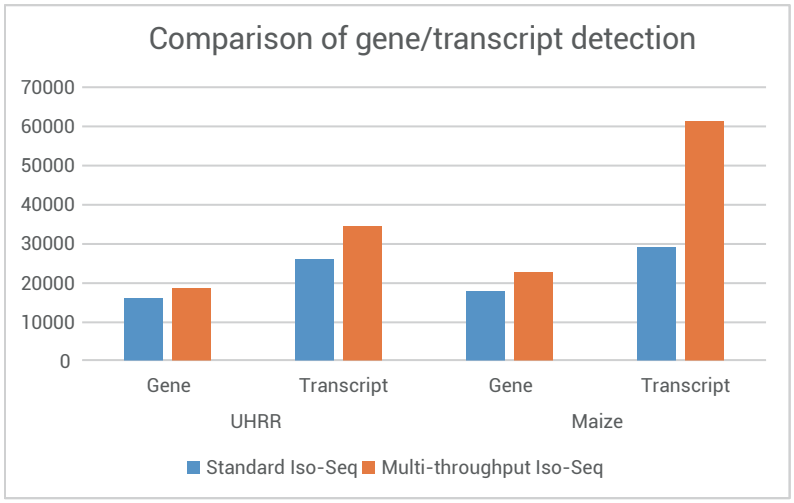
UMI enables more accurate gene quantification

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 reads and allow users to detect double amount of transcripts with the same volume of sequencing data.



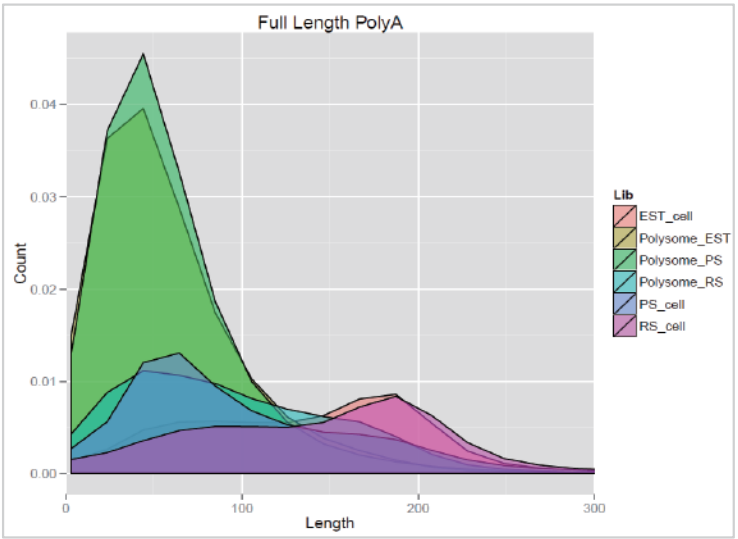
Notes:
Total sequencing data amount as follows:
UHRR-9.47 Gb (Standard Iso-Seq); 9.57 Gb (Multi-throughput Iso-Seq);
Maize-11.04 Gb (Standard Iso-Seq); 19.3 Gb (Multi-throughput Iso-Seq).
Polymerase-reads : The number of polymerases generated high quality reads. Polymerase reads will be then trimmed 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

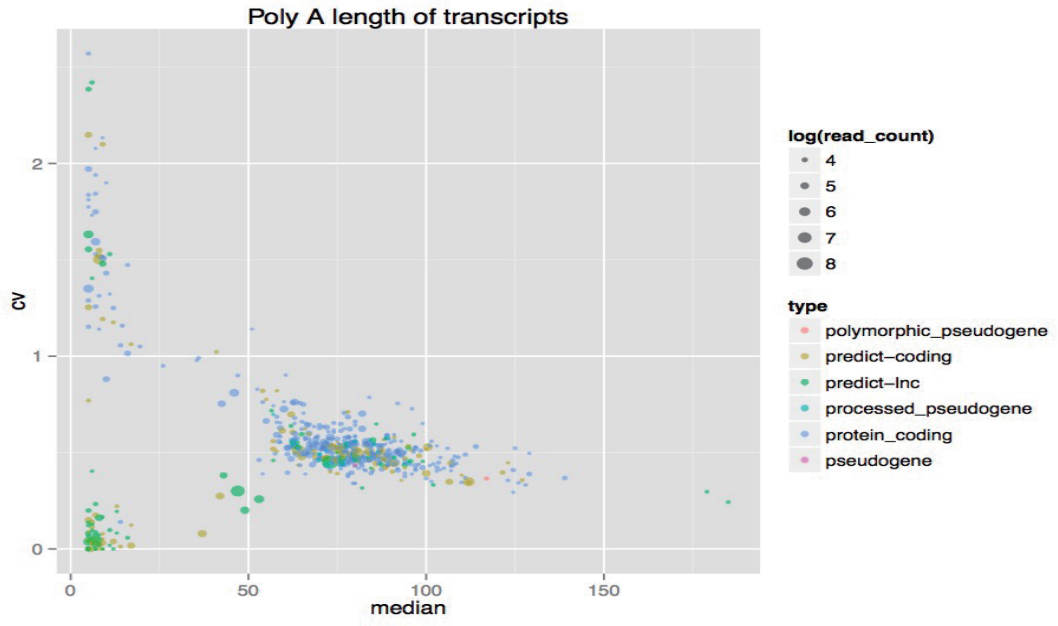
A variety of studies have reported the importance of polyA tail length for gene expression activity.[1-3]. The polyA Iso-Seq service provides additional information of polyA sequence, such as the length distribution of polyA; the occurrence frequency of other bases in polyA and the correlation between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



The length distribution of polyA

AAAAAAAAAAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAA (1);
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA(35);
ACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA(1);
AAAAAAAAAAAAAAAAAAAAAAAAA (5);
ATAAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAA(1);
AAAAAAAAAACAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAA(1);
AAAAAAAAAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA (1);

The occurrence frequency of other bases in polyA



The correlation between polyA length and gene expression

Reference

[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.

Request for Information or Quotation

Contact your BGI account representative for the most affordable rates in the industry and to discuss how we can meet your specific project requirements or for expert advice on experiment design, from sample to bioinformatics.

info@bgi.com
www.bgi.com

BGI Offices

BGI Americas

One Broadway, 14th Floor
Cambridge, MA 02142,
USA

BGI Europe

Jutrzenki 12 A,
02-230 Warszawa,
Poland

BGI Asia

Building NO.7, BGI Park,
Yantian District Shenzhen,
Guangdong Province, China

BGI Australia

L6, CBCRC, 300 Herston
Road, Herston, Brisbane,
Queensland 4006, Australia

BGI Genomics BGI_Genomics

We Sequence, You Discover