Genomic Research

A major medical-scientific Research Center in the Middle East that collaborates internationally with the bio-tech and pharmaceutical industries to develop new drugs, treatments and technologies, needed a new-era of big data analytics solution that could tackle speed- and scalablity related barriers it was continuously challenged with. The Research Center, also a university-affiliated tertiary referral hospital, serves as a national Medical Center in many fields.

The Challenge

The Research Center was working with a file base, running each query of specific chromosome position sample separately, followed by a long complex manual process of comparisons between the findings, with zero flexibility. The final stage involved gathering the massive quantities of collected data, and attaining meaningful insights. The process was a highly time- and resource consuming effort, lasting weeks, months and even years - depending on the size of the data. The Research Center was looking for a more efficient, quick and cost-effective way of loading and analyzing massive quantities of genomic data.

The Solution

SQream provided the Research Center with a first-time ever presented ability: effective data querying of massive-scaling genomic data (BAM/VCF/fasta/fastq files etc.), performed automatically on a database, as opposed to the prior separate, manual process that was used on a file base. With SQream, the Research Center was able to upload 50 BAM files with 30 Billion rows of sequenced data, and run queries on all samples simultaneously - under 1 minute.

Use cases

- Distribution of nucleotides on specific chromosome position from all samples
- Coverage of reads in a specific range
- SNP and InDel analytics
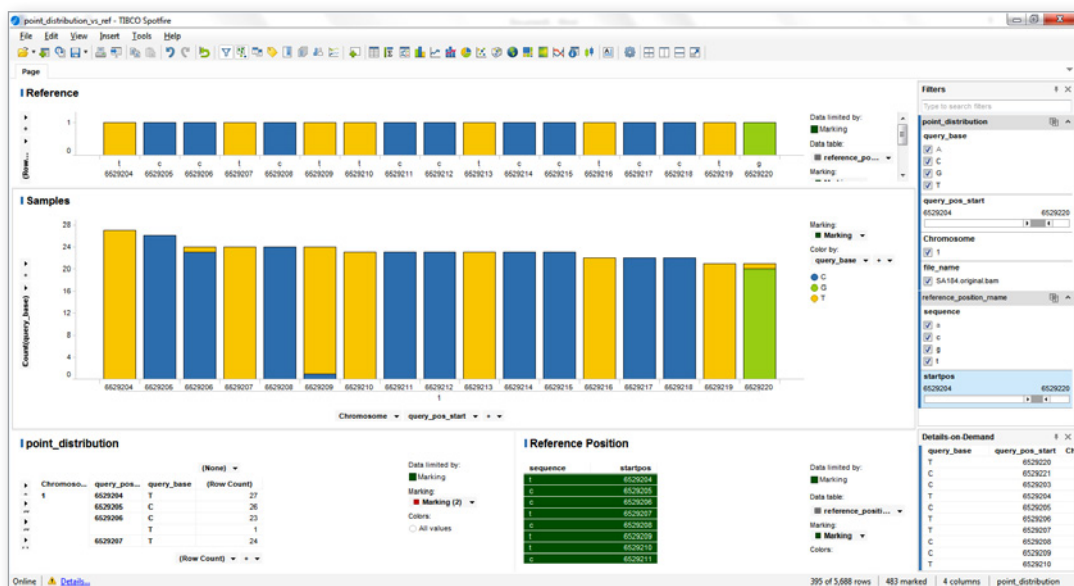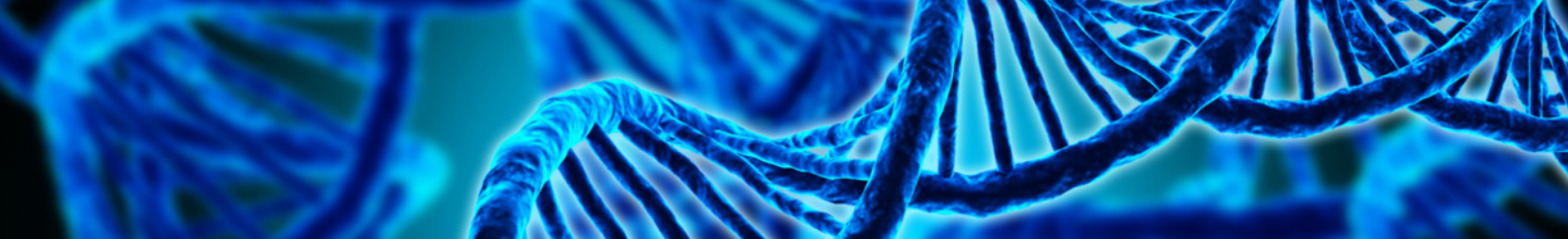
SQREAM
TECHNOLOGIES

Figure 1: The reference genome is displayed on top and the raw data below. Bottom left side: raw data for samples. Bottom right side: raw data for reference

Figure 2: Reads density per chromosome/position/file. Drill down to single read with flag data

Key Benefits

- Significant reduction of research time
- Quicker accomplishments of research phases (measured in hundreds of %)
- Major savings in resources
- Ability to focus on research objectives as opposed to IT

www.sqream.com