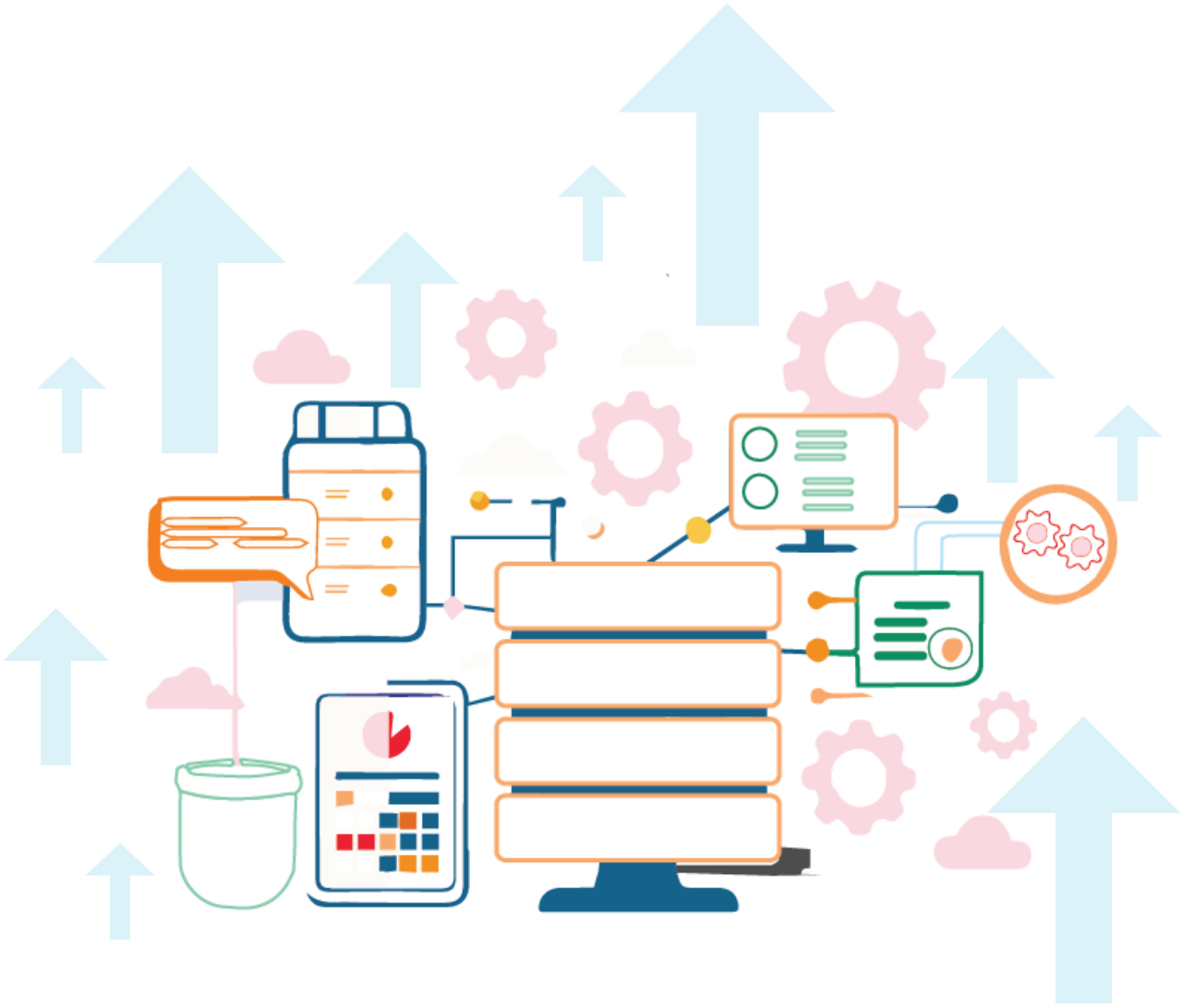
strand

Tools Enabling Data Scaling & Analytics Platforms



#1 Automation Of Time-Consuming WES Steps & Use Of LLMs *In-house solution - festiVAR, addressing variant prioritization at scale*

The Problem

- The Need: Increasing WES workloads in a CAP-accredited laboratory requires an efficient solution for variant prioritization in rare disease diagnosis.
    - Previous variant interpretation algorithms resulted in a high number of genes per case for manual assessment, slowing down the reporting process.

Our Process

- Implemented festiVAR, a variant prioritization algorithm, to streamline the identification of key variants from ~40,000 SNVs, small InDels, and CNVs.
- Enhanced festiVAR with a GPT-3.5 Turbo LLM-based approach to further refine the top 25 genes using free-text clinical notes, significantly improving accuracy in gene-phenotype correlation.

Result / Impact

- festiVAR successfully prioritized all reported variants within the top 25 genes in 99.93% of cases, significantly reducing manual assessment workload.
- The enhanced workflow now enables efficient processing of hundreds of exomes monthly, with automated ACMG criteria assessment and LLM-based literature search integration.

festiVAR is now in production, and accompanied by a user interface, enabling our teams of variant interpreters to process hundreds of exomes every month

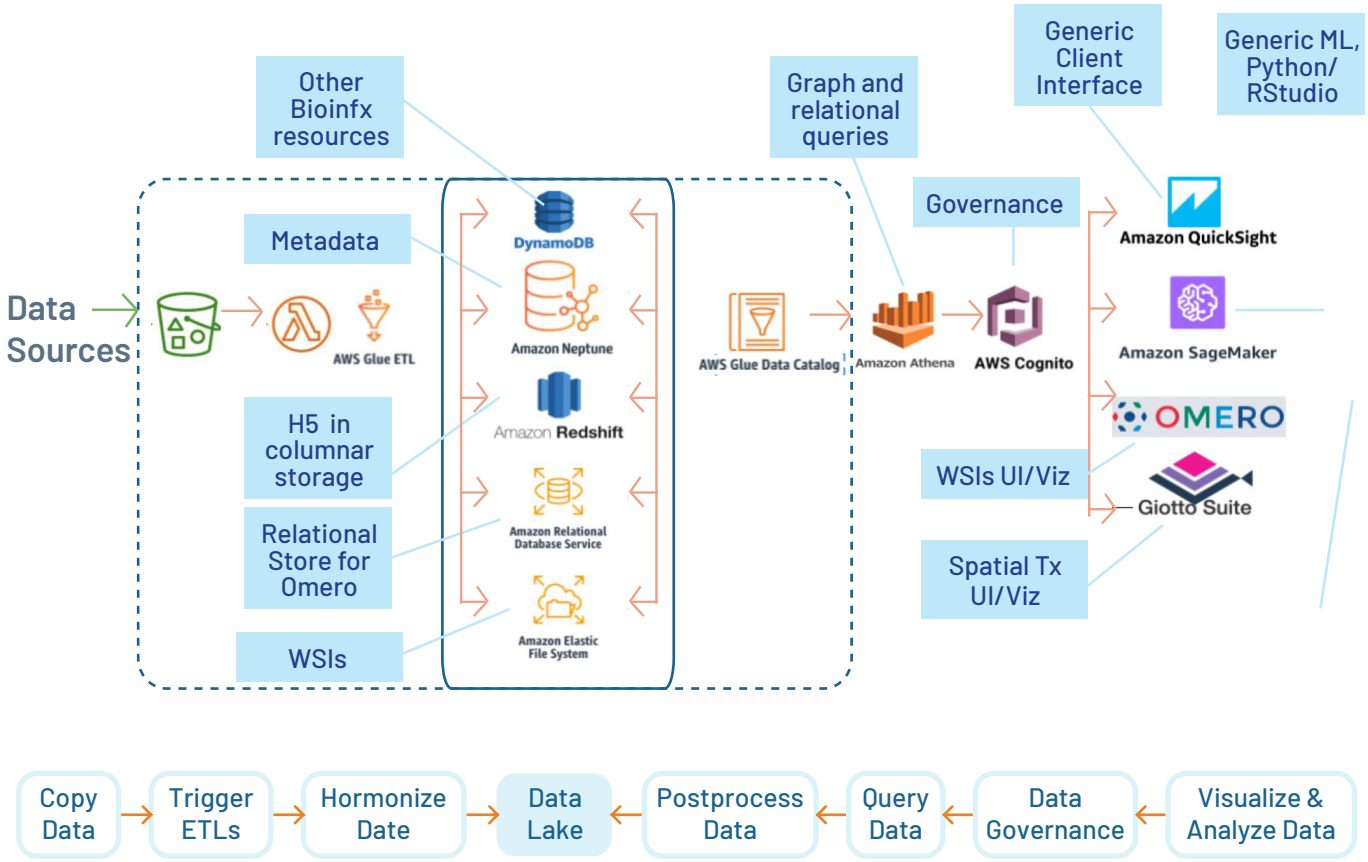
#2 Solution Architecture For An Integrative Data Model

For a large pharmaceutical company

The Problem

- The Need: To integrate data from internal and external sources and ensure the data is FAIR (Findable, Accessible, Interoperable and Reusable) for downstream machine learning and visualization applications.

Our Solution Blueprint



The solution architecture is being developed to enable ETL of all metadata, data, and results from various sources for efficient querying, visualization, and analysis.

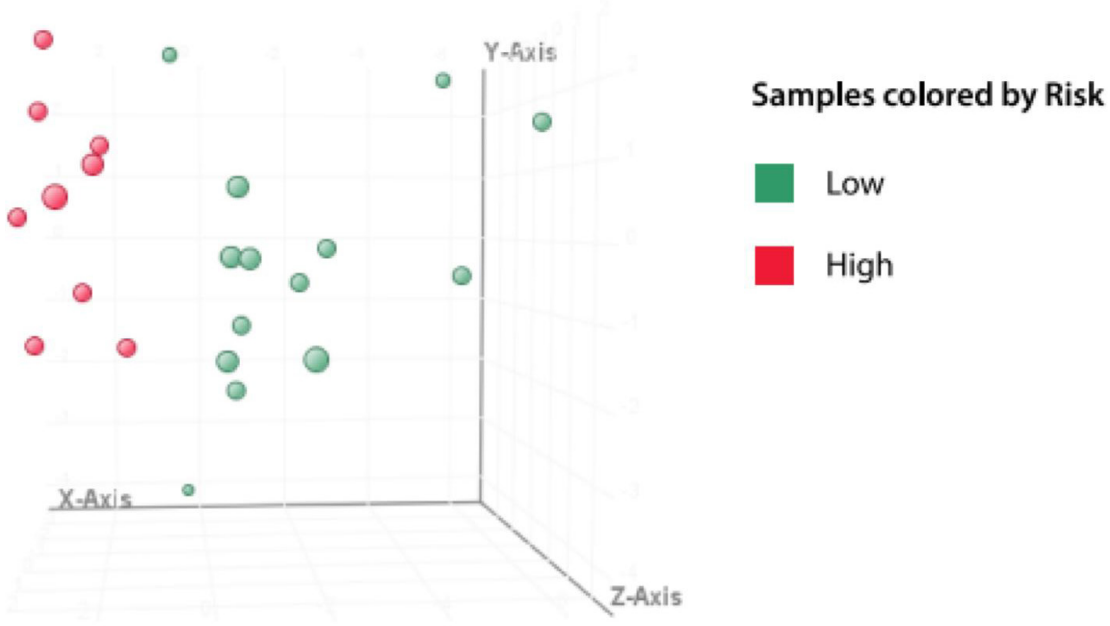
#3 Predictive Biomarkers For Advanced Heart Failure Using StrandNGS

The Problem

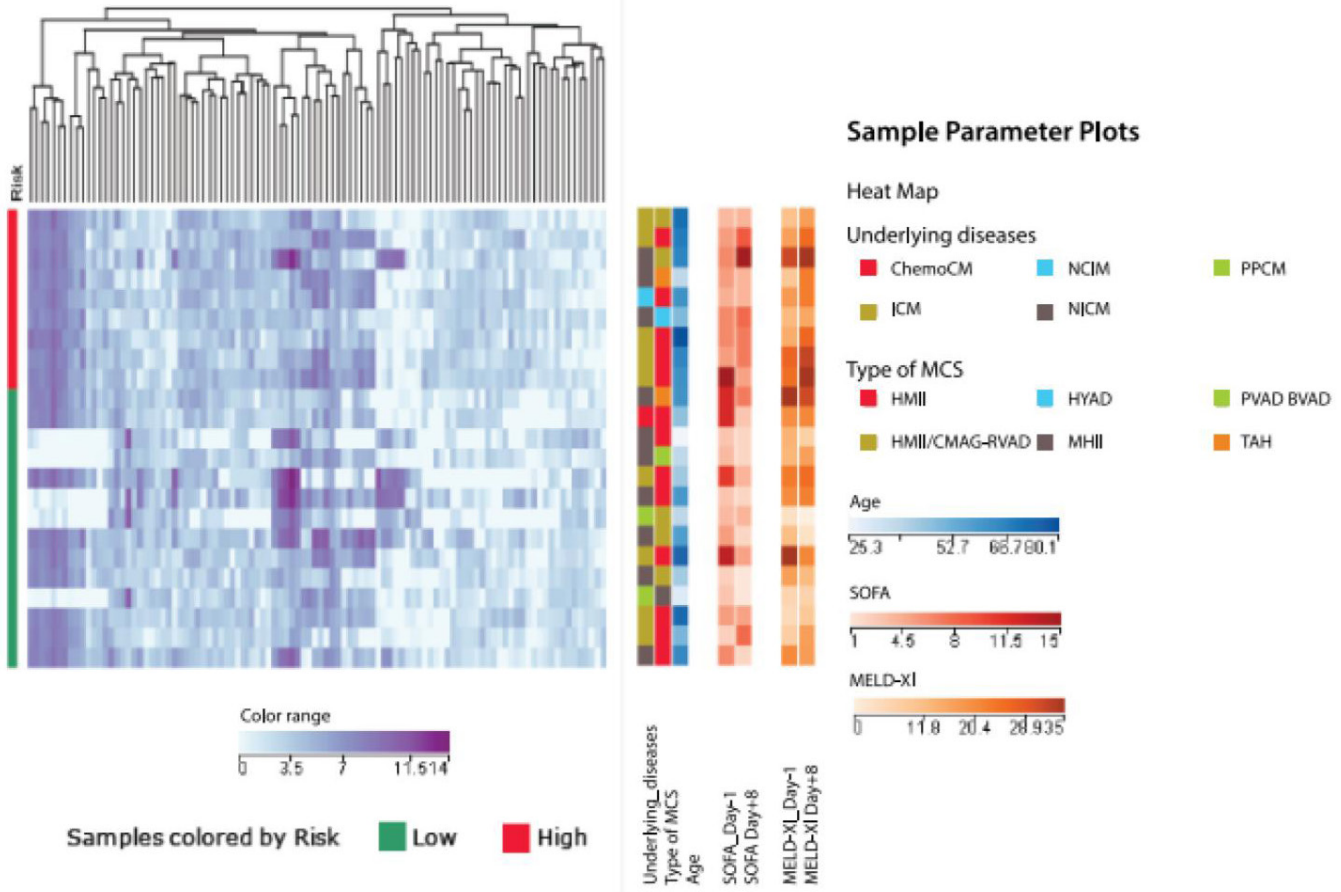
- The Need: Biomarkers for Predicting Postoperative Outcomes in Advanced Heart Failure(AHF) patients to develop a reliable test based on patient-derived data.
    - Transcriptome Profiling: Gene expression profiles of AHF patients
    - Multi-Dimensional Patient Data Analysis: Integrate transcriptome, cytokine, and flow cytometry data to gain a comprehensive understanding of patient profiles

Our Solution

- Transcriptome profiling, cytokine and cytometry analysis data from 29 patient PBMC samples
- Imported into Strand NGS —> Optimal data transformation and normalization
- Quality assessments and filters—> High quality data points—> Expression profiles —> Statistical and fold change analysis
- Class prediction models —> Combinatorial model across platforms



Using the identified list of significantly differentially expressed genes
the patients were successfully stratified as high vs low risk as seen in this PCA plot



Gene expression profiles in preoperative samples from low and high risk patients correlates with their disease information and organ dysfunction

strandngs

Strand NGS provides all tools that would be prerequisites for mining sequencing data to build reliable class prediction models that are suited for varied machine learning goals.

#4 Molecular Subtypes Of Glioblastoma Multiforme Using Genespring

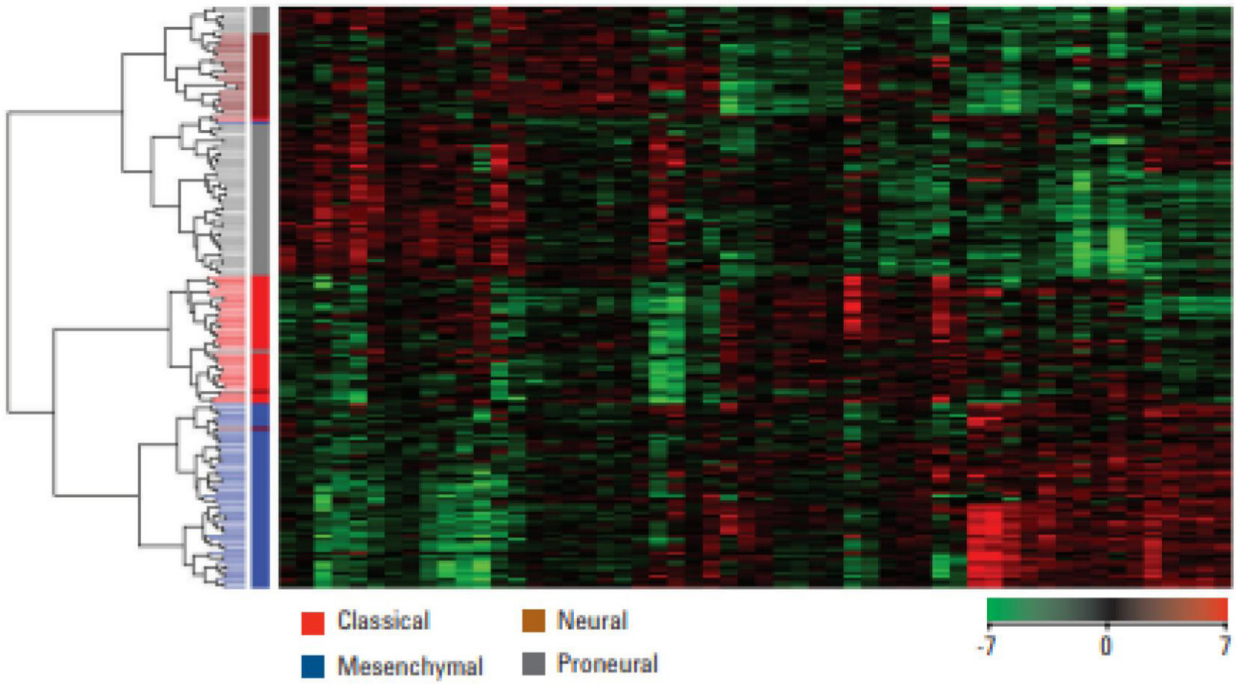
The Problem

- The Need: Biomarkers to classify molecular subtypes of glioblastoma multiforme (GBM).
  - Identify biomarkers leveraging TCGA and patient-derived data
  - Expand TCGA expression data
  - Profile patient proteomic expression
  - Combine transcriptomics and proteomics profiles

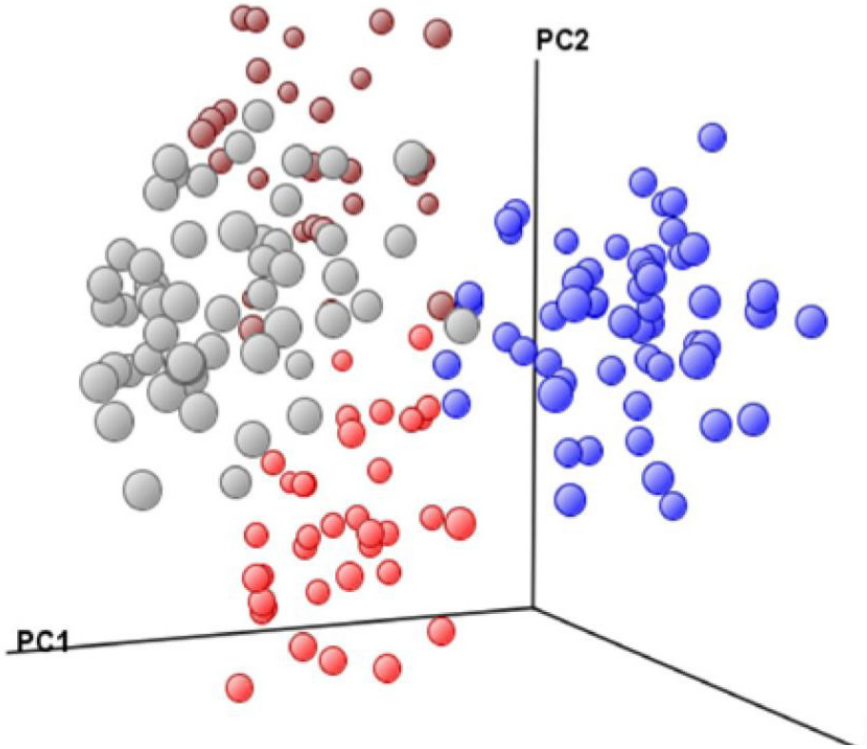
Our Solution

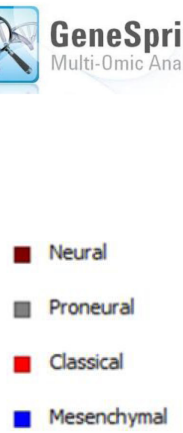
- mRNA/miRNA/proteomic data from multiple array platforms was imported and incorporated into a unified dataset
- Copy number values and other sample parameters were imported as metadata.
- Using the GeneSpring metadata framework batch effects and quality parameters were assessed
- Combined mRNA and proteomics analysis identified GBM subtype signatures that differentiate known subtypes within the larger TCGA cohort
- Clear separation between control and tumor samples, with tumor samples further segregating into subgroups through sample correlation and PCA analysis

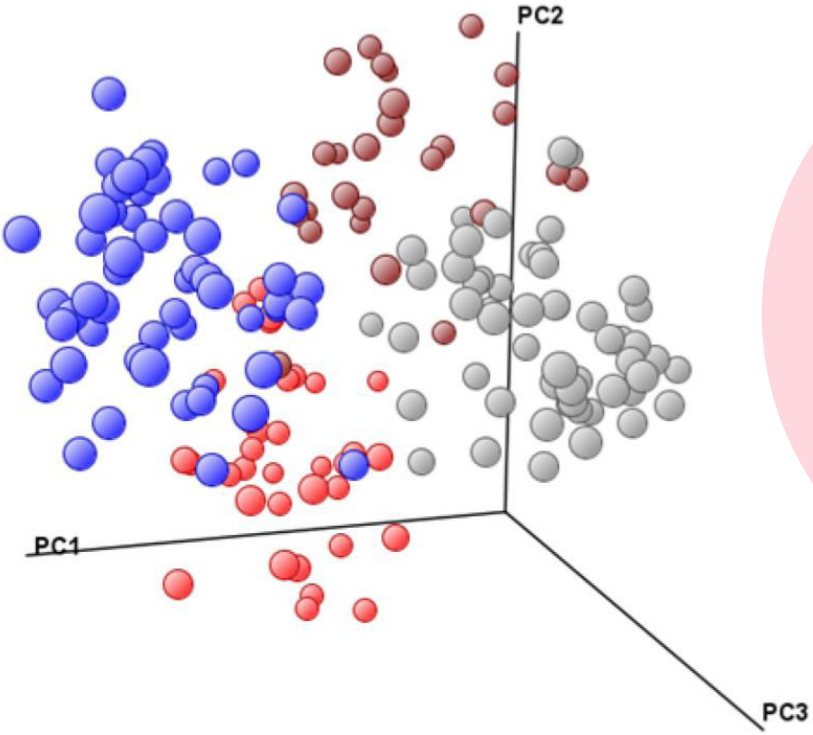
Reported molecular subtypes in GBM based on the expression of 54-gene signatures



Classification of subtypes







GeneSpring/MPP software offers a suite of powerful data analysis tools enabling a thorough investigation of complex biological datasets

Omics CRO

Curation

15 years of experience curating variants, genes, pathways and diseases for clinical reporting and pharma/biotech custom solutions

~50
Molecular Biologists

Bioinformatics and Software

22 years of experience providing bioinformatics solutions to global instrument, diagnostic and pharma companies

~220
SW Engineers, Bioinformaticians

Omics Assays

11 years of experience with sequencing-based diagnostics across oncology and genetics, at our CAP lab in India

~90
Lab Scientists, Clin. Res. Scientists

24+ YEARS OF EXPERIENCE

80,000+ Genetic Tests Reported

500+ Projects Executed for Genomics Majors Globally

Presence in **20+** Countries

strand

7th Floor, MSR North Tower, #144, Outer Ring Road, Nagavara, Bengaluru - 560045

+91 9980448044 hello@strandls.com strandls.com