

## Accurate Assembly of Full-length Consensus for Viral Quasispecies

Jia Tian<sup>1</sup>, Ziyu Gao<sup>1</sup>, Minghao Li<sup>1</sup>, Ergude Bao<sup>2</sup>, and Jin Zhao<sup>1\*</sup>

<sup>1</sup> College of Computer Science and Technology, Qingdao University, China

<sup>2</sup> School of Software Engineering, Beijing Jiaotong University, China

\*Corresponding author. [zhaojin@qdu.edu.cn](mailto:zhaojin@qdu.edu.cn)

### 1. Explanation of consensus refinement

In the consensus refinement phase, FC-Virus utilizes Algorithm 1 to adjust k-mer abundances based on the consensus depth. Algorithm 1 categorizes k-mers at the same position into a block and updates their abundances according to the consensus depth. Note that each block comprises four k-mers: one originating from the consensus, and the remaining three differing from the consensus k-mer only by the final base. If a k-mer's abundance drops to 0, it is marked as used; otherwise, it remains unused. Subsequently, FC-Virus iteratively selects the largest unused k-mer to extend both ends of the consensus sequence, as depicted in Figure S1.

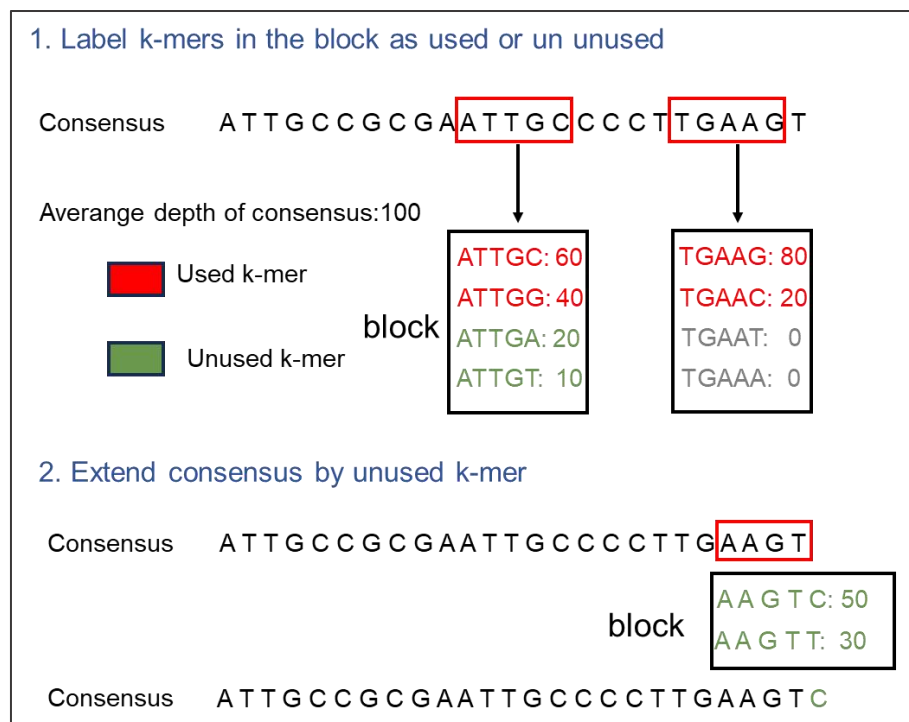


Figure S1. Example of k-mer blocks and consensus refinement process

## 2. Detailed evaluation results

We employed QUAST to compare the contigs (longer than 350 bp) assembled by the algorithms with each viral strain. Detailed results are provided in Tables S1-S6.

**Table S1. Benchmarking results for HIV dataset.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads(%)	Duplication ratio
S1 (9669bp)	IDBA	17.261	515	404	-	3880.35	12.98	1.482
	SOAPdenovo2	4.633	442	445	-	6787.33	4.89	<b>0.987</b>
	ViaDBG	99.959	<b>9664</b>	9268	<b>9664</b>	<b>3783.82</b>	95.98	9.308
	VG-Flow	<b>100.000</b>	9557	9176	9557	4262.36	94.09	4.656
	VStrains	<b>100.000</b>	9991	18333	9991	4615.02	98.59	6.170
	SPAdes	<b>100.000</b>	9557	8818	9557	4267.02	92.06	4.608
	FC-Virus	93.515	9751	<b>9755</b>	9751	4912.32	<b>99.49</b>	1.078
S2 (9719bp)	IDBA	10.588	515	404	-	7630.53	12.98	1.694
	SOAPdenovo2	4.64	439	445	-	6605.92	4.89	0.973
	ViaDBG	99.979	9657	9268	9657	3552.64	95.98	9.293
	VG-Flow	<b>100.000</b>	9546	9176	9546	3456.39	94.09	4.677
	VStrains	<b>100.000</b>	9853	18333	9853	3484.31	98.59	6.154
	SPAdes	<b>100.000</b>	9546	8818	9546	3463.51	92.06	4.634
	FC-Virus	93.477	<b>9755</b>	<b>9755</b>	<b>9755</b>	<b>1599.18</b>	<b>99.49</b>	<b>1.074</b>
S3 (9478bp)	IDBA	13.79	513	404	-	6525.82	12.98	1.630
	SOAPdenovo2	4.695	445	445	-	6516.85	4.89	<b>1.000</b>
	ViaDBG	<b>100.000</b>	<b>9235</b>	<b>9268</b>	<b>9235</b>	4652.68	95.98	9.352
	VG-Flow	<b>100.000</b>	<b>9235</b>	9176	<b>9235</b>	3957.72	94.09	4.761
	VStrains	<b>100.000</b>	9749	18333	9749	<b>3398.54</b>	98.59	6.268
	SPAdes	<b>100.000</b>	9235	8818	<b>9235</b>	3949.4	92.06	4.721
	FC-Virus	95.410	9751	9755	9751	4040.62	<b>99.49</b>	1.078
S4 (9709bp)	IDBA	13.451	515	404	-	7901.00	12.98	1.609
	SOAPdenovo2	4.614	439	445	-	6378.13	4.89	<b>0.980</b>
	ViaDBG	99.979	9654	9268	9654	3946.57	95.98	9.301
	VG-Flow	<b>100.000</b>	9543	9176	9543	3617.24	94.09	4.681
	VStrains	<b>100.000</b>	9853	18333	9853	3624.94	98.59	6.160
	SPAdes	<b>100.000</b>	9543	8818	9543	3692.84	92.06	4.638
	FC-Virus	93.470	<b>9755</b>	<b>9755</b>	<b>9755</b>	<b>1465.91</b>	<b>99.49</b>	1.075
S5 (9659bp)	IDBA	13.531	515	404	-	4831.14	12.98	1.631
	SOAPdenovo2	4.607	445	445	-	-	4.89	<b>1.000</b>
	ViaDBG	99.979	<b>9645</b>	9268	<b>9645</b>	3890.24	95.98	9.343
	VG-Flow	<b>100.000</b>	9534	9176	9534	<b>3817.78</b>	94.09	4.659
	VStrains	<b>100.000</b>	9853	18333	9853	3953.44	98.59	6.191
	SPAdes	<b>100.000</b>	9534	8818	9534	3832.44	92.06	4.644
	FC-Virus	93.478	9755	<b>9755</b>	9755	3905.69	<b>99.49</b>	1.080

**Table S2. Benchmarking results for POLIO dataset.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (7428bp)	IDBA	25.108	630	784	-	4811.64	39.62	<b>1.000</b>
	SOAPdenovo2	-	-	779	-	-	1.81	-
	ViaDBG	<b>100.000</b>	<b>7428</b>	2448	<b>7428</b>	2685.36	63.76	7.615
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>7428</b>	7270	<b>7428</b>	3362.22	<b>100.00</b>	5.389
	SPAdes	<b>100.000</b>	<b>7428</b>	4216	<b>7428</b>	3697.97	78.67	2.261
	FC-Virus	<b>100.000</b>	<b>7428</b>	<b>7460</b>	<b>7428</b>	<b>0.00</b>	98.08	<b>1.000</b>
S2 (7452bp)	IDBA	41.251	973	784	-	<b>2669.14</b>	39.62	<b>1.000</b>
	SOAPdenovo2	-	-	779	-	-	1.81	-
	ViaDBG	99.96	7431	2448	7431	4403.41	63.76	7.939
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>7452</b>	7270	<b>7452</b>	4026.66	<b>100.00</b>	4.792
	SPAdes	<b>100.000</b>	<b>7452</b>	4216	<b>7452</b>	4816.84	78.67	2.187
	FC-Virus	<b>100.000</b>	<b>7452</b>	<b>7460</b>	<b>7452</b>	4482.02	98.08	<b>1.000</b>
S3 (7460bp)	IDBA	25.000	630	784	-	<b>522.10</b>	39.62	<b>1.000</b>
	SOAPdenovo2	-	-	779	-	-	1.81	-
	ViaDBG	99.692	7437	2448	7437	1986.05	63.76	7.326
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>7460</b>	7270	<b>7460</b>	5128.27	<b>100.00</b>	5.377
	SPAdes	<b>100.000</b>	<b>7460</b>	4216	<b>7460</b>	3580.38	78.67	2.246
	FC-Virus	<b>100.000</b>	7457	<b>7460</b>	7457	3795.09	98.08	<b>1.000</b>
S4 (7459bp)	IDBA	46.159	1888	784	-	<b>2934.19</b>	39.62	0.999
	SOAPdenovo2	-	-	779	-	-	1.81	-
	ViaDBG	99.692	7436	2448	7436	4511.1	63.76	8.061
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>99.785</b>	<b>7443</b>	7270	<b>7443</b>	3843.56	<b>100.00</b>	5.383
	SPAdes	<b>99.785</b>	<b>7443</b>	4216	<b>7443</b>	3736.45	78.67	2.021
	FC-Virus	99.732	7439	<b>7460</b>	7439	3400.99	98.08	<b>1.000</b>
S5 (7457bp)	IDBA	46.171	1888	784	-	3457.29	39.62	0.999
	SOAPdenovo2	-	-	779	-	-	1.81	-
	ViaDBG	99.678	7432	2448	7432	4867.39	63.76	8.035
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>99.772</b>	7439	7270	7439	<b>3430.18</b>	<b>100.00</b>	4.788
	SPAdes	<b>99.772</b>	7439	4216	7439	3989.89	78.67	2.021
	FC-Virus	<b>99.772</b>	<b>7440</b>	<b>7460</b>	<b>7440</b>	4448.92	98.08	<b>1.000</b>

**Table S2\_C1. Continuation of Table S2 Benchmarking results for POLIO dataset.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S6 (7454bp)	IDBA	35.538	784	784	-	3806.36	39.62	<b>1.000</b>
	SOAPdenovo2	10.451	779	779	-	-	1.81	<b>1.000</b>
	ViaDBG	99.759	7436	2448	7436	3275.84	63.76	7.439
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>99.785</b>	<b>7438</b>	7270	<b>7438</b>	4838.1	<b>100.00</b>	5.385
	SPAdes	<b>99.785</b>	<b>7438</b>	4216	<b>7438</b>	3149.31	78.67	1.831
	FC-Virus	99.732	7434	<b>7460</b>	7434	<b>2273.34</b>	98.08	<b>1.000</b>

**Table S3. Benchmarking results for HCV dataset.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (9273bp)	IDBA	16.068	516	410	-	8488.96	14.79	1.186
	SOAPdenovo2	14.731	597	405	-	7833.09	14.25	<b>1.000</b>
	ViaDBG	99.978	9266	8959	9266	4669.44	99.10	20.235
	VG-Flow	<b>100.000</b>	<b>9273</b>	8703	<b>9273</b>	4868.92	98.35	10.317
	VStrains	<b>100.000</b>	<b>9273</b>	9151	<b>9273</b>	4814.76	<b>99.73</b>	9.833
	SPAdes	<b>100.000</b>	<b>9273</b>	8680	<b>9273</b>	4981.51	96.99	9.157
	FC-Virus	<b>100.000</b>	<b>9273</b>	<b>9311</b>	<b>9273</b>	<b>4033.21</b>	99.55	<b>1.000</b>
S2 (9284bp)	IDBA	16.168	527	410	-	6239.46	14.79	1.185
	SOAPdenovo2	14.810	597	405	-	<b>3490.90</b>	14.25	<b>1.000</b>
	ViaDBG	99.978	9266	8959	9266	4863.53	99.10	20.218
	VG-Flow	<b>100.000</b>	<b>9284</b>	8703	<b>9284</b>	4871.4	98.35	10.31
	VStrains	<b>100.000</b>	<b>9284</b>	9151	<b>9284</b>	4765.3	<b>99.73</b>	9.826
	SPAdes	<b>100.000</b>	<b>9284</b>	8680	<b>9284</b>	5036.21	96.99	9.147
	FC-Virus	<b>100.000</b>	<b>9284</b>	<b>9311</b>	<b>9284</b>	4319.26	99.55	<b>1.000</b>
S3 (9273bp)	IDBA	11.258	527	410	-	6429.65	14.79	1.266
	SOAPdenovo2	13.383	597	405	-	5882.35	14.25	<b>1.000</b>
	ViaDBG	99.978	9266	8959	9266	4032.51	99.1	20.241
	VG-Flow	<b>100.000</b>	<b>9273</b>	8703	<b>9273</b>	3750.47	98.35	10.317
	VStrains	<b>100.000</b>	<b>9273</b>	9151	<b>9273</b>	3972.45	<b>99.73</b>	9.833
	SPAdes	<b>100.000</b>	<b>9273</b>	8680	<b>9273</b>	4218.97	96.99	9.158
	FC-Virus	<b>100.000</b>	<b>9273</b>	<b>9311</b>	<b>9273</b>	<b>3008.74</b>	99.55	<b>1.000</b>
S4 (9302bp)	IDBA	11.277	527	410	-	6030.53	14.79	1.249
	SOAPdenovo2	13.169	580	405	-	6938.78	14.25	<b>1.000</b>
	ViaDBG	99.968	9266	8959	9266	4843.61	99.1	20.182
	VG-Flow	<b>100.000</b>	<b>9302</b>	8703	<b>9302</b>	4597.44	98.35	10.303
	VStrains	<b>100.000</b>	<b>9302</b>	9151	<b>9302</b>	4513.99	<b>99.73</b>	9.814
	SPAdes	<b>100.000</b>	<b>9302</b>	8680	<b>9302</b>	4736.12	96.99	9.132
	FC-Virus	<b>100.000</b>	<b>9302</b>	<b>9311</b>	<b>9302</b>	<b>3622.88</b>	99.55	<b>1.000</b>

**Table S3\_C1. Continuation of Table S3 Benchmarking results for HCV dataset.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S5 (9298bp)	IDBA	11.196	519	410	-	7587.25	14.79	1.266
	SOAPdenovo2	10.432	597	405	-	5773.20	14.25	<b>1.000</b>
	ViaDBG	99.978	9266	8959	9266	4432.81	99.1	20.188
	VG-Flow	<b>99.989</b>	<b>9297</b>	8703	<b>9297</b>	4356.41	98.35	10.306
	VStrains	<b>99.989</b>	<b>9297</b>	9151	<b>9297</b>	4344.06	<b>99.73</b>	9.818
	SPAdes	<b>99.989</b>	<b>9297</b>	8680	<b>9297</b>	4606.73	96.99	9.136
	FC-Virus	<b>99.989</b>	<b>9297</b>	<b>9311</b>	<b>9297</b>	<b>3474.24</b>	99.55	<b>1.000</b>
S6 (9296bp)	IDBA	15.383	527	410	-	6030.44	14.79	1.194
	SOAPdenovo2	14.619	597	405	-	5445.18	14.25	<b>1.000</b>
	ViaDBG	99.978	9266	8959	<b>9266</b>	4605.89	99.1	20.193
	VG-Flow	<b>100.000</b>	<b>9296</b>	8703	<b>9296</b>	4626.01	98.35	10.306
	VStrains	<b>100.000</b>	<b>9296</b>	9151	<b>9296</b>	4504.22	<b>99.73</b>	9.818
	SPAdes	<b>100.000</b>	<b>9296</b>	8680	<b>9296</b>	4765.32	96.99	9.138
	FC-Virus	<b>100.000</b>	<b>9296</b>	<b>9311</b>	<b>9296</b>	<b>3775.81</b>	99.55	<b>1.000</b>
S7 (9311bp)	IDBA	15.884	519	410	-	8712.98	14.79	1.187
	SOAPdenovo2	13.414	597	405	-	7365.89	14.25	<b>1.000</b>
	ViaDBG	99.871	9266	8959	9266	4674.49	99.1	20.183
	VG-Flow	<b>100.000</b>	<b>9311</b>	8703	<b>9311</b>	4662.34	98.35	10.299
	VStrains	<b>100.000</b>	<b>9311</b>	9151	<b>9311</b>	4559.55	<b>99.73</b>	9.806
	SPAdes	<b>100.000</b>	<b>9311</b>	8680	<b>9311</b>	4816.16	96.99	9.125
	FC-Virus	<b>100.000</b>	<b>9311</b>	<b>9311</b>	<b>9311</b>	<b>3447.54</b>	99.55	<b>1.000</b>
S8 (9302bp)	IDBA	10.052	527	410	-	<b>4385.02</b>	14.79	<b>1.000</b>
	SOAPdenovo2	4.010	373	405	-	9115.28	14.25	<b>1.000</b>
	ViaDBG	99.968	9266	8959	9266	6180.66	99.1	20.181
	VG-Flow	<b>100.000</b>	<b>9302</b>	8703	<b>9302</b>	6427.25	98.35	10.303
	VStrains	<b>100.000</b>	<b>9302</b>	9151	<b>9302</b>	6200.85	<b>99.73</b>	9.814
	SPAdes	<b>100.000</b>	<b>9302</b>	8680	<b>9302</b>	6477.85	96.99	9.133
	FC-Virus	<b>100.000</b>	<b>9302</b>	<b>9311</b>	<b>9302</b>	5471.94	99.55	<b>1.000</b>
S9 (9297bp)	IDBA	16.145	527	410	-	6970.21	14.79	1.185
	SOAPdenovo2	10.390	597	405	-	<b>2795.03</b>	14.25	<b>1.000</b>
	ViaDBG	99.978	9266	8959	9266	4233.28	99.1	20.189
	VG-Flow	<b>100.000</b>	<b>9297</b>	8703	<b>9297</b>	4294.09	98.35	10.305
	VStrains	<b>100.000</b>	<b>9297</b>	9151	<b>9297</b>	4326.54	<b>99.73</b>	9.818
	SPAdes	<b>100.000</b>	<b>9297</b>	8680	<b>9297</b>	4522.44	96.99	9.135
	FC-Virus	<b>100.000</b>	<b>9297</b>	<b>9311</b>	<b>9297</b>	3657.09	99.55	<b>1.000</b>

**Table S3\_C2. Continuation of Table S3 Benchmarking results for HCV dataset.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S10 (9311bp)	IDBA	9.462	527	410	-	7264.47	14.79	<b>1.000</b>
	SOAPdenovo2	-	-	405	-	-	14.25	-
	ViaDBG	99.871	9266	8959	9266	6428.31	99.1	20.177
	VG-Flow	<b>100.000</b>	<b>9311</b>	8703	<b>9311</b>	6236.96	98.35	10.299
	VStrains	<b>100.000</b>	<b>9311</b>	9151	<b>9311</b>	6033.76	<b>99.73</b>	9.806
	SPAdes	<b>100.000</b>	<b>9311</b>	8680	<b>9311</b>	6295.68	96.99	9.125
	FC-Virus	<b>100.000</b>	<b>9311</b>	<b>9311</b>	<b>9311</b>	<b>4790.03</b>	99.55	<b>1.000</b>

**Table S4. Benchmarking results for ZIKV dataset.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (10251bp)	IDBA	6.731	349	375	-	9710.14	7.04	<b>1.000</b>
	SOAPdenovo2	5.239	354	354	-	8380.68	2.18	1.311
	ViaDBG	99.951	<b>10255</b>	3466	<b>10255</b>	4167.46	95.26	29.808
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	10269	10201	10269	4327.35	99.78	14.874
	SPAdes	<b>100.000</b>	10266	2516	10266	4628.86	83.83	9.899
	FC-Virus	99.971	10266	<b>10269</b>	10266	<b>3272.94</b>	<b>99.95</b>	1.002
S2 (10269bp)	IDBA	3.321	341	375	-	8211.14	7.04	<b>1.000</b>
	SOAPdenovo2	3.409	350	354	-	3714.29	2.18	<b>1.000</b>
	ViaDBG	99.951	10255	3466	10255	3979.74	95.26	29.75
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>10269</b>	10201	<b>10269</b>	4325.56	99.78	14.851
	SPAdes	<b>100.000</b>	10266	2516	10266	4417.96	83.83	9.871
	FC-Virus	99.971	10266	<b>10269</b>	10266	<b>3370.35</b>	<b>99.95</b>	<b>1.000</b>
S3 (10269bp)	IDBA	7.089	375	375	-	3296.70	7.04	<b>1.000</b>
	SOAPdenovo2	3.447	354	354	-	<b>3107.34</b>	2.18	<b>1.000</b>
	ViaDBG	99.961	10260	3466	10260	5767.44	95.26	29.612
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>10269</b>	10201	<b>10269</b>	5476.71	99.78	14.849
	SPAdes	<b>100.000</b>	10254	2516	10254	5516.03	83.83	9.786
	FC-Virus	<b>100.000</b>	<b>10269</b>	<b>10269</b>	<b>10269</b>	4791.12	<b>99.95</b>	<b>1.000</b>
S4 (10251bp)	IDBA	3.366	345	375	-	9855.07	7.04	<b>1.000</b>
	SOAPdenovo2	3.414	350	354	-	7714.29	2.18	<b>1.000</b>
	ViaDBG	99.951	<b>10252</b>	3466	<b>10252</b>	4956.6	95.26	29.786
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	10269	10201	10269	5136.23	99.78	14.869
	SPAdes	<b>100.000</b>	10266	2516	10266	5458.54	83.83	9.888
	FC-Virus	99.873	10256	10269	10256	<b>4221.92</b>	<b>99.95</b>	1.002

**Table S4\_C1. Continuation of Table S4 Benchmarking results for ZIKV dataset.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S5 (10269bp)	IDBA	3.321	341	375	-	8797.65	7.04	<b>1.000</b>
	SOAPdenovo2	3.409	350	354	-	<b>4000.00</b>	2.18	<b>1.000</b>
	ViaDBG	99.951	10255	3466	10255	4728.22	95.26	29.722
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>10269</b>	10201	<b>10269</b>	5138.05	99.78	14.847
	SPAdes	<b>100.000</b>	10266	2516	10266	5208.21	83.83	9.748
	FC-Virus	99.971	10266	<b>10269</b>	10266	4315.21	<b>99.95</b>	<b>1.000</b>
S6 (10269bp)	IDBA	7.070	375	375	-	<b>3856.75</b>	7.04	<b>1.000</b>
	SOAPdenovo2	3.447	354	354	-	4237.29	2.18	<b>1.000</b>
	ViaDBG	99.961	10260	3466	10260	6574.71	95.26	29.471
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>10269</b>	10201	<b>10269</b>	6287.79	99.78	14.848
	SPAdes	<b>100.000</b>	10254	2516	10254	6244.79	83.83	9.464
	FC-Virus	<b>100.000</b>	<b>10269</b>	<b>10269</b>	<b>10269</b>	5706.50	<b>99.95</b>	<b>1.000</b>
S7 (10251bp)	IDBA	3.366	345	375	-	9565.22	7.04	<b>1.000</b>
	SOAPdenovo2	5.239	354	354	-	8522.73	2.18	1.311
	ViaDBG	99.951	<b>10255</b>	3466	<b>10255</b>	4847.47	95.26	29.748
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	10269	10201	10269	5126.85	99.78	14.874
	SPAdes	<b>100.000</b>	10266	2516	10266	5447.65	83.83	9.824
	FC-Virus	99.971	10266	<b>10269</b>	10266	<b>4247.03</b>	<b>99.95</b>	1.002
S8 (10269bp)	IDBA	-	-	375	-	-	7.04	-
	SOAPdenovo2	3.409	350	354	-	<b>4285.71</b>	2.18	<b>1.000</b>
	ViaDBG	99.951	10255	3466	10255	4803.55	95.26	29.737
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>10269</b>	10201	<b>10269</b>	5135.45	99.78	14.851
	SPAdes	<b>100.000</b>	10266	2516	10266	5142.59	83.83	9.723
	FC-Virus	99.971	10266	<b>10269</b>	10266	4334.69	<b>99.95</b>	<b>1.000</b>
S9 (10269bp)	IDBA	7.089	375	375	-	4670.33	7.04	<b>1.000</b>
	SOAPdenovo2	3.447	354	354	-	<b>4237.29</b>	2.18	<b>1.000</b>
	ViaDBG	99.961	10260	3466	10260	6445.2	95.26	29.583
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>10269</b>	10201	<b>10269</b>	6203.56	99.78	14.848
	SPAdes	<b>100.000</b>	10254	2516	10254	6215.14	83.83	9.708
	FC-Virus	<b>100.000</b>	<b>10269</b>	<b>10269</b>	<b>10269</b>	5579.90	<b>99.95</b>	<b>1.000</b>

**Table S4\_C2. Continuation of Table S4 Benchmarking results for ZIKV dataset.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S10 (10251bp)	IDBA	-	-	375	-	-	7.04	-
	SOAPdenovo2	3.034	311	354	-	7395.50	2.18	<b>1.000</b>
	ViaDBG	99.932	10181	3466	10181	5198.96	95.26	26.719
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	10269	10201	10269	5555.95	99.78	13.864
	SPAdes	<b>100.000</b>	<b>10251</b>	2516	<b>10251</b>	5360.01	83.83	8.381
	FC-Virus	99.971	10266	<b>10269</b>	10266	<b>5104.22</b>	<b>99.95</b>	1.002
S11 (10269bp)	IDBA	3.321	341	375	-	9384.16	7.04	<b>1.000</b>
	SOAPdenovo2	3.409	350	354	-	<b>5142.86</b>	2.18	<b>1.000</b>
	ViaDBG	99.951	10255	3466	10255	5352.06	95.26	28.987
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>10269</b>	10201	<b>10269</b>	5587.99	99.78	13.845
	SPAdes	<b>100.000</b>	<b>10269</b>	2516	<b>10269</b>	5395.85	83.83	8.985
	FC-Virus	99.942	10263	<b>10269</b>	10263	5242.13	<b>99.95</b>	<b>1.000</b>
S12 (10269bp)	IDBA	7.089	375	375	-	-	7.04	<b>1.000</b>
	SOAPdenovo2	3.447	354	354	-	-	2.18	<b>1.000</b>
	ViaDBG	99.961	10260	3466	10260	6891.24	95.26	25.744
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>10269</b>	10201	<b>10269</b>	6592.31	99.78	12.851
	SPAdes	<b>100.000</b>	10254	2516	10254	<b>5959.74</b>	83.83	7.402
	FC-Virus	<b>100.000</b>	<b>10269</b>	<b>10269</b>	<b>10269</b>	6592.65	<b>99.95</b>	<b>1.000</b>
S13 (10251bp)	IDBA	2.878	295	375	-	10169.49	7.04	<b>1.000</b>
	SOAPdenovo2	3.414	350	354	-	8285.71	2.18	<b>1.000</b>
	ViaDBG	99.941	10198	3466	10198	5473.57	95.26	26.94
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	10269	10201	10269	5328.94	99.78	12.871
	SPAdes	<b>100.000</b>	<b>10266</b>	2516	<b>10266</b>	5449.19	83.83	8.473
	FC-Virus	99.971	<b>10266</b>	<b>10269</b>	<b>10266</b>	<b>4919.15</b>	<b>99.95</b>	1.002
S14 (10269bp)	IDBA	2.873	295	375	-	8135.60	7.04	<b>1.000</b>
	SOAPdenovo2	3.409	350	354	-	<b>285.71</b>	2.18	<b>1.000</b>
	ViaDBG	99.942	10201	3466	10201	4988.14	95.26	26.838
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>10269</b>	10201	<b>10269</b>	5896.82	99.78	14.843
	SPAdes	<b>100.000</b>	10266	2516	10266	5501.12	83.83	9.215
	FC-Virus	99.883	10257	<b>10269</b>	10257	5040.46	<b>99.95</b>	<b>1.000</b>



**Table S4\_C3. Continuation of Table S4 Benchmarking results for ZIKV dataset.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S15 (10269bp)	IDBA	7.089	375	375	-	5769.23	7.04	<b>1.000</b>
	SOAPdenovo2	3.447	354	354	-	6779.66	2.18	<b>1.000</b>
	ViaDBG	99.961	10260	3466	10260	6911.47	95.26	24.689
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>10269</b>	<b>10201</b>	<b>10269</b>	6572.66	99.78	12.851
	SPAdes	<b>100.000</b>	10254	2516	10254	<b>5767.49</b>	83.83	7.196
	FC-Virus	<b>100.000</b>	<b>10269</b>	<b>10269</b>	<b>10269</b>	6407.63	<b>99.95</b>	<b>1.000</b>

**Table S5. Benchmarking results for COVID\_19 dataset in 20000X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (30351bp)	IDBA	98.924	28482	28704	28482	910.99	96.43	0.998
	SOAPdenovo2	94.310	11439	6475	6475	1170.32	36.49	1.001
	ViaDBG	-	-	-	-	-	-	-
	VG-Flow	99.859	28667	28009	28667	484.42	33.67	6.682
	VStrains	99.859	27742	26950	27742	191.34	<b>100.00</b>	4.157
	SPAdes	<b>99.889</b>	29790	29790	29790	201.41	<b>100.00</b>	0.999
	FC-Virus	<b>99.889</b>	<b>29820</b>	<b>30150</b>	<b>29820</b>	<b>157.61</b>	<b>100.00</b>	<b>1.000</b>
S2 (30169bp)	IDBA	98.989	28455	28704	28455	976.39	96.43	0.999
	SOAPdenovo2	94.335	11438	6475	6448	1171.41	36.49	<b>1.000</b>
	ViaDBG	-	-	-	-	-	-	-
	VG-Flow	<b>100.000</b>	28584	28009	28584	557.83	33.67	6.692
	VStrains	<b>100.000</b>	27718	26950	27718	<b>166.71</b>	<b>100.00</b>	4.164
	SPAdes	<b>100.000</b>	29674	29790	29674	252.75	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	<b>29704</b>	<b>30150</b>	<b>29704</b>	178.43	<b>100.00</b>	1.001
S3 (30178bp)	IDBA	99.080	28456	28704	28456	785.50	96.43	0.998
	SOAPdenovo2	94.341	11439	6475	6449	1071.34	36.49	<b>1.000</b>
	ViaDBG	-	-	-	-	-	-	-
	VG-Flow	<b>100.000</b>	28594	28009	28594	492.83	33.67	6.692
	VStrains	<b>100.000</b>	27719	26950	27719	160.19	<b>100.00</b>	4.164
	SPAdes	<b>100.000</b>	29685	29790	29685	<b>10.11</b>	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	<b>29715</b>	<b>30150</b>	<b>29715</b>	87.50	<b>100.00</b>	1.001

**Table S6. Benchmarking results for HIV-labmix dataset.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Error per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (9669bp)	IDBA	-	-	-	-	-	-	-
	SOAPdenovo2	-	-	-	-	-	-	-
	ViaDBG	94.064	9129	5510	9129	3807.18	87.14	62.473
	VG-Flow	97.528	3919	1433	3610	<b>3361.55</b>	90.65	12.35
	VStrains	<b>99.938</b>	9001	5659	9001	3951.70	96.25	5.630
	SPAdes	95.077	1951	642	1359	4290.88	79.37	3.547
	FC-Virus	93.515	<b>9306</b>	<b>9665</b>	<b>9306</b>	4545.46	<b>97.69</b>	<b>1.029</b>
S2 (9719bp)	IDBA	-	-	-	-	-	-	-
	SOAPdenovo2	-	-	-	-	-	-	-
	ViaDBG	94.022	9129	5510	9129	3654.80	87.14	62.184
	VG-Flow	95.329	4096	1433	3610	<b>3078.09</b>	90.65	12.662
	VStrains	<b>100.000</b>	9001	5659	9001	3748.81	96.25	5.602
	SPAdes	95.257	1951	642	1391	3731.48	79.37	3.514
	FC-Virus	93.477	<b>9303</b>	<b>9665</b>	<b>9303</b>	3429.00	<b>97.69</b>	<b>1.024</b>
S3 (9478bp)	IDBA	-	-	-	-	-	-	-
	SOAPdenovo2	-	-	-	-	-	-	-
	ViaDBG	95.410	9048	5510	9048	4146.09	87.14	62.590
	VG-Flow	95.410	4096	1433	3610	3117.54	90.65	12.972
	VStrains	<b>99.831</b>	8970	5659	8970	3764.66	96.25	5.741
	SPAdes	96.107	1951	642	1360	3852.23	79.37	3.614
	FC-Virus	95.410	<b>9301</b>	<b>9665</b>	<b>9301</b>	<b>2601.88</b>	<b>97.69</b>	<b>1.029</b>
S4 (9709bp)	IDBA	-	-	-	-	-	-	-
	SOAPdenovo2	-	-	-	-	-	-	-
	ViaDBG	94.016	9129	5510	9129	3301.27	87.14	62.249
	VG-Flow	95.324	4096	1433	3610	<b>2879.64</b>	90.65	12.679
	VStrains	<b>100.000</b>	9001	5659	9001	3848.56	96.25	5.607
	SPAdes	95.252	1951	642	1391	3799.19	79.37	3.418
	FC-Virus	93.470	<b>9306</b>	<b>9665</b>	<b>9306</b>	2987.32	<b>97.69</b>	<b>1.025</b>
S5 (9659bp)	IDBA	-	-	-	-	-	-	-
	SOAPdenovo2	-	-	-	-	-	-	-
	ViaDBG	94.026	9129	5510	9129	4440.45	87.14	62.599
	VG-Flow	95.341	4091	1433	3610	<b>3199.08</b>	90.65	12.726
	VStrains	<b>100.000</b>	9001	5659	9001	4154.58	96.25	5.637
	SPAdes	95.196	1951	642	1391	3841.37	79.37	3.584
	FC-Virus	93.478	<b>9303</b>	<b>9665</b>	<b>9303</b>	4063.20	<b>97.69</b>	<b>1.030</b>

**Table S7. Benchmarking results for COVID-19 dataset in 50X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (30351bp)	IDBA	98.606	28383	28679	28383	1139.10	96.13	0.997
	SOAPdenovo2	95.627	9555	20716	7126	14194.49	70.38	1.004
	ViaDBG	-	-	-	-	-	-	-
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>99.819</b>	<b>29783</b>	<b>29762</b>	<b>29783</b>	<b>137.15</b>	<b>100.00</b>	2.497
	SPAdes	<b>99.819</b>	29762	<b>29762</b>	29762	141.12	<b>100.00</b>	0.999
	FC-Virus	99.742	29755	29755	29755	181.48	<b>100.00</b>	<b>1.000</b>
S2 (30169bp)	IDBA	98.878	28377	28679	28377	1245.34	96.13	0.997
	SOAPdenovo2	96.040	9554	20716	7126	14281.89	70.38	1.005
	ViaDBG	-	-	-	-	-	-	-
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>29688</b>	<b>29762</b>	<b>29688</b>	235.89	<b>100.00</b>	2.500
	SPAdes	<b>100.000</b>	29667	<b>29762</b>	29667	259.55	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	29683	29755	29683	262.78	<b>100.00</b>	<b>1.000</b>
S3 (30178bp)	IDBA	98.939	28378	28679	-	1060.78	96.13	0.997
	SOAPdenovo2	96.139	9556	20716	-	14093.27	70.38	1.004
	ViaDBG	-	-	-	-	-	-	-
	VG-Flow	-	-	-	-	-	-	-
	VStrains	<b>100.000</b>	<b>29699</b>	<b>29762</b>	<b>29699</b>	63.34	<b>100.00</b>	2.500
	SPAdes	<b>100.000</b>	29678	<b>29762</b>	29678	70.76	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	29694	29755	29694	<b>30.31</b>	<b>100.00</b>	<b>1.000</b>

**Table S8. Benchmarking results for COVID-19 dataset in 2000X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (30351bp)	IDBA	98.928	29424	29643	29424	922.31	99.56	0.997
	SOAPdenovo2	93.794	13344	5386	5245	8135.99	26.35	1.188
	ViaDBG	93.489	2068	929	1454	<b>141.89</b>	<b>100.00</b>	3.032
	VG-Flow	<b>99.886</b>	<b>30020</b>	<b>30116</b>	<b>30020</b>	655.8	<b>100.00</b>	2.011
	VStrains	<b>99.886</b>	29810	29810	29810	199.67	<b>100.00</b>	2.000
	SPAdes	<b>99.886</b>	29789	29789	29789	191.34	<b>100.00</b>	<b>0.999</b>
	FC-Virus	<b>99.886</b>	29789	29789	29789	187.98	<b>100.00</b>	<b>0.999</b>
S2 (30169bp)	IDBA	98.999	29310	29643	29310	977.62	99.56	0.998
	SOAPdenovo2	93.830	13344	5386	5245	8151.86	26.35	1.189
	ViaDBG	93.708	2068	929	1487	<b>158.63</b>	<b>100.00</b>	3.038
	VG-Flow	<b>100.000</b>	<b>29900</b>	<b>30116</b>	<b>29900</b>	710.13	<b>100.00</b>	2.012
	VStrains	<b>100.000</b>	29695	29810	29695	252.66	<b>100.00</b>	2.001
	SPAdes	<b>100.000</b>	29674	29789	29674	246.01	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	29674	29789	29674	239.27	<b>100.00</b>	<b>1.000</b>

**Table S8\_C1. Continuation of Table S8 for COVID-19 dataset in 2000X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S3 (30178bp)	IDBA	99.090	29321	29643	29321	766.07	99.56	0.997
	SOAPdenovo2	93.566	13344	5386	5245	8076.71	26.35	1.187
	ViaDBG	93.684	2068	929	1487	110.06	<b>100.00</b>	3.038
	VG-Flow	<b>100.000</b>	<b>29916</b>	<b>30116</b>	<b>29916</b>	509.1	<b>100.00</b>	2.012
	VStrains	<b>100.000</b>	29706	29810	29706	<b>11.78</b>	<b>100.00</b>	2.001
	SPAdes	<b>100.000</b>	29685	29789	29685	20.21	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	29685	29789	29685	23.58	<b>100.00</b>	<b>1.000</b>

**Table S9. Benchmarking results for COVID-19 dataset in 4000X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (30351bp)	IDBA	98.934	29412	29634	29412	939.54	99.50	0.996
	SOAPdenovo2	94.021	11439	6475	6475	1082.22	36.36	<b>1.000</b>
	ViaDBG	97.225	2068	904	1338	<b>134.47</b>	<b>100.00</b>	3.307
	VG-Flow	<b>99.889</b>	29790	29790	29790	194.69	<b>100.00</b>	0.999
	VStrains	-	-	-	-	-	-	-
	SPAdes	<b>99.889</b>	29790	29790	29790	191.34	<b>100.00</b>	0.999
	FC-Virus	<b>99.889</b>	<b>29799</b>	<b>29799</b>	<b>29799</b>	177.86	<b>100.00</b>	<b>1.000</b>
S2 (30169bp)	IDBA	98.999	29296	29634	29296	994.90	99.50	0.997
	SOAPdenovo2	94.018	11439	6475	6448	1075.78	36.36	<b>1.000</b>
	ViaDBG	97.442	2068	904	1338	<b>167.99</b>	<b>100.00</b>	3.314
	VG-Flow	<b>100.000</b>	29674	29790	29674	246.01	<b>100.00</b>	<b>1.000</b>
	VStrains	-	-	-	-	-	-	-
	SPAdes	<b>100.000</b>	29674	29790	29674	246.01	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	<b>29683</b>	<b>29799</b>	<b>29683</b>	239.19	<b>100.00</b>	<b>1.000</b>
S3 (30178bp)	IDBA	99.090	29307	29634	29307	769.61	99.50	0.996
	SOAPdenovo2	94.021	11439	6475	6449	982.58	36.36	<b>1.000</b>
	ViaDBG	97.389	2068	904	1338	122.06	<b>100.00</b>	3.315
	VG-Flow	<b>100.000</b>	29685	29790	29685	<b>16.84</b>	<b>100.00</b>	<b>1.000</b>
	VStrains	-	-	-	-	-	-	-
	SPAdes	<b>100.000</b>	29685	29790	29685	20.21	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	<b>29694</b>	<b>29799</b>	<b>29694</b>	33.68	<b>100.00</b>	<b>1.000</b>

**Table S10. Benchmarking results for COVID-19 dataset in 6000X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (30351bp)	IDBA	98.924	28477	28698	28477	921.37	96.48	0.998
	SOAPdenovo2	94.340	11439	6475	6475	1266.08	36.54	<b>1.000</b>
	ViaDBG	95.439	2068	988	1338	<b>108.65</b>	<b>100.00</b>	3.587
	VG-Flow	<b>99.889</b>	<b>30069</b>	<b>30265</b>	<b>30069</b>	496.6	<b>100.00</b>	2.009
	VStrains	<b>99.889</b>	29811	29811	29811	201.27	<b>100.00</b>	2.000
	SPAdes	<b>99.889</b>	30065	<b>30265</b>	30065	857.07	<b>100.00</b>	1.009
	FC-Virus	<b>99.889</b>	29811	29811	29811	177.79	<b>100.00</b>	<b>1.000</b>
S2 (30169bp)	IDBA	98.989	28450	28698	28450	986.84	96.48	0.999
	SOAPdenovo2	94.338	11439	6475	6448	1266.83	36.54	<b>1.000</b>
	ViaDBG	95.646	2068	988	1347	<b>132.27</b>	<b>100.00</b>	3.596
	VG-Flow	<b>100.000</b>	<b>29953</b>	<b>30265</b>	<b>29953</b>	549.25	<b>100.00</b>	2.010
	VStrains	<b>100.000</b>	29695	29811	29695	252.57	<b>100.00</b>	2.001
	SPAdes	<b>100.000</b>	29949	<b>30265</b>	29949	914.59	<b>100.00</b>	1.009
	FC-Virus	<b>100.000</b>	29695	29811	29695	235.73	<b>100.00</b>	1.001
S3 (30178bp)	IDBA	99.080	28451	28698	28451	768.67	96.48	0.998
	SOAPdenovo2	94.341	11439	6475	6449	1152.54	36.54	<b>1.000</b>
	ViaDBG	95.594	2068	988	1347	104.83	<b>100.00</b>	3.597
	VG-Flow	<b>100.000</b>	<b>29964</b>	<b>30265</b>	<b>29964</b>	359.79	<b>100.00</b>	2.01
	VStrains	<b>100.000</b>	29706	29811	29706	<b>10.1</b>	<b>100.00</b>	2.001
	SPAdes	<b>100.000</b>	29960	<b>30265</b>	29960	674.18	<b>100.00</b>	1.009
	FC-Virus	<b>100.000</b>	29706	29811	29706	33.67	<b>100.00</b>	1.001

**Table S11. Benchmarking results for COVID-19 dataset in 8000X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (30351bp)	IDBA	98.924	28482	28705	28482	927.90	96.45	0.998
	SOAPdenovo2	93.676	6662	6475	4567	558.90	36.72	1.003
	ViaDBG	95.623	2068	943	1329	<b>109.4</b>	<b>100.00</b>	3.428
	VG-Flow	<b>99.889</b>	<b>30073</b>	<b>30265</b>	<b>30073</b>	814.25	<b>100.00</b>	2.017
	VStrains	<b>99.889</b>	29811	29811	29811	201.34	<b>100.00</b>	2
	SPAdes	<b>99.889</b>	30044	30244	30044	854.34	<b>100.00</b>	1.008
	FC-Virus	<b>99.889</b>	29811	29811	29811	181.14	<b>100.00</b>	<b>1.000</b>
S2 (30169bp)	IDBA	98.989	28455	28705	28455	993.36	96.45	<b>0.999</b>
	SOAPdenovo2	93.705	6662	6475	4567	548.94	36.72	1.002
	ViaDBG	95.831	2068	943	1329	<b>110.53</b>	<b>100.00</b>	3.436
	VG-Flow	<b>100.000</b>	<b>30032</b>	<b>30265</b>	<b>30032</b>	878	<b>100.00</b>	2.022
	VStrains	<b>100.000</b>	29695	29811	29695	252.66	<b>100.00</b>	2.001
	SPAdes	<b>100.000</b>	30028	30244	30028	934.37	<b>100.00</b>	1.012
	FC-Virus	<b>100.000</b>	29695	29811	29695	239.10	<b>100.00</b>	<b>1.001</b>

**Table S11\_C1. Continuation of Table S11 for COVID-19 dataset in 8000X depth.**

S3 (30178bp)	IDBA	99.080	28456	28705	28456	775.22	96.45	0.998
	SOAPdenovo2	93.721	6662	6475	4567	434.11	36.72	1.002
	ViaDBG	95.779	2068	943	1329	113.59	<b>100.00</b>	3.437
	VG-Flow	<b>100.000</b>	<b>29964</b>	<b>30265</b>	<b>29964</b>	664.67	<b>100.00</b>	2.018
	VStrains	<b>100.000</b>	29706	29811	29706	<b>10.1</b>	<b>100.00</b>	2.001
	SPAdes	<b>100.000</b>	29939	30244	29939	677.99	<b>100.00</b>	1.009
	FC-Virus	<b>100.000</b>	29706	29811	29706	30.30	<b>100.00</b>	<b>1.001</b>

**Table S12. Benchmarking results for COVID-19 dataset in 10000X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (30351bp)	IDBA	98.924	28479	28702	28479	917.82	96.45	0.998
	SOAPdenovo2	94.648	6725	6475	4714	1367.31	36.14	1.005
	ViaDBG	95.607	2068	950	1308	<b>136.97</b>	<b>100.00</b>	3.429
	VG-Flow	<b>99.889</b>	<b>29811</b>	<b>29811</b>	<b>29811</b>	184.5	<b>100.00</b>	<b>1.000</b>
	VStrains	-	-	-	-	-	-	-
	SPAdes	<b>99.889</b>	<b>29811</b>	<b>29811</b>	<b>29811</b>	191.21	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>99.889</b>	29790	29790	29790	171.19	<b>100.00</b>	0.999
S2 (30169bp)	IDBA	98.989	28452	28702	28452	983.24	96.45	0.999
	SOAPdenovo2	94.679	6725	6475	4714	1386.25	36.14	1.005
	ViaDBG	95.784	2068	950	1308	<b>156.52</b>	<b>100.00</b>	3.439
	VG-Flow	<b>100.000</b>	<b>29695</b>	<b>29811</b>	<b>29695</b>	235.73	<b>100.00</b>	1.001
	VStrains	-	-	-	-	-	-	-
	SPAdes	<b>100.000</b>	<b>29695</b>	<b>29811</b>	<b>29695</b>	245.84	<b>100.00</b>	1.001
	FC-Virus	<b>100.000</b>	29674	29790	29674	229.16	<b>100.00</b>	<b>1.000</b>
S3 (30178bp)	IDBA	99.080	28453	28702	28453	785.52	96.45	0.998
	SOAPdenovo2	94.694	6725	6475	4714	1262.20	36.14	1.005
	ViaDBG	95.782	2068	950	1308	134.01	<b>100.00</b>	3.438
	VG-Flow	<b>100.000</b>	<b>29706</b>	<b>29811</b>	<b>29706</b>	26.93	<b>100.00</b>	1.001
	VStrains	-	-	-	-	-	-	-
	SPAdes	<b>100.000</b>	<b>29706</b>	<b>29811</b>	<b>29706</b>	<b>20.2</b>	<b>100.00</b>	1.001
	FC-Virus	<b>100.000</b>	29685	29790	29685	40.42	<b>100.00</b>	<b>1.000</b>

**Table S13. Benchmarking results for COVID-19 dataset in 12000X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (30351bp)	IDBA	98.924	28480	28701	28480	911.09	96.48	0.998
	SOAPdenovo2	94.648	11439	6475	6475	2275.66	36.27	1.005
	ViaDBG	94.608	2068	995	1377	<b>122.83</b>	<b>100.00</b>	3.345
	VG-Flow	<b>99.889</b>	<b>29790</b>	<b>29790</b>	<b>29790</b>	181.26	<b>100.00</b>	<b>0.999</b>
	VStrains	<b>99.889</b>	<b>29790</b>	<b>29790</b>	<b>29790</b>	198.05	<b>100.00</b>	1.999
	SPAdes	<b>99.889</b>	<b>29790</b>	<b>29790</b>	<b>29790</b>	201.41	<b>100.00</b>	<b>0.999</b>
	FC-Virus	<b>99.889</b>	<b>29790</b>	<b>29790</b>	<b>29790</b>	177.91	<b>100.00</b>	<b>0.999</b>

**Table S13\_C1. Continuation of Table S13 for COVID-19 dataset in 12000X depth.**

S2 (30169bp)	IDBA	98.989	28453	28701	28453	973.10	96.48	0.999
	SOAPdenovo2	94.675	11438	6475	6448	2294.51	36.27	1.005
	ViaDBG	94.81	2068	995	1377	<b>133.55</b>	<b>100.00</b>	3.353
	VG-Flow	<b>100.000</b>	<b>29674</b>	<b>29790</b>	<b>29674</b>	232.53	<b>100.00</b>	<b>1.000</b>
	VStrains	<b>100.000</b>	<b>29674</b>	<b>29790</b>	<b>29674</b>	252.75	<b>100.00</b>	2.000
	SPAdes	<b>100.000</b>	<b>29674</b>	<b>29790</b>	<b>29674</b>	256.12	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	<b>29674</b>	<b>29790</b>	<b>29674</b>	235.90	<b>100.00</b>	<b>1.000</b>
S3 (30178bp)	IDBA	99.080	28454	28701	28454	778.80	96.48	0.998
	SOAPdenovo2	94.677	11439	6475	6449	2177.54	36.27	1.005
	ViaDBG	94.758	2068	995	1377	121.88	<b>100.00</b>	3.354
	VG-Flow	<b>100.000</b>	<b>29685</b>	<b>29790</b>	<b>29685</b>	30.32	<b>100.00</b>	<b>1.000</b>
	VStrains	<b>100.000</b>	<b>29685</b>	<b>29790</b>	<b>29685</b>	13.47	<b>100.00</b>	2.000
	SPAdes	<b>100.000</b>	<b>29685</b>	<b>29790</b>	<b>29685</b>	<b>10.11</b>	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	<b>29685</b>	<b>29790</b>	<b>29685</b>	33.69	<b>100.00</b>	<b>1.000</b>

**Table S14. Benchmarking results for COVID-19 dataset in 14000X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (30351bp)	IDBA	98.924	28479	28700	28479	921.31	96.47	0.998
	SOAPdenovo2	94.856	11439	6475	6475	2401.18	36.41	1.003
	ViaDBG	94.487	2068	987	1349	<b>117.22</b>	<b>100.00</b>	3.570
	VG-Flow	<b>99.889</b>	<b>29790</b>	<b>29790</b>	<b>29790</b>	194.69	<b>100.00</b>	<b>0.999</b>
	VStrains	-	-	-	-	-	-	-
	SPAdes	<b>99.889</b>	<b>29790</b>	<b>29790</b>	<b>29790</b>	<b>201.41</b>	<b>100.00</b>	<b>0.999</b>
	FC-Virus	<b>99.889</b>	<b>29790</b>	<b>29790</b>	<b>29790</b>	<b>181.26</b>	<b>100.00</b>	<b>0.999</b>
S2 (30169bp)	IDBA	98.989	28452	28700	28452	986.77	96.47	0.999
	SOAPdenovo2	94.884	11438	6475	6448	2412.99	36.41	1.002
	ViaDBG	94.686	2068	987	1352	<b>153.13</b>	<b>100.00</b>	3.579
	VG-Flow	<b>100.000</b>	<b>29674</b>	<b>29790</b>	<b>29674</b>	246.01	<b>100.00</b>	<b>1.000</b>
	VStrains	-	-	-	-	-	-	-
	SPAdes	<b>100.000</b>	<b>29674</b>	<b>29790</b>	<b>29674</b>	252.75	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	<b>29674</b>	<b>29790</b>	<b>29674</b>	239.27	<b>100.00</b>	<b>1.000</b>
S3 (30178bp)	IDBA	99.080	28453	28700	28453	768.61	96.47	0.998
	SOAPdenovo2	94.890	11439	6475	6449	2292.46	36.41	1.003
	ViaDBG	94.637	2068	987	1352	106.38	<b>100.00</b>	3.581
	VG-Flow	<b>100.000</b>	<b>29685</b>	<b>29790</b>	<b>29685</b>	<b>16.84</b>	<b>100.00</b>	<b>1.000</b>
	VStrains	-	-	-	-	-	-	-
	SPAdes	<b>100.000</b>	<b>29685</b>	<b>29790</b>	<b>29685</b>	10.11	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	<b>29685</b>	<b>29790</b>	<b>29685</b>	30.32	<b>100.00</b>	<b>1.000</b>

**Table S15. Benchmarking results for COVID-19 dataset in 16000X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (30351bp)	IDBA	98.934	29421	29641	29421	925.77	99.5	0.997
	SOAPdenovo2	94.648	11439	6475	6475	1991.80	36.43	1.005
	ViaDBG	96.23	2068	1011	1354	<b>132.56</b>	<b>100.00</b>	3.888
	VG-Flow	<b>99.889</b>	29790	29790	29790	181.26	<b>100.00</b>	0.999
	VStrains	<b>99.889</b>	29790	29790	29790	203.08	<b>100.00</b>	1.999
	SPAdes	<b>99.889</b>	29790	29790	29790	204.76	<b>100.00</b>	0.999
	FC-Virus	<b>99.889</b>	<b>29820</b>	<b>29820</b>	<b>29820</b>	181.09	<b>100.00</b>	<b>1.000</b>
S2 (30169bp)	IDBA	98.999	29305	29641	29305	981.08	99.50	0.998
	SOAPdenovo2	94.675	11438	6475	6448	2010.65	36.43	1.005
	ViaDBG	96.441	2068	1011	1354	<b>127.29</b>	<b>100.00</b>	3.898
	VG-Flow	<b>100.000</b>	29674	29790	29674	232.53	<b>100.00</b>	<b>1.000</b>
	VStrains	<b>100.000</b>	29674	29790	29674	254.43	<b>100.00</b>	2.000
	SPAdes	<b>100.000</b>	29674	29790	29674	256.12	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	<b>29704</b>	<b>29820</b>	<b>29704</b>	242.40	<b>100.00</b>	1.001
S3 (30178bp)	IDBA	99.090	29316	29641	29316	769.51	99.50	0.997
	SOAPdenovo2	94.677	11439	6475	6449	1893.68	36.43	1.005
	ViaDBG	96.389	2068	1011	1354	129.96	<b>100.00</b>	3.899
	VG-Flow	<b>100.000</b>	29685	29790	29685	30.32	<b>100.00</b>	<b>1.000</b>
	VStrains	<b>100.000</b>	29685	29790	29685	8.42	<b>100.00</b>	2.000
	SPAdes	<b>100.000</b>	29685	29790	29685	<b>6.74</b>	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	<b>29715</b>	<b>29820</b>	<b>29715</b>	30.29	<b>100.00</b>	1.001

**Table S16. Benchmarking results for COVID-19 dataset in 18000X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S1 (30351bp)	IDBA	98.924	28479	28701	28479	921.26	96.49	0.998
	SOAPdenovo2	93.200	6662	6475	4567	586.09	40.83	<b>1.000</b>
	ViaDBG	96.424	2068	987	1508	<b>135.35</b>	<b>100.00</b>	3.441
	VG-Flow	<b>99.889</b>	29790	29790	29790	169.52	<b>100.00</b>	1.999
	VStrains	<b>99.889</b>	29790	29790	29790	203.08	<b>100.00</b>	1.999
	SPAdes	<b>99.889</b>	29790	29790	29790	204.76	<b>100.00</b>	0.999
	FC-Virus	<b>99.889</b>	<b>29820</b>	<b>29820</b>	<b>29820</b>	171.03	<b>100.00</b>	<b>1.000</b>
S2 (30169bp)	IDBA	98.989	28452	28701	28452	983.30	96.49	0.999
	SOAPdenovo2	93.223	6662	6475	4567	586.78	40.83	<b>1.000</b>
	ViaDBG	96.637	2068	987	1508	<b>151.65</b>	<b>100.00</b>	3.449
	VG-Flow	<b>100.000</b>	29674	29790	29674	220.73	<b>100.00</b>	2.000
	VStrains	<b>100.000</b>	29674	29790	29674	254.43	<b>100.00</b>	2.000
	SPAdes	<b>100.000</b>	29674	29790	29674	256.12	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	<b>29704</b>	<b>29820</b>	<b>29704</b>	232.30	<b>100.00</b>	1.001



**Table S16\_C1. Continuation of Table S16 for COVID-19 dataset in 18000X depth.**

Strain ID	Assembler	Genome fraction	Largest alignment	N50	NGA50	Errors per 100 kbp	Aligned reads (%)	Duplication ratio
S3 (30178bp)	IDBA	99.080	28453	28701	28453	775.35	96.49	0.998
	SOAPdenovo2	93.242	6662	6475	4567	485.55	40.83	<b>1.000</b>
	ViaDBG	96.584	2068	987	1508	117.26	<b>100.00</b>	3.45
	VG-Flow	<b>100.000</b>	29685	29790	29685	42.11	<b>100.00</b>	2.000
	VStrains	<b>100.000</b>	29685	29790	29685	8.42	<b>100.00</b>	2.000
	SPAdes	<b>100.000</b>	29685	29790	29685	<b>6.74</b>	<b>100.00</b>	<b>1.000</b>
	FC-Virus	<b>100.000</b>	<b>29715</b>	<b>29820</b>	<b>29715</b>	40.38	<b>100.00</b>	1.001

Similar to previous studies on strain-level genome assembly, we use metaQUAST to assess the assembly results of SPAdes. The performance of SPAdes in recovering strain-level genomes is summarized in Table S17.

**Table S17. Performance of SPAdes in recovering all strains**

Dataset	Number of contigs	Genome fraction	Duplication ratio	N50	NGA50	Error rate (mis+indel+N's)
HIV	8	90.104	1.001	8818	8170	6.89
POLIO	8	44.106	1.000	4216	5828	20.28
HCV	10	90.591	1.000	8680	8420	4.35
ZIKV	50	57.836	1.005	3473	6212	24.58
COVID-19	27	34.709	1.040	1366	-	105.19
HIV_LABMIX	33	50.108	1.049	727	637	197.29

### 3. Command Lines used for benchmarking

For benchmarking, we used several existing tools, which are listed below. All tools were tested with their default settings unless stated otherwise.

- **QUAST:** version 5.2.0

```
python quast.py contig.fasta -r reference.fasta -l forward.fastq -2 reverse.fastq -o
<output dir> --min-contig 350 --min-identity 90
```

```
python metaquast.py <contig1.fasta> <contig2.fasta> . . . <contigm.fasta> -r  
<reference1.fasta>,<reference2.fasta>, . . . ,<referencen.fasta> -o <output dir>  
--unique-mapping
```

- **IDBA:** version 1.1.1

```
fq2fa --merge --filter read_1.fq read_2.fq read.fa
```

```
idba -r read.fa -o <output dir>
```

- **SOAPdenovo2:** version 2.0.4

```
SOAPdenovo-63mer all -s example.config -K 63 -R -o N50 1>ass.log 2>ass.err
```

- **viaDBG:** version 1.0

```
viaDBG -p <paired end dir> -o <output> -u <unitig> -k 127 -c dsk -n -t 16  
--postprocess
```

- **VG-Flow:** version 0.0.4

```
python build_graph_msga.py -f reads1.fastq -r reads2.fastq -c scaffolds.fasta -t 32 -vg  
vg
```

```
python vg-flow.py -m 100 -c 200 --greedy mode=all node abundance.txt contig  
graph.final.gfa
```

- **SPAdes:** version 3.15.5

```
python spades.py -1 reads1.fastq -2 reads2.fastq -o <output dir>
```

- **FC-Virus:** version 1.0

```
fc-virus -t fq --left reads1.fastq --right reads2.fastq -o <output dir>
```