

Comparison of Optical Genome Mapping With Conventional Diagnostic Methods for Structural Variant Detection in Hematologic Malignancies

Yeeun Shim, M.S.^{1,2*}, Yu-Kyung Koo, M.D.^{3*}, Saeam Shin, M.D., Ph.D.³, Seung-Tae Lee, M.D., Ph.D.^{3,4}, Kyung-A Lee, M.D., Ph.D.³, and Jong Rak Choi, M.D., Ph.D.^{3,4}

¹Brain Korea 21 PLUS Project for Medical Science, Yonsei University, Seoul, Korea;

²MDxK (Molecular Diagnostics Korea), Inc., Gwacheon, Korea; ³Department of Laboratory Medicine, Yonsei University College of Medicine, Seoul, Korea; ⁴Dxome Co., Ltd., Seongnam, Korea.

Supplemental Data Table S1. Numbers of CNVs and structural variants pre- and post-user filtering

Sample ID	CNV Manufacturer default filtering-min size: 500 kbp-confidence: ≥ 0.99	No user filtering	Deletion		Duplication		Insertion		Inversion		Translocation	
			Manufacturer default filtering-self molecule count: ≥ 5	User filtering-confidence: ≥ 0.9 -% in control database: 1%	Manufacturer default filtering-self molecule count: ≥ 5	User filtering-% in control database: 1%	Manufacturer default filtering-self molecule count: ≥ 5	User filtering-confidence: ≥ 0.9 -% in control database: 1%	Manufacturer default filtering-self molecule count: ≥ 5	User filtering-% in control database: 1%	Manufacturer default filtering-self molecule count: ≥ 5	Manual filtering-% in control database: 1%
1	28	-	1,682	60	69	9	4,401	27	60	0	6	6
2	2	-	1,727	49	51	7	4,410	23	62	2	2	2
3	2	-	1,627	58	60	1	4,453	31	55	1	2	2
4	92	-	1,677	57	70	5	4,475	46	73	3	15	15
5	111	-	1,382	43	23	2	3,280	19	39	3	2	2
6	74	-	1,379	50	41	6	3,555	20	42	4	1	1
7	20	-	1,628	50	63	2	4,395	34	51	0	2	2
8	51	-	1,612	53	55	5	4,425	32	54	6	0	0
9	21	-	1,563	49	59	8	4,266	27	58	3	1	1
10	3	-	1,623	57	65	4	4,315	44	65	0	3	3
11	35	-	1,671	48	93	6	4,511	29	61	2	3	3
12	520	-	1,359	49	31	1	3,368	30	46	2	5	5
13	29	-	1,637	46	58	10	4,392	39	58	0	2	2
14	80	-	1,630	43	80	7	4,376	32	61	5	0	0
15	24	-	1,706	59	97	10	4,386	25	52	1	0	0
16	88	-	1,310	41	32	1	3,309	26	45	3	0	0
17	7	-	1,740	64	95	8	4,511	31	74	6	2	2
18	12	-	1,701	64	68	4	4,408	36	73	2	3	3
19	32	-	1,681	60	42	2	4,425	31	65	2	1	1
20	65	-	1,633	75	51	3	4,404	36	68	4	5	5

21	5	-	1,678	75	70	4	4,446	26	54	5	8	8
22	76	-	1,605	63	49	3	4,217	25	61	2	1	1
23	1	-	1,637	60	65	5	4,378	31	58	1	0	0
24	30	-	1,490	59	29	0	3,928	20	57	2	3	3
25	8	-	1,630	44	67	6	4,391	34	58	0	1	1
26	49	-	1,663	69	51	3	4,275	43	58	5	7	7
27	37	-	1,634	48	74	4	4,369	28	60	2	1	1
Mean (range)	55.63 (1 – 520)		1,603.89 (1,310 – 1,740)	55.30 (41 – 75)	59.56 (23 – 97)	4.67 (0 – 10)	4,224.78 (3,280 – 4,511)	30.56 (19 – 46)	58.07 (39 – 74)	2.44 (0 – 6)	2.81 (0 – 15)	2.81 (0 – 15)

Abbreviation: CNV, copy number variant.

Supplemental Data Table S2. OGM data quality

Sample ID	Total DNA (Gbp)	Average N50 (≥ 150 kbp) (Mbp)	Average label density	Average mapping rate	Estimated effective coverage	PLV (%)	NLV (%)
	>320*	NA*	14–17*	$\geq 60\%*$	>80 \times or >100 $\times*$	<10%*	<15%*
1	1,269.8	0.227	16.03	68.7%	265.39	4%	8%
2	1,307.1	0.249	16.24	78.8%	326.06	3%	9%
3	1,307.7	0.238	15.63	81.3%	343.43	3%	8%
4	1,327.7	0.212	13.62†	68.9%	285.22	3%	18%
5	717.9	0.204	15.54	30.6%	66.46	3%	9%
6	1,080.0	0.213	15.45	29.9%	102.65	3%	12%
7	1,328.8	0.234	14.83	67.8%	282.79	3%	13%
8	1,250.9	0.199	15.59	43.1%	174.16	4%	13%
9	1,645.3	0.286	19.85	37.4%	198.78	3%	15%
10	979.7	0.181	14.67	77.9%	234.79	3%	10%
11	1,301.9	0.217	14.47	63.1%	256.59	3%	18%
12	1,306.8	0.204	11.84	24.5%	99.11	2%	30%
13	1,510.7	0.211	15.40	65.0%	308.02	4%	10%
14	1,303.3	0.208	13.28	50.9%	203.75	3%	15%
15	1,308.2	0.210	14.68	83.5%	337.92	4%	10%
16	956.5	0.222	16.76	28.4%	87.75	3%	10%
17	1,333.9	0.241	14.89	87.3%	361.99	2%	11%
18	1,310.3	0.205	14.89	87.6%	352.74	3%	10%
19	1,099.6	0.207	16.59	61.1%	217.03	4%	11%

20	1,409.3	0.191	14.48	73.1%	322.79	3%	11%
21	1,515.4	0.219	15.30	77.2%	367.14	4%	10%
22	1,517.1	0.194	15.40	58.2%	277.26	3%	9%
23	589.7	0.200	14.97	70.4%	128.24	2%	9%
24	786.7	0.204	14.05	38.9%	98.85	3%	16%
25	1,300.9	0.198	14.94	79.8%	323.24	3%	9%
26	916.6	0.228	17.02	48.0%	139.86	4%	8%
27	1,280.0	0.202	15.17	43.0%	173.59	3%	14%
Mean	1,187.2	0.208	15.25	65.9%	251.16	3%	11%
SD	3,03.72	0.02	0.86	0.17	102.44	0.01	0.02

*QC criteria as per the manufacturer's guidelines.

†Values outside the optimal range mentioned in the manufacturer's guidelines are indicated in bold.

Total DNA (Gbp): total amount of DNA.

Average N50 (≥ 150 kbp) (Mbp): the molecule length N50 for all molecules that are ≥ 150 kbp in length.

Average label density (/100 kbp): the number of labels detected by the image detection algorithm per 100-kbp DNA length for molecules ≥ 150 kbp.

Average mapping rate: the percentage of molecules that map to the reference for molecules ≥ 150 kbp.

Estimated effective coverage: average mapping rate \times total DNA/length of the reference.

PLV %: percentage of molecule labels absent in reference labels.

NLV %: percentage of reference labels absent in molecule labels.

Abbreviations: OGM, optical genome mapping; NA, not available.