RNA-Seq FAQ

2011-04-21 (Version 1)

Contents

1. What are the advantages of RNA-Seq and what are some of the research articles about RNA-Seq?

   RNA-seq is used to analyze gene expression of certain biological objects in specific conditions. This technology combines the library construction strategy of transcriptome sequencing and the bioinformatics analysis method of digital gene expression (DGE). It is a highly accurate quantification method and has high repeatability. It also has wide detection range and is economical. It can be widely used in agricultural research, biomarker identification, environmental improvement, disease related studies, drug screening, and many other fields.

   References:

   [1] Sultan, M., M. H. Schulz, et al. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* 2008,321(5891): 956-60.

   [2] Mortazavi, A., B. A. Williams, et al. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 2008,5(7) : 621-8.

2. What are the advantages of RNA-Seq comparing to DGE?

   RNA-Seq uses the library construction strategy of transcriptome sequencing and can harvest a maximum level of poly A transcripts. The detection of transcripts in DGE is limited to the genes containing CATG recognition sites. When the sequences do not contain the recognition sites, no signal can be detected in DGE, resulting in the loss of valid information.

3. What kinds of samples are accepted for RNA-Seq? What are the requirements for total RNA sample?

   We accept total RNA, mRNA, tissue, suspension cell lines, etc.

   For plant samples:

   - Concentration ≥ 400ng/μl
   - Total quantity ≥ 20μg

   For mammalian species samples (human or mouse):

   - Concentration ≥ 80 ng/μl
   - Total quantity ≥5 μg

   For other animal species:

   - Concentration ≥200 ng/μl

- Total quantity ≥10 μg
- Sample purity: OD260/280=1.8-2.2; animal samples: RNA 28S:18S≥1.0, RIN≥7.0

Please avoid protein contamination during RNA isolation.

4. Can we use our library for sequencing?

   You can use your library. If the adapters you use are same as what BGI uses, the library can be used in Illumina sequencing directly; otherwise, you are required to provide primers. Agilent 2100 Bioanalyzer is used to test the segment size. The accurate quantification is determined by Q-PCR.

5. How can you guarantee the stability of library construction? Will the results be reliable?

   Because of the variation of the efficiency of mRNA combing to oligo(dT) magnetic beads, the results cannot be very constant in different experiments. But since we have standardized the library construction procedure and perform quality checks when the library is constructed, the results reliability is always assured.

6. What bioinformatics analysis results are provided?

   Standard bioinformatics analysis includes:

   - Statistics of raw reads
   - Sequencing saturation analysis
   - Statistics of reads randomness
   - Analysis of the distribution of reads on reference genome
   - Gene expression annotation
   - Screening of differentially expressed genes
   - Experimental repeatability analysis
   - Expression pattern analysis
   - GO enrichment analysis
   - Pathway enrichment analysis
   - Protein-protein interaction network analysis

7. How do you assure accuracy when there are large numbers of differential expressed genes?

   We use a false-discovery-rate (FDR) ≤ 0.001 and use at least two-times gene expression difference is the criterion for calling significantly different gene expression, which is similar to the standard used in microarray analysis. BGI can always reduce the number of differential expression results by lowering the FDR value.

8. How do we access data files from the project, and how do you ensure the security of the data?

   The customer is informed by email the ftp address, the user name, and the password. The data can be accessed via ftp server. The data are stored on the ftp server for one week. If you need a longer storage time, you will have to register and pay expenses during any extra week of storage.

9. How long does an RNA-Seq analysis project take?

   In general, it requires 30 working days to finish a entire RNA-Seq analysis once the samples are tested and qualified. However, if the library construction fails, the completion time will be longer. If you need expedited service and are willing to pay the related expenses, we can prioritize your request. BGI project management staff will provide timely reporting to customers to keep you informed about the testing progress.