

Supplementary Tables

technology	instrument	Sample ID	Depth of Coverage			Read Length distribution		
			median	mean	sd	median	mean	sd
Illumina	Xten	ILMN1	42	40.9	12.7	293	308.8	117.2
Illumina	Xten	ILMN2	119	122.6	667.2	301	319.6	117.0
MGI	G400	MGI1	45	44.6	12.1	371	378.3	84.8
Nanopore	PromethION R9.4	ONT1	333	27.3	7.8	4000	7177.8	10531.9
Nanopore	PromethION R9.4	ONT2	68	53.2	153.6	4860	17316.0	36739.6
Nanopore	PromethION R10.4.1	ONT3	54	42.7	192.7	5941	9572.0	10785.6
Nanopore	PromethION R10.4.1	ONT4	52	58.7	234.1	9017	14245.4	14033.9
PacBio	Sequel II	PB1	24	28.1	99.0	13385	13481.7	1192.3
PacBio	Sequel II	PB2	62	52.5	164.8	13283	14268.1	2874.3
PacBio	Revio	PB3	45	37.6	138.4	13617	14235.3	3224.4
10X Genomics		10X	26	27.3	53.2	34000	45884.7	41749.8

Table S1 - Coverage, Insert Size, and Read Length Metrics for HG002

Summary of sequencing metrics for the HG002 sample across all technologies, including mean genome coverage, insert size distributions (for paired-end technologies), and read length distributions (for long-read platforms). These metrics provide a comparison of library characteristics and sequencing performance across Illumina, MGI, ONT, PacBio, and synthetic long-read (10X Genomics) technologies.

Regions	Variant Type	Zygosity	T2T v1.1	CMRG v1.0
Easy-to-map	SNVs	Heterozygous	1726233	7017
		Homozygous alt	1149487	4658
		Total	2875720	11675
	Indels	Heterozygous	103716	465
		Homozygous alt	75310	327
		Total	179026	792
Difficult-to-map	SNVs	Heterozygous	460312	3273
		Homozygous alt	303912	2921

	Indels	Total	764224	6194
		Heterozygous	538693	1492
		Homozygous alt	228518	1109
		Total	767211	2601

Table S2 - Variant Counts in T2T and CMRG Validation Sets

Counts of variants in the T2T and CMRG validation sets are stratified by variant type (SNVs and indels), genomic context (easy-to-map vs. difficult-to-map regions), and zygosity (heterozygous vs. homozygous alternative variants). This breakdown highlights the distribution of variant types across genomic complexity and allelic state.

caller	Technology	T2T version 1.1			CMRG version 1.0.0		
		Precision	Recall	F1 score	Precision	Recall	F1 score
Combined							
longranger	10X	96.77%	90.96%	93.77%	96.69%	89.39%	92.90%
clair3	ilmn	98.07%	96.65%	97.35%	96.67%	93.20%	94.90%
deepvariant	ilmn	99.33%	97.59%	98.45%	99.02%	94.19%	96.54%
dragen	ilmn	99.41%	98.83%	99.12%	98.93%	96.24%	97.57%
gatk3	ilmn	98.38%	97.44%	97.91%	96.38%	93.94%	95.14%
gatk4	ilmn	98.52%	97.22%	97.87%	96.65%	93.76%	95.18%
megabolt	ilmn	98.40%	96.32%	97.34%	96.88%	93.35%	95.08%
parabricks	ilmn	98.39%	96.80%	97.59%	96.66%	93.60%	95.10%
clair3	mgi	98.19%	96.99%	97.58%	96.78%	93.00%	94.85%
deepvariant	mgi	99.46%	97.90%	98.68%	98.89%	94.22%	96.50%
dragen	mgi	99.45%	98.54%	98.99%	98.69%	95.65%	97.15%
gatk3	mgi	99.01%	97.99%	98.50%	96.72%	93.58%	95.12%
gatk4	mgi	99.05%	97.82%	98.43%	96.93%	93.38%	95.12%
megabolt	mgi	98.91%	97.35%	98.12%	97.11%	93.21%	95.12%
clair3	R10	97.45%	92.33%	94.82%	94.58%	93.96%	94.27%
deepvariant	R10	96.67%	93.58%	95.10%	97.18%	94.97%	96.06%
clair3	R9	90.38%	85.15%	87.68%	89.90%	86.75%	88.30%
deepvariant	R9	72.72%	83.67%	77.81%	74.29%	84.89%	79.23%
nanocaller	R9	52.79%	78.12%	63.01%	50.59%	78.24%	61.45%
clair3	revio	98.65%	98.67%	98.66%	96.64%	98.40%	97.51%

deepvariant	revio	99.43%	99.24%	99.34%	99.09%	98.65%	98.87%
clair3	sequel	94.81%	96.32%	95.56%	94.86%	96.73%	95.79%
deepvariant	sequel	94.57%	95.77%	95.17%	96.28%	96.37%	96.33%
gatk4	sequel	76.91%	93.95%	84.58%	78.89%	93.75%	85.68%
		T2T version 1.1			CMRG version 1.0.0		
caller	Technology	Precision	Recall	F1 score	Precision	Recall	F1 score
SNVs							
longranger	10X	99.14%	96.98%	98.04%	98.39%	93.30%	95.78%
clair3	ilmn	99.08%	98.01%	98.54%	97.31%	94.06%	95.66%
deepvariant	ilmn	99.71%	98.31%	99.00%	99.24%	94.58%	96.86%
dragen	ilmn	99.72%	99.09%	99.40%	99.19%	96.52%	97.83%
gatk3	ilmn	98.94%	98.32%	98.63%	96.39%	94.23%	95.30%
gatk4	ilmn	99.09%	98.22%	98.65%	96.71%	94.11%	95.39%
megabolt	ilmn	99.21%	98.08%	98.64%	97.08%	93.97%	95.50%
parabrick	ilmn	99.13%	98.20%	98.66%	96.76%	94.06%	95.39%
clair3	mgi	99.01%	98.09%	98.55%	97.44%	93.91%	95.65%
deepvariant	mgi	99.73%	98.44%	99.08%	99.06%	94.76%	96.86%
dragen	mgi	99.70%	98.88%	99.29%	98.92%	96.11%	97.50%
gatk3	mgi	99.20%	98.48%	98.84%	96.72%	94.01%	95.35%
gatk4	mgi	99.30%	98.40%	98.85%	97.00%	93.85%	95.40%
megabolt	mgi	99.38%	98.35%	98.86%	97.17%	93.80%	95.45%
clair3	R10	99.24%	98.31%	98.77%	95.70%	97.81%	96.74%
deepvariant	R10	99.53%	98.87%	99.20%	99.37%	98.34%	98.85%
clair3	R9	98.56%	97.00%	97.78%	96.28%	96.28%	96.28%
deepvariant	R9	93.68%	95.63%	94.65%	92.96%	94.56%	93.75%
nanocaller	R9	88.31%	91.75%	90.00%	84.79%	89.94%	87.29%
clair3	revio	99.47%	99.60%	99.53%	96.85%	98.98%	97.90%
deepvariant	revio	99.77%	99.67%	99.72%	99.40%	99.03%	99.21%
clair3	sequel	99.23%	99.54%	99.38%	97.32%	98.68%	97.99%
deepvariant	sequel	99.55%	99.49%	99.52%	99.24%	98.61%	98.92%
gatk4	sequel	98.68%	98.78%	98.73%	98.83%	96.94%	97.88%
		T2T version 1.1			CMRG version 1.0.0		
caller	Technology	Precision	Recall	F1 score	Precision	Recall	F1 score

Indels							
longranger	10X	84.25%	65.62%	73.78%	86.91%	70.26%	77.70%
clair3	ilmn	93.74%	90.90%	92.30%	93.49%	88.98%	91.18%
deepvariant	ilmn	97.70%	94.56%	96.10%	97.89%	92.25%	94.99%
dragen	ilmn	98.14%	97.78%	97.96%	97.70%	94.89%	96.27%
gatk3	ilmn	95.94%	93.76%	94.84%	96.33%	92.49%	94.37%
gatk4	ilmn	96.06%	92.99%	94.50%	96.37%	92.06%	94.16%
megabolt	ilmn	94.78%	88.90%	91.74%	95.89%	90.34%	93.03%
parabricks	ilmn	95.15%	90.93%	92.99%	96.12%	91.35%	93.67%
clair3	mgi	94.69%	92.32%	93.49%	93.49%	88.54%	90.95%
deepvariant	mgi	98.33%	95.63%	96.96%	98.04%	91.60%	94.71%
dragen	mgi	98.38%	97.12%	97.75%	97.55%	93.39%	95.43%
gatk3	mgi	98.19%	95.89%	97.02%	96.68%	91.46%	94.00%
gatk4	mgi	97.99%	95.36%	96.66%	96.58%	91.10%	93.76%
megabolt	mgi	96.87%	93.17%	94.98%	96.81%	90.34%	93.46%
clair3	R10	87.67%	67.15%	76.05%	88.04%	75.09%	81.05%
deepvariant	R10	82.80%	71.30%	76.62%	85.65%	78.47%	81.90%
clair3	R9	46.03%	35.21%	39.90%	50.55%	40.11%	44.73%
deepvariant	R9	19.60%	33.25%	24.66%	21.40%	37.60%	27.28%
nanocaller	R9	6.21%	20.73%	9.56%	5.36%	21.04%	8.55%
clair3	revio	95.18%	94.75%	94.97%	95.61%	95.58%	95.59%
deepvariant	revio	98.03%	97.43%	97.73%	97.58%	96.75%	97.16%
clair3	sequel	77.34%	82.73%	79.94%	83.22%	87.18%	85.15%
deepvariant	sequel	74.95%	80.14%	77.45%	82.44%	85.43%	83.91%
gatk4	sequel	34.24%	73.64%	46.74%	35.46%	78.16%	48.79%

Table S3 - Precision, Recall, and F1-Score Performance Across Sequencing Technologies

Benchmarking metrics were calculated using hap.py for all analyzed sequencing technologies based on the T2T and CMRG validation datasets. For each metric, precision, recall, and F1-score, the top-performing variant caller is highlighted in bold.

		Processing walltime
--	--	---------------------

Technology	Caller	bam processing	variant calling	total time
Illumina				
	GATK3	30:43:46	5:01:58	35:45:44
	GATK4	30:43:46	2:40:35	33:24:21
	DeepVariant	30:43:46	1:01:38	31:45:24
	Dragen	0:19:54	0:13:18	0:33:12
	MegaBolt	3:07:39	1:36:38	4:44:17
	Parabricks	0:56:50	0:20:03	1:16:53
MGI				
	GATK3	41:57:11	4:32:00	46:29:11
	GATK4	41:57:11	7:05:47	49:02:58
	DeepVariant	41:57:11	0:52:58	42:50:09
	Dragen	0:20:16	0:13:41	0:33:57
	MegaBolt	2:11:54	5:28:38	7:40:32

Table S4 - Computational Resource Usage for Short-Read Sequencing Technologies

Summary of computational requirements for short-read sequencing (SRS) platforms, including CPU usage, memory consumption, and runtime for alignment and variant calling workflows. Metrics are reported for Illumina and MGI datasets across the analyzed benchmarks.

Ref size ranges	Deletions		Insertions	
	T2T	CMRG	T2T	CMRG
[0,30)	13579	0	13992	0
[30, 50)	7446	0	7483	0
[50, 250)	8781	56	11773	53
[250, 500)	2487	19	4450	27
[500, 6000)	1739	14	5268	25
[6000, 150000)	314	0	666	0
All ranges	34346	89	43632	105

Table S5 - Stratification of Structural Variant Counts by Size

Counts of deletions and insertions (SVs) are shown for GIAB's T2T and CMRG datasets, stratified into six size categories: 0-30 bp, 31-50 bp, 51-250 bp, 251-500 bp, 501-6,000 bp, and 6,000-150,000 bp. This detailed breakdown highlights the distribution of SVs across size ranges and benchmark datasets.

type	technology	platform	callers	Precision	Recall	F1
T2T						
del	R10	LRS	cutesv2	98.95%	90.43%	94.50%
del	R10	LRS	dysgu	99.80%	80.20%	88.93%
del	R10	LRS	sniffles2	98.61%	90.25%	94.25%
del	R10	LRS	svim	98.69%	90.59%	94.47%
del	R9	LRS	cutesv2	98.87%	82.31%	89.83%
del	R9	LRS	dysgu	99.29%	89.01%	93.87%
del	R9	LRS	sniffles2	98.56%	88.34%	93.17%
del	R9	LRS	svim	98.09%	90.79%	94.30%
del	revio	LRS	cutesv2	99.36%	88.09%	93.39%
del	revio	LRS	dysgu	99.43%	90.94%	95.00%
del	revio	LRS	pbsv	98.40%	89.14%	93.54%
del	revio	LRS	sawfish	98.68%	92.79%	95.65%
del	revio	LRS	sniffles2	99.13%	91.69%	95.27%
del	revio	LRS	svim	99.13%	91.99%	95.43%
del	sequel	LRS	cutesv2	99.40%	82.38%	90.09%
del	sequel	LRS	dysgu	99.60%	89.89%	94.49%
del	sequel	LRS	pbsv	97.82%	89.83%	93.66%
del	sequel	LRS	sawfish	98.67%	92.42%	95.44%
del	sequel	LRS	sniffles2	99.08%	91.02%	94.88%
del	sequel	LRS	svim	99.25%	91.56%	95.25%
del	10X	synthetic	longranger	96.25%	35.65%	52.02%
del	ilmn	SRS	dragen	91.90%	63.31%	74.97%
del	ilmn	SRS	dysgu	96.11%	38.50%	54.98%
del	ilmn	SRS	metasv	97.05%	37.42%	54.01%

del	mgi	SRS	dragen	94.78%	50.31%	65.73%
del	mgi	SRS	dysgu	96.23%	39.48%	55.99%
del	mgi	SRS	metasv	96.25%	39.33%	55.84%
ins	R10	LRS	cutesv2	98.69%	93.52%	96.04%
ins	R10	LRS	dysgu	99.44%	62.37%	76.66%
ins	R10	LRS	sniffles2	99.15%	93.01%	95.98%
ins	R10	LRS	svim	99.06%	91.83%	95.31%
ins	R9	LRS	cutesv2	98.01%	84.82%	90.94%
ins	R9	LRS	dysgu	97.62%	86.09%	91.49%
ins	R9	LRS	sniffles2	98.40%	88.99%	93.46%
ins	R9	LRS	svim	98.12%	83.93%	90.47%
ins	revio	LRS	cutesv2	99.17%	86.85%	92.60%
ins	revio	LRS	dysgu	98.59%	93.26%	95.85%
ins	revio	LRS	pbsv	99.39%	92.70%	95.93%
ins	revio	LRS	sawfish	99.27%	95.70%	97.45%
ins	revio	LRS	sniffles2	99.27%	93.58%	96.34%
ins	revio	LRS	svim	99.47%	93.05%	96.15%
ins	sequel	LRS	cutesv2	98.98%	81.29%	89.27%
ins	sequel	LRS	dysgu	98.79%	92.54%	95.56%
ins	sequel	LRS	pbsv	99.11%	90.33%	94.52%
ins	sequel	LRS	sawfish	99.23%	95.73%	97.45%
ins	sequel	LRS	sniffles2	99.20%	92.84%	95.91%
ins	sequel	LRS	svim	99.25%	92.33%	95.67%
ins	ilmn	SRS	dragen	96.93%	36.81%	53.36%
ins	ilmn	SRS	dysgu	87.34%	10.82%	19.26%
ins	ilmn	SRS	metasv	98.04%	10.16%	18.42%
ins	mgi	SRS	dragen	98.08%	23.08%	37.37%
ins	mgi	SRS	dysgu	87.66%	11.19%	19.84%
ins	mgi	SRS	metasv	99.00%	13.38%	23.58%
CMRG						
del	R10	LRS	cutesv2	93.41%	92.55%	92.98%
del	R10	LRS	dysgu	98.72%	82.98%	90.17%
del	R10	LRS	sniffles2	90.53%	93.62%	92.05%
del	R10	LRS	svim	91.58%	94.68%	93.10%
del	R9	LRS	cutesv2	92.86%	85.11%	88.81%

del	R9	LRS	dysgu	96.63%	92.55%	94.55%
del	R9	LRS	sniffles2	95.60%	94.68%	95.14%
del	R9	LRS	svim	91.49%	93.62%	92.54%
del	revio	LRS	cutesv2	96.43%	88.30%	92.18%
del	revio	LRS	dysgu	100.00%	93.62%	96.70%
del	revio	LRS	pbsv	97.59%	88.30%	92.71%
del	revio	LRS	sawfish	100.00%	94.68%	97.27%
del	revio	LRS	sniffles2	94.57%	93.62%	94.09%
del	revio	LRS	svim	95.60%	94.68%	95.14%
del	sequel	LRS	cutesv2	96.10%	80.85%	87.82%
del	sequel	LRS	dysgu	100.00%	91.49%	95.56%
del	sequel	LRS	pbsv	95.60%	93.62%	94.60%
del	sequel	LRS	sawfish	100.00%	93.62%	96.70%
del	sequel	LRS	sniffles2	95.40%	90.43%	92.85%
del	sequel	LRS	svim	96.63%	93.62%	95.10%
del	10X	synthetic	longranger	87.50%	38.30%	53.28%
del	ilmn	SRS	dragen	96.23%	54.26%	69.39%
del	ilmn	SRS	dysgu	93.33%	45.74%	61.40%
del	ilmn	SRS	metasv	87.76%	45.74%	60.14%
del	mgi	SRS	dragen	91.49%	45.74%	60.99%
del	mgi	SRS	dysgu	95.56%	46.81%	62.84%
del	mgi	SRS	metasv	90.91%	42.55%	57.97%
ins	R10	LRS	cutesv2	89.52%	97.25%	93.23%
ins	R10	LRS	dysgu	94.19%	74.31%	83.08%
ins	R10	LRS	sniffles2	90.38%	96.33%	93.26%
ins	R10	LRS	svim	90.10%	92.66%	91.36%
ins	R9	LRS	cutesv2	85.56%	81.65%	83.56%
ins	R9	LRS	dysgu	81.63%	80.73%	81.18%
ins	R9	LRS	sniffles2	88.12%	92.66%	90.33%
ins	R9	LRS	svim	67.42%	81.65%	73.86%
ins	revio	LRS	cutesv2	92.94%	85.32%	88.97%
ins	revio	LRS	dysgu	93.81%	92.66%	93.23%
ins	revio	LRS	pbsv	90.10%	93.58%	91.81%
ins	revio	LRS	sawfish	95.24%	97.25%	96.23%
ins	revio	LRS	sniffles2	93.14%	97.25%	95.15%

ins	revio	LRS	svim	92.63%	91.74%	92.19%
ins	sequel	LRS	cutesv2	91.95%	86.24%	89.00%
ins	sequel	LRS	dysgu	93.75%	93.58%	93.66%
ins	sequel	LRS	pbsv	94.85%	93.58%	94.21%
ins	sequel	LRS	sawfish	95.19%	97.25%	96.21%
ins	sequel	LRS	sniffles2	92.00%	94.50%	93.23%
ins	sequel	LRS	svim	91.49%	88.07%	89.75%
ins	ilmn	SRS	dragen	72.88%	39.45%	51.19%
ins	ilmn	SRS	dysgu	44.44%	14.68%	22.07%
ins	ilmn	SRS	metasv	66.67%	15.60%	25.28%
ins	mgi	SRS	dragen	71.79%	25.69%	37.84%
ins	mgi	SRS	dysgu	51.43%	16.51%	25.00%
ins	mgi	SRS	metasv	79.31%	19.27%	31.00%

Table S6 - Discovery Metrics for Structural Variants Across Technologies

Precision, recall, and F1-score metrics were calculated for deletions (DEL) and insertions (INS) in the T2T and CMRG validation sets across all tested SRS and LRS SV callers. These metrics reflect the discovery performance of each caller for both SV types and validation datasets.

Type	Technology	Platform	Callers	Precision	Recall	F1
T2T						
del	R10	LRS	cutesv2	92.35%	88.27%	90.27%
del	R10	LRS	dysgu	87.47%	85.26%	86.35%
del	R10	LRS	sniffles2	92.42%	88.54%	90.44%
del	R10	LRS	svim	91.13%	87.38%	89.22%
del	R9	LRS	cutesv2	89.80%	84.98%	87.32%
del	R9	LRS	dysgu	86.19%	84.48%	85.33%
del	R9	LRS	sniffles2	90.51%	86.59%	88.51%
del	R9	LRS	svim	90.76%	87.20%	88.94%
del	revio	LRS	cutesv2	89.60%	84.17%	86.80%
del	revio	LRS	dysgu	91.50%	88.55%	90.00%
del	revio	LRS	pbsv	91.59%	87.72%	89.61%
del	revio	LRS	sawfish	95.40%	93.25%	94.31%

del	revio	LRS	sniffles2	93.16%	89.62%	91.36%
del	revio	LRS	svim	91.72%	87.21%	89.41%
del	sequel	LRS	cutesv2	88.75%	82.99%	85.77%
del	sequel	LRS	dysgu	90.72%	87.83%	89.25%
del	sequel	LRS	pbsv	90.47%	86.73%	88.56%
del	sequel	LRS	sawfish	95.60%	93.66%	94.62%
del	sequel	LRS	sniffles2	92.42%	88.84%	90.59%
del	sequel	LRS	svim	91.50%	86.94%	89.16%
del	ilmn	SRS	dragen	89.32%	86.47%	87.87%
del	ilmn	SRS	dysgu	77.75%	76.43%	77.08%
del	ilmn	SRS	metasv	85.32%	84.89%	85.11%
del	mgi	SRS	dragen	89.96%	87.04%	88.48%
del	mgi	SRS	dysgu	73.14%	71.46%	72.29%
del	mgi	SRS	metasv	85.43%	84.90%	85.16%
del	10X	synthetic	longranger	86.61%	85.29%	85.95%
ins	R10	LRS	cutesv2	75.83%	63.19%	68.94%
ins	R10	LRS	dysgu	75.03%	74.85%	74.94%
ins	R10	LRS	sniffles2	75.76%	63.69%	69.20%
ins	R10	LRS	svim	78.94%	69.07%	73.68%
ins	R9	LRS	cutesv2	68.38%	55.68%	61.38%
ins	R9	LRS	dysgu	71.74%	60.61%	65.71%
ins	R9	LRS	sniffles2	73.19%	61.25%	66.69%
ins	R9	LRS	svim	77.78%	71.02%	74.25%
ins	revio	LRS	cutesv2	67.07%	53.51%	59.53%
ins	revio	LRS	dysgu	72.65%	61.36%	66.53%
ins	revio	LRS	pbsv	79.39%	68.21%	73.37%
ins	revio	LRS	sawfish	87.44%	79.28%	83.16%
ins	revio	LRS	sniffles2	76.72%	65.10%	70.44%
ins	revio	LRS	svim	77.88%	66.67%	71.84%
ins	sequel	LRS	cutesv2	65.83%	52.53%	58.43%
ins	sequel	LRS	dysgu	72.50%	61.56%	66.58%
ins	sequel	LRS	pbsv	78.55%	68.45%	73.15%
ins	sequel	LRS	sawfish	87.63%	79.48%	83.35%
ins	sequel	LRS	sniffles2	77.05%	65.47%	70.79%
ins	sequel	LRS	svim	77.62%	66.57%	71.67%

ins	ilmn	SRS	dragen	71.34%	61.98%	66.33%
ins	ilmn	SRS	dysgu	54.47%	50.79%	52.57%
ins	ilmn	SRS	metasv	70.69%	67.00%	68.80%
ins	mgj	SRS	dragen	73.39%	64.38%	68.59%
ins	mgj	SRS	dysgu	52.23%	48.34%	50.21%
ins	mgj	SRS	metasv	72.54%	67.99%	70.19%
CMRG						
del	R10	LRS	cutesv2	96.47%	94.25%	95.35%
del	R10	LRS	dysgu	85.71%	84.62%	85.16%
del	R10	LRS	sniffles2	95.35%	93.18%	94.25%
del	R10	LRS	svim	95.40%	93.26%	94.32%
del	R9	LRS	cutesv2	94.87%	92.50%	93.67%
del	R9	LRS	dysgu	83.72%	82.76%	83.24%
del	R9	LRS	sniffles2	96.55%	94.38%	95.45%
del	R9	LRS	svim	96.51%	94.32%	95.40%
del	revio	LRS	cutesv2	95.06%	92.77%	93.90%
del	revio	LRS	dysgu	91.86%	89.77%	90.80%
del	revio	LRS	pbsv	90.12%	87.95%	89.02%
del	revio	LRS	sawfish	97.75%	97.75%	97.75%
del	revio	LRS	sniffles2	96.55%	95.45%	96.00%
del	revio	LRS	svim	95.40%	93.26%	94.32%
del	sequel	LRS	cutesv2	97.30%	94.74%	96.00%
del	sequel	LRS	dysgu	90.59%	89.53%	90.06%
del	sequel	LRS	pbsv	90.80%	89.77%	90.29%
del	sequel	LRS	sawfish	98.86%	98.86%	98.86%
del	sequel	LRS	sniffles2	96.39%	94.12%	95.24%
del	sequel	LRS	svim	95.35%	93.18%	94.25%
del	ilmn	SRS	dragen	94.12%	94.12%	94.12%
del	ilmn	SRS	dysgu	78.57%	76.74%	77.65%
del	ilmn	SRS	metasv	62.79%	62.79%	62.79%
del	mgj	SRS	dragen	97.67%	97.67%	97.67%
del	mgj	SRS	dysgu	74.42%	72.73%	73.56%
del	mgj	SRS	metasv	62.50%	62.50%	62.50%
del	10X	synthetic	longranger	77.14%	75.00%	76.06%
ins	R10	LRS	cutesv2	87.23%	77.36%	82.00%

ins	R10	LRS	dysgu	66.67%	66.67%	66.67%
ins	R10	LRS	sniffles2	87.23%	78.10%	82.41%
ins	R10	LRS	svim	83.52%	75.25%	79.17%
ins	R9	LRS	cutesv2	79.22%	68.54%	73.49%
ins	R9	LRS	dysgu	86.25%	78.41%	82.14%
ins	R9	LRS	sniffles2	86.52%	76.24%	81.05%
ins	R9	LRS	svim	79.78%	79.78%	79.78%
ins	revio	LRS	cutesv2	78.48%	66.67%	72.09%
ins	revio	LRS	dysgu	83.52%	75.25%	79.17%
ins	revio	LRS	pbsv	86.81%	77.45%	81.87%
ins	revio	LRS	sawfish	94.00%	88.68%	91.26%
ins	revio	LRS	sniffles2	78.95%	70.75%	74.63%
ins	revio	LRS	svim	85.23%	75.00%	79.79%
ins	sequel	LRS	cutesv2	80.00%	68.09%	73.56%
ins	sequel	LRS	dysgu	84.44%	74.51%	79.17%
ins	sequel	LRS	pbsv	85.87%	77.45%	81.44%
ins	sequel	LRS	sawfish	92.93%	86.79%	89.76%
ins	sequel	LRS	sniffles2	86.96%	77.67%	82.05%
ins	sequel	LRS	svim	87.21%	78.13%	82.42%
ins	ilmn	SRS	dragen	86.05%	86.05%	86.05%
ins	ilmn	SRS	dysgu	68.75%	68.75%	68.75%
ins	ilmn	SRS	metasv	55.56%	58.82%	57.14%
ins	mgi	SRS	dragen	85.71%	85.71%	85.71%
ins	mgi	SRS	dysgu	83.33%	83.33%	83.33%
ins	mgi	SRS	metasv	56.52%	61.90%	59.09%

Table S7 - Genotyping Metrics for Structural Variants Across Technologies

Precision, recall, and F1-score metrics for genotyping deletions (DEL) and insertions (INS) were calculated for the T2T and CMRG validation sets across all tested SRS and LRS SV callers. These metrics reflect the accuracy of each caller in correctly assigning genotypes for both SV types and validation datasets.

Truth Set	Variant Type	Mean F1 Type	LRS	SRS	synthetic	LRS_vs_SRS_ratio	LRS_vs_synthetic_ratio	SRS_sv_synthetic_ratio
T2T	del	Discovery	93.77%	60.25%	52.02%	1.56	1.80	1.16

T2T	del	Genotyping	89.29%	82.66%	85.95%	1.08	1.04	0.96
T2T	ins	Discovery	93.65%	28.64%	NA	3.27	NA	NA
T2T	ins	Genotyping	70.18%	62.78%	NA	1.12	NA	NA
CMRG	del	Discovery	93.50%	62.12%	53.28%	1.51	1.76	1.17
CMRG	del	Genotyping	93.17%	78.05%	76.06%	1.19	1.22	1.03
CMRG	ins	Discovery	90.17%	32.06%	NA	2.81	NA	NA
CMRG	ins	Genotyping	79.70%	73.35%	NA	1.09	NA	NA

Table S8 - Comparison of Mean Discovery and Genotyping F1 Scores Across Technologies

Mean F1 scores for both discovery and genotyping were calculated for SV callers across three sequencing technology types: synthetic long reads (10X Genomics: LongRanger), short-read sequencing (SRS: Dragen, Dysgu, and Manta/MetaSV), and long-read sequencing (LRS: CuteSV, Dysgu, pbsv, Sawfish, Sniffles2, and SVIM). Ratios of mean F1 scores were computed to facilitate direct comparison of performance across technologies.

Truth Set	SV type	ranges	LRS mean TP	SRS mean TP	synthetic mean TP	LRS_vs_SRS_ratio	LRS_vs_synthetic_ratio	SRS_sv_synthetic_ratio
T2T	del	[30, 50)	4785.1	1029.8	327.0	4.65	14.63	3.15
T2T	del	[50, 250)	4525.6	1883.5	1197.0	2.40	3.78	1.57
T2T	del	[250, 500)	1606.2	1164.5	1193.0	1.38	1.35	0.98
T2T	del	[500, 6000)	851.9	611.2	618.0	1.39	1.38	0.99
T2T	del	6000+	39.4	31.8	78.0	1.24	0.50	0.41
T2T	ins	[30, 50)	4900.7	801.2	NA	6.12	NA	NA
T2T	ins	[50, 250)	6398.1	1615.0	NA	3.96	NA	NA
T2T	ins	[250, 500)	2713.5	671.2	NA	4.04	NA	NA
T2T	ins	[500, 6000)	2516.5	87.8	NA	28.65	NA	NA
T2T	ins	6000+	71.3	2.0	NA	35.65	NA	NA
CMRG	del	[50, 250)	53.8	17.3	12.0	3.10	4.48	1.44
CMRG	del	[250, 500)	17.0	15.7	15.0	1.09	1.13	1.04
CMRG	del	[500, 6000)	13.2	10.3	9.0	1.28	1.47	1.15
CMRG	ins	[50, 250)	49.2	16.7	NA	2.95	NA	NA
CMRG	ins	[250, 500)	23.2	6.2	NA	3.76	NA	NA

CMRG	ins	[500, 6000)	24.1	0.2	NA	144.30	NA	NA
------	-----	-------------	------	-----	----	--------	----	----

Table S9 - Comparison of True Positive Counts Across Sequencing Technologies

Average true positive (TP) counts were calculated for structural variant calls across seven size categories: 0-30 bp, 31-50 bp, 51-250 bp, 251-500 bp, 501-6,000 bp, and 6,000+ bp. Results include one synthetic long-read caller (10X Genomics: LongRanger), three short-read callers (SRS: Dragen, Dysgu, and MetaSV), and four or five long-read callers (LRS: CuteSV, Dysgu, pbsv, Sniffles2, and SVIM). Ratios of mean TP counts were computed to facilitate comparison of performance across technologies and SV size categories.