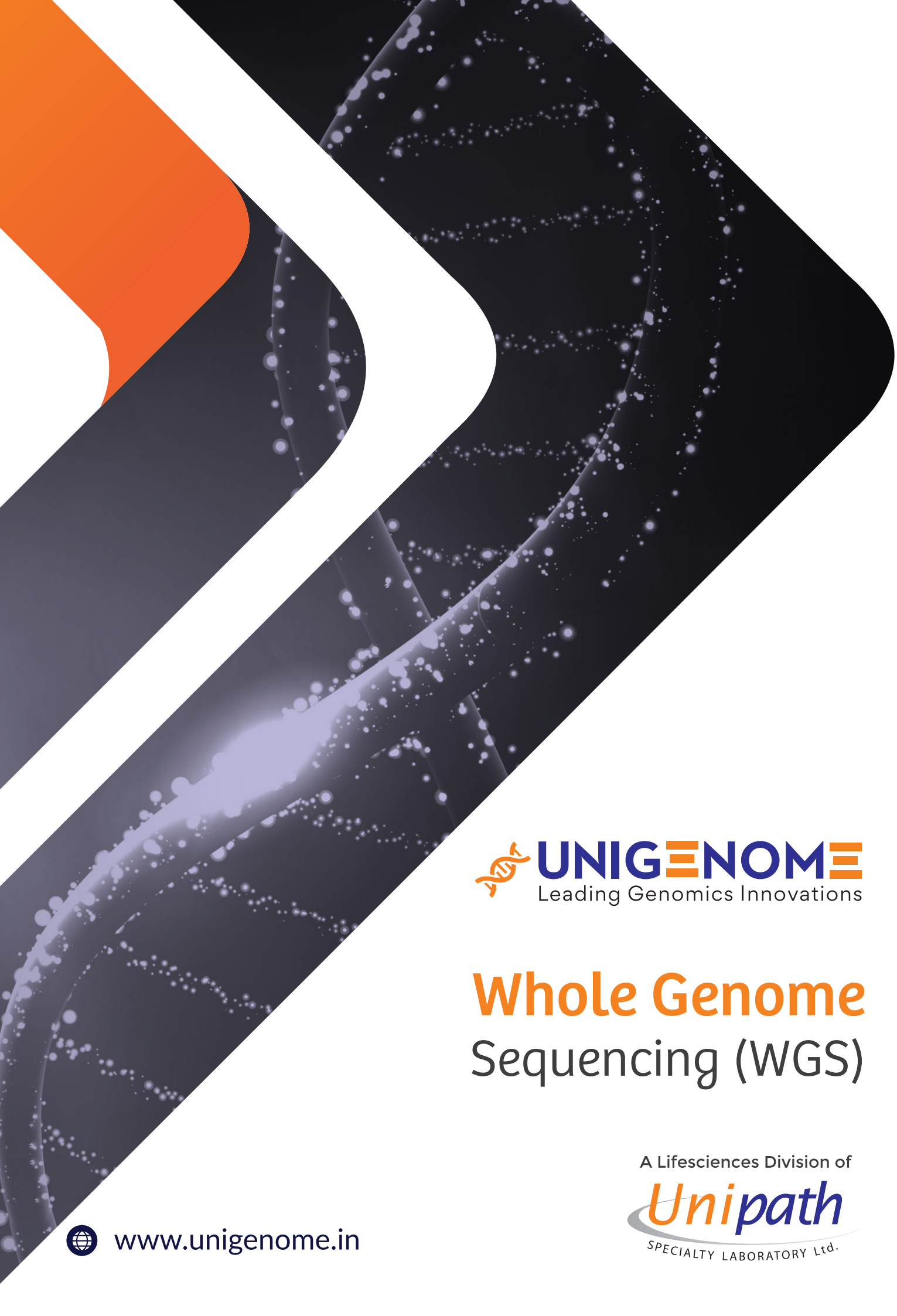


UNIGENOME
Leading Genomics Innovations

Whole Genome Sequencing (WGS)

A Lifesciences Division of
Unipath
SPECIALTY LABORATORY Ltd.

www.unigenome.in

Whole Genome Sequencing (WGS)

Background

- The term 'genomics' was first coined in 1986 by Dr. Thomas Roderick, a geneticist in Bar Harbour, Maine, and was initially intended as a term to encompass the study and comparison of genomes of various species, including their evolution and relationships. Essentially, genomics involves the application of DNA sequencing and the subsequent analyses using in vitro experiments and bioinformatics approaches to study the structure and function of genes for all organisms like humans, plants, pathogens, animals etc. Genome sequencing uses combined approaches like mate-pair, paired end, shotgun, etc. Shotgun and paired end sequencing are useful for small genomes like bacterial genomes, viral genomes etc whereas for large plant genomes hybrid approach of mate paired, pair end and long read sequencing is also required. WGS can be used for studying genic and structural variations like, Single Nucleotide Polymorphisms (SNPs), Copy Number Variations (CNVs), insertions, deletions, gene ontology, pathway analysis and other genome-wide association studies (GWAS) that can affect gene regulatory pathways and signaling networks.

- Next generation sequencing technology has significantly improved the capacity to perform low-cost, efficient WGS, and has made it a feasible tool to enhance clinical diagnostic investigations in near real-time. Next-generation processes generally involve parallel sequencing, producing vast quantities of data that require modern computation methods to assemble the sequence reads.

- Unigenome uses high throughput sequencing platform Novaseq 6000 for sequencing projects. The NovaSeq 6000 System brings a new era in sequencing with groundbreaking innovations to provide users with the throughput, speed, and flexibility required to complete projects faster and more economically than ever before. Combining the best features of previous Illumina platforms, the Novaseq 6000 System incorporates additional innovations to deliver tunable output of up to 6 Tb and 20 billion reads—all in about 2 days. The Novaseq 6000 System leverages proven Illumina sequencing by synthesis (SBS) chemistry—the most widely adopted next-generation sequencing (NGS) technology. SBS chemistry delivers exceptional data accuracy, the highest yield of error-free reads, and the highest percentage of base calls above Q30 in the industry. At Unigenome, denovo and reference based sequencing using multiple set of libraries like Paired End and Mate Paired libraries is carried out using various NGS technologies. For non-model organism denovo genome assembly is performed on High Performing Computing Cluster (HPCC) using latest available assembly tools to assemble genome into contigs and scaffolds.

- Further, pipeline is used to reduce N's in assembled scaffolds by using the paired reads. Comprehensive annotation pipeline integrated with various functional databases like NCBI nr, uniprotKB, Pfamand COG databases is used to decipher the genome. For model organism reference based sequencing is perform on HPCC using DNA aligner to map reads on reference genome , variant discovery and its annotation.

Sample requirement for WGS

- We accept genomic DNA (gDNA), microbial cultures, plant tissues, animal/human tissue etc.

Genomic DNA sample

- gDNA should be of high molecular weight intact double stranded genomic DNA, free from RNA contamination with quantity as mentioned below:
  - For PE libraries, 50 µg of gDNA
  - For Mate paired library, 100-200 µg of gDNA
- gDNA should have an absorbance ratio (A260/280) of ~1.8 to 2.0 with minimum 300-500ng/µL concentration.

Quality control of gDNA samples

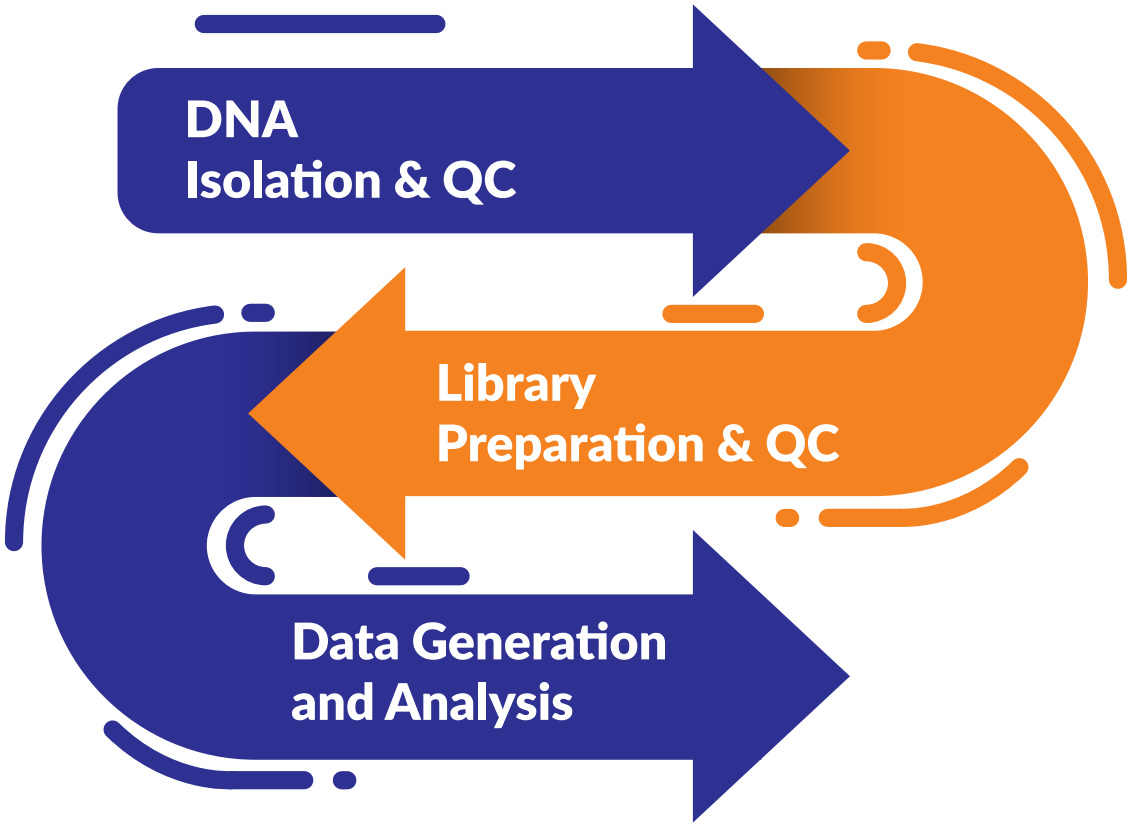
- gDNA samples will be subjected to both qualification and quantification by 1% agarose gel electrophoresis & Qubit/Nanodrop respectively.

Sample type and requirements

- gDNA samples will be subjected to both qualification and quantification by 1% agarose gel electrophoresis & Qubit/Nanodrop respectively.
  - **Microbial culture:** Cell Pellets are preferred. Glycerol stocks or culture plate of pure isolates can also be sent.
  - **Plant tissues:** (seedling sample, leaf, stem, flower, fruit, grain etc)
  - Minimum 3-5 g of specimen sample should be provided in -20^{o}C.
  - **Human Tissue sample:** FFPE block with at-least 30% Cellularity of cancer cells. Liquid biopsy sample should be transferred in normal saline or 1X PBS solution in 1.5ml nuclease free tubes.
  - **Human/Animal Mitochondrial Genome Sequencing:** 5-10ml of K2-EDTA blood
  - **Mitochondrial/Chloroplast Genome Sequencing:** Isolated Mitochondrial DNA or Chloroplast DNA

Note: All types of samples should be transported in cool packs to Unigenome Ahmedabad, Gujarat, India.

Workflow of Whole Genome Sequencing



Bioinformatics Pipeline of Genome sequencing Deliverables

Whole Genome Resequencing/Reference based analysis on Illumina Platform (Eukaryotes/Prokaryotes/Virus)

- Clean reads after filtration of adapter and low quality bases
- Alignment file obtained after mapping of reads to reference genome
- PCR duplication removal from alignment file
- Coverage analysis and its distribution
- Variant calling (SNPs and INDELS)
- Variant filtration
- Graphical representation of genome wide distribution of variant
- Variant annotation
- Comprehensiveness compiled report and data deliverables

De novo Genome Sequencing of bacterial genome & Analysis On Illumina Platform

- Clean reads after filtration of adapter and low quality bases
- Genome assembly stats and summary
- tRNA prediction
- rRNA prediction
- SSR prediction
- Coding gene prediction
- Gene annotation against NR, Uniprot, COG and Pfam
- Gene ontology (GO) of annotated genes
- Gene KEGG pathway analysis
- Antibiotic resistance genes identification
- Comprehensiveness compiled report and data deliverables

Denovo Mitochondrial DNA Sequencing and analysis on Illumina Platform

- Clean reads after filtration of adapter and low quality bases
- De novo mitochondrial genome assembly
- A,T,G, C content analysis
- Protein coding genes prediction
- rRNAs genes prediction
- tRNAs genes prediction
- Probable D-loop identification
- Circular representation of mitochondrial genome
- Comprehensiveness compiled report and data deliverables

Denovo Whole Genome Sequencing of Bacteriophage

- Clean reads after filtration of adapter and low quality bases
- De novo genome assembly
- Gene prediction of the phage
- Functional annotation
- Comprehensiveness compiled report and data deliverables

Denovo Whole Genome Sequencing of Fungus on Illumina Platform

- Clean reads after filtration of adapter and low quality bases
- Genome assembly stats and summary
- tRNA prediction
- rRNA prediction
- SSR prediction
- Coding gene prediction
- Gene annotation against NR, Uniprot, COG and Pfam
- Gene ontology (GO) of annotated genes
- Gene KEGG pathway analysis
- Antibiotic resistance genes identification
- Comprehensiveness compiled report and data deliverables

Denovo Whole Genome Sequencing of Bacteriophage

- Clean reads after filtration of adapter and low quality bases
- De novo genome assembly
- Gene prediction of the phage
- Functional annotation
- Comprehensiveness compiled report and data deliverables

Denovo Whole Genome Sequencing of Fungus on Illumina Platform

- Clean reads after filtration of adapter and low quality bases
- Genome assembly stats
  - Total number of scaffolds
  - Total number of contigs
  - Average scaffolds size
  - Maximum scaffold size
  - Minimum scaffold size
  - Percentage of A, T, G, C and N in assembly
  - Scaffold N50
- tRNA prediction
- rRNA prediction

- SSR prediction
- Transposon identification
- Coding gene prediction
- Gene annotation against NR, Uniprot, KOG and Pfam
- Gene ontology (GO) of annotated genes
- Gene KEGG pathway analysis
- Antibiotic resistance genes identification
- Non-coding RNA prediction
- Comprehensiveness compiled report and data deliverables

Denovo Whole viral genome analysis (DNA virus and RNA virus)

- Clean reads after filtration of adapter and low quality bases
- Denovo assembly of reads to generate scaffolds
- Gene prediction from scaffolds
- Functional annotation of predicted genes
- Comprehensiveness compiled report and data deliverables

Note: Customized analysis as per client's need is provided for any of the above mentioned services with additional charges

UNIGENOME
Leading Genomics Innovations

A Lifesciences Division of
Unipath
SPECIALTY LABORATORY Ltd.

UNIGENOME
Unipath House Behind Sahajanand College,
Opposite Kamdhenu Complex, Panjarapole,
Ambawadi, Ahmedabad-380015

www.unigenome.in
info@unigenome.in