GenomeBoost Server™

Genome4me
Pioneering AI-Powered Cancer Genomics

Accelerating Crucial Discovery in Genome Research and Medicine

As whole genome sequencing (WGS) emerges as a standard technology in genome research and medicine, our **GenomeBoost Server™ (GBS™)** is designed to empower researchers and clinicians with faster, more efficient WGS data analysis. By integrating cutting-edge hardware design with highly optimized software, **GBS™** ensures rapid processing of WGS data without compromising even the slightest accuracy. This innovative solution facilitates the identification of genetic variations and the discovery of critical insights that drive advancements in genome research and medicine.

GBS™ achieves an impressive processing time of 4.9 samples per hour for 30x WGS variant calling.

GBS™ produces 100% identical results to those of GATK™ best practices.

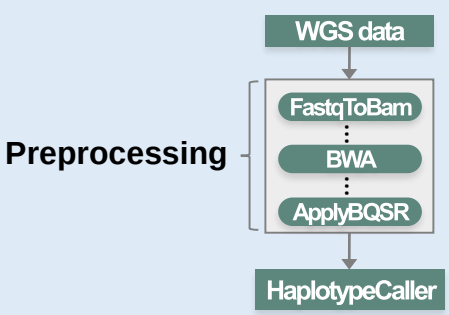
Processing Throughput (30x WGS)

Analysis pipelines	Preprocessing	Germline variant calling	Somatic variant calling	Throughput unit
16-sample batch run*	8	**12**	22	minutes/sample
	68,300	**43,100**	23,800	samples/year
Single-sample run	19	35	40	minutes/sample
	28,300	15,200	13,100	samples/year

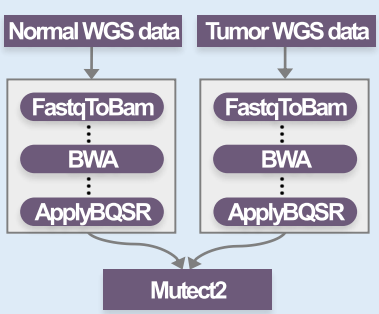
* 64 WGS samples were analyzed for preprocessing and germline calling, while 32 pairs of WGS samples were analyzed for somatic calling.

WGS Analysis Pipelines Supported by GBS™

Germline variant calling

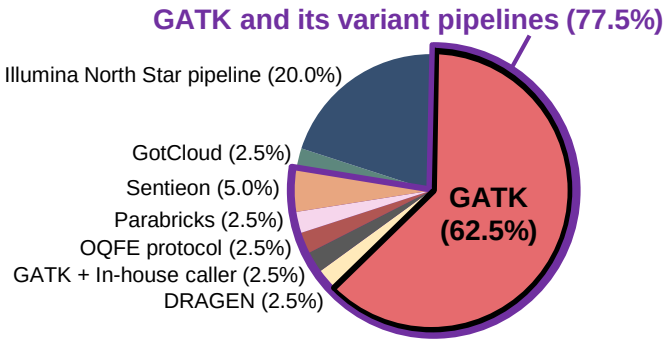


Somatic variant calling



Genome Analysis Tool Kit™ (GATK™) is the gold standard for WGS analysis

For dozens of years, GATK™ has been the gold standard software for genome data analysis. The vast majority of recently published studies* (77.5%) have employed the standard GATK™ pipeline (62.5%) or variant pipelines derived from GATK™ (15.0%).



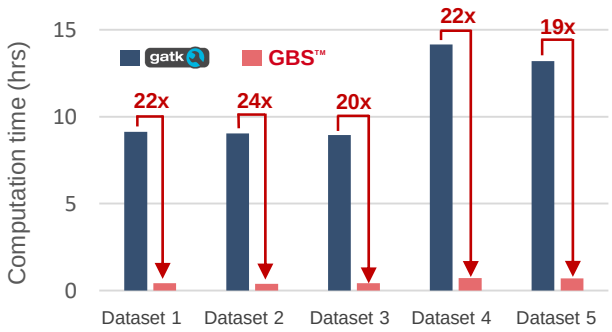
*Based on 40 WGS publications on high-profile journals published between 2015 and 2024.

GATK™ WARP reference implementation

Recently, the GATK™ team has noticed that many workflows referred to as GATK™ best practices diverge significantly from their recommendations. To reduce guesswork, GATK™ has unveiled WDL (a scripting language maintained by the OpenWDL community) Analysis Research Pipelines (WARP) as its reference implementations, designed specifically for human genome research with Illumina sequencers. GBS™ adheres strictly to the WARP reference implementation to ensure users experience the most accurate and consistent analysis.

Super-Fast Performance of GBS™

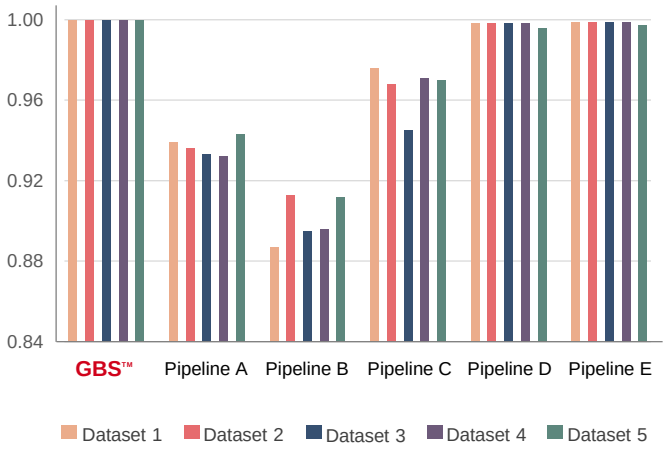
19-24 times faster than GATK™ for 30x WGS preprocessing



Most Accurate Analysis by GBS™

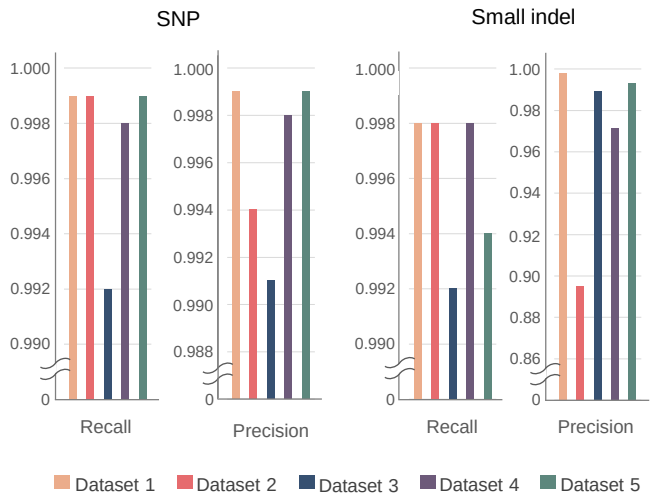
- 100% identical results to the GATK™ Best Practices WARP reference implementation.
- Competing pipelines (denoted as A, B, C, D, and E below) achieved fast analysis time but failed to maintain compatibility with GATK™.

Consistency with GATK™: Recall



- Even pipeline E, that has the highest accuracy except GBS™, showed discrepancy with GATK™ in all of the 155 WGS samples tested. While most samples showed high concordance with GATK™, some samples showed substantial discrepancies.

Consistency with GATK™: Pipeline E



GBS™ Enabling Seamless Scale-Up for Massive-Scale Discoveries

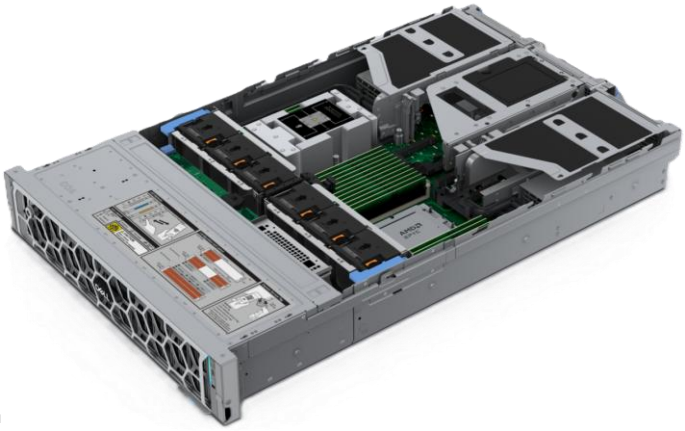
GBS™ offers unmatched scalability and performance for growing demands of genomic workloads with a flexible architecture that enables efficient multi-node clustering. Its scalable architecture supports the full spectrum of genomic research, from individual studies to national-scale initiatives, enabling faster breakthroughs in diagnostics and precision medicine.

- **High-Throughput Capacity**: Processes up to **43,100 whole genomes annually per node**, ensuring fast and efficient analysis.
- **Multi-Node Scalability**: Distribute workloads across nodes for efficient processing of large projects.
- **Flexible Deployment**: Scales easily from small labs to massive-scale facilities with multi-node solutions.

GBS™, The Ultimate Solution for Complete Genome Analysis

Equipped with cutting-edge AMD EPYC™ 9755 CPUs, ample main memory, and high-speed NVMe SSDs, GBS™ delivers exceptional performance for demanding workloads.

Hardware	Specification
Server	Dell PowerEdge™ R7725
Processor	2 x AMD EPYC™ 9755 CPUs (2x128 cores, 512M cache, DDR5)
Main Memory	1.5 TB RAM (24 x 64GB 6400 MT/s DDR5)
Storage	2 x 7.7 TB NVMe SSD



List of Pre-Installed Software in GBS™

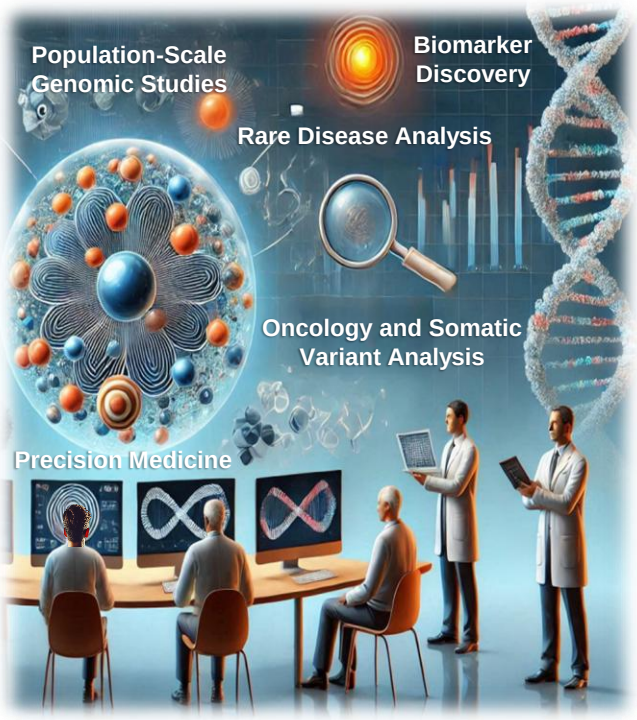
Software	Version	Description
GenomeBoost™	1.2	Optimized software derived from BWA-MEM, integrating various GATK™ tools
GenomeBoost™ Tool	1.2	An automated batch tool for running the GATK best practices with GenomeBoost™
GATK™	4.6.0.0	Genome Analysis Tool Kit
Picard	3.2.0	A set of command line tools for manipulating sequencing data and formats
BWA	0.7.18	A software package for mapping DNA sequences against a reference genome
samtools	1.20	A suite of utilities for interacting with high-throughput sequencing data
bcftools	1.20	A set of utilities that manipulate variant calls in VCF/BCF
Snakemake	8.29	Workflow management system
OpenPBS	23.06.06	A fast, powerful workload manager and job scheduler
OpenJDK	17.0.11	Open-source Java Development Kit
Python	3.12.3	Widely-used high-level programming language
Rocky Linux™	8.9	Open-source operating system compatible with Red Hat Enterprise Linux™

Accelerate, Streamline, and Scale Your Genome Analysis

- **High-Throughput Capacity**: Processes up to 43,100 whole genomes annually per node, enabling fast and efficient analysis also with GATK™-equivalent accuracy
- **Automated Pipelines**: Simplifies workflows with user-friendly automation, reducing manual intervention and human errors
- **Scalable Architecture**: Supports multi-node clusters for seamless workload distribution, handling projects of any size
- **Cost-Effective Solution**: Combines speed, accuracy, and scalability to deliver impactful genomic insights at reduced costs

Diverse Applications of GBS™

- **Precision Medicine**: Rapid analysis of individual patient genomes to support real-time clinical decision-making and personalized treatment plans
- **Oncology and Somatic Variant Analysis**: Detection of somatic mutations in tumor samples to guide targeted cancer therapies and monitor tumor progression
- **Biomarker Discovery**: Identifying genetic markers and mutations linked to diseases for diagnostic, therapeutic insights and drug development
- **Rare Disease Analysis**: Detecting rare genetic variants to support research and diagnosis of orphan diseases
- **Population-Scale Genomic Studies**: Efficient processing of large datasets for identifying genetic variations and understanding population-wide genetic patterns



GBS™, Accurate Ultra-Fast Genome Analysis Server

The Dell PowerEdge R7725 is a high-performance, dual-socket server offering exceptional scalability, energy efficiency, and advanced features to support data-intensive applications and virtualization.

The AMD EPYC 9755 is a high-performance processor with 128 cores and a clock speed up to 4.10 GHz, designed to deliver exceptional computational power.

GBS™ includes a 3-year warranty with dedicated support and service coverage from a highly qualified Dell partner, Dasan and Genome4me.

Last update: 10/13/2025

Contact: sales@genome4me.com