

# Benchmark Testing: Long-Read Assembly and Variant Calling of Eukaryotes Using Lasergene Genomics

Long-read datasets are computationally intensive, but Lasergene Genomics simplifies long-read analysis for data sets of any size, including complete human genomes.

Lasergene Genomics excels in **speed, accuracy, and ease of use**. Customizable parameters are automatically optimized for each sequencing technology: Illumina, PacBio SMRT, PacBio HiFi, Oxford Nanopore, or Oxford Nanopore “Duplex”. Select the variant calling algorithm you prefer, including GATK options. During the run, simultaneous alignment & variant calling—plus multicore parallelization across chromosomes—quickly and efficiently takes you from raw FASTQ reads to variant analysis results.

## SeqMan NGen v18.1 Benchmark Testing Results

We used SeqMan NGen v18.1, released September 2025, to perform reference-guided assemblies for multiple eukaryotic genomes. Three read technologies were tested: PacBio HiFi, ONT Duplex, and Illumina (Figure 1). Assembly times correlated strongly with input read volume for all platforms.

(Please see the reverse for our testing setup.)

Data Platform	Organism Name	Genome Length (Mbp)	Input reads (Gbp)	Assembly Time (min)
PacBio HiFi	<i>Aspergillus niger</i>	35	25.2	19
	<i>Arabidopsis</i>	119	14.8	32
	Flax	316	21.7	57
	Corn	2180	55.9	408
	Human	3200	112.0	372
ONT Duplex	Tomato	900	35.1	246
	Human	3200	132.0	516
Illumina	<i>Aspergillus fumigatus</i>	30	5.8	7
	Zebrafish	1400	47.0	120
	Human (HG002)	3200	189.0	414

Figure 1: SeqMan NGen 18.1 assembly times for various eukaryotic genomes using three sequencing data platforms.

# Benchmark Testing Setup

## Software:

- SeqMan NGen v18.1 - reference-guided workflow
- Default parameters were used except for the adjustable parameter "Minimum match percentage," which sets the overall alignment stringency. This value was set at 85-98% depending on how close the sequenced strain was to the reference genome. This relationship was determined by the percent of reads aligned and the number of variants (SNPs) detected.

## Hardware (including minimum recommendations):

- 2025 Mac Studio
- 28-core Apple M3 Ultra (we recommend at least 8 cores)
- 96 GB RAM ( $\geq 32$ GB RAM is needed for data sets with  $>50$ Gbp input reads)
- 4TB high-speed NVMe SSD (we recommend using an external data drive)

## Data:

- Publicly available genomic data from eukaryotic species ranging in size from 30Mbp (*Aspergillus*) to  $\sim 3.2$  Gbp (human)
- Separate assemblies were run using PacBio HiFi, Oxford Nanopore Duplex, and Illumina data

## Additional Gains Coming in Spring 2026

Comparative benchmarks are in progress, but preliminary testing shows a speed improvement of approximately 50% using SeqMan NGen v19.0 compared to v18.1 (Figure 2). Version 19.0 is in development for release in spring 2026.

Figure 2: Genome assembly performance gains comparing SeqMan NGen v19.0 beta (green) to v.18.1 (orange). All assemblies used the same benchmark testing setup described above.

