DNASTAR
SOFTWARE FOR LIFE SCIENTISTS

LASERGENE

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.

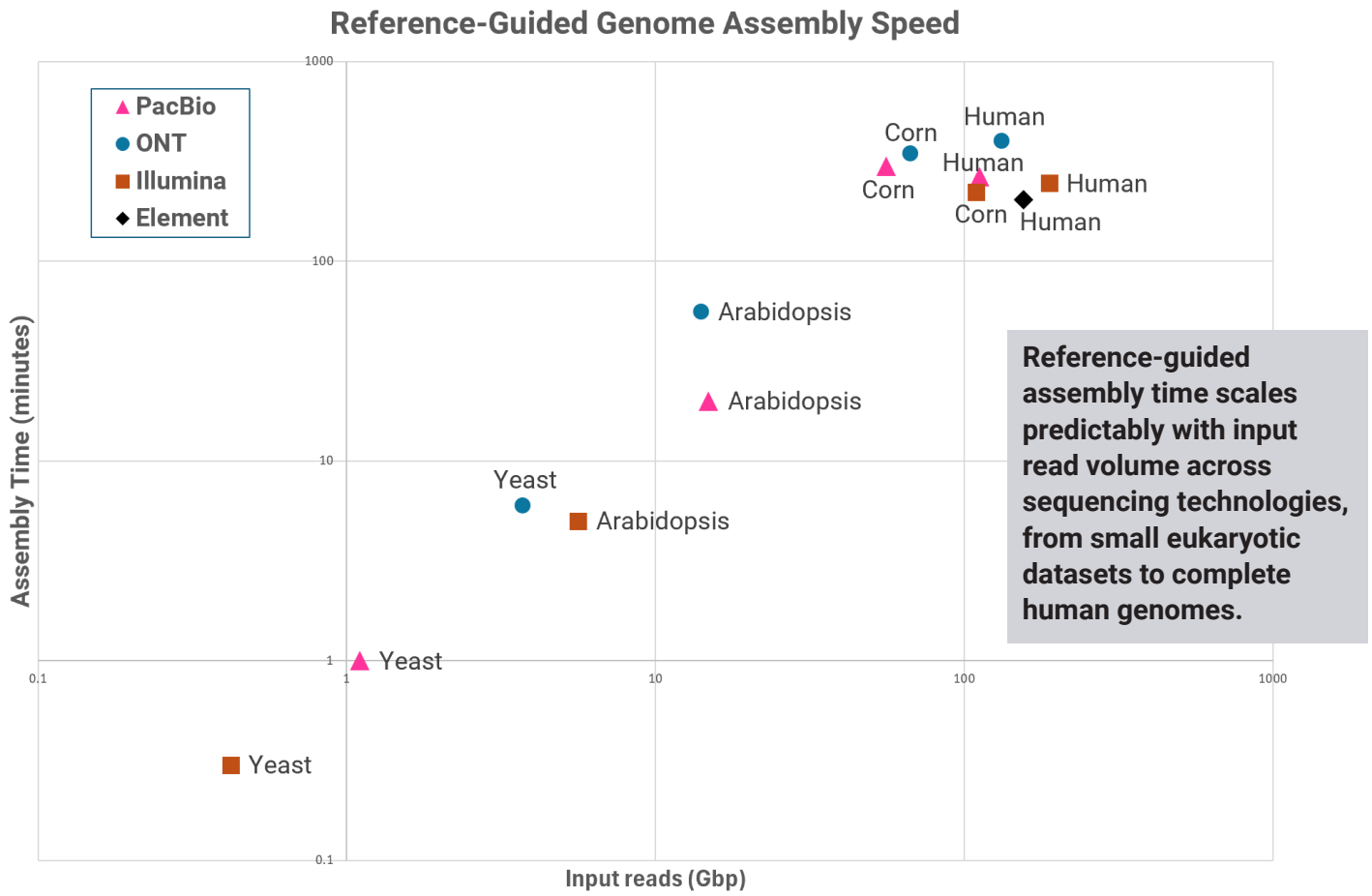
Lasergene Genomics excels in **speed, accuracy, and ease of use**. Customizable parameters are automatically optimized for each sequencing technology: Illumina, PacBio SMRT, PacBio HiFi, Oxford Nanopore, or Oxford Nanopore Duplex. Select the variant calling algorithm you prefer, including GATK options.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

Lasergene Genomics Benchmark Results

Assembly speeds correlated strongly with input read volume for all platforms.

Data Platform	Organism / Reference	Genome Length (Mbp)	Input reads (Gbp)	Assembly Time (min)
PacBio HiFi	Yeast (*S. cerevisiae* S288C)	12	1.1	1
	Arabidopsis (*A. thaliana* TAIR 10.1)	119	14.8	20
	Corn (Zm-B73)	2180	55.9	300
	Human (HG002)	3200	112	264
ONT Duplex	Yeast (*S. pombe* 972h)	13	3.7	6
	Arabidopsis (*A. thaliana* TAIR 10.1)	119	14.0	56
	Corn (Zm-B73)	2180	66.5	348
	Human (HG002)	3200	132	402
Illumina	Yeast (*S. cerevisiae* S288C)	12	0.4	0.3
	Arabidopsis (*A. thaliana* TAIR 10.1)	119	5.6	5
	Corn (Zm-B73)	2180	109	222
	Human (HG002)	3200	189	246
Element	Human (HG002)	3200	156	204

The figure on the reverse of this page shows a graphical representation of these assembly times on a log scale.



Benchmark Testing Setup

We used the SeqMan NGen application in Lasergene Genomics 19.0 to perform reference-guided assemblies for multiple eukaryotic genomes. Two long-read and two short-read technologies were tested: PacBio HiFi, ONT Duplex, Illumina, and Element Biosciences AVITI24.

Software:

- SeqMan NGen 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 used:

- 2025 Mac Studio
- 28-core Apple M3 Ultra
- 96 GB RAM
- 4TB high-speed NVMe SSD

Minimum recommendation:

- (N/A)
- ≥ 8 cores
- ≥ 32GB RAM is needed for data sets with > 50Gbp input reads
- We recommend using an external data drive

Data:

- Publicly available genomic data from eukaryotic species ranging in size from 12Mbp (yeast) to ~3.2 Gbp (human)
- Separate assemblies were run using PacBio HiFi, ONT Duplex, Illumina, and Element data

608.258.7420 USA
866.511.5090 toll free
0.808.271.1041 UK
0.800.182.4747 Germany
1202 Ann Street
Madison, WI 53713
www.dnastar.com
info@dnastar.com