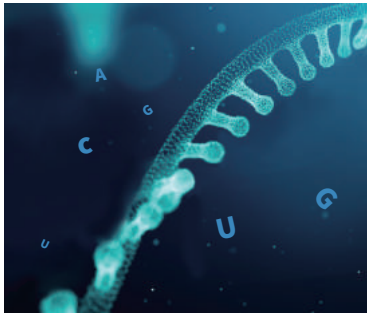# Novogene

Eukaryotic mRNA Sequencing

mRNA-Seq uses next-generation sequencing (NGS) to reveal the presence and quantity of messenger RNA in a biological sample at a given moment, analyzing the continuously changing cellular transcriptome. Novogene's mRNA-Seq, based on state-of-the-art Illumina NovaSeq platforms with paired-end 150 bp sequencing strategy, offers complete solutions for gene expression quantification and differential gene expression analysis among groups of samples, as well as for identification of novel transcripts, alternative splicing, and gene fusion events, etc. Our experienced bioinformaticians work closely with customers to provide standard and customized data analysis and publication ready results for species with and without a reference genome.
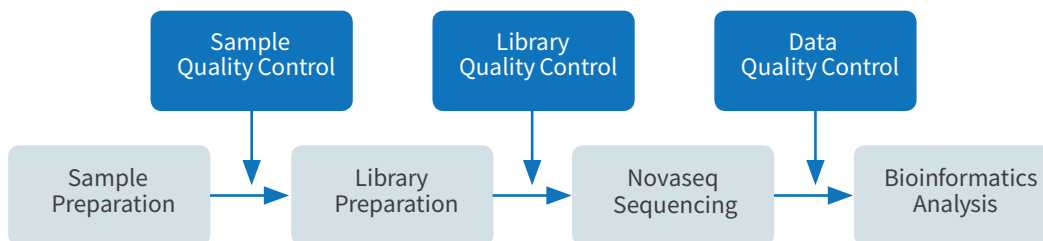
Extensive experience with over 30,000 samples

Data quality exceeds Illumina's official guarantee

In house pipeline to meet different analysis requirement

Project workflow

Sample Quality Control →

Library Quality Control

Data Quality Control

Sample Preparation → Library Preparation → Novaseq Sequencing → Bioinformatics Analysis

Sequencing parameter

Platform	Illumina NovaSeq 6000	
Read length	Pair-end 150	
Recommended Sequencing Depth	Recommended: 15G for animals and plants 3G for fungi	Minimum: 6G for animals and plants 2G for fungi
Data quality	Guaranteed Q30 ≥ 80%, exceeding Illumina's official benchmark of ≥ 75%	
Turnaround time	Express service: 15 working days from sample delivery to data delivery （≤24 samples)	

Samples requirement

Library Type	Sample Type	Amount Required	Volume	Concentration	RNA Integrity Number (Agilent 2100)	Purity (NanoDrop™)
Eukaryotic RNA-Seq (cDNA library)	Total RNA	≥ 400 ng	≥ 20 µL	≥ 20 ng/µL	≥ 6.8 (Animal), smooth base line ≥ 6.3 (Plant and Fungus), smooth base line	OD260/280 ≥ 2.0, OD260/230 ≥ 2.0, No degradation, No contamination

Analysis Pipeline

Species with reference genome

Raw Data → Quality Control → Mapping to Reference Genome → Novel Gene Prediction → Quantification → Differential Expression Analysis → Enrichment Analysis → GO Enrichment / KEGG Pathway Enrichment → Alternative Splicing Analysis → Mutation → Analysis Report

Species without reference genome

Raw Data → Quality Control → Transcriptome Reconstruction → Annotation → Quantification → Differential Expression Analysis → Enrichment Analysis → GO Enrichment / KEGG Pathway Enrichment → CDS Prediction → SSR Detection → Mutation → Analysis Report
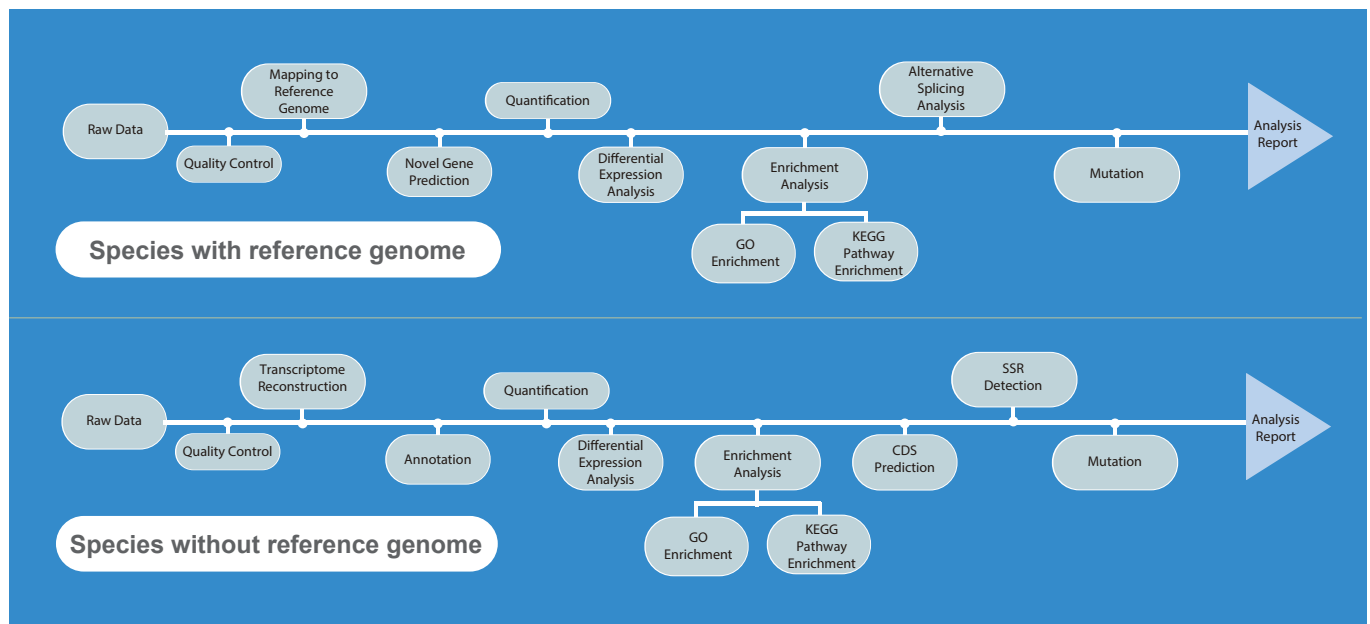
Project experience

Background

Neuroblastoma (NB), one of the most common malignant solid tumors in pediatric population that arises from neural crest-derived cells, constitutes 15% of cancer related mortality in childhood. Poor clinical outcome in patients suffering from high risk NB. The mechanisms essential for the aggressiveness and progression of NB still warrant further investigation.

Sample and sequencing strategy

Tumor cells with/ without MYCN amplification
Library preparation: RNA-seq library
2 Sequencing: Illumina platform

Conclusion

ARMC12 plays a crucial role in tumor progression and could be a potential therapeutic approach for NB. Mechanistically, ARMC12 physically interacts with retinoblastoma binding protein 4 (RBBP4) to facilitate the formation and activity of polycomb repressive complex 2, resulting in transcriptional repression of tumor suppressive genes
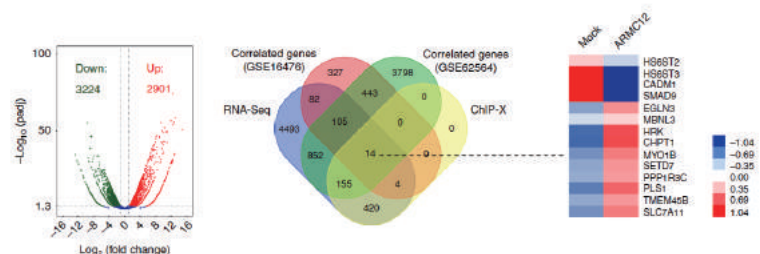
Ectopic expression of ARMC12 represses the expression of PRC2 downstream tumor suppressive genes in NB cells.

Publications using Novogene's expertise

Year	Journal	Article
2019	Applied and Environmental Microbiology	l-Rhamnose metabolism in clostridium beijerinckii strain DSM 6423
2018	Scientific Reports	Csde1 binds transcripts involved in protein homeostasis and controls their expression in an erythroid cell line
2018	Int J Mol Sci	Overexpression of the Rybp Gene inhibits differentiation of bovine myoblasts into myotubes
2018	Lung Cancer	Identification of MET exon14 skipping by targeted DNA- and RNA-based next-generation sequencing in pulmonary sarcomatoid carcinomas
2018	Nature communication	Armadillo repeat containing 12 promotes neuroblastoma progression through interaction with retinoblastoma binding protein 4
2017	PeerJ	Integration of lncRNA-miRNA-mRNA reveals novel insights into oviposition regulation in honey bees
2017	Mol Ecol	Insights into deep-sea adaptations and host-symbiont interactions: A comparative transcriptome study on Bathymodiolus mussels and their coastal relatives

Eukaryotic lncRNA Sequencing
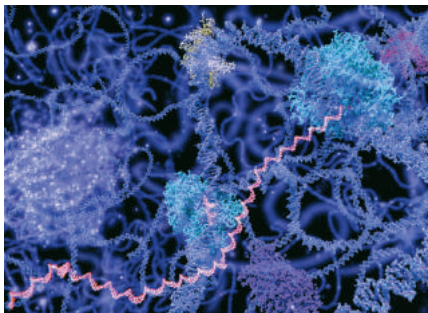
Long non-coding RNAs (lncRNAs) are transcripts longer than 200 nt that do not encode for proteins. lncRNAs are extensively reported to be involved in transcriptional regulations by targeting mRNAs. In Novogene, the progressive library preparation enables information enrichment and gene expression profiling for both coding and non-coding transcripts in a single sequencing run. This allows our bioinformatics team provides comprehensive analysis for both lncRNAs and mRNAs, releasing not only the quantification and functional enrichment of interested transcripts, but also the regulatory relation between lncRNAs and mRNAs.

Extensive experience with over 30,000 samples

Data quality exceeds Illumina's official guarantee

In house pipeline to meet different analysis requirement

Project workflow

Sample Quality Control → Total RNA → Library Construction

Library Quality Control → Sequencing

Data Quality Control → Bioinformatics Analysis
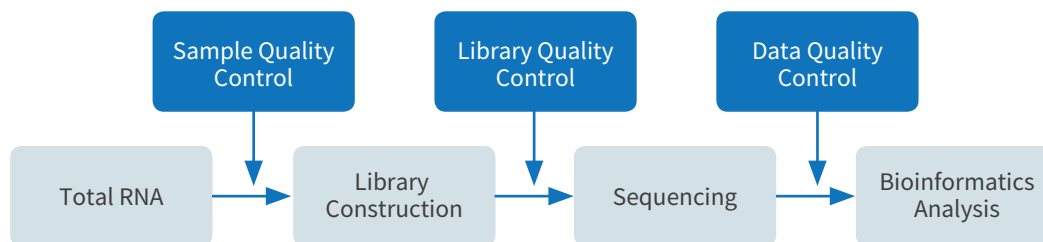
Sequencing parameter

Platform	Illumina NovaSeq 6000	
Read length	Pair-end 150	
Recommended Sequencing Depth	Recommended: 12 Gb ~ 15 Gb data per sample (40 Mb ~ 50 Mb paired reads per sample)	Minimal: 9 Gb per data per sample (~ 30 Mb paired reads per sample)
Data quality	Guaranteed Q30 ≥ 80%, exceeding Illumina's official benchmark of ≥ 75%	
Turnaround time	18 working days from samples arrival to data analysis (≤ 24 samples)	

Samples requirement

Library Type	Sample Type	Amount	Volume	Concentration	RNA Integrity Number (Agilent 2100)	Purity (NanoDrop™)
Eukaryotic long non-coding RNA-Seq (Directional & rRNA removal library)	Total RNA	≥ 500 ng	≥ 10 µL	≥ 50 ng/µL	≥ 6.5 (Animal), ≥ 6.0 (Plant and Fungus) smooth base line	OD260/280 ≥ 2.0, OD260/230 ≥ 2.0, No degradation, No contamination
	Exosomal RNA	≥ 20 ng	≥ 10 µL		Fragments distributing between 25-200nt (by high sensitive Agilent Bioanalyzer 2100), FU>10, with no peak>2000nt	

Analysis Pipeline

- Raw Data
- Quality Control
- Mapping to Reference Genome
- LncRNA Identification
- Novel Gene Prediction
- Quantification
- LncRNA Target Prediction
- Differential Expression Analysis
- Enrichment Analysis
  - GO Enrichment
  - KEGG Pathway Enrichment
- Alternative Splicing Analysis
- Mutation
- Analysis Report
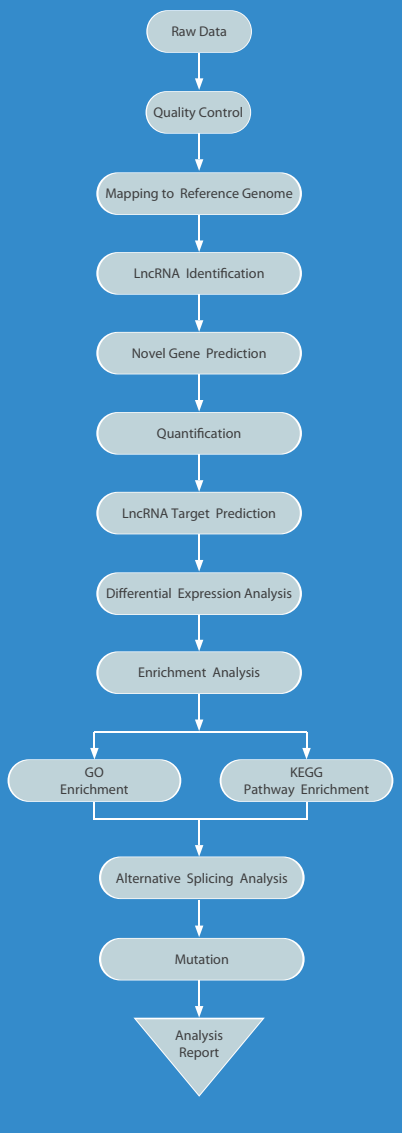
Novogene Data

REPRESENTATIVE DATA QUALITY RESULTS OF lncRNA SEQUENCING (PE 150) FROM NOVOGENE:

Sample Name	# of Raw Reads	# of Clean Reads	Clean Bases	Error Rate (%)	Q20 (%)	Q30 (%)	GC Content (%)
Sample1	118726766	117424068	17.61G	0.01	97.70	94.09	52.33
Sample2	133948474	132481702	19.87G	0.01	97.85	94.40	50.66
Sample3	144285430	142696176	21.4G	0.01	97.84	94.42	52.15
Sample4	133066914	132086794	19.81G	0.01	98.27	95.50	49.99
Sample5	133570278	131801718	19.77G	0.01	98.28	95.49	53.05
Sample6	111400698	107732776	16.16G	0.01	97.82	94.31	51.56

Project Example

Systematic Analysis of Long Noncoding RNAs in the Senescence-accelerated Mouse Prone 8 Brain Using RNA Sequencing

Molecular Therapy—Nucleic Acids (2016) 5, e343

This project was to understand the significance of lncRNAs in Alzheimer's disease pathogenesis by using mice models. In this paper, 8,422 known lncRNA and 19,423 presumed lncRNAs were detected. Differential expression analysis and enrichment analysis were performed on both lncRNAs and mRNAs. Potential regulatory relation between lncRNAs and mRNAs were also predicted.
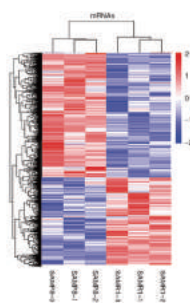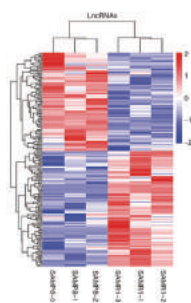

Publications using Novogene's expertise

Year	Journal	Article
2019	BioRxiv	An impact of HP1γ on the fidelity of pre-mRNA splicing arises from its ability to bind RNA via intronic repeated sequences
2019	Frontiers in genetics	Long read single-molecule real-time sequencing elucidates transcriptome-wide heterogeneity and complexity in esophageal squamous cells
2018	BMC Genomics	The temporal expression patterns of brain transcriptome during chicken development and ageing
2018	Cancer Research	Long noncoding RNA pancEts-1 promotes neuroblastoma progression through hnRNPK-mediated β-catenin stabilization
2018	Cell	Self-recognition of an inducible host lncRNA by RIG-I feedback restricts innate immune response
2017	Journal of agricultural and food chemistry	Zearalenone exposure induces the apoptosis of porcine granulosa cells and changes long noncoding RNA expression to promote antiapoptosis by activating the JAK2–STAT3 pathway
2017	Peer J	Integration of lncRNA–miRNA–mRNA reveals novel insights into oviposition regulation in honey bees
2016	BMC Genomics	Genome-wide analysis of long non-coding RNAs at early stage of skin pigmentation in goats (Capra hircus)
2016	Biology of Reproduction	Systematic identification of long non-coding RNAs in immature and mature porcine testes