

# SV-JIM, supplementary information including detailed configurations, tables, and figures.\*

C. Malcolm Todd<sup>1</sup>[0000-0003-2522-2833], Lingling Jin<sup>1</sup>[0000-0002-4586-2347], and Ian McQuillan<sup>1</sup>[0000-0002-7998-4430]

Department of Computer Science, University of Saskatchewan,  
Saskatoon, SK, Canada  
malcolm.todd@usask.ca, {lingling.jin, mcquillain}@cs.usask.ca

## 1 Software details and supplementary information

### 1.1 Sequence alignment configurations

SV-JIM executes Minimap2 (version 2.24-r1122) using the software’s available ‘map-pb’ or ‘map-ont’ sequencing technology pre-sets, and with a Z drop score of 600 and inversion Z drop score of 200 by default. These settings were chosen according to the original cuteSV publication [6]; however, these settings are configurable. Likewise, SV-JIM uses the ‘asm5’ preset Minimap2 configuration by default when aligning the input genome assemblies given its use in SVIM-asm’s publication [5] and the assumption that the samples will be closely related. After alignment, SV-JIM performs post-processing using Samtools (version 1.16.1) [1] to sort, index, and convert the output alignment files to BAM format. Version information for each software used by SV-JIM is also available at <https://github.com/USask-BINFO/SV-JIM>.

### 1.2 SV caller details and configurations

Name	Input Type	Detectable SV Types
cuteSV [6]	Long Read	DEL, INS, INV, DUP, BND
Sniffles [8]	Long Read	DEL, INS, INV, DUP, BND, nested SVs
SVIM [4]	Long Read	DEL, INS, INV, DUP, BND
PAV [2]	Assembly	DEL, INS, INV
SVIM-asm [5]	Assembly	DEL, INS, INV, DUP, BND

**Table 1.** SV callers executed by SV-JIM. Abbreviations used in the table: DEL (deletions), INS (insertions), INV (inversions), DUP (duplications), BND (translocation breakends).

---

\* This research is supported by Natural Sciences and Engineering Research Council of Canada (NSERC), grant number 2019-06424 (LJ), 2022-05092 (IM), and a Canadian Graduate Scholarship - Doctoral (CMT).

SV caller	Published	Version Used	Citations*	Reported Performance in Release Publication
cuteSV	2020	2.0.2	154	Recall and F1 scores above 90% using the Human Pangenome HG002 data [6]
PAV	2021	2.2.4	348	Low false discovery rate of approx. 4% using GIAB's HG002 sample [2]
Sniffles & Sniffles2	2018 & 2022	2.0.7**	1121 & 47	Accuracy above 74% on simulated SVs in <i>A. thaliana</i> and <i>Homo sapiens</i> [8]
SVIM	2019	1.4.2	204	Accuracy above 83% on simulated data [4]
SVIM-asm	2020	1.0.3	59	F1 scores above 90% using the GIAB SV benchmark [5]

**Table 2.** Details on the SV callers included in SV-JIM. \*Citation total reflects the number of citations reported on Google Scholar as of October 2023. \*\*Version information is for Sniffles2 because it was the SV caller used by SV-JIM.

Maximum SV Size	Minimum MAPQ Score	Minimum Supporting Reads
300,000 bps (300 Kbps)	20	10

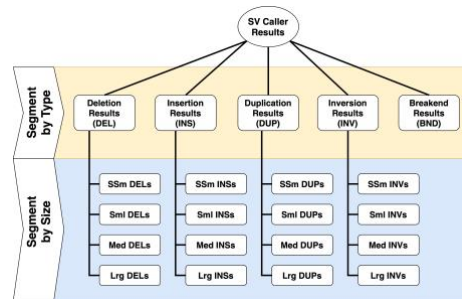
**Table 3.** SV-JIM default configurations and filtering for SV callers based on size, quality, and the amount of supporting evidence.

Tool	Provided fields for filtering	Missing/Placeholder fields for filtering
CuteSV	POS, END, SVLEN, RE	QUAL
Sniffles2	POS, END, SVLEN, SUPPORT, QUAL	N/A
SVIM	POS, END, SUPPORT, QUAL	SVLEN (for INVs)
PAV	POS, SVLEN	QUAL, SUPPORT, END
SVIM-ASM	POS, END, SVLEN	QUAL, SVLEN (for INVs), SUPPORT

**Table 4.** Differences in the reported VCF fields used for filtering between tools. (Definitions: POS=Start position, END=End position, QUAL=Phred-scaled quality score, SVLEN=SV's length, SUPPORT & RE=Amount of read evidence)

Size Class	Minimum Size (bps)	Maximum Size (bps)
Super small (SSm)	50	99
Small (Sml)	100	1000
Medium (Med)	1001	100,000
Large (Lrg)	100,001	300,000

**Table 5.** Sizing thresholds used when segmenting SV-JIM's results. The range of possible SV sizes is bounded based on filtering used during the experiments.



**Fig.1.** SV-JIM's SV caller result segmentation process to separate deletion (DEL), insertion (INS), duplication (DUP), inversion (INV), and translocation breakends (BND) by type and by size.

## 2 Results, supplementary tables and figures

### 2.1 SV-JIM computation demands

SV-JIM's only overhead above the demands of its included tools is the management and reformatting of the data files required by these tools. Consequently, the demands of SV-JIM's own scripts are negligible relative to more memory intensive tasks like the sequence alignment and SV calling performed by the included tools. This is also observable in the experiment's wall clock details where four programs responsible for sequence alignment and SV calling tasks are responsible for up to 90% of SV-JIM's total wall clock time. These programs are PAV, Minimap2, SVIM, and Samtools, and details for each program's wall clock times are detailed in Table 6. Therefore, we recommend that users concerned about their computational resources should consult these tools' publications and online resources for more detailed assessments of their resource demands to determine if they have the necessary resources to run the tools included within SV-JIM.

Software	<i>A. thaliana</i>	<i>B. nigra</i>	<i>H. sapiens</i>
PAV	01H:56M:38S (69.37%)	01D:05H:43M:13S (65.34%)	01H:48M:11S (4.75%)
Minimap2	30M:36S (18.20%)	08H:01M:30S (17.64%)	14H:49M:53S (39.07%)
SVIM	05M:02S (2.99%)	04H:08M:31S (9.11%)	10H:50M:35S (23.84%)
Samtools	04M:29S (2.67%)	02H:45M:00S (6.05%)	08H:28M:30S (18.64%)
Subtotal	02H:36M:45S (93.23%)	01D:20H:38M:14S (98.14%)	01D:11H:57M:09S (94.70%)

**Table 6.** Wall clock details and their percentage of the experiment's total execution time for the most time-consuming tools included in SV-JIM.

## 2.2 *Brassica nigra* results

Using the NI100 and C2 genotypes, SV-JIM generated multiple SV sets, including the original tool results, multiple SV intersection sets, and several aggregated SV sets for each SV caller. SV-JIM's outputs were filtered over several passes as described in SV-JIM's Methodology section, including:

1. Removal of any results reporting unresolved (N) nucleotides.
2. Removal of results that were low quality, oversized, or offered insufficient read evidence.

SV Caller	Filtering	DEL	INS	DUP	INV	BND	Total
cuteSV	Unfiltered	54,251	40,895	1,075	572	9,840	106,633
	Removed Ns	54,144	40,895	1,075	572	9,840	106,526
	Final Filtering	48,312	37,865	1,075	572	8,498	96,322
PAV	Unfiltered*	20,591	18,915	N/A	145	N/A	39,651
	Removed Ns	20,334	18,466	N/A	125	N/A	38,925
	Final Filtering	20,237	18,466	N/A	125	N/A	38,828
Sniffles2	Unfiltered	44,416	36,998	231	391	11,125	93,161
	Removed Ns	44,343	36,998	231	391	11,125	93,088
	Final Filtering	44,270	36,998	189	153	11,124	92,734
SVIM	Unfiltered**	672,631	257,931	9,686	662	119,072	1,059,982
	Removed Ns	672,631	257,931	9,686	662	119,072	1,059,982
	Final Filtering	26,095	19,805	543	11	1,418	47,872
SVIM-asm	Unfiltered	14,169	12,642	14	23	96	26,994
	Removed Ns	14,163	12,220	14	23	96	26,516
	Final Filtering	14,163	12,220	14	23	96	26,516

**Table 7.** *Brassica nigra* result totals during filtering for deletion (DEL), insertion (INS), duplication (DUP), inversion (INV), and translocation breakends (BND). Cells containing N/A denote types that were undetectable by that SV tool. Note: \*PAV unfiltered totals reflect some filtering to remove any SNPs or small variations to ensure all tools are compared using the common SV definition featured in the current literature. \*\*SVIM does not provide initial filtering based on the total supporting reads, causing its initial results to be inflated due to its reporting of SVs with even one supporting read.

SV Caller	Size	DEL	INS	DUP	INV	BND	Subtotal	Total
cuteSV	SSm	11,824	10,063	0	3	8,498	30,388	96,322
	Sml	<b>22,731</b>	<b>19,744</b>	99	118	X	<b>42,692</b>	
	Med	13,735	8,058	<b>949</b>	<b>431</b>	X	23,173	
	Lrg	22	0	27	20	X	69	
PAV	SSm	4,180	3,908	N/A	0	N/A	8,088	38,828
	Sml	<b>9,374</b>	<b>8,458</b>	N/A	11	N/A	<b>17,843</b>	
	Med	6,665	6,100	N/A	<b>111</b>	N/A	12,876	
	Lrg	18	0	N/A	3	N/A	21	
Sniffles2	SSm	11,555	10,154	0	0	11,124	32,833	92,734
	Sml	<b>20,372</b>	<b>19,501</b>	5	48	X	<b>39,926</b>	
	Med	12,334	7,343	<b>177</b>	<b>93</b>	X	19,947	
	Lrg	9	0	7	12	X	28	
SVIM	SSm	6,829	6,366	0	0	1,418	14,613	47,872
	Sml	<b>12,278</b>	<b>11,282</b>	106	4	X	<b>23,670</b>	
	Med	6,985	2,157	<b>427</b>	<b>7</b>	X	9,576	
	Lrg	3	0	10	0	X	13	
SVIM-asm	SSm	2,971	2,760	3	0	96	5,830	26,516
	Sml	<b>6,276</b>	<b>5,211</b>	<b>5</b>	10	X	<b>11,502</b>	
	Med	4,912	4,248	<b>5</b>	<b>12</b>	X	9,177	
	Lrg	4	1	1	1	X	7	

**Table 8.** *Brassica nigra* filtered SV caller outputs for deletion (DEL), insertion (INS), duplication (DUP), inversion (INV), and translocation break ends (BND). Results are segmented into super small (SSm), small (Sml), medium (Med), and large (Lrg) categories. SV size categories are defined as: SSm  $\leq$  100 bps, 100 bps < Sml  $\leq$  1 Kbps, 1 Kbps < Med  $\leq$  100 Kbps, Lrg > 100 Kbps. Cells containing N/A denote types that were undetectable by that SV tool, and cells containing 'X' refer to SV sizes with no results for that SV type.

Tools	Type	A - B	B - A	$A \cap B$	$A \cup B$	$A \cap B / \text{Size}(A)$	$A \cap B / \text{Size}(B)$	Jaccard Index
A: cuteSV	DEL	33,772	5,588	14,540	53,900	30.10%	72.24%	26.98%
	INS	26,904	7,420	10,961	45,285	28.95%	59.63%	24.20%
B: PAV	INV	557	110	15	682	2.62%	12.00%	2.20%
A: cuteSV	DEL	7,170	3,128	41,142	51,440	85.16%	92.93%	<b>79.98%</b>
	INS	4,140	3,273	33,725	41,138	89.07%	91.15%	<b>81.98%</b>
	DUP	892	6	183	1,081	17.02%	96.83%	<b>16.93%</b>
B: Sniffles2	INV	441	22	131	594	22.90%	85.62%	<b>22.05%</b>
	BND	3,655	6,285	4,839	14,779	56.97%	43.50%	<b>32.74%</b>
A: cuteSV	DEL	22,643	426	25,669	48,738	53.13%	98.37%	52.67%
	INS	19,325	1,265	18,540	39,130	48.96%	93.61%	47.38%
	DUP	853	321	222	1,396	20.65%	40.88%	15.90%
B: SVIM	INV	561	0	11	572	1.92%	100.00%	1.92%
	BND	7,651	575	843	9,069	9.92%	59.45%	9.30%
A: cuteSV	DEL	35,850	1,701	12,462	50,013	25.79%	87.99%	24.92%
	INS	28,652	3,007	9,213	40,872	24.33%	75.39%	22.54%
	DUP	1,071	10	4	1,085	0.37%	28.57%	0.37%
B: SVIM-asm	INV	572	23	0	595	0.00%	0.00%	0.00%
	BND	8,448	50	46	8,544	0.54%	47.92%	0.54%
A: PAV	DEL	6,097	30,294	13,976	50,367	69.63%	31.57%	27.75%
	INS	7,770	26,471	10,527	44,768	57.53%	28.45%	23.51%
B: Sniffles2	INV	113	141	12	266	9.60%	7.84%	4.51%
A: PAV	DEL	8,722	14,744	11,351	34,817	56.55%	43.50%	32.60%
	INS	10,470	11,978	7,827	30,275	42.78%	39.52%	25.85%
B: SVIM	INV	121	7	4	132	3.20%	36.36%	3.03%
A: PAV	DEL	12,655	6,745	7,418	26,818	36.96%	52.38%	27.66%
	INS	11,736	5,659	6,561	23,956	35.86%	53.69%	27.39%
B: SVIM-asm	INV	125	23	0	148	0.00%	0.00%	0.00%
A: Sniffles2	DEL	19,127	952	25,143	45,222	56.79%	96.35%	55.60%
	INS	18,860	1,667	18,138	38,665	49.02%	91.58%	46.91%
	DUP	74	428	115	617	60.85%	21.18%	18.64%
B: SVIM	INV	142	0	11	153	7.19%	100.00%	7.19%
	BND	10,278	572	846	11,696	7.61%	59.66%	7.23%
	DEL	32,060	1,953	12,210	46,223	27.58%	86.21%	26.42%

A: Sniffles2	INS	28,173	3,395	8,825	40,393	23.85%	72.22%	21.85%
	DUP	185	10	4	199	2.12%	28.57%	2.01%
B: SVIM-asm	INV	153	23	0	176	0.00%	0.00%	0.00%
	BND	11,083	55	41	11,179	0.37%	42.71%	0.37%
A: SVIM	DEL	15,494	3,562	10,601	29,657	40.62%	74.85%	35.75%
	INS	13,042	5,457	6,763	25,262	34.15%	55.34%	26.77%
B: SVIM-asm	DUP	539	10	4	553	0.74%	28.57%	0.72%
	INV	11	23	0	34	0.00%	0.00%	0.00%
	BND	1,404	82	14	1,500	0.99%	14.58%	0.93%

**Table 9.** Intersected pairs of SV caller outputs with Jaccard index rates for *Brassica nigra*. Results are segmented by SV type into deletions (DEL), insertions (INS), duplications (DUP), inversions (INV), and translocation break ends (BND).

Tools	Type	A - B	B - A	A ∩ B	A ∪ B	A ∩ B / Size(A)	A ∩ B / Size(B)	Jaccard Index
A: cuteSV	DEL	34,576	240	13,736	48,552	28.43%	98.28%	28.29%
	INS	27,704	366	10,161	38,231	26.83%	96.52%	26.58%
B: PAV ∩ Sniffles2	INV	560	0	12	572	2.10%	100.00%	2.10%
A: cuteSV	DEL	37,034	73	11,278	48,385	23.34%	99.36%	23.31%
	INS	30,294	256	7,571	38,121	19.99%	96.73%	19.86%
B: PAV ∩ SVIM	INV	568	0	4	572	0.70%	100.00%	0.70%
A: cuteSV	DEL	41,399	505	6,913	48,817	14.31%	93.19%	14.16%
	INS	32,617	1,313	5,248	39,178	13.86%	79.99%	13.40%
B: PAV ∩ SVIM-asm	INV	572	0	0	572	0.00%	0.00%	0.00%
A: cuteSV	DEL	23,357	188	24,955	48,500	51.65%	99.25%	<b>51.45%</b>
	INS	19,943	216	17,922	38,081	47.33%	98.81%	<b>47.06%</b>
	DUP	962	2	113	1,077	10.51%	98.26%	<b>10.49%</b>
B: Sniffles2 ∩ SVIM	INV	561	0	11	572	1.92%	100.00%	1.92%
	BND	7,912	264	582	8,758	6.85%	68.79%	<b>6.65%</b>
A: cuteSV	DEL	36,207	105	12,105	48,417	25.06%	99.14%	25.00%
	INS	29,249	209	8,616	38,074	22.75%	97.63%	22.63%
	DUP	1,071	0	4	1,075	0.37%	100.00%	0.37%
B: Sniffles2 ∩ SVIM-asm	INV	572	0	0	572	0.00%	0.00%	0.00%
	BND	8,467	14	27	8,508	0.32%	65.85%	0.32%
A: cuteSV	DEL	37,753	42	10,559	48,354	21.86%	99.60%	21.84%
	INS	31,225	123	6,640	37,988	17.54%	98.18%	17.48%
	DUP	1,071	0	4	1,075	0.37%	100.00%	0.37%
B: SVIM ∩ SVIM-asm	INV	572	0	0	572	0.00%	0.00%	0.00%
	BND	8,487	7	7	8,501	0.08%	50.00%	0.08%
A: PAV	DEL	8,932	14,002	11,141	34,075	55.50%	44.31%	32.70%
	INS	10,826	10,667	7,471	28,964	40.83%	41.19%	25.76%
B: Sniffles2 ∩ SVIM	INV	121	7	4	132	3.20%	36.36%	<b>3.03%</b>
A: PAV	DEL	13,267	5,404	6,806	25,477	33.91%	55.74%	26.71%
	INS	13,230	3,758	5,067	22,055	27.69%	57.42%	22.97%
B: Sniffles2 ∩ SVIM-asm	INV	125	0	0	125	0.00%	0.00%	0.00%
A: PAV	DEL	14,031	4,559	6,042	24,632	30.10%	56.99%	24.53%
	INS	14,372	2,838	3,925	21,135	21.45%	58.04%	18.57%
B: SVIM ∩ SVIM-asm	INV	125	0	0	125	0.00%	0.00%	0.00%
A: Sniffles2	DEL	33,804	135	10,466	44,405	23.64%	98.73%	23.57%
	INS	30,431	196	6,567	37,194	17.75%	97.10%	17.66%



B: SVIM $\cap$ SVIM-asm	DUP	185	0	4	189	2.12%	100.00%	2.12%
	INV	153	0	0	153	0.00%	0.00%	0.00%
	BND	11,115	5	9	11,129	0.08%	64.29%	0.08%

**Table 10.** Intersected trios of SV caller outputs with Jaccard index rates for *Brassica nigra*. Results are segmented by SV type into deletions (DEL), insertions (INS), duplications (DUP), inversions (INV), and translocation break ends (BND).

Tools	SV Type	A - B	B - A	A $\cap$ B	A $\cup$ B	A $\cap$ B / Size(A)	A $\cap$ B / Size(B)	Jaccard Index
A: cuteSV $\cap$ PAV	DEL	3,405	14,063	11,080	28,548	76.49%	44.07%	<b>38.81%</b>
	INS	3,482	10,743	7,395	21,620	67.99%	40.77%	<b>34.20%</b>
B: Sniffles2 $\cap$ SVIM	INV	11	7	4	22	26.67%	36.36%	<b>18.18%</b>
A: cuteSV $\cap$ PAV	DEL	7,726	5,451	6,759	19,936	46.66%	55.36%	33.90%
	INS	5,923	3,871	4,954	14,748	45.55%	56.14%	33.59%
B: Sniffles2 $\cap$ SVIM-asm	INV	15	0	0	15	0.00%	0.00%	0.00%
A: cuteSV $\cap$ PAV	DEL	8,462	4,578	6,023	19,063	41.58%	56.82%	31.60%
	INS	7,019	2,905	3,858	13,782	35.47%	57.05%	27.99%
B: SVIM $\cap$ SVIM-asm	INV	15	0	0	15	0.00%	0.00%	0.00%
A: cuteSV $\cap$ Sniffles2	DEL	30,699	158	10,443	41,300	25.38%	98.51%	25.29%
	INS	27,188	226	6,537	33,951	19.38%	96.66%	19.25%
	DUP	179	0	4	183	2.19%	100.00%	<b>2.19%</b>
B: SVIM $\cap$ SVIM-asm	INV	131	0	0	131	0.00%	0.00%	0.00%
	BND	4,834	9	5	4,848	0.10%	35.71%	<b>0.10%</b>
A: PAV $\cap$ Sniffles2	DEL	7,998	4,623	5,978	18,599	42.77%	56.39%	32.14%
	INS	6,705	2,941	3,822	13,468	36.31%	56.51%	28.38%
B: SVIM $\cap$ SVIM-asm	INV	12	0	0	12	0.00%	0.00%	0.00%
A: cuteSV $\cap$ PAV	DEL	8,520	4,501	5,965	18,986	41.18%	56.99%	31.42%
	INS	7,077	2,767	3,800	13,644	34.94%	57.87%	27.85%
B: Sniffles2 $\cap$ (SVIM $\cap$ SVIM-asm)	INV	15	0	0	15	0.00%	0.00%	0.00%

**Table 11.** Intersected combinations of four and five SV caller outputs with Jaccard index rates for *Brassica nigra*. Results are segmented by SV type into deletions (DEL), insertions (INS), duplications (DUP), inversions (INV), and translocation break ends (BND).

SV Caller	Type	Filtered	Support from 1+ tools	Support from 2+ tools	Support from 3+ tools
cuteSV	DEL	48,312	42,627	28,871	16,460
	INS	37,865	35,192	22,101	11,408
	DUP	1,075	308	114	4
	INV	572	137	20	4
	BND	8,498	4,967	598	7
PAV	DEL	20,237	15,294	14,119	11,957
	INS	18,466	12,724	10,785	8,635
	INV	125	15	12	4
Sniffles2	DEL	44,270	41,586	28,615	16,381
	INS	36,998	34,332	21,801	11,335
	DUP	189	185	113	4
	INV	153	131	19	4
	BND	11,124	5,236	603	5
SVIM	DEL	26,095	25,931	25,283	15,640
	INS	19,805	19,028	18,277	10,219
	DUP	543	224	113	4
	INV	11	11	11	4
	BND	1,418	1,127	584	5
SVIM-asm	DEL	14,163	13,048	12,399	11,311
	INS	12,220	10,693	9,139	7,777
	DUP	14	5	4	4
	INV	23	0	0	0
	BND	96	65	29	5

**Table 12.** *Brassica nigra* aggregated SV caller outputs with minimum support from N other tools. Results are segmented by deletion (DEL), insertion (INS), duplication (DUP), inversion (INV), and translocation break end (BND) types. N+ columns denote SV counts with support from N or more other tools.

### 2.3 *Arabidopsis thaliana* results

SV-JIM generated multiple SV sets using the TAIR10 and Ler genotypes, including the original tool results, several aggregated SV sets for each SV tool, and multiple SV intersection sets.

SV Caller	Filtering	DEL	INS	DUP	INV	BND	Total
cuteSV	Unfiltered	3,723	3,068	118	53	440	7,402
	Removed Ns	3,708	3,068	118	53	440	7,387
	Final Filtering	3,495	2,944	118	53	393	7,003
PAV	Unfiltered*	4,360	4,358	N/A	17	N/A	8,735
	Removed Ns	4,351	4,345	N/A	17	N/A	8,713
	Final Filtering	4,351	4,345	N/A	17	N/A	8,713
Sniffles2	Unfiltered	3,368	2,867	24	39	462	6,760
	Removed Ns	3,358	2,867	24	39	462	6,750
	Final Filtering	3,353	2,867	21	17	462	6,720
SVIM	Unfiltered**	13,797	118,148	643	140	7,846	140,574
	Removed Ns	13,797	118,148	643	140	7,846	140,574
	Final Filtering	1,854	1,502	34	0	124	3,514
SVIM-asm	Unfiltered	2,156	2,215	0	6	14	4,391
	Removed Ns	2,148	2,208	0	6	14	4,376
	Final Filtering	2,148	2,208	0	6	14	4,376

**Table 13.** *Arabidopsis thaliana* result totals during filtering for deletion (DEL), insertion (INS), duplication (DUP), inversion (INV), and translocation break ends (BND). Cells containing N/A denote types not detectable by the tool they relate to. Notes: \*PAV unfiltered totals reflect some filtering to remove any SNPs or small variations below 50bps that were reported to ensure all tools are compared using the common SV definition featured in the current literature. \*\*SVIM does not provide initial filtering based on the total supporting reads, so the large number of initial results is due to the reporting of SVs with even one supporting read.

SV Caller	Size	DEL	INS	DUP	INV	BND	Subtotal	Total
cuteSV	SSm	995	1,031	0	0	393	2,419	7,003
	Sml	<b>1,759</b>	<b>1,558</b>	6	7	X	<b>3,330</b>	
	Med	740	355	<b>102</b>	<b>27</b>	X	1,224	
	Lrg	1	0	10	19	X	30	
PAV	SSm	1,075	1,070	N/A	0	N/A	2,145	8,713
	Sml	<b>2,147</b>	<b>2,131</b>	N/A	2	N/A	<b>4,280</b>	
	Med	1,129	1,143	N/A	<b>14</b>	N/A	2,286	
	Lrg	0	1	N/A	1	N/A	2	
Sniffles2	SSm	985	1,017	0	0	462	2,464	6,720
	Sml	<b>1,710</b>	<b>1,511</b>	0	2	X	<b>3,223</b>	
	Med	658	339	<b>21</b>	<b>10</b>	X	1,028	
	Lrg	0	0	0	5	X	5	
SVIM	SSm	666	647	0	0	124	1,437	3,514
	Sml	<b>997</b>	<b>833</b>	10	0	X	<b>1,840</b>	
	Med	191	22	<b>21</b>	0	X	234	
	Lrg	0	0	3	0	X	3	
SVIM-asm	SSm	601	639	0	0	14	1,254	4,376
	Sml	<b>997</b>	<b>971</b>	0	<b>3</b>	X	<b>1,971</b>	
	Med	550	598	0	2	X	1,150	
	Lrg	0	0	0	1	X	1	

**Table 14.** *Arabidopsis thaliana* filtered SV caller outputs for deletion (DEL), insertion (INS), duplication (DUP), inversion (INV), and translocation break ends (BND). Results are further segmented into super small (SSm), small (Sml), medium (Med), and large (Lrg) categories. SV size categories are defined as: SSm  $\leq$  100 bps, 100 bps < Sml  $\leq$  1 Kbps, 1 Kbps < Med  $\leq$  100 Kbps, Lrg > 100 Kbps. Cells containing N/A denote types not detectable by the SV caller to which they relate, and cells containing 'X' refer to SV sizes that do not exist for that particular type. Cells highlighted in grey represent missing SV types for which the tool reported no results or had no results of sufficient confidence to survive filtering.

Tools	SV Type	A - B	B - A	A ∩ B	A ∪ B	A ∩ B / Size(A)	A ∩ B / Size(B)	Jaccard Index	
A: cuteSV	DEL	630	1,448	2,865	4,943	81.97%	66.43%	57.96%	
	INS	486	1,845	2,458	4,798	83.49%	57.00%	51.23%	
B: PAV	INV	46	10	7	63	13.21%	41.18%	11.11%	
A: cuteSV	DEL	290	148	3,205	3,643	91.70%	95.59%	<b>87.98%</b>	
	INS	304	227	2,640	3,171	89.67%	92.08%	<b>83.25%</b>	
	DUP	100	3	18	121	15.25%	85.71%	14.88%	
	B: Sniffles2	INV	41	5	12	58	22.64%	70.59%	<b>20.69%</b>
	BND	187	256	206	649	52.42%	44.59%	<b>31.74%</b>	
A: cuteSV	DEL	1,653	12	1,842	3,507	52.70%	99.35%	52.52%	
	INS	1,481	39	1,463	2,983	49.69%	97.40%	49.04%	
	DUP	99	15	19	133	16.10%	55.88%	14.29%	
	B: SVIM	INV	53	0	0	53	0.00%	0.00%	0.00%
	BND	333	64	60	457	15.27%	48.39%	13.13%	
A: cuteSV	DEL	1,634	287	1,861	3,782	53.25%	86.64%	49.21%	
	INS	1,353	617	1,591	3,561	54.04%	72.06%	44.68%	
	DUP	118	0	0	118	0.00%	0.00%	0.00%	
	B: SVIM-asm	INV	53	6	0	59	0.00%	0.00%	0.00%
	BND	385	6	8	399	2.04%	57.14%	2.01%	
A: PAV	DEL	1,490	534	2,819	4,843	65.42%	84.07%	59.19%	
	INS	1,917	490	2,377	4,784	55.36%	82.91%	50.60%	
B: Sniffles2	INV	14	14	3	31	17.65%	17.65%	9.68%	
A: PAV	DEL	2,639	184	1,670	4,493	38.76%	90.08%	37.17%	
	INS INV	2,952	160	1,342	4,454	31.25%	89.35%	30.13%	
B: SVIM		17	0	0	17	0.00%	0.00%	0.00%	
A: PAV	DEL	2,293	132	2,016	4,441	46.79%	93.85%	45.40%	
	INS	2,241	155	2,053	4,449	47.81%	92.98%	46.15%	
B: SVIM-asm	INV	17	6	0	23	0.00%	0.00%	0.00%	
A: Sniffles2	DEL	1,542	43	1,811	3,396	54.01%	97.68%	53.33%	
	INS	1,442	77	1,425	2,944	49.70%	94.87%	48.40%	
	DUP	13	26	8	47	38.10%	23.53%	<b>17.02%</b>	
	B: SVIM	INV	17	0	0	17	0.00%	0.00%	0.00%
	BND	394	56	68	518	14.72%	54.84%	13.13%	
A: Sniffles2	DEL	1,530	325	1,823	3,678	54.37%	84.87%	49.56%	
	INS	1,300	641	1,567	3,508	54.66%	70.97%	44.67%	
	DUP	21	0	0	21	0.00%	0.00%	0.00%	

B: SVIM-asm	INV	17	6	0	23	0.00%	0.00%	0.00%
	BND	458	10	4	472	0.87%	28.57%	0.85%
A: SVIM	DEL	704	998	1,150	2,852	62.03%	53.54%	40.32%
	INS	581	1,287	921	2,789	61.32%	41.71%	33.02%
B: SVIM-asm	DUP	34	0	0	34	0.00%	0.00%	0.00%
	INV	0	6	0	6	0.00%	0.00%	0.00%
	BND	119	9	5	133	4.03%	35.71%	3.76%

**Table 15.** Intersected pairs of SV caller outputs with Jaccard index rates for *Arabidopsis thaliana*. Results are segmented by SV type into deletions (DEL), insertions (INS), duplications (DUP), inversions (INV), and translocation break ends (BND). Cells highlighted in grey represent SV types restricted to zero due to their absence from one of the SV callers being considered.

Tools	SV Type	A - B	B - A	A ∩ B	A ∪ B	A ∩ B / Size(A)	A ∩ B / Size(B)	Jaccard Index
A: cuteSV	DEL	744	68	2,751	3,563	78.71%	97.59%	<b>77.21%</b>
	INS	658	91	2,286	3,035	77.65%	96.17%	<b>75.32%</b>
B: PAV ∩ Sniffles2	INV	50	0	3	53	5.66%	100.00%	<b>5.66%</b>
A: cuteSV	DEL	1,830	5	1,665	3,500	47.64%	99.70%	47.57%
	INS INV	1,620	18	1,324	2,962	44.97%	98.66%	44.70%
B: PAV ∩ SVIM		53	0	0	53	0.00%	0.00%	0.00%
A: cuteSV	DEL	1,668	189	1,827	3,684	52.27%	90.62%	49.59%
	INS INV	1,392	501	1,552	3,445	52.72%	75.60%	45.05%
B: PAV ∩ SVIM-asm		53	0	0	53	0.00%	0.00%	0.00%
A: cuteSV	DEL	1,689	5	1,806	3,500	51.67%	99.72%	51.60%
	INS	1,524	5	1,420	2,949	48.23%	99.65%	48.15%
	DUP	110	0	8	118	6.78%	100.00%	<b>6.78%</b>
	INV	53	0	0	53	0.00%	0.00%	0.00%
B: Sniffles2 ∩ SVIM	BND	355	30	38	423	9.67%	55.88%	<b>8.98%</b>
A: cuteSV	DEL	1,692	20	1,803	3,515	51.59%	98.90%	51.29%
	INS	1,441	64	1,503	3,008	51.05%	95.92%	49.97%
	DUP	118	0	0	118	0.00%	0.00%	0.00%
	INV	53	0	0	53	0.00%	0.00%	0.00%
B: Sniffles2 ∩ SVIM-asm	BND	390	1	3	394	0.76%	75.00%	0.76%
A: cuteSV	DEL	2,347	2	1,148	3,497	32.85%	99.83%	32.83%
	INS	2,031	8	913	2,952	31.01%	99.13%	30.93%
	DUP	118	0	0	118	0.00%	0.00%	0.00%
	INV	53	0	0	53	0.00%	0.00%	0.00%
B: SVIM ∩ SVIM-asm	BND	390	2	3	395	0.76%	60.00%	0.76%
A: PAV	DEL	2,650	152	1,659	4,461	38.50%	91.61%	37.19%
	INS INV	2,990	121	1,304	4,415	30.37%	91.51%	29.54%
B: Sniffles2 ∩ SVIM		17	0	0	17	0.00%	0.00%	0.00%
A: PAV	DEL	2,513	27	1,796	4,336	41.68%	98.52%	41.42%
	INS INV	2,765	38	1,529	4,332	35.61%	97.57%	35.30%
B: Sniffles2 ∩ SVIM-asm		17	0	0	17	0.00%	0.00%	0.00%
A: PAV	DEL	3,172	13	1,137	4,322	26.39%	98.87%	26.31%
	INS INV	3,398	25	896	4,319	20.87%	97.29%	20.75%
B: SVIM ∩ SVIM-asm		17	0	0	17	0.00%	0.00%	0.00%
A: Sniffles2	DEL	2,206	3	1,147	3,356	34.21%	99.74%	34.18%
	INS	1,962	16	905	2,883	31.57%	98.26%	31.39%
	DUP	21	0	0	21	0.00%	0.00%	0.00%
	INV	17	0	0	17	0.00%	0.00%	0.00%
B: SVIM ∩ SVIM-asm	BND	462	3	2	465	0.43%	40.00%	0.43%

**Table 16.** Intersected trios of SV caller outputs with Jaccard index rates for *Arabidopsis thaliana*. Results are segmented by SV type into deletions (DEL), insertions (INS), duplications (DUP), inversions (INV), and translocation break ends (BND). Cells highlighted in grey represent SV types restricted to zero due to their absence from one of the SV callers being considered.

Tools	Type	A - B	B - A	A ∩ B	A ∪ B	A ∩ B / Size(A)	A ∩ B / Size(B)	Jaccard Index
A: cuteSV ∩ PAV	DEL	1,207	157	1,654	3,018	57.81%	91.33%	54.80%
	INS	1,139	124	1,301	2,564	53.32%	91.30%	50.74%
	INV	7	0	0	7	0.00%	0.00%	0.00%
B: Sniffles2 ∩ SVIM	DEL	1,084	46	1,777	2,907	62.11%	97.48%	<b>61.13%</b>
A: cuteSV ∩ PAV	INS	971	98	1,469	2,538	60.20%	93.75%	<b>57.88%</b>
B: Sniffles2 ∩ SVIM-asm	INV	7	0	0	7	0.00%	0.00%	0.00%
A: cuteSV ∩ PAV	DEL	1,726	15	1,135	2,876	39.67%	98.70%	39.46%
A: cuteSV ∩ PAV	INS	1,550	31	890	2,471	36.48%	96.63%	36.02%
B: SVIM ∩ SVIM-asm	INV	7	0	0	7	0.00%	0.00%	0.00%
A: cuteSV ∩ Sniffles2	DEL	2,060	5	1,145	3,210	35.73%	99.57%	35.67%
	INS	1,738	19	902	2,659	34.17%	97.94%	33.92%
	DUP	18	0	0	18	0.00%	0.00%	0.00%
	INV	12	0	0	12	0.00%	0.00%	0.00%
B: SVIM ∩ SVIM-asm	BND	205	4	1	210	0.49%	20.00%	<b>0.48%</b>
A: PAV ∩ Sniffles2	DEL	1,683	14	1,136	2,833	40.30%	98.78%	40.10%
A: PAV ∩ Sniffles2	INS	1,493	37	884	2,414	37.19%	95.98%	36.62%
B: SVIM ∩ SVIM-asm	INV	3	0	0	3	0.00%	0.00%	0.00%
A: cuteSV ∩ PAV	DEL	1,727	13	1,134	2,874	39.64%	98.87%	<b>39.46%</b>
A: cuteSV ∩ PAV	INS	1,559	24	881	2,464	36.11%	97.35%	<b>35.75%</b>
B: Sniffles2 ∩ (SVIM ∩ SVIM-asm)	INV	7	0	0	7	0.00%	0.00%	0.00%

**Table 17.** Intersected combinations of four and five SV caller outputs with Jaccard index rates for *Arabidopsis thaliana*. Results are segmented by SV type into deletions (DEL), insertions (INS), duplications (DUP), inversions (INV), and translocation break ends (BND). Cells highlighted in grey represent SV types restricted to zero due to their absence from one of the SV callers being considered.



SV Caller	Type	Filtered	Support from 1+ tools	Support from 2+ tools	Support from 3+ tools
cuteSV	DEL	3,495	3,346	2,981	2,312
	INS	2,944	2,823	2,516	1,927
	DUP	118	30	8	0
	INV	53	17	3	0
	BND	393	220	39	3
PAV	DEL	4,351	3,107	2,832	2,301
	INS	4,345	3,009	2,446	1,901
	INV	17	7	3	0
Sniffles2	DEL	3,353	3,274	2,941	2,310
	INS	2,867	2,738	2,482	1,918
	DUP	21	18	8	0
	INV	17	12	3	0
	BND	462	245	41	1
SVIM	DEL	1,854	1,847	1,824	1,669
	INS	1,502	1,484	1,455	1,330
	DUP	34	19	8	0
	INV				0
	BND				
		0	0	0	
	122	94	41	1	
SVIM-asm	DEL	2,148	2,051	1,874	1,791
	INS	2,208	2,097	1,653	1,501
	DUP	0	0	0	0
	INV	6	0	0	0
	BND	14	10	6	1

**Table 18.** *Arabidopsis thaliana* aggregated SV caller outputs with minimum support from any N other tools. Results are segmented by deletion (DEL), insertion (INS), duplication (DUP), inversion (INV), and translocation break end (BND) types. N+ columns denote SV counts with support from N or more other tools. Cells highlighted in grey represent SV types restricted to zero due to their absence in the SV caller's original result.

## 2.4 *Homo sapiens* results

Using the hs37d5 reference genome and HG002 assembly from GIAB's Ashkenazim trio, SV-JIM generated multiple SV sets, including the individual SV caller results, several aggregated SV sets for each caller, and multiple SV intersection sets.

SV Caller	Filtering	DEL	INS	DUP	INV	BND	Total
cuteSV	Unfiltered	11,503	16,027	1,385	253	594	29,762
	Removed	11,490	16,027	1,385	253	594	29,749
	Ns Final Filtering	10,463	14,220	1,385	253	498	26,819
PAV	Unfiltered*	7,167	8,954	N/A	39	N/A	16,160
	Removed	7,116	8,954	N/A	39	N/A	16,109
	Ns Final Filtering	7,114	8,954	N/A	39	N/A	16,107
Sniffles2	Unfiltered	8,904	22,043	27	59	259	31,292
	Removed	8,898	22,043	27	59	259	31,286
	Ns Final Filtering	8,894	22,043	27	56	259	31,279
SVIM	Unfiltered**	228,859	7,744,013	82,481	15,293	692,796	8,763,442
	Removed	228,859	7,744,013	82,481	15,293	692,796	8,763,442
	Ns Final Filtering	7,597	8,333	98	8	68	16,104
SVIM-asm	Unfiltered	8,806	11,250	48	63	472	20,639
	Removed	8,795	11,250	48	62	472	20,627
	Ns Final Filtering	8,795	11,250	48	62	472	20,627

**Table 19.** Human genome result totals during filtering for deletion (DEL), insertion (INS), duplication (DUP), inversion (INV), and translocation break ends (BND). Cells containing N/A denote types that were not detectable by the SV tool. \*Note: PAV unfiltered totals reflect some filtering to remove any SNPs or small variations below 50bps that were reported to ensure all tools are compared using the common SV definition featured in the current literature. \*\*High SVIM totals appear to result from the large volume of reads used in the experiment and the SVIM's inclusion of SV candidates with just one supporting read.

SV Caller	Size	DEL	INS	DUP	INV	BND	Subtotal	Total
cuteSV	SSm	4,025	5,116	4	3	498	9,646	26,819
	Sml	<b>5,479</b>	<b>8,136</b>	335	77	X	<b>14,027</b>	
	Med	946	968	<b>1,032</b>	<b>159</b>	X	3,105	
	Lrg	13	0	14	14	X	41	
PAV	SSm	2,801	2,796	N/A	0	N/A	5,597	16,107
	Sml	<b>3,679</b>	<b>5,099</b>	N/A	4	N/A	<b>8,782</b>	
	Med	631	1,058	N/A	<b>34</b>	N/A	1,723	
	Lrg	3	1	N/A	1	N/A	5	
Sniffles2	SSm	3,463	5,983	0	0	259	9,705	31,279
	Sml	<b>4,595</b>	7,229	0	14	X	<b>11,838</b>	
	Med	833	<b>8,831</b>	<b>26</b>	<b>40</b>	X	9,730	
	Lrg	3	0	1	2	X	6	
SVIM	SSm	3,043	3,095	1	0	68	6,207	16,113
	Sml	<b>3,981</b>	<b>4,829</b>	<b>50</b>	3	X	<b>8,863</b>	
	Med	580	409	46	<b>5</b>	X	1,040	
	Lrg	2	0	1	0	X	3	
SVIM-asm	SSm	3,101	3,190	2	0	472	6,765	20,627
	Sml	<b>4,727</b>	<b>6,748</b>	7	13	X	<b>11,495</b>	
	Med	960	1,312	<b>36</b>	<b>46</b>	X	2,354	
	Lrg	7	0	3	3	X	13	

**Table 20.** Human genome original filtered SV caller outputs for deletion (DEL), insertion (INS), duplication (DUP), inversion (INV), and translocation break ends (BND). Results are further segmented into super small (SSm), small (Sml), medium (Med), and large (Lrg) categories. SV size categories are defined as: SSm  $\leq$  100 bps, 100 bps < Sml  $\leq$  1 Kbps, 1 Kbps < Med  $\leq$  100 Kbps, Lrg > 100 Kbps. Cells containing 'N/A' denote types not detectable by the SV tool, and cells containing 'X' refer to SV sizes that do not exist for that SV type.

Tools	Type	A - B	B - A	$A \cap B$	$A \cup B$	$A \cap B / \text{Size}(A)$	$A \cap B / \text{Size}(B)$	Jaccard Index
A: cuteSV	DEL	4,988	1,565	5,475	12,028	52.33%	77.77%	45.52%
	INS	7,938	2,618	6,282	16,838	44.18%	70.58%	37.31%
B: PAV	INV	240	26	13	279	5.14%	33.33%	4.66%
A: cuteSV	DEL	2,174	605	8,289	11,068	79.22%	93.20%	<b>74.89%</b>
	INS	3,767	11,590	10,453	25,810	73.51%	47.42%	40.50%
	DUP	1,360	2	25	1,387	1.81%	92.59%	1.80%
B: Sniffles2	INV	208	11	45	264	17.79%	80.36%	<b>17.05%</b>
	BND	388	170	89	647	18.66%	34.36%	<b>13.76%</b>
A: cuteSV	DEL	3,195	329	7,268	10,792	69.46%	95.67%	67.35%
	INS	6,728	841	7,492	15,061	52.69%	89.91%	49.74%
	DUP	1,298	11	87	1,396	6.28%	88.78%	6.23%
B: SVIM	INV	245	0	8	253	3.16%	100.00%	3.16%
	BND	455	46	22	523	4.61%	32.35%	4.21%
A: cuteSV	DEL	4,493	2,825	5,970	13,288	57.06%	67.88%	44.93%
	INS	7,585	4,615	6,635	18,835	46.66%	58.98%	35.23%
	DUP	1,367	30	18	1,415	1.30%	37.50%	1.27%
B: SVIM-asm	INV	253	62	0	315	0.00%	0.00%	0.00%
	BND	414	409	63	886	13.21%	13.35%	7.11%
A: PAV	DEL	1,904	3,806	5,088	10,798	72.77%	57.21%	47.12%
	INS	2,980	16,181	5,862	25,023	66.30%	26.59%	23.43%
B: Sniffles2	INV	32	49	7	88	17.95%	12.50%	7.95%
A: PAV	DEL	2,143	2,748	4,849	9,740	69.35%	63.83%	49.78%
	INS	4,150	3,641	4,692	12,483	53.06%	56.31%	37.59%
B: SVIM	INV	39	8	0	47	0.00%	0.00%	0.00%
A: PAV	DEL	1,375	3,178	5,617	10,170	80.33%	63.87%	55.23%
	INS	1,1512	3,920	7,330	12,762	82.90%	65.16%	<b>57.44%</b>
B: SVIM-asm	INV	39	62	0	101	0.00%	0.00%	0.00%
A: Sniffles2	DEL	1,881	584	7,013	9,478	78.85%	92.31%	73.99%
	INS	14,728	1,018	7,315	23,061	33.19%	87.78%	31.72%
	DUP	17	88	10	115	37.04%	10.20%	8.70%
B: SVIM	INV	49	1	7	57	12.50%	87.50%	12.28%
	BND	233	42	26	301	10.04%	38.24%	8.64%
	DEL	3,278	3,179	5,616	12,073	63.14%	63.85%	46.52%

A: Sniffles2	INS	15,708	4,915	6,335	26,958	28.74%	56.31%	23.50%
	DUP	20	41	7	68	25.93%	14.58%	<b>10.29%</b>
B: SVIM-asm	INV	56	62	0	118	0.00%	0.00%	0.00%
	BND	231	444	28	703	10.81%	5.93%	3.98%
A: SVIM	DEL	2,292	3,490	5,305	11,087	69.83%	60.32%	47.85%
	INS	3,325	6,242	5,008	14,575	60.10%	44.52%	34.36%
B: SVIM-asm	DUP	96	46	2	144	2.04%	4.17%	1.39%
	INV	8	62	0	70	0.00%	0.00%	0.00%
	BND	44	448	24	516	35.29%	5.08%	4.65%

**Table 21.** Intersected pairs of SV caller outputs with Jaccard index rates for the Human genome. Results are segmented by SV type into deletions (DEL), insertions (INS), duplications (DUP), inversions (INV), and translocation break ends (BND).

Tools	Type	A - B	B - A	$A \cap B$	$A \cup B$	$A \cap B / \text{Size}(A)$	$A \cap B / \text{Size}(B)$	Jaccard Index
A: cuteSV	DEL	5,514	139	4,949	10,602	47.30%	97.27%	46.68%
	INS	8,828	470	5,392	14,690	37.92%	91.98%	36.71%
B: PAV $\cap$ Sniffles2	INV	247	1	6	254	2.37%	85.71%	2.36%
A: cuteSV	DEL	5,689	75	4,774	10,538	45.63%	98.45%	45.30%
	INS	9,770	242	4,450	14,462	31.29%	94.84%	30.77%
B: PAV $\cap$ SVIM	INV	253	0	0	253	0.00%	0.00%	0.00%
A: cuteSV	DEL	5,510	664	4,953	11,127	47.34%	88.18%	44.51%
	INS	8,531	1,641	5,689	15,861	40.01%	77.61%	35.87%
B: PAV $\cap$ SVIM-asm	INV	253	0	0	253	0.00%	0.00%	0.00%
A: cuteSV	DEL	3,564	114	6,899	10,577	65.94%	98.37%	<b>65.23%</b>
	INS	7,167	262	7,053	14,482	49.60%	96.42%	48.70%
	DUP	1,375	0	10	1,385	0.72%	100.00%	0.72%
	B: Sniffles2 $\cap$ SVIM	INV	246	0	7	253	2.77%	100.00%
	BND	469	18	8	495	1.68%	30.77%	1.62%
A: cuteSV	DEL	5,001	154	5,462	10,617	52.20%	97.26%	51.45%
	INS	8,491	606	5,729	14,826	40.29%	90.43%	38.64%
	DUP	1,379	1	6	1,386	0.43%	85.71%	0.43%
	B: Sniffles2 $\cap$ SVIM-asm	INV	253	0	0	253	0.00%	0.00%
	BND	469	20	8	497	1.68%	28.57%	1.61%
A: cuteSV	DEL	5,256	98	5,207	10,561	49.77%	98.15%	49.30%
	INS	9,519	307	4,701	14,527	33.06%	93.87%	32.36%
	DUP	1,383	0	2	1,385	0.14%	100.00%	0.14%
B: SVIM $\cap$ SVIM-asm	INV	253	0	0	253	0.00%	0.00%	0.00%
	BND	474	21	3	498	0.63%	12.50%	0.60%
A: PAV	DEL	2,363	2,384	4,629	9,376	66.20%	66.01%	49.37%
	INS	4,419	2,892	4,423	11,734	50.02%	60.46%	37.69%
B: Sniffles2 $\cap$ SVIM	INV	39	7	0	46	0.00%	0.00%	0.00%
A: PAV	DEL	2,272	896	4,720	7,888	67.51%	84.05%	59.84%
	INS	3,371	864	5,471	9,706	61.88%	86.36%	<b>56.37%</b>
B: Sniffles2 $\cap$ SVIM-asm	INV	39	0	0	39	0.00%	0.00%	0.00%
A: PAV	DEL	2,436	749	4,556	7,741	65.16%	85.88%	58.86%
	INS	4,359	525	4,483	9,367	50.70%	89.52%	47.86%
B: SVIM $\cap$ SVIM-asm	INV	39	0	0	39	0.00%	0.00%	0.00%
	DEL	3,822	233	5,072	9,127	57.03%	95.61%	55.57%
A: Sniffles2	INS	17,341	306	4,702	22,349	21.33%	93.89%	21.04%

B: SVIM $\cap$ SVIM-asm	DUP	25	0	2	27	7.41%	100.00%	<b>7.41%</b>
	INV	56	0	0	56	0.00%	0.00%	0.00%
	BND	251	16	8	275	3.09%	33.33%	<b>2.91%</b>

**Table 22.** Intersected trios of SV caller outputs with Jaccard index rates for Human. Results are segmented by SV type into deletions (DEL), insertions (INS), duplications (DUP), inversions (INV), and translocation break ends (BND).

Tools	SV Type	A - B	B - A	A $\cap$ B	A $\cup$ B	A $\cap$ B / Size(A)	A $\cap$ B / Size(B)	Jaccard Index
A: cuteSV $\cap$ PAV	DEL	848	2,434	4,579	7,861	84.37%	65.29%	58.25%
	INS	1,952	3,043	4,272	9,267	68.64%	58.40%	46.10%
B: Sniffles2 $\cap$ SVIM	INV	13	7	0	20	0.00%	0.00%	<b>0.00%</b>
A: cuteSV $\cap$ PAV	DEL	811	1,000	4,616	6,427	85.06%	82.19%	71.82%
	INS	1,159	1,270	5,065	7,494	81.38%	79.95%	<b>67.59%</b>
B: Sniffles2 $\cap$ SVIM-asm	INV	13	0	0	13	0.00%	0.00%	0.00%
A: cuteSV $\cap$ PAV	DEL	936	814	4,491	6,241	82.75%	84.66%	71.96%
	INS	1,969	753	4,255	6,977	68.36%	84.96%	60.99%
B: SVIM $\cap$ SVIM-asm	INV	13	0	0	13	0.00%	0.00%	<b>0.00%</b>
A: cuteSV $\cap$ Sniffles2	DEL	3,264	280	5,025	8,569	60.62%	94.72%	58.64%
	INS	5,898	453	4,555	10,906	43.58%	90.95%	41.77%
	DUP	23	0	2	25	8.00%	100.00%	<b>8.00%</b>
B: SVIM $\cap$ SVIM-asm	INV	45	0	0	45	0.00%	0.00%	<b>0.00%</b>
	BND	88	23	1	112	1.12%	4.17%	<b>0.89%</b>
A: PAV $\cap$ Sniffles2	DEL	717	934	4,371	6,022	85.91%	82.39%	<b>72.58%</b>
	INS	1,608	754	4,254	6,616	72.57%	84.94%	64.30%
B: SVIM $\cap$ SVIM-asm	INV	7	0	0	7	0.00%	0.00%	<b>0.00%</b>
A: cuteSV $\cap$ PAV	DEL	1,090	735	4,337	6,162	79.92%	85.51%	70.38%
	INS	2,105	583	4,119	6,807	66.18%	87.60%	60.51%
B: Sniffles2 $\cap$ (SVIM $\cap$ SVIM-asm)	INV	13	0	0	13	0.00%	0.00%	<b>0.00%</b>

**Table 23.** Intersected combinations of four and five SV caller outputs with Jaccard index rates for Human. Results are segmented by SV type into deletions (DEL), insertions (INS), duplications (DUP), inversions (INV), and translocation break ends (BND).

SV Caller	Type	Filtered	Support from 1+ tools	Support from 2+ tools	Support from 3+ tools
cuteSV	DEL	10,463	9,107	7,826	5,717
	INS	14,220	11,768	9,095	5,828
	DUP	1,385	124	15	2
	INV	253	56	13	0
	BND	498	122	18	3
PAV	DEL	7,114	6,204	5,464	5,048
	INS	8,954	8,037	6,556	5,499
	INV	39	14	6	0
Sniffles2	DEL	8,894	8,544	7,553	5,580
	INS	22,043	11,199	8,834	5,781
	DUP	27	27	13	2
	INV	56	46	13	0
	BND	259	123	22	1
SVIM	DEL	7,597	7,487	7,218	5,465
	INS	8,333	7,977	7,533	5,000
	DUP	98	87	10	2
	INV	8	8	7	0
	BND	68	54	17	1
SVIM-asm	DEL	8,795	6,763	5,999	5,502
	INS	11,250	8,590	6,955	5,771
	DUP	48	19	6	2
	INV	62	0	0	0
	BND	472	99	17	1

**Table 24.** Human aggregated SV caller outputs with minimum support from N other tools. Results are segmented by deletion (DEL), insertion (INS), duplication (DUP), inversion (INV), and translocation break end (BND) types. N+ columns denote SV counts with support from N or more other tools. Cells highlighted in grey represent SV types restricted to zero due to their absence in the original result of the tool being evaluated.



## 2.5 Human Genome Benchmarking results

The performance of each SV Caller, all tool combinations, and their aggregated sets were assessed using Truvari [3] ‘bench’ and the GIAB benchmarking set [9] to calculate their precision, recall, and F1 scores. The definitions and formulas for these metrics are provided in Table 25. The GIAB benchmark set defines several results tiers based on the SV location’s confidence, including Tiers I and II. For example, Tier I defines the SV locations with the highest confidence regarding their location and predicted length. By comparison, Tier II details the regions with strong SV evidence but whose sequence and size could not be reported confidently. The decision was made to include both tiers when benchmarking because both provide strong evidence of the SV locations. In addition, Truvari was not configured to analyze the reported ALT and REF sequences, so the lack of confidence in Tier II’s reported sequences would not negatively affect the assessment. The benchmark also provided several BED files that could be used to pre-filter the SV caller and benchmark results. Therefore, the benchmark was pre-filtered using bedtools [7] ‘intersect’ and these BED files to isolate the SVs in the Tier I and Tier II regions that overlapped the BED file locations by 100%. The choice to pre-filter the benchmark outside of Truvari was made because Truvari pre-filters both the benchmark and the SV caller inputs when provided a BED file. Additionally, cuteSV [6] and SVIM-asm [5] detail the Truvari commands they used in their publications’ supplementary materials, and both tools provided BED file arguments during their execution.

Metric	Definition	Formula
Precision	The percentage of the predicted results that were correct	$\text{True Positives} / (\text{True Positives} + \text{False Positives})$
Recall	The percentage of all the true results that the SV caller detected	$\text{True Positives} / (\text{True Positives} + \text{False Negatives})$
F1 Scores	The harmonic mean of the precision and recall scores	$(2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$

**Table 25.** Common accuracy metrics used to assess the performance of SV callers during benchmarking.

SV caller	Supporting Callers	True Positives	False Positives	False Negatives	Precision	Recall	F1 ratio
cuteSV	Individual Caller	11,497	13,186	1,234	46.58%	<b>90.31%</b>	61.46%
	1+ Other Callers	11,307	9,568	1,424	54.17%	88.81%	67.29%
	2+ Other Callers	10,377	6,544	2,354	61.33%	81.51%	69.99%
	3+ Other Callers	7,518	4,027	5,213	65.12%	59.05%	61.93%
PAV	Individual Caller	7,732	8,123	4,999	48.77%	60.73%	54.10%
	1+ Other Callers	7,623	6,527	5,108	53.87%	59.88%	56.72%
	2+ Other Callers	7,284	4,736	5,447	60.60%	57.21%	58.86%
	3+ Other Callers	6,898	3,649	5,833	65.40%	54.18%	59.27%
Sniffles2	Individual Caller	11,036	16,578	1,695	39.97%	86.69%	54.71%
	1+ Other Callers	10,945	7,661	1,786	58.83%	85.97%	69.85%
	2+ Other Callers	10,212	5,727	2,519	64.07%	<b>80.21%</b>	<b>71.24%</b>
	3+ Other Callers	7,439	3,774	5,292	66.34%	58.43%	62.14%
SVIM	Individual Caller	9,683	5,912	3,048	62.09%	76.06%	68.37%
	1+ Other Callers	9,642	5,514	3,089	63.62%	75.74%	69.15%
	2+ Other Callers	9,493	5,012	3,238	65.45%	74.57%	69.71%
	3+ Other Callers	6,961	3,418	5,770	<b>67.07%</b>	54.68%	60.24%
SVIM-asm	Individual Caller	8,328	11,717	4,403	41.55%	65.42%	50.82%
	1+ Other Callers	8,195	7,158	4,536	53.38%	64.37%	58.36%
	2+ Other Callers	7,866	5,088	4,865	60.72%	61.79%	61.25%
	3+ Other Callers	7,384	3,889	5,347	65.50%	58.00%	61.52%

**Table 26.** Benchmarking the individual and aggregated SV caller outputs against the Tier 1 and Tier 2 entries of the GIABv0.6 benchmark. Due to the benchmark set's contents, an evaluation was only possible for the deletion (DEL) and insertion (INS) type totals. N+ value in column two denotes required support from N or more other tools.

Intersection of	True Positives	False Positives	False Negatives	Precision	Recall	F1 ratio
cuteSV $\cap$ PAV	7,125	4,542	5,606	61.07%	55.97%	58.41%
cuteSV $\cap$ Sniffles2	10,672	7,024	2,059	60.31%	<b>83.83%</b>	<b>70.15%</b>
cuteSV $\cap$ SVIM	9,504	4,977	3,227	65.63%	74.65%	69.85%
cuteSV $\cap$ SVIM-asm	7,678	4,927	5,053	60.91%	60.31%	60.61%
PAV $\cap$ Sniffles2	6,972	3,827	5,759	64.56%	54.76%	59.26%
PAV $\cap$ SVIM	6,334	3,161	6,379	<b>66.71%</b>	49.75%	56.99%
PAV $\cap$ SVIM-asm	7,311	5,636	5,420	56.47%	57.43%	56.94%
Sniffles2 $\cap$ SVIM	9,400	4,686	3,331	<b>66.73%</b>	73.84%	<b>70.10%</b>
Sniffles2 $\cap$ SVIM-asm	7,595	4,356	5,136	63.55%	59.66%	61.54%
SVIM $\cap$ SVIM-asm	6,834	3,479	5,897	66.27%	53.68%	59.31%

**Table 27.** Benchmarking of intersected pairs of SV caller outputs against the Tier 1 and Tier 2 entries of the GIABv0.6 benchmark. Due to the benchmark set's contents, an evaluation was only possible for the deletion (DEL) and insertion (INS) type totals.

Intersection of	True Positives	False Positives	False Negatives	Precision	Recall	F1 ratio
cuteSV $\cap$ (PAV $\cap$ Sniffles2)	6,765	3,464	5,966	66.14%	53.14%	58.93%
cuteSV $\cap$ (PAV $\cap$ SVIM)	6,250	2,932	6,481	68.07%	49.09%	57.04%
cuteSV $\cap$ (PAV $\cap$ SVIM-asm)	6,839	3,803	5,892	64.26%	53.72%	58.52%
cuteSV $\cap$ (Sniffles2 $\cap$ SVIM)	9,297	4,431	3,434	67.72%	<b>73.03%</b>	<b>70.27%</b>
cuteSV $\cap$ (Sniffles2 $\cap$ SVIM-asm)	7,356	3,835	5,375	65.73%	57.78%	61.50%
cuteSV $\cap$ (SVIM $\cap$ SVIM-asm)	6,732	3,176	5,999	67.95%	52.88%	59.47%
PAV $\cap$ (Sniffles2 $\cap$ SVIM)	6,215	2,803	6,516	<b>68.92%</b>	48.82%	57.15%
PAV $\cap$ (Sniffles2 $\cap$ SVIM-asm)	6,768	3,423	5,963	66.41%	53.16%	59.05%
PAV $\cap$ (SVIM $\cap$ SVIM-asm)	6,159	2,880	6,572	68.14%	48.38%	56.58%
Sniffles2 $\cap$ (SVIM $\cap$ SVIM-asm)	6,700	3,074	6,031	68.55%	52.63%	59.54%

**Table 28.** Benchmarking of intersected combinations of three SV caller outputs against the Tier 1 and Tier 2 entries of the GIABv0.6 benchmark. Due to the benchmark set's contents, an evaluation was only possible for the deletion (DEL) and insertion (INS) type totals.

Intersection of	True Positives	False Positives	False Negatives	Precision	Recall	F1 ratio
(cuteSV $\cap$ PAV) $\cap$ (Sniffles2 $\cap$ SVIM)	6,141	2,679	6,590	69.63%	48.24%	56.99%
(cuteSV $\cap$ PAV) $\cap$ (Sniffles2 $\cap$ SVIM-asm)	6,565	3,116	6,166	67.81%	51.57%	58.58%
(cuteSV $\cap$ PAV) $\cap$ (SVIM $\cap$ SVIM-asm)	6,070	2,676	6,661	69.40%	47.68%	56.53%
(cuteSV $\cap$ Sniffles2) $\cap$ (SVIM $\cap$ SVIM-asm)	6,633	2,947	6,098	69.24%	<b>52.10%</b>	<b>59.46%</b>
(PAV $\cap$ Sniffles2) $\cap$ (SVIM $\cap$ SVIM-asm)	6,039	2,586	6,692	70.02%	47.44%	56.56%
(cuteSV $\cap$ PAV) $\cap$ (Sniffles2 $\cap$ (SVIM $\cap$ SVIM-asm))	5,972	2,484	6,759	<b>70.62%</b>	46.91%	56.37%

**Table 29.** Benchmarking of intersected combinations of four and five SV caller outputs against the Tier 1 and Tier 2 entries of the GIABv0.6 benchmark. Due to the benchmark set's contents, an evaluation was only possible for the deletion (DEL) and insertion (INS) type totals.

SV caller	Prefilter	Benchmark Total	SV Caller Total	True Positives	False Positives	False Negatives	Precision	Recall	F1 ratio
cuteSV	Bench only (T1&2)	12,731	24,683	11,497	13,186	1,234	46.58%	90.31%	61.46%
	Truvari (T1&2)	12,731	21,608	11,488	10,120	1,243	53.17%	90.24%	66.91%
	Truvari (T1)	9,646	9,804	9,059	745	587	92.40%	93.91%	93.15%
PAV	Bench only Truvari (T1&2)	12,731	15,855	7,732	8,123	4,999	48.77%	60.73%	54.10%
	Truvari (T1&2)	12,731	14,701	7,721	6,980	5,010	52.52%	60.65%	56.29%
	Truvari (T1)	9,646	6,708	6,048	660	3,598	90.16%	62.70%	73.96%
Sniffles2	Bench only (T1&2)	12,731	27,614	11,036	16,578	1,695	39.97%	86.69%	54.71%
	Truvari (T1&2)	12,731	18,639	11,029	7,610	1,702	59.17%	86.63%	70.32%
	Truvari (T1)	9,646	16,443	8,853	7,590	793	53.84%	91.78%	67.87%
SVIM	Bench only (T1&2)	12,731	15,595	9,683	5,912	3,048	62.09%	76.06%	68.37%
	Truvari (T1&2)	12,731	14,888	9,678	5,210	3,053	65.01%	76.02%	70.08%
	Truvari (T1)	9,646	8,220	7,928	292	1,718	96.45%	82.19%	88.75%
SVIM-asm	Bench only (T1&2)	12,731	20,045	8,328	11,717	4,403	41.55%	65.42%	50.82%
	Truvari (T1&2)	12,731	18,262	8,320	9,942	4,411	45.56%	65.35%	53.69%
	Truvari (T1)	9,646	7,430	6,635	795	3,011	89.30%	68.78%	77.41%

**Table 30.** Benchmarking of the individual caller outputs comparing the effect of prefiltering just the benchmark with BEDTools or both the SV output and the benchmark with Truvari. Truvari 'bench' pre-filtering tested using BED files provided in the GIABv0.6 benchmark for entries from both Tier 1 and Tier 2 combined (T1&2) and from Tier 1 only (T1). These

results demonstrate how much SV caller precision can be overstated depending on how Truvari is configured during benchmarking. Due to the benchmark set's contents, an evaluation was only possible for the deletion (DEL) and insertion (INS) type totals.

## References

1. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H.: Twelve years of SAMtools and BCFtools. *Giga-Science* **10**(2) (02 2021). <https://doi.org/10.1093/gigascience/giab008>, <https://doi.org/10.1093/gigascience/giab008>, giab008
2. Ebert, P., Audano, P.A., Zhu, Q., Rodriguez-Martin, B., Porubsky, D., Bonder, M.J., Sulovari, A., Ebler, J., Zhou, W., Serra Mari, R., et al.: Haplotype-resolved diverse human genomes and integrated analysis of structural variation. *Science* **372**(6537), eabf7117 (2021)
3. English, A.C., Menon, V.K., Gibbs, R.A., Metcalf, G.A., Sedlazeck, F.J.: Truvari: refined structural variant comparison preserves allelic diversity. *Genome Biology* **23**(1), 1–20 (2022)
4. Heller, D., Vingron, M.: SVIM: structural variant identification using mapped long reads. *Bioinformatics* **35**(17), 2907–2915 (2019)
5. Heller, D., Vingron, M.: SVIM-asm: structural variant detection from haploid and diploid genome assemblies. *Bioinformatics* **36**(22-23), 5519–5521 (2020)
6. Jiang, T., Liu, Y., Jiang, Y., Li, J., Gao, Y., Cui, Z., Liu, Y., Liu, B., Wang, Y.: Long-read-based human genomic structural variation detection with cuteSV. *Genome Biology* **21**(1), 1–24 (2020)
7. Quinlan, A.R., Hall, I.M.: BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**(6), 841–842 (2010)
8. Sedlazeck, F.J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., Von Haeseler, A., Schatz, M.C.: Accurate detection of complex structural variations using singlemolecule sequencing. *Nature Methods* **15**(6), 461–468 (2018)
9. Zook, J.M., Hansen, N.F., Olson, N.D., Chapman, L., Mullikin, J.C., Xiao, C., Sherry, S., Koren, S., Phillippy, A.M., Boutros, P.C., et al.: A robust benchmark for detection of germline large deletions and insertions. *Nature Biotechnology* **38**(11), 1347–1355 (2020)