Whole Exome Sequencing

Exome sequencing provides a cost-effective alternative to whole genome sequencing as it targets only the protein coding region of the human genome responsible for a majority of known disease related variants. Whether you are conducting studies in rare Mendelian disorders, complex diseases, cancer research, or human population studies, Novogene's comprehensive human whole exome sequencing service provides a high-quality, affordable and convenient solution.

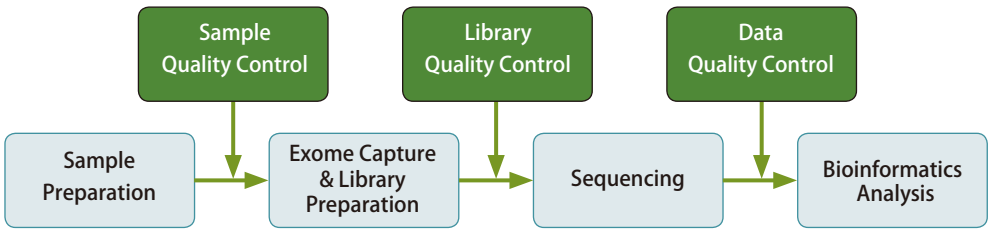
Novogene's bioinformatics analysis includes data QC, mapping with reference genome, SNP/InDel, somatic SNP/InDel calling, statistics and annotation. Novogene utilizes internationally recognized softwares in bioinformatics analysis, e.g. BWA, SAMtools, GATK, etc.

In particular, Novogene bioinformatics pipeline includes annotation with the exome aggregation consortium (ExAC). ExAC dataset spans 60,706 unrelated individuals sequenced as part of various disease-specific and population genetic studies. This population scale database greatly facilitates research of disease pathogenesis.

The Novogene Advantage

- Unsurpassed data quality: We guarantee a Q30 score ⩾ 80%, exceeding Illumina's official guarantee of ⩾ 75%.
- State-of-the-art exome capture: Agilent SureSelect Human All Exome V6 (58 M) are used.
- Accurate variant calling with longer read length up to 150 bp.
- Extraordinary informatics expertise: Novogene uses its cutting-edge bioinformatics pipeline and internationally recognized best-in-class software to provide customers with "publication-ready data".

Project Workflow



EXOME CAPTURE

- Agilent SureSelect Human All Exon V6 Kit

SEQUENCING STRATEGY

- 180 ~ 280 bp insert DNA library
- Illumina platform, paired-end 150pb

DATA QUALITY GUARANTEE

- We guarantee that ⩾ 80% of bases have a sequencing quality score ⩾ Q30, which exceeds Illumina's official guarantee of ⩾ 75%

TURNAROUND TIME

- 22 working days after verification of sample quality (without data analysis)
- Additional 5 working days for data analysis

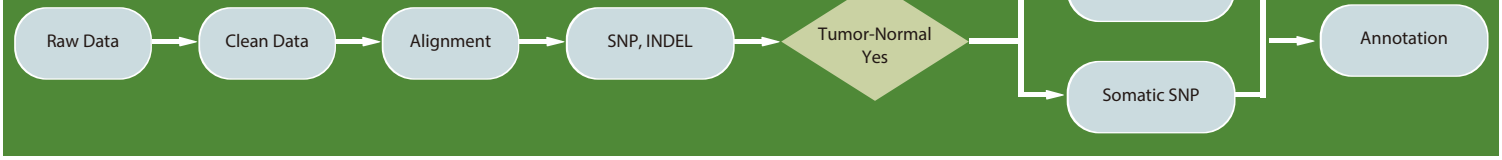
SAMPLE REQUIREMENTS

- Input DNA:
  - For fresh sample: ⩾ 0.6 μg (a minimum of 200 ng can be accepted with risk)
  - For FFPE sample: ⩾ 1 μg
- DNA concentration: ⩾ 20 ng/ μl
- DNA Volume: ⩾ 20 μl
- OD260/280 = 1.8 - 2.0 without degradation or contamination

RECOMMENDED SEQUENCING DEPTH

- For Mendelian disorder/rare disease: effective sequencing depth above 50x
- For tumor sample: effective sequencing depth above 100x

Analysis Pipeline



NOVOGENE (UK) COMPANY LTD.
25 Cambridge Science Park
Milton Road
Cambridge, CB4 0FW
United Kingdom

Tel: +44(0)1223 628750
Eml: europe@novogene.com
Web: www.novogene.com
China · China Hong Kong · Singapore · UK · USA

Follow us on LinkedIn

Novogene Data

Novogene provides highly qualitied NGS services. In standard practice, Novogene guarantees an average Q30 of 80% for WGS, exceeding Illumina's official guarantee of 75%.

The following table includes actual Novogene data from human exome sequencing projects, and demonstrates the quality of our sequencing. Alignment of the results to the reference genome (UCSC hg38) showed an average mapping ratio of 99.83%.

REPRESENTATIVE WHOLE EXOME SEQUENCING DATA FROM NOVOGENE

Sample Name	Raw Data (Gb) [1]	Effective (%)[2]	Error (%) [3]	Q20 (%) [4]	Q30 (%)[5]	GC (%)[6]	Mapped (%) [7]	Average sequencing depth (x) [8]	Coverage of target region [9]	Percentage of target with > 4x coverage [10]	Percentage of target with > 10x coverage [11]	Percentage of target with > 20x coverage [12]
Novo 1	6.5	98.17	0.03	95.38	88.74	51.98	99.72%	67.30	99.70%	99.40%	98.10%	92.20%
Novo 2	7.9	98.23	0.03	95.87	89.70	52.47	99.79%	83.26	99.60%	99.40%	98.60%	94.90%
Novo 3	8.4	98.39	0.03	96.00	90.09	52.10	99.83%	88.24	99.60%	99.50%	98.90%	96.00%
Novo 4	8.8	98.19	0.03	97.23	92.83	50.59	99.85%	87.94	99.70%	99.40%	98.80%	96.60%
Novo 5	9	98.99	0.03	96.08	90.20	51.30	99.86%	91.54	99.90%	99.70%	99.20%	97.10%
Novo 6	12.3	98.57	0.03	97.49	93.19	51.85	99.88%	117.46	99.90%	99.50%	98.20%	94.80%
Novo 7	14.5	98.54	0.03	96.66	91.38	52.62	99.89%	212.67	99.70%	99.30%	98.20%	95.50%
Novo 8	14.9	99.08	0.03	97.25	92.88	53.63	99.78%	188.48	99.60%	99.30%	98.60%	97.10%
Novo 9	15	98.81	0.03	97.50	93.36	51.73	99.86%	149.42	99.70%	99.40%	98.30%	95.50%
Novo 10	15.2	98.71	0.03	97.39	93.22	52.52	99.81%	179.56	99.60%	99.30%	98.50%	96.60%
Novo 11	15.6	99.10	0.03	97.28	92.61	52.64	99.81%	222.68	99.80%	99.50%	98.80%	97.10%
Novo 12	15.7	99.18	0.03	97.39	93.11	52.75	99.84%	206.79	99.80%	99.60%	99.10%	97.90%
Novo 13	16.3	98.81	0.03	97.16	92.27	51.27	99.78%	173.35	99.60%	99.30%	98.50%	96.50%
Novo 14	16.6	98.70	0.04	96.99	92.40	51.77	99.84%	155.91	99.70%	99.40%	98.70%	96.70%
Novo 15	16.9	99.16	0.03	96.12	90.45	52.35	99.75%	208.09	99.60%	99.20%	98.20%	95.70%
Novo 16	17.1	98.71	0.03	97.22	92.78	52.20	99.83%	176.51	99.60%	99.40%	98.80%	97.30%
Novo 17	17.2	98.85	0.03	97.72	93.91	53.47	99.84%	176.91	99.60%	99.30%	98.70%	97.00%
Novo 18	17.6	99.21	0.03	97.06	92.18	53.09	99.83%	193.89	99.80%	99.60%	98.90%	97.30%
Novo 19	18	99.19	0.03	97.13	92.60	52.62	99.87%	388.68	99.60%	99.50%	99.30%	98.70%
Novo 20	18.7	98.86	0.03	97.56	93.53	52.47	99.85%	188.92	99.80%	99.50%	98.60%	96.40%
Novo 21	19	99.05	0.03	97.23	92.46	52.86	99.90%	209.24	99.60%	99.30%	98.70%	97.30%
Novo 22	19.6	99.11	0.03	97.20	92.47	53.53	99.89%	196.20	99.80%	99.60%	99.00%	97.50%
Novo 23	19.8	98.81	0.03	97.10	92.41	51.31	99.84%	215.30	99.60%	99.40%	98.80%	97.50%

1 Original sequencing data (in gigabases).
2 Percentage of clean reads from all raw reads.
3 Average error rate of all bases in read1 and read2.
4 Percentage of reads with an average quality greater than Q20.
5 Percentage of reads with an average quality greater than Q30.
6 Percentage of G and C bases from total bases.
7 Percentage of total reads that mapped to the reference genome.
8 Average sequencing depth (times coverage) per reference genome target region.
9 Percentage of target region covered by sequencing.
10 Percentage of bases in target region with a sequencing depth ⩾ 4x.
11 Percentage of bases in target region with a sequencing depth ⩾ 10x.
12 Percentage of bases in target region with a sequencing depth ⩾ 20x.

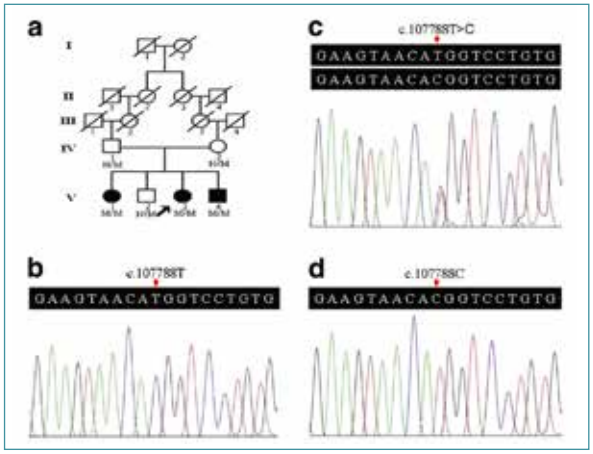
Project Example

The following study utilized Novogene's expert exome sequencing service.

Identification of a Novel Mutation in the Titin Gene in a Chinese Family with Limb-Girdle Muscular Dystrophy 2J

Molecular Neurobiology (2015):1-6.

This study explored the genetic basis for an inherited myopathy, limb-girdle muscular dystrophy 2J (LGMD2J), in a multigenerational Chinese family. An exome library from a family member with LGMD2J was prepared and sequenced at Novogene, and bioinformatics analysis revealed a novel homozygous point mutation in the titin gene, which is known to be associated with LGMD2J. Further analyses revealed that the homozygous variant was present in family members with this form of muscular dystrophy, but individuals who did not show the same phenotype either bore heterozygous mutation or did not have the mutation. These findings will contribute to the understanding of the disease and improve genetic testing.



Screening for the novel titin gene mutation associated with LGMD2J.
a) Relationships and presence/absence of LGMD2J in the family under study. N and M indicated, respectively, the normal and mutated titin gene. Panels b-d, partial sequence of the titin gene from different individuals showing the region of the novel mutation. b) no mutation, c) heterozygous variant, d) homozygous variant.

EXAMPLES OF PUBLICATIONS USING NOVOGENE'S EXPERTISE

Year	Journal	Article
2018	Nature Communications	Whole-Exome Sequencing Reveals the Origin and Evolution of Hepato-Cholangiocarcinoma
2018	Cancer Research	Multiregion Sequencing Reveals the Genetic Heterogeneity and Evolutionary History of Osteosarcoma and Matched Pulmonary Metastases
2017	Gastroenterology	Genetic Alterations in Esophageal Tissues From Squamous Dysplasia to Carcinoma
2017	Nature Communications	Simultaneous Evolutionary Expansion and Constraint of Genomic Heterogeneity in Multifocal Lung Cancer
2017	Cancer Research	Clonality, Heterogeneity and Evolution of Synchronous Bilateral Ovarian Cancer
2016	American Journal of Human Genetics	Biallelic SUN5 Mutations Cause Autosomal-Recessive Acephalic Spermatozoa Syndrome
2016	Cell Research	Single-Cell Exome Sequencing Identifies Mutations in KCP, LOC440040, and LOC440563 as Drivers in Renal Cell Carcinoma Stem Cells