Novogene

Novogene Product and Services FAQs

Leading Edge Genomic Services & Solutions

contents

1. DNA sequencing

1.1 Human whole genome sequencing (hWGS)

1.1.1 Sample requirements

· What is your recommended kit to carry out DNA extraction for WGS projects?

We recommend Qiagen DNeasy series kit for extraction of genomic DNA for human WGS project. Alternatively, you can also use other commercially available DNA extraction kits.

· Could we send the samples in 96-well plates or PCR stripe tubes?

No. For sample submission, 96 wells plate format is not supported. Samples should be shipped in 1.5 ml or 2.0 ml flip-top microcentrifuge tubes.

· How will Novogene perform sample QC on DNA samples?

Novogene utilizes 2 major QC methods for DNA sample qualification:

① Agarose gel electrophoresis analysis for DNA purity and integrity

② Qubit® 2.0 flurometer for DNA quantification.

· What is the average fragment size the shearing of gDNA samples ?

The genomic DNA of each sample will be randomly sheared into short fragments of about 350 bp.

· Can Novogene construct longer insert fragment libraries?

Yes, we have successfully constructed libraries with up to 800 bp long. Please contact your sales representative / technical support for more details.

· What sequencing strategy does Novogene use for hWGS services?

For WGS, we use the Illumina NovaSeq 6000 sequencer to carry out paired-end 150 bp sequencing.

· What is the sequencing coverage do you recommend for hWGS analysis?

For SNP and Indel calling, our bioinformatics team recommend at least 20X sequencing coverage.

For SV and CNV detection we recommend 30X sequencing coverage.

· Can Novogene provide any customized analysis based on client's requirements?

Yes. For customized analysis, please contact your sales representative / technical support member. If feasible, our Bioinformatics team will create a customized analysis workflow for your research's needs.

· What files do Novogene also provide along with the analysis report?

For hWGS, we provide FASTQ, BAM, annotation and VCF files with the analysis report.

· Can Novogene provide hWGS standard analysis for data from other sequencing companies?

Yes. please contact your sales representative / technical support for more details. Our Bioinformatics team will assess the project before beginning.

1.2 Plant and animal re-sequencing (PAWGS)

1.2.1 Sample requirements

· What is your recommended kit to carry out DNA extraction for PAWGS projects?

We recommend the Qiagen DNeasy series kit for extraction of genomic DNA for plant and animal re-sequencing projects. Alternatively, you can also use other commercially available DNA extraction kits.

1.2.2 Sample shipping

· Could I send lyophilized-DNA to your company?

Yes. Please note that lyophilized-DNA requires transfer and dilution service. There will be a small fee for this service.

· Since gDNA samples are relative stable, can we ship out samples at room temperature?

No. We do not advise to ship out your samples at room temperature. Even though DNA is more stable than RNA, the quality can possibly be affected by external environment. Please submit your package with sufficient dry ice (5kg per day) and ship the samples via FedEx or other preferred courier services.

· What sequencing strategy do Novogene use for plant & animal WGS services?

For PAWGS, we use Illumina NovaSeq 6000 sequencer to conduct paired-end 150 bp sequencing. We also provide PacBio Sequel and Oxford Nanopore platform if your research requires. Please contact your sales representative / technical support for more details.

· What is the sequencing coverage do you recommend for each analysis?

a) For SNP and Indel detection using Illumina platform, we recommend at least 20X sequencing coverage.
For SV and CNV detection using illumina platform, we recommend at least 30X sequencing coverage
b) For SV detection using PacBio Sequel I/II or Nanopore platform, at least 10X sequencing coverage is required.

· Can Novogene provide CRISPR/Cas9 On-target and Off-target detection analysis?

Yes. We can perform the analysis if the following information is available:
① The diploids and chromosome-level reference genome.
② The gRNA sequence of case/control paired samples.
③ Results of standard analysis.
Please contact your sales representative / technical support for more details.

· What is the sequencing coverage do you recommend for CRISPR/Cas9 on-target and off-target detection analysis?

Our recommended sequencing depth of on-target and off-target detection is greater than 30×.

· Can Novogene provide genome integration & insertion detection analysis?

Yes. This service includes:

① Vector Evaluation (not required for commercial vector)

② Insertion Calling: include SNP, Indel, SV, insertion point detection, and 2 kb flank sequence

③ Copy Number Detection & Visualization of Recombination (optional)

Please contact your sales representative / technical support for more details.

1.3 Whole exome sequencing (WES)

1.3.1 Sample requirements

· Does Novogene have recommended kits or protocols for sequencing degraded samples for hWES ?

We do not recommend further processing of degraded samples if the QC report indicate the following:

① Main bands of about 500bp. This indicates the quality of samples has been severely degraded.

② Library preparation will be difficult, as this may lead to failed libraries or lower data output.

In this situation we recommend that you re-extract samples, and re-send samples.

· Can Novogene construct libraries for low input gDNA samples from FFPE for hWES service?

Yes, our lab has successful experience in library preparation for low input samples from FFPE. The success of library preparation depends on DNA quality and amount.

1.3.2 Can Novogene combine the sequencing data from different exome capture kits and perform combined data analysis altogether?

We do not recommend performing data analysis on data from different exome capture kits due to differences in the capture regions. As we cannot confirm if the observed variants are due to differences in the samples or the capture kits themselves, we recommend to use the same kit for library preparation on all the samples.

2. Regular RNA products

2.1 Eukaryotic mRNA sequencing

2.1.1 Sample requirements

· I have only NanoDrop to assess the quantity and quality of my samples, how will it be compared against your in-house QC?

We recommend running QC check using Bioanalyzer2100 for RNA integrity (degradation). We prefer a high quality sample with lesser amount, than a bad quality sample with more amount.

* 260/280: A ratio of ~1.8 is accepted as "pure" for DNA. A ratio of ~2.0 is accepted as "pure" for RNA. If the ratio is lower in either case, it may indicate the presence of protein, phenol, or other contaminants that absorb strongly at or near 280 nm.

* 260/230: This ratio is used as a secondary measure of nucleic acid purity. The 260/230 values for "pure" nucleic acid are often higher than the respective 260/280 values. Expected 260/230 values are commonly in the range of 2.0-2.2. If the ratio is much lower than expected, it may indicate the presence of contaminants which absorb at 230 nm. If the 260/230 is not ideal, but the Qubit and the BioAnalyzer 2100 numbers are good, then a sub par 260/230 number most likely will not affect library prep and sequencing.

· Why do I get such a low RIN number when I run my total RNA samples on the Bioanalyzer 2100 even though I know that my insect RNA is not degraded?

In certain species (Insects and Arthropoda), their 28S rRNA has been described to harbor "hidden breaks" in which 28S rRNA is cleaved under denaturing conditions and co-migrate with 18S rRNA, which may appear "degraded" on native agarose gels. Laboratories that rely on gel electrophoresis to determine RNA integrity will thus be confounding. Therefore, your RIN number will always be low regardless of the quality of total RNA

· Can tissues or body fluids be sent to you to perform RNA extraction?

We provide RNA extraction services for tissues or body fluids. However, we recommend that you send us purified RNA samples. If you have a need for RNA extraction, please contact your Novogene representative to discuss your specific needs.

2.1.2 Library preparation

· What information will a "strand-specific" library show, as compared to a non-directional library?

Strand specific library will allow you to know which strand of DNA (+ or -) is the gene transcribed from.

· Would it be possible to perform RNA sequencing for low-input RNA samples?

① If total RNA ⩾ 400 ng, we will proceed with regular library preparation protocol.

② If total RNA < 400 ng, we will use the same kit, but using a modified protocol for low-input RNA (with more PCR cycles). An additional QC fee for the use of high sensitive Agilent 2100 will be changed.

③ If total RNA < 10 ng, we will recommend you to amplify the whole transcriptome RNA .

2.1.3 Analysis

· Can I perform whole transcriptome amplification for part of my samples, and compare the results with the rest of the samples that went through regular mRNA library construction?

It is not advised to compare data obtained from different library construction methods. Different treatment on the total RNA will introduce certain biasness to the sequencing data and affect the validity of the comparison results. The same applies when comparing data from stranded vs. non-stranded library, or comparing data from mRNA library vs lncRNA library etc.

· Can we only have mapping services for my project to be analyzed? If so, at what is the price?

Yes, you can choose only mapping analysis. Please contact your sales representative for pricing information.

· I have performed RNA-seq with a different vendor. Can I send the data to you for analysis? can I also combine the data with my current project, and perform differential expression analysis?

Yes, this is possible. We would simply require a few pieces of information from the project:

Sample number
Species (Latin name and reference genome)
Lib prep method (Kit name)
Sequencing platform & strategy
Data amount (per sample),
Data format. Ex. Raw data in fastq format. BAM files etc. (If the data has been processed, please let us know the software and parameters used),
Detailed analysis requirement.

For comparison with a dataset generated by Novogene, the library preparation method and sequencing strategy should be the same.

2.2 Prokaryotic mRNA sequencing

2.2.1 Library preparation

· What library preparation method will you use for prokaryotic RNA?

For prokaryotic RNA samples, we will prepare 250~300 bp insert , strand-specific libraries with rRNA removal.

3. Pre-made library

3.1 Library preparation

· What are PhiX control libraries?

PhiX serve as control libraries, generated from the PhiX virus. They are composed of 45%AT and 55%GC bases. Adding PhiX library could be added to enhance the library diversity for lane-sequencing and improve data quality.

· Why should we spike-in PhiX while sequencing?

For most libraries, Illumina recommends using PhiX control libraries with a low-concentration spike-in (1%) to monitor sequencing quality control for cluster generation, sequencing, and alignment.

For unbalanced samples such as WGBS libraries, use a higher concentration PhiX spike-in to improve crosstalk and phasing calculations.

For samples with low diversity, use a high-concentration spike-in (5% or higher) of PhiX to create a more diverse set of clusters. See below for the list of library types which require Phix:

Library Type	Characteristics
Restriction-site Associated DNA sequencing (RAD) Library	Low Diversity
Genotyping by Sequencing (GBS) Library	Low Diversity
10× Single Cell RNA Library	Low Diversity
Single-cell DNA /RNA Library	Low Diversity
Amplicon Library	Low Diversity
Microbe Re-sequencing Library	Unbalanced
Whole Genome Bisulfite Sequencing (WGBS) Library	Unbalanced

· If my library has adapter contamination, how would this affect my data quality?

Libraries with adapter contamination will affect cluster generation, which will lead to high adapter rate in sequencing data.

Novogene recommends using AMPure beads for adapter removal before submitting samples

· What standard approach would be used for pooling?

Pooling will be based on the qPCR quantification.

3.2 Sequencing

· What kinds of sequencing strategies are available?

There are several sequencing strategies.

① PE150 on NovaSeq platform.

② PE250 on NovaSeq SP
③ SE50/PE50 on NovaSeq SP platform.

· Can I use custom sequencing primers for sequencing?
Yes. However, samples with custom primers will be shipped to China for sequencing. , as our lab in China will be able to provide better customized solutions.

· Can we order one lane sequencing with custom read1 primer?
For NovaSeq S2/S4 projects with custom sequencing primers, an entire flowcell has to be ordered.
Please provide the sequence of the primer, form of delivery (lyophilized etc.), amount in nmoles, and desired the volume of buffer to be added for your desired concentration.

· We are having a 3 single-cell ATAC libraries (10k cells per library) for sequencing. Is Novogene experienced with this type of project? Which sequencing platform is better? What is the optimal sequencing depth per cell? And what else I have to know?
Yes, Novogene has experience with these types of libraries. Due to the unique read length of 16bp i5 index, we will need to change the settings of the sequencer for a whole flowcell. Therefore, you will need to purchase an entire flowcell must be purchased for this project. The recommended sequencing depth is 25,000 read pairs per nucleus. So, the total data output you need is 25,000 read pairs/nucleus*10,000 cells/lib*3lib=750M read pairs.

Depending on the data output that you need, NovaSeq SP platform may be the best choice. The output of the NovaSeq SP is 800M/flow cell. The sequencing strategy is PE50. But read2 will be 49bp, 1 bp shorter than the kit requires, because of the length of index5. It has been verified that it does not affect the results analysis. A 1% PhIX spike-in is strongly recommended.

You may refer to the link for more details. https://support.10xgenomics.com/single-cell-atac/sequencing/doc/specifications-sequencing-requirements-for-single-cell-atac?tdsourcetag=s_pcqq_aiomsg

Sequencing Depth & Run Parameters

Sequencing Depth	25,000 read pairs per nucleus (25,000 reads for Read 1N; 25,000 reads for Read 2N)
Sequencing Type	Paired-end, dual indexing
Sequencing Read	Recommended Number of Cycles
Read 1N	50 cycles
i7 Index	8 cycles
i5 Index	16 cycles
Read 2N	50 cycles

3.3 Data analysis

· Can I get BCL file for my libraries?
BCL file can only be provided for lane sequencing.

4. Epigenetic products

4.1 Whole genome bisulfite sequencing (WGBS)

4.1.1 Sample requirements

· Could we perform bisulfite sequencing on whole tissues (such as placenta with different cell composition) is correct or not?

WGBS is helpful for investigating parental expression bias. However, placental tissue is not optimal, as it has a mixed origin of parent and offspring. PBMCs (peripheral blood mononuclear cells) may be more appropriate.

4.1.2 Sequencing

·What would be the appropriate sequencing depth?

We recommend at least 30× coverage.

4.2 Reduced representation bisulfite sequencing (RRBS)

4.2.1 Sequencing

· I need to perform an epigenome analysis of 58 human samples. with bioinformatics analysis. We have previously used EPIC bead chip from Illumina. What sequencing recommendations do you have?

We are able to provide the EPIC bead chip sequencing.

In this situation, we would recommend RRBS (Reduced Representation Bisulfite Sequencing). RRBS was developed primarily for animals. RRBS is an accurate, efficient and economical method for DNA methylation research. Enrichment of promoter and CpG island regions by enzymatic cleavage (Mspl), combined with bisulfite sequencing, achieved high resolution of DNA methylation status detection and high utilization of sequencing data.

4.3 Chromatin immunoprecipitation sequencing (CHIP-seq)

4.3.1 Sequencing

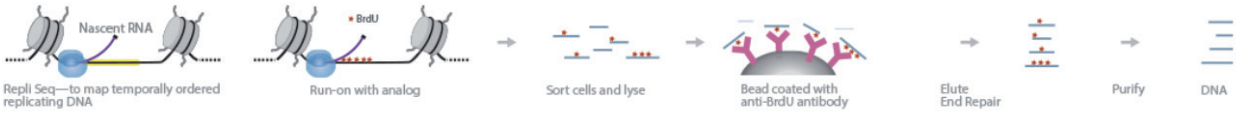
· Does Novogene have protocols for Repli-seq and what is the analysis content for this sequencing?

Novogene does not offer Repli-seq services. However, the principle of Repli-seq is similar to ChIP-seq, where our lab is very experienced in. We may be able to customize the workflow of ChIP-seq into this project.

As for bioinformatic analysis, our team will construct a new mode for Repli-seq analysis based on your project. Some aspects of analysis for ChIP-Seq will also be used in Repli-seq, such as strand cross correlation, peak calling, motif, differential analysis, annotation, and enrichment., based on your analysis requirements.

* Note: Cycling cells duplicate their DNA content during S phase, shortly after a process called replication timing (RT). Early and late replicating regions differ in terms of mutation rates, transcriptional activity, chromatin marks and sub-nuclear position. Moreover, RT is regulated during development and is altered in disease. Exploring mechanisms linking RT to other cellular processes in normal and diseased cells will be facilitated by rapid and robust methods with which to measure RT genome wide.

Repli-Seq



Repli-Seq maps the sequences of nascent DNA replication strands throughout the whole genome during each of the six cell cycle phases. This is achieved by growing cells in bromouridine triphosphate (BrdU) media to replace thymidine. The cells are then sorted to their current state in cell division using fluorescent-activated cell sorting (FACS). BrdU-labelled DNA strands are immunoprecipitated by anti-BrdU antibodies on magnetic beads. These immunoprecipitated strands can be prepared for sequencing following TruSeq DNA library preparation protocol.

Pros:

Maps sequences of newly replicated DNA to the phases of cell division.

Low sample input required (5000 cells) makes it suitable for studying rare cell populations.

Streamlined DNA library prep step.

Cons:

The BrdU labeling requirement limits this approach to cultured cells.

· Can you provide sequencing service for 24 yeast samples with 0.04Gb per sample for CHIP-seq?

We generally do not accept such low data output per sample. However, we may be able to sequence the samples using NovaSeq SP SE50 under special circumstances, since the minimum amount of data for the NovaSeq S4 PE150 strategy is 1G per sample, the NovaSeq SP SE50 strategy is recommended for this yeast Chip-seq project.

4.3.2 Analysis

· We have opted for a "spike-in" approach using is a kit purchased from Active Motif that addresses normalization issues related to the target we have ChIP-ed (H3K27me3). It involves adding a small amount of Drosophila chromatin into each sample, which will be immunoprecipitated along with our target of interest. For the analysis, we would need the sequences to be mapped to both the human/mouse and Drosophila genomes, and then the Drosophila reads can be used to normalize the human/mouse data. Will it be feasible?

We can use the Drosophila reads to normalize the human/mouse data. We can include this normalization in the standard analysis.

Our protein (Trib2) is a pseudokinase and its interaction with DNA is not well known. Our immediate aimAt least one negative control is needed to eliminate the noise/background. If the target of interest with DNA is not well known, a positive control could help to determine the effect of the ChIP assay.

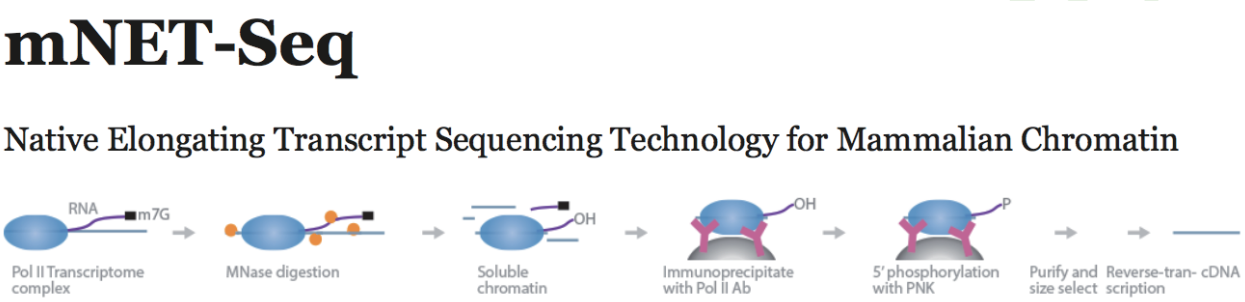
4.4 RNA immunoprecipitation sequencing (RIP-seq)

4.4.1 Analysis

· Does Novogene have protocols for mNET-seq and what is the analysis content for this sequencing?

Novogene does not offer mNET-seq services. However, the principle of mNET-seq is similar to that of RIP-seq , where we are very experienced in.

Note:



mNET-Seq generates profiles of nascent RNA and co-transcriptional RNA processing associated with different C-terminal domain (CTD) phosphorylation states throughout the whole genome (Nojima et al., 2015). mNET-Seq is able to provide precise sequence reads of RNAPII active sites during transcript elongation and also RNA processing intermediates. First, elongating RNAPII complexes are isolated through chromatin fractionation. They are digested with MNase, breaking down all exposed DNA while leaving RNA strands protected by RNAPII or spliceosomes intact. The RNAPII complexes are immunoprecipitated using RNAPII antibodies and 5′ phosphorylated by T4 PNK. Next, 3′ linkers are ligated to the 3′ hydroxyl end of the RNA strand still embedded within RNAPII. They are also incubated with radioactive ATP to facilitate size selection. Nascent RNA strands are size-selected for 35_100 nt, processed into cDNA sequencing libraries, and sequenced. The use of various RNAPII antibodies during purification raises the versatility and specificity of the technique in targeting CTDs of RNAPII.

Advantages:

Maps nascent RNA strands and co-transcriptional RNA processing during RNAPII elongation with phosphorylated CTDs;
Able to detect sense and antisense transcripts at TSS (Nojima et al., 2016);
No crosslinking eliminates introduction of artifactual interactions;
MNase digestion is specific and efficient;
Various RNAPII-specific antibodies can be used to increase targeting accuracy.

Disadvantages:

Nascent RNAs shorter than 35 nt cannot be detected reliably;
RNA can degrade during RNAPII immunoprecipitation;
mNET-Seq peaks might be obscured by peaks from co-transcriptional RNA cleavage;
PCR amplification may give rise to peaks from amplification bias.

5 Whole transriptome sequencing (WTS)

5.1 Sequencing

· What is the recommended sequencing depth of whole trancriptome sequencing?

WTS consist of two library type, lncRNA and small RNA. lncRNA-seq will propduce 12G raw data/sample, strand specific libraries with rRNA removal. Small RNA libraries will be 20M raw reads/sample.

5.2 Analysis

· What analysis contents do you provide for WTS?

We provide analysis for WTS, including analysis for lncRNA & mRNA, circular RNA, small RNA, and association analysis.

6. Microbial sequencing

6.1 Metagenomic amplicon sequencing

6.1.1 Sequencing

· Does Novogene provide a service for primer synthesis, or utilize those already in the system for a repeat? Could you provide me only 100,000 raw reads per sample?

According to our policy, we can safekeep your samples for up to 3 months after data release (additional charges might apply). Unfortunately, we don't provide primer synthesis services, since our current platform is incompatible with customized PCR products. If the primer can be synthesized by 3rd parties company, we can provide the following two options:

① If you can send us the PCR product after amplification, we can sequence it on the Novaseq, with PE250 strategy with a minimum output of 1Gb.

② Alternatively, you can add barcodes on each of your amplicon and mix them into one or two tubes, then demultiplex by barcode after sequencing. This will most likely be more cost-effective.

If you only require 0.1M reads per sample, according to our new price list, we can do mixed library for your amplicons. In this case, each sample can get 100,000 raw reads, but you need to add 8bp barcode for each amplicon and provide us the barcode sequence and position.

· We are interested to survey variations in 10 target genes between 50 different samples/strains of wheat. We will be amplifying 10 different genomic regions from each sample using PCR and are considering whether to send us pre-prepared libraries for each sample or amplicons for library prep and sequencing at Novogene. The size of the amplicons/target regions we are interested in is around 4000bp. Here are my two questions:

① For your PCR amplicon sequencing service, would it be possible for them to combine 10 separate amplicons (unique genomic regions) from each sample into a single tube for library prep at Novogene? (50 samples. 10

amplicons per sample -> pool amplicons; 50 library preps (one per sample))
② For SNP/InDel calling in these target regions, what do you think would be an appropriate coverage?

Yes, those amplicons could be pooled into one single tube after adding 8bp barcode. The size of each amplicon should not be more than 450bp. In addition, the barcodes may not be helpful if the client wants the full sequence for a PCR product that is 1kb+. Therefore, we recommend that the client prepares libraries for each sample rather than to pool them together. In terms of coverage, 30X should be sufficient for SNP/InDel calling in these target regions. As our client did not exactly require bioinformatic, here are four strategies. The first two are for sequencing on Pacbio, and the last two are sequenced by Illumina.

① PCR product with barcode sequence
Sequencing strategy:
Library prep - Pool PCR products together and construct one PacBio DNA library.
Seq – Being sequenced on one Sequel SMRT cell.

② PCR product without barcode sequence
Sequencing strategy:
Library prep – Do not pool the PCR product together; one PCR product per PacBio DNA library.
Seq - Be sequenced on one Sequel SMRT cell.

③ Premade library (Illumina platform)
Sequencing strategy:
Library prep - Pool of 50-amplicon libraries
Seq - 1 lane HiSeq

④ PCR-product (Illumina platform)
Sequencing strategy:
Library prep - one PCR product per library
Seq – Ilumina PE150

PacBio sequencing is not optimal for SNP/Indel. On Illumina platform, it is too difficult to assembly with 4000bp. There will be several gaps or blankets. Our team cannot accept PCR product with this data output. Therefore, , we do not recommend samples being sequenced by Illumina over 2000bp.

However, if you have requested Illumina seq and you have not prepared primers/barcodes for their existing amplicon preps, solution 4 could meet your needs.

· We need to test the stability of recombinant virus obtained from chicken embryo eggs. We passage the recombinant virus up to 8-10 passages and check HA and NA gene sequence at each passage. What we need is to identify mutations at each nucleotide position of influenza HA and NA genes, also, we need to know the

frequency of minor variant at each nucleotide position. It is virus PCR product lead for 1800 samples. And we will ship them in batches of 24 samples over the next 2 years. What kind of analysis could we have to meet our requirements?

This will most likely go through our WGS pipeline. We will align the PCR product sequences to the reference genome sequence and then identify mutations such as SNPs and variant frequencies.

Therefore,

① We need to know the length of the PCR products.

② We need a reference genome of the species.

· The genes we amplified are HA and NA genes from influenza A. We will have different influenza A viruses each time. The length of HA is between 1.7-1.8kb and NA is between 1.5-1.6kb. The length may vary with different viruses but will be in that range. Out target sample size is 1800 in two years. We will try to run 24 samples each time. Yes, we do want to assemble sequence for each gene based on the sequencing data.

We can perform this customized analysis.

6.2 Metagenomics

· What sequencing depth is sufficient to identify certain genes, for instance, all the genes of antimicrobial resistance in the samples?

This can vary depending on the project. The microbe content is different from different samples and thus will require different depth of sequencing. We can refer to papers for more information.

In short, there is no easy way to estimate read depth required for shotgun metagenomics sequencing. Environmental samples have a large distribution of species; each species would have to be accounted for individually. You would need to know the number of total species in the sample, the genome size, and relative abundance for each species. In most cases this is not possible when you are sequencing a sample for the first time.

Let's assume you were dealing with a simple sample that had 10 bacterial species and wanted 100x coverage depth for de no assembly. If your 10 bacterial species had an estimated genome size of 2 Mb, you'd aim for around 2 Gb of sequencing data per sample.

10 dominant bacterial species * 100× * 2 Mb = 2 Gb

Reference https://genohub.com/shotgun-metagenomics-sequencing/

Our data demonstrate that D0.5 with ⩾ 50 million reads would be a suitable compromise for sequencing bovine fecal samples and adequately inferring their resistome, considering that no further classes were discovered by the D1 sequencing depth and only a single unique mechanism was discovered as compared to the D0.5 level.

Reference https://www.nature.com/articles/s41598-018-24280-8

So 50 Mb reads is the optimized data amount.

· Can we use metagenomics data to perform amplicon analysis or vice versa?

No, we cannot. Metagenomics involve sequencing the whole genome, while amplicons are sequenced at specific

target regions in the genome, which is usually very short. Even though meta-sequencing may contain the desired target region, the sequencing depth of metagenomics and amplicon sequencing is not the same, and it is not possible to analyze the target region.

In addition, the purpose of metagenomics is to study the genome and its function of a microorganism. The purpose of amplicon sequencing is to study the microbial community structure in a specific environment. The purposes are different, and the intermediate process from sample acquisition to final data acquisition is different. Therefore, metagenomics data cannot be used for amplicon analysis and vice versa.

6.3 Small genome sequencing

6.3.1 Sequencing

·I am very interested in small genome sequencing (mitochondria, chloroplast). I would like to perform genome sequencing for 6 samples of chloroplast in mixed tomato and eggplant. I would like to send you fresh leaves for each sample and extracting chloroplast DNA. Would it be feasible?

We recommend that you extract the chloroplast DNA yourselves and send DNA samples to us. we will be able to proceed with the library preparation subsequently.

7. Long-read sequencing

7.1 Pacbio sequel DNA sequencing

7.1.1 Sequencing

· What are my options for long reads sequencing?

We have four options:

① PacBio DNA sequel II (1 SMRT cell /sample) + Iso-seq (15 G/sample)

② PacBio DNA sequel II (1 SMRT cell/ /sample) + RNA-seq (Illunima, 6 G/sample)

③ Human WGS (Illumina, 30×/sample) + RNA-seq (Illunima, 6 G/sample)

It is worth noting that PacBio platform is not very accurate in SNP detection because of the machine's characterization. You could also choose the fourth option.

④ Pacbio DNA sequel II (1 SMRT cell/ /sample) + Human WGS (Illumina, 30×/sample) + RNA-seq (Illunima, 6 G/sample)