

Additional file 1 — Supplement for: Partitioned Multi-MUM finding for scalable pangenomics with MumemtoM

Vikram Shivakumar^{1,*} and Ben Langmead^{1,*}

¹Department of Computer Science, Johns Hopkins University
**corresponding author; langmea@cs.jhu.edu*

October 15, 2025

Supplementary Tables

Table S1: Runtime and memory usage comparison between serial and parallel multi-MUM computation for both anchor and string-based merging algorithms. The time and memory for the merging step is reported separately and not included in the partition computation time.

Dataset	Partitions	Seqs per partition	Serial (no merge)				Parallel (no merge)				Merge step			
			Anchor		String		Anchor		String		Anchor		String	
			Time (hrs)	Mem (GB)	Time (hrs)	Mem (GB)	Time (hrs)	Mem (GB)	Time (hrs)	Mem (GB)	Time (min)	Mem (GB)	Time (min)	Mem (GB)
Chr 19 HPRC ($N = 474$)	1	474	4.77	47.30	–	–	–	–	–	–	–	–	–	–
	5	96	4.99	14.99	5.13	14.99	1.09	71.71	1.07	71.34	1.00	2.95	15.28	22.36
	10	48	5.90	10.27	5.65	10.16	0.62	95.16	0.60	94.25	0.68	3.32	17.33	22.32
	20	24	7.20	7.69	6.70	7.43	0.38	134.53	0.37	132.25	0.86	3.49	27.87	24.12
	40	12	8.30	5.74	7.81	5.80	0.23	204.43	0.21	207.34	1.59	3.62	48.49	29.52
<i>A. thaliana</i> ($N = 69$)	1	69	2.58	75.53	–	–	–	–	–	–	–	–	–	–
	5	15	3.74	29.15	3.63	28.20	0.79	137.52	0.76	133.97	0.57	2.25	19.65	16.14
	10	7	4.91	20.18	4.91	19.04	0.56	184.68	0.62	175.22	0.80	2.39	45.19	26.20
	20	4	6.46	14.58	5.16	13.89	0.38	260.23	0.33	231.90	1.25	2.44	78.30	51.19