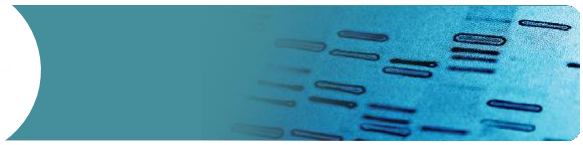
MEDIKONIA
仁 邁 樂 康
Genomic Services & Solutions



Whole Human Genome Sequencing

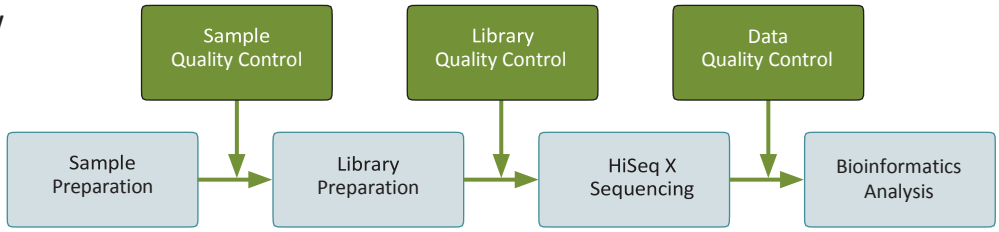
Human whole genome sequencing enables researchers to catalog the genetic constitution of individuals and capture all the variants present in a single assay. It is applied to the study of cancer and a variety of diseases, as well as human population evolution studies and pharmacogenomics.

Medikonia is one of the first few companies in the world based on the powerful Illumina HiSeq X Ten, capable of sequencing up to 18,000 human genomes per year at the lowest cost per genome. We have extensive experience providing whole genome sequencing service on this powerful system, having successfully sequenced thousands of genomes with high quality results. With the throughput and capacity of the HiSeq X Ten, our deep experience with the system, and our advanced bioinformatics capabilities, Medikonia is able to expertly meet customer needs for executing large projects with timely turn-around and the highest quality results.

The Advantages

- State-of-the-art NGS technologies: Medikonia uses state-of-the-art technology, including the latest generation Illumina HiSeq X Ten.
- Highest data quality: We guarantee a Q30 score ≥ 80%, exceeding Illumina's official guarantee of ≥ 75%.
- Extraordinary informatics expertise: Medikonia uses its cutting-edge bioinformatics pipeline and internationally recognized best-in-class software to provide customers highly reliable "publication-ready data".

Project Workflow



SEQUENCING STRATEGY

- 350 bp insert DNA library
- HiSeq X platform, paired-end 150 bp

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

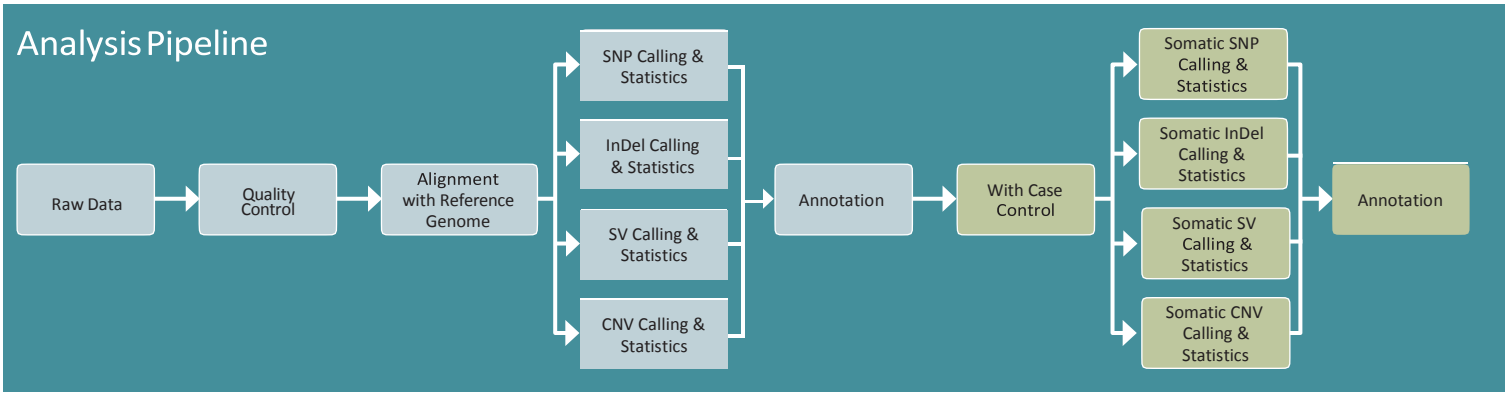
- 15 working days after verification of sample quality (without data analysis)
- Additional 8 working days for data analysis

RECOMMENDED SEQUENCING DEPTH

- For tumor tissues: 50×, adjacent normal tissues and blood 30×
- For rare diseases: 30~50×

SAMPLE REQUIREMENTS

- Input DNA:
  - For fresh sample: ≥ 1.0 µg (a minimum of 200 ng can be accepted with risk)
  - For FFPE sample: ≥ 1.5 µg
- DNA concentration: ≥ 20 ng/µl
- DNA Volume: ≥ 10 µl
- Purity: OD260/280 = 1.8-2.0 without degradation or RNA contamination



BIOINFORMATICS ANALYSIS INCLUDES:

- Data quality control: filtering out reads containing adapters or with low quality
- Alignment with reference genome, statistics of sequencing depth and coverage
- SNP/InDel/SV/CNV calling, annotation and statistics
- Somatic SNP/InDel/SV/CNV calling, annotation and statistics (paired tumor samples)

Medikonia Data

Medikonia provides the highest quality NGS services. We guarantee that over 80% of bases will have a sequencing quality score ≥ Q30. In standard practice, Medikonia achieves an average Q30 of 87.89% for WGS, exceeding Illumina's official guarantee of 75%. Additionally, an average of 98.4% of our raw sequencing data passes the quality control standards for effective clean data.

The following table includes actual Medikonia data from human genome sequencing projects, and demonstrates the quality of our sequencing. Alignment of the results to the reference genome (UCSC hg19) showed an average mapping ratio of 99.44%.

REPRESENTATIVE WHOLE EXOME SEQUENCING DATA FROM MEDIKONIA

Sample Name	Raw Data (Gb)[1]	Effective (%)[2]	Error (%)[3]	Q20 (%)[4]	Q30 (%)[5]	GC (%)[6]	Mapped (%)[7]	Average Sequencing Depth (x)[8]	Genome Coverage[9]	Percentage of Genome with ≥ 4x coverage[10]	Percentage of Genome with ≥ 10x coverage[11]	Percentage of Genome with ≥ 20x coverage[12]
Novo 1	98.35	99.08%	0.04%	95.42%	88.97%	42.64%	99.82%	30.37	99.70%	99.11%	95.49%	78.45%
Novo 2	94.02	99.72%	0.03%	96.83%	91.14%	42.69%	99.76%	28.37	99.67%	99.00%	94.77%	74.43%
Novo 3	96.21	99.92%	0.03%	97.60%	92.57%	41.20%	99.68%	27.97	98.90%	98.32%	95.65%	81.02%
Novo 4	99.50	99.53%	0.03%	95.76%	89.77%	40.49%	99.67%	29.30	99.76%	99.45%	97.37%	83.66%
Novo 5	95.27	99.95%	0.03%	92.26%	90.30%	40.48%	99.74%	28.28	99.67%	99.17%	96.29%	80.92%
Novo 12	95.24	99.90%	0.03%	97.68%	93.00%	41.75%	99.89%	28.97	99.65%	99.10%	95.95%	79.37%
Novo 13	110.90	99.68%	0.03%	95.88%	89.93%	40.40%	99.61%	32.94	99.77%	99.51%	97.90%	87.23%
Novo 14	92.89	98.85%	0.05%	93.04%	85.48%	40.95%	99.54%	26.72	99.79%	99.54%	97.56%	79.68%
Novo 15	107.26	97.90%	0.06%	91.81%	83.74%	42.62%	99.43%	31.70	99.74%	99.36%	97.29%	84.21%
Novo 16	161.17	99.69%	0.04%	95.02%	87.93%	41.99%	99.81%	48.49	99.06%	98.77%	97.91%	93.85%
Novo 17	173.39	97.11%	0.03%	93.49%	87.30%	42.28%	99.12%	49.62	99.78%	99.45%	98.27%	94.22%
Novo 18	160.21	98.46%	0.05%	93.26%	85.73%	41.37%	99.64%	46.85	99.79%	99.60%	98.91%	94.44%
Novo 19	153.97	98.04%	0.04%	93.64%	87.27%	42.00%	99.48%	44.48	99.78%	99.58%	98.92%	94.82%
Novo 20	161.17	99.41%	0.03%	95.57%	89.24%	42.89%	98.64%	47.86	99.02%	98.72%	97.75%	93.27%
Novo 21	152.27	96.16%	0.04%	92.29%	85.85%	43.02%	99.01%	43.45	99.77%	99.53%	98.47%	91.79%
Novo 22	156.40	97.30%	0.04%	93.15%	87.05%	43.60%	99.09%	44.87	99.73%	99.33%	97.43%	89.33%
Novo 23	168.98	98.56%	0.03%	95.46%	89.32%	43.37%	99.77%	50.36	99.74%	99.29%	97.44%	90.59%
Novo 24	153.61	95.29%	0.05%	91.93%	84.63%	43.50%	99.36%	42.94	99.77%	99.44%	97.76%	89.72%
Novo 25	152.13	95.45%	0.08%	90.28%	81.78%	43.83%	99.34%	42.72	99.73%	99.28%	97.10%	87.37%
Novo 26	156.35	99.59%	0.03%	96.95%	91.69%	42.71%	99.90%	47.28	99.75%	99.44%	98.13%	92.12%
Novo 27	155.05	99.58%	0.03%	97.15%	92.12%	43.23%	99.84%	46.44	99.72%	99.37%	97.89%	91.14%
Novo 28	156.54	98.29%	0.03%	95.04%	89.26%	42.64%	99.59%	46.06	99.75%	99.52%	98.65%	93.59%
Novo 29	176.26	97.35%	0.07%	91.45%	83.18%	43.35%	99.41%	49.56	99.77%	99.54%	98.67%	94.14%
Novo 30	152.38	95.57%	0.06%	91.33%	83.81%	43.63%	99.28%	43.49	99.76%	99.44%	98.00%	91.02%
Novo 31	180.69	96.45%	0.07%	91.24%	83.07%	41.72%	99.42%	50.54	98.81%	99.65%	99.08%	96.22%
Novo 32	96.51	99.82%	0.05%	93.05%	85.87%	42.56%	98.49%	32.16	99.73%	99.08%	94.74%	75.55%
Novo 33	95.85	99.82%	0.04%	93.24%	86.20%	42.55%	99.87%	31.97	99.74%	99.16%	95.30%	77.47%

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 (UCSC hg19).
8 Average sequencing depth.
9 Percentage of genome covered by sequencing.
10 Percentage of bases in genome with a sequencing depth ≥ 4x.
11 Percentage of bases in genome with a sequencing depth ≥ 10x.
12 Percentage of bases in genome with a sequencing depth ≥ 20x

EXAMPLES OF PUBLICATIONS

Journal	Title
2016-Scientific Reports	Elimination of HIV-1 Genomes from Human T-lymphoid Cells by CRISPR/Cas9 Gene Editing
2016-European Journal of Human Genetics	Genome-wide linkage analysis and whole-genome sequencing identify a recurrent SMARCAD1 variant in a unique Chinese family with Basan syndrome.