scientific reports

Check for updates

OPEN

# Benchmarking long-read structural variant calling tools and combinations for detecting somatic variants in cancer genomes

Safa Kerem Aydin, Kubra Celikbas Yilmaz & Ahmet Acar✉

**Cancer genomes have a complicated landscape of mutations, including large-scale rearrangements known as structural variants (SVs). These SVs can disrupt genes or regulatory elements, playing a critical role in cancer development and progression. Despite their importance, accurate identification of somatic structural variants (SVs) remains a significant bottleneck in cancer genomics. Long-read sequencing technologies hold great promise in SV discovery, and there is an increasing number of efforts to develop new tools to detect them. In this study, we employ eight widely used SV callers on paired tumor and matched normal samples from both the NCI-H2009 lung cancer cell line and the COLO829 melanoma cell line, the latter of which has a well-established somatic SV truth set. Following separate variation detection in both tumor and normal DNA, the VCF merging procedure and a subtraction method were used to identify candidate somatic SVs. Additionally, we explored different combinations of the tools to enhance the accuracy of true somatic SV detection. Our analysis adopts a comprehensive approach, evaluating the performance of each SV caller across a spectrum of variant types and numbers in finding cancer-related somatic SVs. This study, by comparing eight different tools and their combinations, not only reveals the benefits and limitations of various techniques but also establishes a framework for developing more robust SV calling pipelines. Our findings highlight the strengths and weaknesses of current SV calling tools and suggest that combining multiple tools and testing different combinations can significantly enhance the validation of somatic alterations.**

**Keywords** Tool benchmarking, Cancer genomics, Long-read sequencing, Somatic structural variants, SV calling tools, Whole-genome sequencing data

Structural variants (SVs) are substantial chromosomal rearrangements that can arise due to various mechanisms, including errors in DNA replication, repair processes, and random somatic events[1]. These alterations can span from small-scale changes of 50 base pairs to large-scale chromosomal abnormalities[2]. The impact of SVs on an organism's phenotype varies widely depending on their size, location, and functional consequences[3]. While some SVs may have minimal phenotypic effects, others can lead to severe genetic disorders or contribute significantly to the pathogenesis of complex diseases, including cancer[4].

Cancer, in particular, is characterized by a high prevalence of structural variations, which play a crucial role in shaping its complex genomic landscape[5]. SVs, such as insertions, deletions, duplications, inversions, and translocations, are key drivers of genomic instability in cancer cells[6]. These mutations directly influence gene function and regulatory networks, leading to dysregulation of critical cellular processes and promoting tumorigenesis and disease progression[7]. Notably, SVs have been linked to modifying the activity of tumor suppressor genes, activating oncogenes, and producing fusion genes that drive uncontrolled cell growth and proliferation, all of which are hallmarks of cancer development[8].

The study of structural variations is an emerging and rapidly evolving field within genomics, gaining increasing significance as we uncover their intricate roles in health and disease. In addition to cancer, SVs play a significant role in the genesis of several hereditary illnesses and complex human features[9]. Our capacity to identify and describe SVs with greater sensitivity and precision has been completely transformed by recent developments in DNA sequencing technology, such as the improvements in second and third-generation sequencing platforms[10].

Department of Biological Sciences, Middle East Technical University, Universiteler Mah. Dumlupınar Bulvarı 1, 06800 Çankaya, Ankara, Turkey. ✉email: acara@metu.edu.tr

Given the importance of SVs in cancer and disease development, new approaches and algorithms are being developed specifically for finding structural variations (SVs) in genomic data. There is a growing amount of research examining the usefulness of long-read sequencing technologies, such as those provided by PacBio and Oxford Nanopore, in finding SVs[11]. Long-read sequencing technologies have considerable advantages over standard short-read sequencing methods, including their capacity to cover larger genomic regions and capture complex structural variants with reliable accuracy[12,13]. As a result, researchers are actively investigating and testing the efficacy of long-read sequencing for SV detection in a variety of applications[14]. While existing tools primarily focus on standard germline variant detection, current approaches for identifying somatic structural variants (SVs) have limitations due to limited knowledge regarding their frequencies and phenotype effects[15]. Typically, the process of finding somatic variants involves merging VCF files and applying defined parameters to identify candidate SVs present in the tumor data but absent in the matched normal data, indicating a potential somatic event[16,17]. Although this approach is widely used, it often necessitates manual curation and validation using tools like IGV (Integrative Genomics Viewer), which can be time-consuming and prone to false positives such as germline variants due to differences in sequencing depth or random reading errors between datasets[18]. However, recent developments have led to the generation of tools specifically designed for directly detecting somatic SVs by analyzing tumor and matched-normal samples, offering a more streamlined and accurate approach to somatic variant identification[19].

Within the field of genomic research, notable progress has been made in identifying single nucleotide variations (SNVs) by the use of short-read methods on Illumina data, especially when considering clinical studies, which have historically been the main focus. Even though a lot of methods have been generated and evaluated for the detection of structural variants (SVs) as well as SNVs using a variety of sequencing technologies, finding a superior tool that performs well in every aspect is still challenging. Extensive benchmarking studies have improved our knowledge of SNV techniques, and there are studies that focus on combining SNV data from various tools into a multi-caller strategy[20]. This combined approach has demonstrated increased accuracy in outcomes, demonstrating the possible advantages of multiple tool integration[21]. It is interesting to note that while SV methods have advanced, there aren't many comprehensive studies in the literature, similar to SNV caller combination, on combining SV calls of tools for determining final SV calls, indicating a topic that needs more investigation and study.

In this study, we aimed to evaluate the performance of several long-read structural variant (SV) calling tools, including Sniffles[11], cuteSV[22], Delly[23], DeBreak[24], Dysgu[25], NanoVar[26], SVIM[27], and Severus[28]. Given the complexity of identifying somatic SVs, particularly in high-depth sequencing datasets from different platforms, we also explored how combining multiple tools could enhance the detection of these variants. By assessing various tool combinations, our goal was to improve the accuracy of somatic SV detection and refine the identification of candidate SVs. To support our evaluation, we leveraged a somatic structural variation (SV) truth set derived from the COLO829 melanoma cancer cell line[29]. Overall, integrating this robust truth set allowed us to confidently assess the performance of these tools in whole-genome sequencing studies, particularly in the context of somatic SV identification.

## Materials and methods

### The tumor/normal pair

We obtained the nanopore long-read data, sequenced on the MinION, GridION and PromethION platforms with R9.4 flow cells (Oxford Nanopore Technologies), for tumor and normal samples from publicly available repositories accessed via Bioproject IDs PRJDB10898 and PRJEB27698. PRJDB10898 corresponds to lung cancer adenocarcinoma, while PRJEB27698 belongs to melanoma cancer cell line and includes a well-established truth set which was used to test our current approach. Our selection criteria were based on identifying a lung cancer dataset recognized for its substantial presence of somatic structural variants in which a stage 4 adenocarcinoma cell line derived from lung tissue was selected. By incorporating the long-read dataset (PRJEB27698) with its truth set into our analysis, we aimed to validate the findings and the robustness of our approach.

An extensive assessment of eight distinct tools, each specified with its version above, was conducted with an aim to identify structural variations (SVs) in cancer using long-read sequencing data (Table 1). These tools offer a variety of functionalities tailored to address specific challenges in SV detection: cuteSV (v2.1.0) focuses on sensitive SV detection in long-read data; Sniffles2 (v2.2) is versatile in analyzing various data types and variant levels; NanoVar (v1.5.1) prioritizes accuracy, particularly on low-depth long reads; dysgu (v1.6.2) supports both short-read and long-read data and provides extensive control over filtering and merging processes; delly (v1.2.6) incorporates multiple signals for SV identification and enables copy number alteration analyses; DeBreak (v1.0.2) specializes in long-read SV discovery; SVIM (v2.0.0) excels at distinguishing similar SV types; and Severus (v0.1.1) focuses on somatic SV calling in tumor-normal analysis, utilizing long-read phasing capabilities[11,22–28].

### Separate SV calling on tumor and matched normal data

A quality assessment using the FASTQC (version 0.12.1) tool on both the tumor samples and the matched normal samples was performed[30]. This initial step allowed for the evaluation and documentation of per-sequence quality scores and total bases. Subsequent to this quality assessment, reference genome alignment was carried out using minimap2 (v2.22) and *GCA_000001405.15_GRCh38_no_alt_analysis_set.fa* reference genome[31], utilizing the long-read alignment parameter (-ax map-ont) to ensure accurate alignment of the sequences. Quality control measures of bam files were assessed using the Qualimap BAMQC tool (version 2.2.2)[32], enabling the extraction of coverage and mapping quality information essential for downstream analyses. Following these processes, both the fastq and aligned bam files were processed for downstream analyses, particularly the variant calling analysis (Fig. 1).

| Tool | Version | References | GitHub link |
|---|---|---|---|
| Sniffles | 2.2 | [26] | https://github.com/fritzsedlazeck/Sniffles |
| NanoVar | 1.5.1 | [25] | https://github.com/benoukraflab/NanoVar |
| Dysgu | 1.6.2 | [23] | https://github.com/kcleal/dysgu |
| Delly | 1.2.6 | [24] | https://github.com/dellytools/delly |
| DeBreak | 1.0.2 | [27] | https://github.com/Maggi-Chen/DeBreak |
| SVIM | 2.0.0 | [28] | https://github.com/eldariont/svim |
| Severus | 0.1.1 | [29] | https://github.com/KolmogorovLab/Severus |
| cuteSV | 2.1.0 | [11] | https://github.com/tjiangHIT/cuteSV |

**Table 1**. Structural variant calling tools used.

For the structural variant (SV) calling analysis, a standardized approach was employed with the specified minimum SV length set to at least 50 base pairs or more. Except for Severus, SV calling for all other tools was performed separately on both tumor and normal data using the following tools and parameters: *cuteSV --min_size 50*, *debreak --min_size 50*, *delly lr --minrefsep 50*, *dysgu run --min-size 50 -x*, *nanovar --minlen 50*, *sniffles*, *svim alignment --min_sv_size 50*, *severus --min-sv-size 50*.

The SV calling process was executed separately for all tools except Severus, which accepts both tumor and matched normal data and directly generates somatic SV calls.

## Filtering and merging VCF files

To perform filtering of VCFs based on SV length and quality, bcftools (version 1.8) with *bcftools view* command was used[33]. Variants not marked as PASS were filtered out using bcftools, and only PASS variants were considered. Also, for SVIM and delly, post-filtering was carried out via *bcftools filter* command. After the postprocessing steps, VCF files merging was performed using the SURVIVOR tool (version 1.0.7)[34]. The merge command operates using tumor and matched normal input VCF files as arguments and compares the SV calls across the files. It identifies overlapping or concordant SVs between tumor and matched normal and combines them into a single entry in the output VCF file. This process helps to consolidate SV calls and generate a comprehensive view of the detected structural variations as part of filtering the candidate somatic structural variants. We defined the parameters as following: *SURVIVOR merge 1000 1 1 0 0 50.* This step was designed to merge VCF files from the tumor and matched-normal samples in order to identify somatic structural variant (SV) candidates. The parameters specified in the command included a maximum distance of 1000 base pairs between breakpoints, requiring at least one supporting caller for an SV to be considered, taking the SV type into account, and setting a minimum SV size of 50 base pairs. Moreover, in merging parameters; by not considering the strands of SVs or estimating distance based on SV size, the command focuses on identifying large structural variants with high confidence levels. The merged output file contains the consolidated SV calls from both tumor and normal samples found in close proximity in the genome (within 1000 bp), facilitating further analysis and characterization of somatic SVs in the studied dataset. After the merging step, the number of SVs for tumor and matched-normal was counted, and the SV types were listed for both datasets and tools for comparison in the next steps. Annotation of the merged VCF files was carried out using ANNOVAR (version 2020-06-08)[35], a functional gene annotation tool, before applying a somatic filtering.

## Obtaining candidate somatic SVs

After merging tumor and normal SV calls based on a maximum 1000 bp distance, the variants that exist in the tumor but not in matched-normal for a defined region was considered as somatic candidates. Number of merging events were checked on IGV if they are present in the defined windows as listed in the VCF files.

Following data manipulation was carried out using R (v4.3.2) to identify somatic variant candidates from a merged tumor and normal VCF files, annotations were carried out using the ANNOVAR. Rows where the variant is not present in the tumor were removed based on genotype (GT) information. This step ensured that only variants with potential somatic significance, i.e., present in the tumor but not in the normal sample, were retained. After filtering, the filtered SVs were written to an excel file, representing potential somatic structural variations.

## Merging all somatic candidates of different tools

All somatic variant calls from various tools have been assembled into a single VCF file for easier handling and comparison during downstream analysis. The merge parameters employed in this step are the same as those used in the previous stage to merge tumor and normal VCFs, ensuring that the merging process is uniform. The combinatorial strategy applied in our study categorizes structural variants (SVs) based on the number of tools detecting them, using overlaps in genomic regions with the same SV type and similar lengths. Variants are grouped into categories called "combX," where "X" refers to the minimum number of tools that support the detection. For example, comb8 includes variants identified by all 8 tools, comb7 includes those detected by at least 7 tools (including those in comb8), and this approach continues down to comb2, which includes all variants detected by at least 2 tools. Variants in comb2 are the most inclusive, covering those supported by any 2 tools as well as all variants from higher categories like comb3, comb4, and so on. Our approach ensures that as the number of supporting tools increases, so does the confidence in the accuracy and reliability of the identified

variants. By classifying variants in this way, our study systematically assesses their consistency across different tools.

## Comparison of SVs derived from the outputs of different tools

Using both lung cancer and melanoma datasets, the number and distribution of somatic variants were assessed. The candidate somatic SVs identified by each tool were quantified using bcftools and visualized using R. The counts of various SV types—including insertion, deletion, inversion, duplication, and translocation—were identified for each tool. These curated data were then used to create bar graphs, illustrating the SV type distribution across the different tools.

For the melanoma datasets with truthset, the identified somatic variants were rigorously compared against the truthset. Initially, the same somatic variants detected by at least two variant callers were extracted, and the threshold was gradually raised to include those identified by three, four, five, six, seven, and all eight variant callers. These variants were subsequently organized into separate VCF files. Subsequently, we employed R scripts to match the identified variants with those in the somatic truth set. The matching criteria were defined by several parameters, including position, chromosome, structural variant (SV) length, and SV type, which were aligned with merging parameters of the SURVIVOR. The variants smaller than 50 base pairs were excluded from the truth set, resulting in a total of 62 somatic structural variants being considered for the analysis. Variants that met these exact criteria were recognized as present in the somatic truth set VCF file. This approach allowed us to systematically evaluate the effectiveness of various tool combinations in capturing variants validated by the truth set. Through the generation of comprehensive statistics, we assessed the performance of each tool and their combination in detecting known somatic variants.

## Generating SV intersection visualizations, circos plots, and statistical analyses

Intersection plots were generated to visualize the overlap among somatic variants using UpSetR package (version 1.4.0)[36]. Somatic outputs were visualized using circos plots using Circlize R package (version 0.4.16)[37]. In cases where no information was available in a row, the "NA" (not available) flag was used to denote the absence of variants for that particular tool. Following the merging of all somatic outputs, annotation was carried out to determine insights into the functional significance of the identified variants.

The melanoma somatic structural variant truth set was generated from comprehensive whole-genome sequencing data using multiple platforms—10× Genomics, Illumina HiSeq, Oxford Nanopore, and Pacific Biosciences. A gold-standard SV call set was curated through meticulous validation techniques, including PCR, hybrid capture, and Bionano optical mapping, ensuring a high-confidence reference for benchmarking tool performance[29].

For benchmarking the melanoma dataset, we evaluated the performance of various variant-calling tools using key metrics: true positives (TP), false positives (FP), and false negatives (FN). We calculated precision as the ratio of the number of TPs to the total sum of TP and FP numbers, which measures the accuracy of the tool's positive predictions. Recall, calculated as the number TPs divided by the sum of TP and FN numbers, reflects the tool's ability to identify all relevant positive t. The F1 score, the harmonic mean of precision and recall, combines these two metrics to provide a balanced measure of the tool's performance. This comprehensive evaluation allowed us to assess the effectiveness of each tool and their combinations in detecting somatic variants.

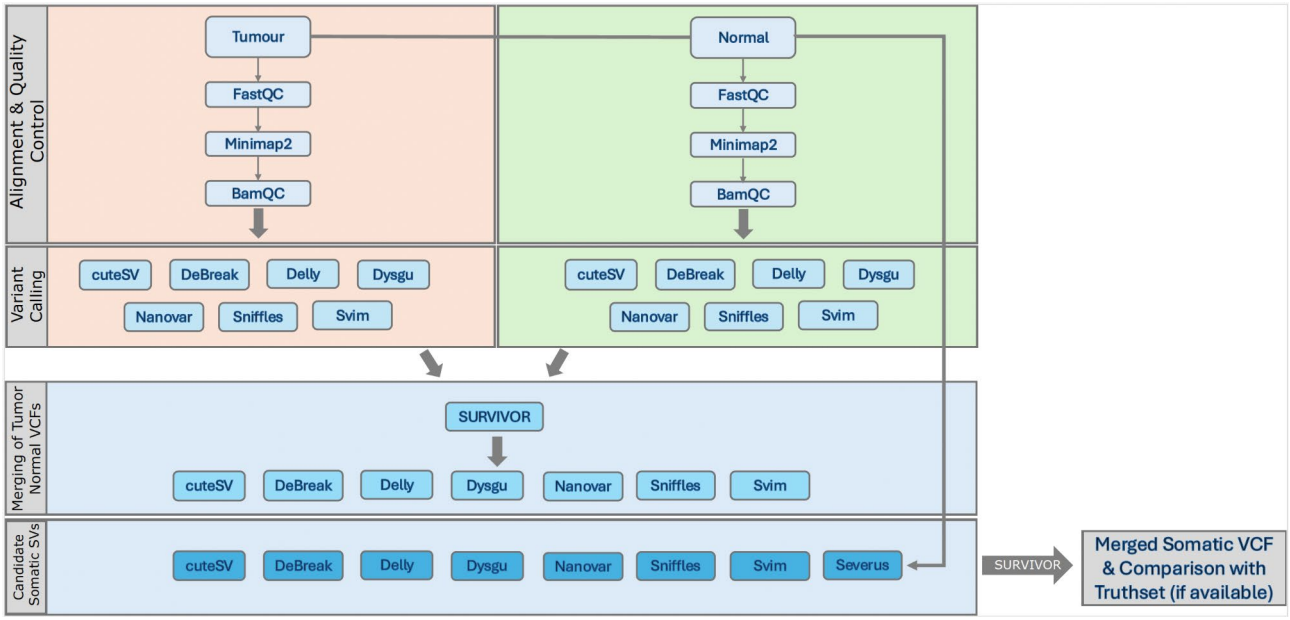


**Fig. 1**. Overall workflow of obtaining candidate somatic structural variants.

$$coverage\,percentage = \frac{TP}{total\ variants\ in\ truthset} \quad (62)$$

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 * precision * recall}{precision + recall}$$

## Results

### Quality Metrics and sequencing statistics

Firstly, the quality statistics for the tumor (H2009 and COLO829) and their matched-normal (BL2009 and COLO829BL) samples were obtained (Fig. 2). The mean sequence quality for H2009 and BL2009 was found as 19 and 18, respectively, suggesting nearly 99% accuracy according to Phred scores close to 20[38]. COLO829 and its matched-normal, COLO829BL, exhibited mean sequence qualities as 13 and 19, respectively. The H2009 dataset yielded 11,814,963 reads covering 118.3 gigabases (Gb) of data, with a mean mapping quality of 26.59 and a coverage of 36.717. In comparison, BL2009 produced 17,514,279 reads, covering 158.8 Gb with a mean mapping quality of 27.59 and a coverage of 49.8379. The COLO829 dataset has 18,688,039 reads, spanning 196 Gb, with a mean mapping quality of 43.22 and a coverage of 60.2679, while COLO829BL achieved 16,693,484 reads over 143.8 Gb, with a mean mapping quality of 43.14 and a coverage of 44.225. Despite the variation in coverage values across the datasets, all samples met the standards for long-read whole-genome sequencing. The higher coverage was achieved by the datasets which is particularly advantageous for the subtraction method filtering, as it relies on a dependable matched-normal baseline. Moreover, consistency between these datasets was confirmed and it has demonstrated no apparent discrepancies in the quality assessments of the FASTQ and BAM files.

*Evaluating speed and resource usage in SV callers*

To provide a comprehensive evaluation of structural variant (SV) calling tools, it is important to assess not only their accuracy but also their computational efficiency, as runtime performance significantly impacts their use in real-world applications. Fort his reason, we conducted an analysis of the runtime characteristics of seven widely-used SV callers—CuteSV, Debreak, Delly, Dysgu, Sniffles, SVIM, and Nanovar—using a normal dataset for each tool. Additionally, since Severus was uniquely developed for testing both tumor and normal datasets, its performance was evaluated using tumor and normal datasets.

Our evaluation incorporated multiple metrics to provide a broader understanding of computational performance. These include user time, system time, elapsed time, CPU usage, and memory consumption. By testing all tools in a computational environment and using the same datasets, we ensured fair comparability of runtime performance.

A detailed breakdown of the performance results of various tools used in our study are provided in Table 2, including their command execution times, CPU usage, and elapsed times. CuteSV had a user time of 5740.93 s and a system time of 85.12 s, with a total elapsed time of 7 h, 0 min, and 33 s, indicating a high CPU usage of 1386%. Debreak exhibited significantly higher resource consumption, with a user time of 14,751.88 s, a system time of 153.94 s, and an elapsed time of 37 h, 13 min, and 41 s, resulting in a CPU usage of 667%. Delly required 11,868.11 s of user time and 253.96 s of system time, completing in 3 h, 22 min, and 38 s with 99% CPU usage. Dysgu had a user time of 8535.81 s, system time of 540.59 s, and finished in 2 h, 31 min, and 37 s, using 99% CPU. Severus consumed 23,774.12 s of user time and 925.37 s of system time, completing in 1 h, 9 min, and 49 s with a CPU usage of 589% by processing both datasets and directly output somatic variants. Sniffles had a user time of 4346.48 s, system time of 67.27 s, and completed in 20 h, 29 min, and 76 s with 358% CPU usage. SVIM used 9895.59 s of user time, 443.67 s of system time, and finished in 2 h, 52 min, and 49 s, with 99% CPU usage. Finally, Nanovar had a user time of 7098.97 s and a system time of 495.14 s, with a total elapsed time of 1 h, 53 min, and 18 s, showing 111% CPU usage (Table 2). These performance metrics provide insights into the computational demands and efficiency of each tool used in our study.

### Structural variant count distributions across tumor and matched-normal samples

Next, the distribution of variant counts across tumor and normal samples together with somatic candidates for multiple tools within the COLO829/COLO829BL and H2009/BL2009 datasets was checked. In both datasets, normal samples consistently exhibited the highest number of variants detected across all tools, followed by tumor samples (Fig. 3). This trend was evident regardless of the tool used for structural variant (SV) calling. The structural variant (SV) calls for both the COLO829 and H2009 datasets across all tools exhibited distinct patterns in SV type distributions, which varied depending on the dataset and the tool used. Each tool exhibited unique characteristics in its output, reflecting differences in sensitivity and specificity toward certain SV types, particularly when comparing tumor and matched-normal samples.

The insertions and deletions were the most consistently detected SV types across all tools in both matched-normal datasets. For example, in the COLO829 dataset, CuteSV identified 10,936 insertions and 15,958 deletions in the tumor sample, compared to 10,911 insertions and 16,380 deletions in the normal sample (Supplementary File). A similar pattern was observed in the H2009 dataset, where Nanovar and SVIM consistently found higher counts of insertions and deletions in the normal samples. This finding suggests that normal samples might have a higher baseline of certain types of SVs. On the other hand, significant variability was detected across all

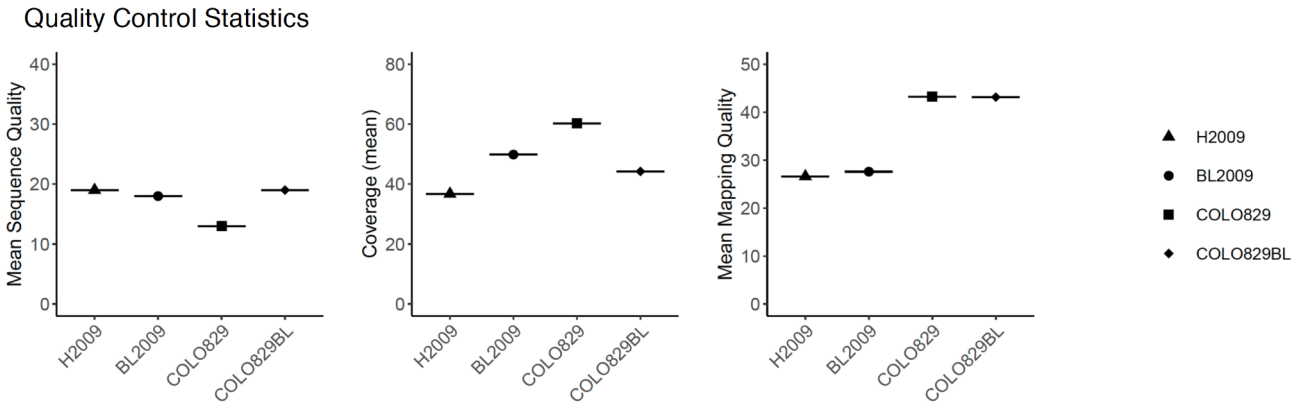


**Fig. 2**. Quality statistics for tumor and matched-normal samples. This figure compares the sequencing quality metrics, including mean sequence quality, coverage, and mean mapping quality, for tumor samples (H2009, COLO829) and their matched-normal counterparts (BL2009, COLO829BL).

| Tool | User Time (s) | System Time (s) | Elapsed Time (h:mm:ss) | CPU Usage (%) | Max Resident Set Size (KB) | File System Outputs (KB) |
|---|---|---|---|---|---|---|
| CuteSV | 5740.93 | 85.12 | 7:00.33 | 1386 | 7894208 | 8470288 |
| Debreak | 14751.88 | 153.94 | 37:13.41 | 667 | 2606972 | 1024840 |
| Delly | 11868.11 | 253.96 | 03:22:38 | 99 | 2605328 | 130096 |
| Dysgu | 8535.81 | 540.59 | 02:31:37 | 99 | 6683628 | 534109392 |
| Severus | 23774.12 | 925.37 | 01:09:49 | 589 | 64296924 | 125808 |
| Sniffles | 4346.48 | 67.27 | 20:29.76 | 358 | 2147636 | 39976 |
| SVIM | 9895.59 | 443.67 | 02:52:49 | 99 | 8603964 | 4361048 |
| Nanovar | 7098.97 | 495.14 | 01:53:18 | 111 | 49293888 | 302730104 |

**Table 2**. Runtime performance metrics and resource utilization of structural variant calling tools evaluated across consistent datasets and computational environments.

of the tools when more complex SV types, such as inversions, duplications, and translocations were analyzed. For instance, the inversion counts detected by Dysgu and DeBreak were relatively low across both datasets, often under 100 for both tumor and normal samples, indicating these tools may have stricter criteria for calling inversions or may be less sensitive to this SV type. In contrast, Nanovar and SVIM reported higher inversion counts, particularly in the COLO829 dataset, where Nanovar detected 131 inversions in the tumor sample and 108 in the normal sample. This discrepancy highlights the difference between the SV callers in detecting more complex rearrangements (Supplementary File). Of note, the translocations were detected as the most varying among the tools. In the COLO829 dataset, Nanovar detected an exceptionally high number of translocations, with counts exceeding 1000 for both tumor and normal samples, reaching up to 3002 in the normal sample. This contrasts sharply with Dysgu and DeBreak, which identified fewer than 100 translocations in the same samples (Fig. 4, Supplementary File). The variance suggests that Nanovar might have a broader definition of translocations or might capture a higher number of false positives, especially since it distinguishes between transpositions and translocations—the categories which we combined for uniformity. The large number of translocations reported by Nanovar, especially in the normal samples, could imply either an increased sensitivity to breakends or an overestimation of this SV type. Importantly, Severus, which specializes in directly identifying somatic variants, offered a streamlined output by combining its tumor and normal data into a single somatic variant count. In the COLO829 dataset, Severus identified 1714 somatic SVs, showing a more conservative and targeted approach compared to the other tools. This contrasts with Dysgu and Nanovar, which reported significantly higher numbers of candidate somatic SVs after applying the subtraction method. Specifically, in the H2009 dataset, Nanovar identified 3736 somatic SVs, while Dysgu identified 3813, indicating these tools might be more prone to detecting a broader range of candidate somatic variants, potentially including more false positives.

To obtain a detailed analysis of the different structural variant (SV) types—insertions (INS), deletions (DEL), inversions (INV), duplications (DUP), and translocations (TRA)—across various SV calling tools for the COLO829/COLO829BL and H2009/BL2009 datasets, the findings were plotted using barchart graph (Fig. 4). The significant differences were observed in the outputs, particularly for complex SV types like translocations, underscore the necessity for careful interpretation of SV calls. These findings highlight the importance of cross-tool validation to ensure the accuracy and reliability of results. In cancer genomics, where precise identification of somatic variants is essential, leveraging multiple tools and validating their outputs against each other are important. This approach helps mitigating the risks of false positives or negatives and ensures a more comprehensive and accurate characterization of the genomic landscape.

## Overlaps and unique patterns across SV tools via intersection and circos plots

For the analysis of the overlapping and unique somatic SVs detected by multiple tools across the datasets, the interception plot was utilized (Fig. 5). In the H2009/BL2009 dataset, Dysgu, Nanovar and Delly identified a much higher number of distinct somatic SVs compared to Severus and Svim. On the other hand, only a very small subset of the SVs (284) were shared across all tools. The commonly detected variants, shown in the filled dots within the same column of the matrix, represented the most consistent and reliable somatic SVs. The overlap suggests that they are the strongest somatic candidates, benefiting from consistent detection across various tools. The majority of common variants in this set are primarily insertions and deletions, indicating that these SV types are more robustly detected by all tools. Moreover, 83 of variants were shared with all seven tools except Dysgu. For the COLO829/COLO829BL dataset, a similar pattern was found in which CuteSV, Dysgu, Delly and Nanovar detected a large number of unique SVs. On the other hand, Debreak, Sniffles and Svim detected significantly fewer variants. Despite the large number of variants identified by each tool individually, the number of variants common to all tools remains limited to 5, which is very lower than the H2009 dataset. This reflects that most variants are not uniformly detected across the tools, reinforcing the value of combining different tools to increase the reliability of detected somatic SVs. Moreover, tools such as Dysgu and Nanovar detected a higher unique SVs compated to any other tool. This performance suggests the possibility of false positives, as these variants are not corroborated by other tools. Interestingly, the common somatic SVs are largely insertions and deletions, with no duplications, inversions, or translocations detected by all tools. This discrepancy likely stems from methodological differences in how each tool detects these more complex SV types (e.g., DUP, INV, and TRA). Since each tool uses a unique algorithm to classify structural variants, they tend to focus on different parts of the genome, leading to a lack of agreement for certain SV categories. This absence of uniformity across complex SV types further underscores the importance of using multiple tools in parallel to enhance the sensitivity and ensure comprehensive detection of all somatic variants.

Notably, there were no consensus variants found by all tools for duplication, inversion, and translocation types. The number of variants detected by only one single SV caller was significantly higher than the common variants across all tools. However, there were numerous common variants observed for insertion and deletion types across all tools. Specifically, 125 insertion variants and 30 deletion variants were common among all 8 tools for H2009/BL2009 dataset (Supplementary Fig. 1). In the COLO829 dataset there were few overlapping variants for insertion and deletion, yet the number of shared variants for inversion, duplication and translocation were fewer compared to the H2009/BL2009 dataset (Supplementary Fig. 2).

To visualize somatic structural variant types compherensively, circus plots were generated. CuteSV, Severus, Sniffles, SVIM, and DeBreak findings exhibited similar patterns for insertion and deletion variants (Fig. 6A). In contrast, Delly, Dysgu, and Nanovar displayed more intense lines, particularly Nanovar, which exhibited a higher number of translocation events, making it challenging to track (Fig. 6A,B). Additionally, DeBreak, SVIM, and Dysgu tools produced fewer translocation events, (Fig. 6B). Notably, translocations between chromosome X and chromosome 1 were visible in all SV callers except the DeBreak (Fig. 6).

## Performance metrics of tools and combinations

The detection capabilities and coverage percentages of various SV calling tools when compared to a somatic truth set were calculated and summarized (Supplementary File and Table 3). Among the tools, Severus demonstrated the highest coverage percentage at 91.94%, detecting 57 variants present in the truth set out of

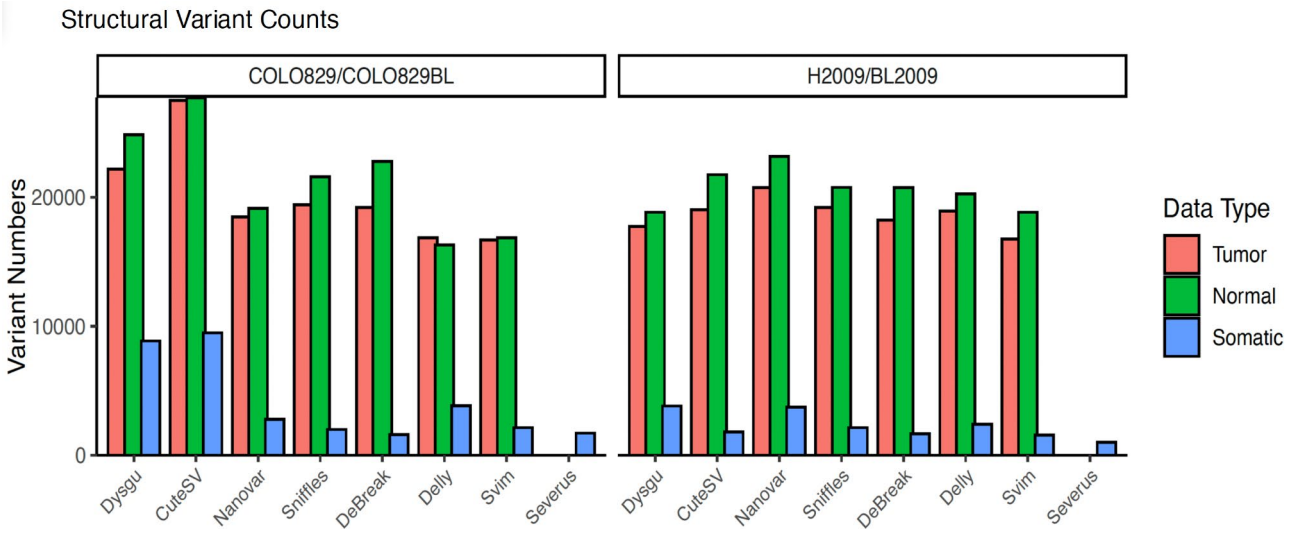


**Fig. 3**. Tumor, normal, and somatic SV counts across all tools for both lung adenocarcinoma and melanoma datasets. The graph illustrates the number of structural variants (SVs) detected by each tool, split into tumor (red), normal (green), and somatic (blue) categories. These counts provide insight into the detection capabilities of the tools and their effectiveness in distinguishing somatic variants across both cancer types.

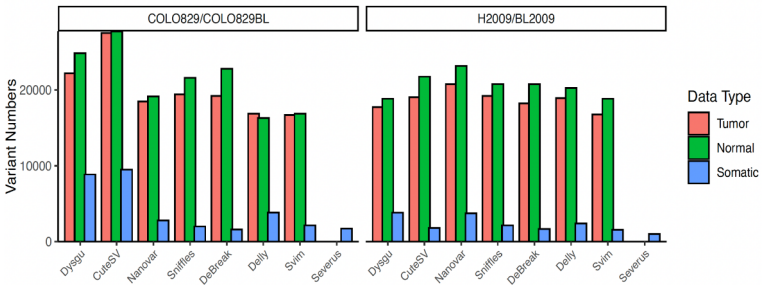


**Fig. 4**. Stacked bar plot of somatic structural variant counts across allSV calling tools datasets. This figure illustrates the distribution of structural variant (SV) types—insertions (INS), deletions (DEL), duplications (DUP), inversions (INV), and translocations (TRA)—across multiple SV calling tools for both the COLO829/COLO829BL and H2009/BL2009 datasets.

1,714 total calls, resulting in a detection rate of 3.33%. It was followed by comb5, which, despite detecting only 153 variants, achieved a coverage percentage of 38.71% and a notably high detection rate of 15.69%. In contrast, tools like Dysgu and Nanovar had lower coverage percentages, 20.97% and 30.65% respectively, reflecting less comprehensive detection of the truth set variants. Specifically, Dysgu identified 13 out of 8860 variants with a coverage of 20.97% and a detection rate of 0.15%, while Nanovar detected 19 out of 2792 variants with a coverage of 30.65% and a detection rate of 0.68%. Tools like cutesv and svim showed moderate coverage percentages of 45.16%, with detection rates of 0.29% and 1.31%, respectively. Notably, comb8 achieved a 100% detection rate, but this result was based on only 5 variants detected, which may indicate either a very precise or a very selective calling approach. These statistics highlight the variability in tool performance, emphasizing the need for careful consideration of tool capabilities based on specific research requirements.

Severus achieved the highest recall rate at 0.919, identifying 57 true positives out of 62 total variants (TP + FN), with a precision of 0.033 and an F1 score of 0.064 (Fig. 7). This shows Severus's strong capability in identifying true somatic variants, although its precision remains relatively low. On the other hand, comb8, while having an ideal precision of 100% due to no false positives, yielded a very low recall rate of 0.081 and an F1 score of 0.149, reflecting its potential limitations in comprehensive variant detection. Comb6 showed a high precision of 0.294 and a reasonable F1 score of 0.265, despite a relatively low recall rate of 0.242. Notably, comb4 exhibited the highest F1 score of 0.140, with a balanced precision of 0.080 and a recall rate of 0.548. In contrast, tools like Dysgu and Nanovar, with F1 scores of 0.003 and 0.013 respectively (Table 4, Fig. 7), had lower overall performance, particularly in precision and recall, indicating challenges in accurately identifying somatic variants. These results underscore the importance of evaluating both precision and recall when selecting SV calling tools, as well as considering the trade-offs between comprehensiveness and accuracy in variant detection.

To investigate the capability of structural variant (SV) calling tools in detecting specific variant types validated in the truthset, we focused on analyzing the counts of insertions (INS), deletions (DEL), and other SV types (e.g., inversions, duplications, and translocations). Using both single tools and their combinations, we extracted the raw numbers of each SV type detected and compared these with the validated truthset (Supplementary File). Furthermore, we compared the number of validated INDELs detected by each single and combination of tools (Supplementary Fig. 3). These findings highlight the strengths of specific tools and combinations in detecting validated SV types, offering insights into their potential applicability for studies focused on structural variant detection. For example, we examined which tools are better at detecting translocations by evaluating their accuracy in identifying these variants in the truth set. In our analysis, we observed that Severus identified 16 translocation variants and SVIM identified translocation 14 variants that were confirmed in the truth set. In addition, a detailed information on the length of the structural variants and the tools that are most successful in detecting the larger variants are provided (Supplementary File). This allowed us to identify which tools excel at detecting larger rearrangements, which may involve significant portions of the genome. By focusing on these specific metrics, the most suitable tools based on the type of structural variant such as small insertions, deletions, or more complex rearrangements like inversions and translocations can be obtained.

The recall (blue line) values are consistently higher than precision and F1 scores across most tools and combinations, with tools like Delly, Severus, and some combinations (e.g., comb2, comb3) achieving recall values around or above 0.5. This indicates that these tools are particularly good at identifying a higher proportion of true positives out of the total actual positives. However, the precision (green line), which measures the accuracy of the positive predictions, tends to be much lower for individual tools like CuteSV, Dysgu, and Nanovar, indicating a higher number of false positives among the detected SVs (Table 4, Fig. 7). Interestingly, the combination strategies (comb5, comb6, comb7, comb8) show a significant increase in precision, especially in comb8, where precision reaches 1.0, meaning all positive predictions made were correct. This suggests that combining outputs from multiple tools effectively reduces false positives. However, this comes at the cost of recall, which decreases sharply in these combinations, particularly in comb7 and comb8. The F1 score (red line), which balances precision and recall, remains low for most tools and combinations, highlighting the difficulty in achieving both high recall and high precision simultaneously. The highest F1 scores are observed in the middle-range combinations (e.g., comb3, comb4) and reaching the highest point at comb6, suggesting a trade-off between precision and recall is necessary to optimize overall performance. Overall, results underscore the

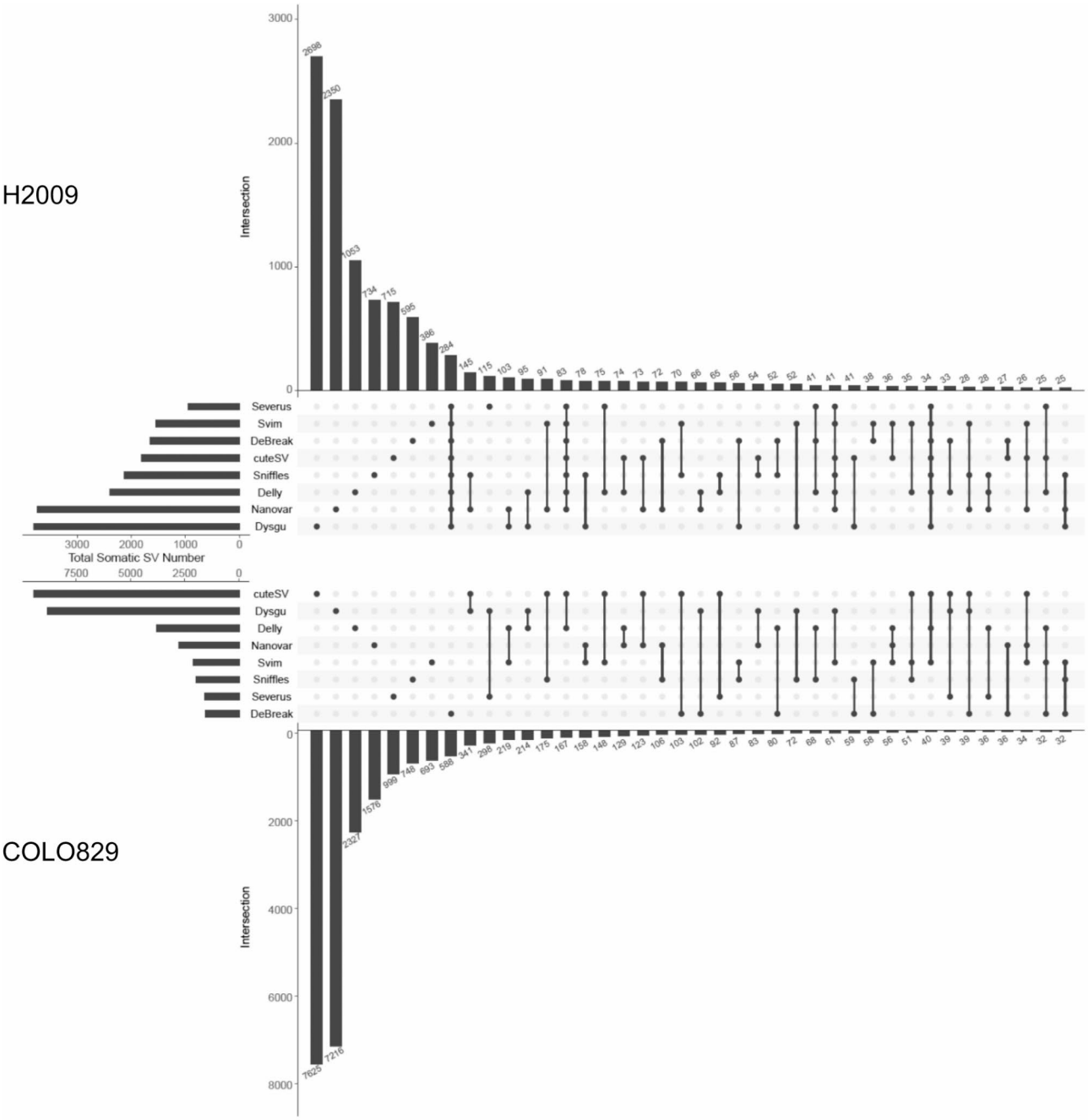


**Fig. 5**. Intersection of somatic structural variants detected by multiple tools in H2009/BL2009 and COLO829/COLO829BL pairs. The overlap of somatic structural variants (SVs) detected by multiple tools, with only 284 and 5 common variants across all tools in H2009/BL2009 and COLO829/COLO829BL datasets, respectively These common variants, detected by several SV callers, represent the strongest somatic candidates due to their consistent identification, increasing confidence in their validity.

variability in performance across different SV detection tools and combinations, emphasizing the importance of selecting the right approach depending on whether the focus is on maximizing recall, precision, or balancing both.

## Discussion

The study conducted a comprehensive comparison of eight long-read structural variant (SV) calling tools, including Sniffles, cuteSV, Delly, DeBreak, Dysgu, NanoVar, SVIM, and Severus, focusing on their efficacy in detecting cancer-related somatic SVs. This analysis yielded valuable insights into how these tools perform in identifying similar variants across various SV types and counts. Additionally, the study explored random combinations of SV callers that could effectively uncover somatic SVs. The findings highlighted the inadequacy

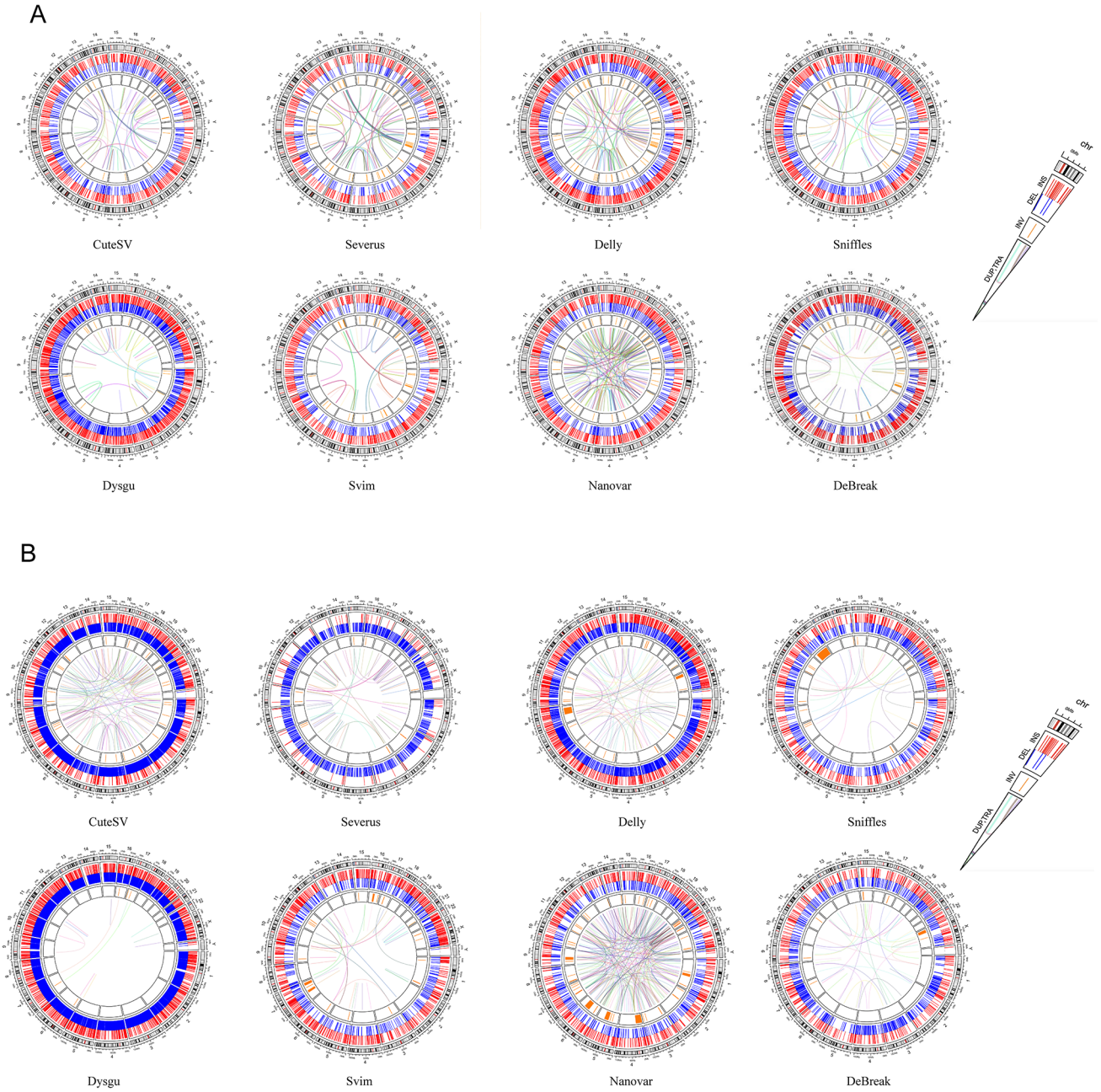


**Fig. 6**. Circos plots illustrating somatic structural variants detected by each tool, categorized by variant types (INS, DEL, INV, DUP, and TRA) across both datasets for H2009/BL2009 somatic structural variant counts (A), and for the COLO829/COLO829BL somatic structural variant counts (B). The color-coding and arc positioning indicate the SV type and location of each variant, providing a comprehensive way to observe structural rearrangements, the frequency of specific variants, and their distribution across different datasets. The outer side of the plot displays chromosome numbers, while the legend from inside to outside showcases duplication-translocation, inversion, deletion, and insertion events. Each red line represents a distinct insertion event, blue lines signify separate deletion events, and orange lines denote inversion events. Translocation events between chromosomes are displayed in the middle with connecting lines, while duplication variants are shown as individual lines in the center.

of relying solely on a single tool for somatic SV detection in research, as it may result in false positive germline SVs or miss important SVs that can be detected by other advanced tools employing distinct methodologies. Given the ongoing development of long-read somatic SV callers, approaches such as combining tools and SV calls are crucial for achieving reliable results.

The combinations included in Supplementary File (e.g., comb2, comb3, comb4…comb8) were generated based on the overlapping variants detected by the eight tools included in our analysis. These combinations were created without any assumptions, predefined criteria, or prior knowledge about the performance of individual

| Tool | Variants Detected | Present in Truth Set | Coverage Percentage | Detection Rate |
|---|---|---|---|---|
| cutesv | 9520 | 28 | 45.16% | 0.29% |
| sniffles | 1998 | 31 | 50.00% | 1.55% |
| delly | 3841 | 43 | 69.35% | 1.12% |
| dysgu | 8860 | 13 | 20.97% | 0.15% |
| debreak | 1595 | 43 | 69.35% | 2.70% |
| nanovar | 2792 | 19 | 30.65% | 0.68% |
| svim | 2139 | 28 | 45.16% | 1.31% |
| severus | 1714 | 57 | 91.94% | 3.33% |
| comb2 | 4310 | 52 | 83.87% | 1.21% |
| comb3 | 1257 | 46 | 74.19% | 3.66% |
| comb4 | 423 | 34 | 54.84% | 8.04% |
| comb5 | 153 | 24 | 38.71% | 15.69% |
| comb6 | 51 | 15 | 24.19% | 29.41% |
| comb7 | 20 | 9 | 14.52% | 45.00% |
| comb8 | 5 | 5 | 8.06% | 100.00% |

**Table 3**. Detection summary of SV calling tools and tool combinations.
Overview of the detection results from various structural variant (SV) calling tools and their combinations, including the total number of variants detected, the number of variants in the truth set, coverage percentage, and detection rate.

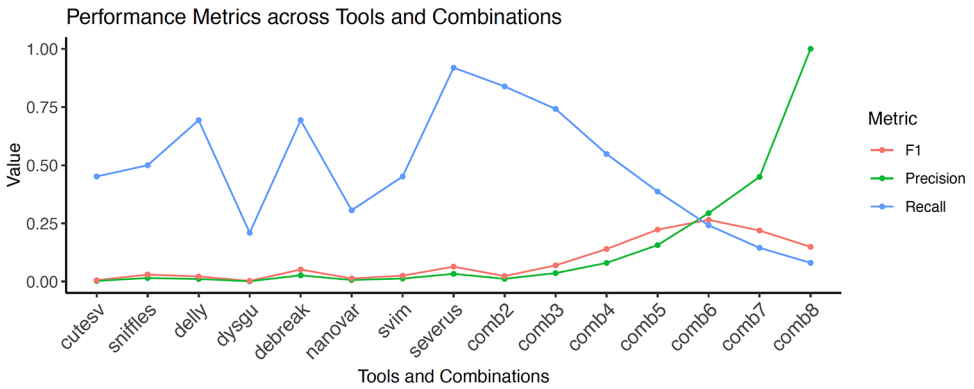


**Fig. 7**. F1 score, precision, and recall performance line plot. The performance metrics—F1 score, precision, and recall—across different tools and their combinations in detecting structural variants (SVs).

tools. The process was entirely random, with the focus solely on identifying shared variants across different tools rather than selecting tools based on specific characteristics or metrics. This unbiased approach ensures that the combinations reflect the collective performance of the tools in detecting overlapping structural variants, providing a comprehensive and impartial analysis.

Additionally, while the combination strategy involves integrating more than one tool, we also calculated and evaluated the individual performance of each tool using metrics such as the F1 score. This could provide making informed decisions when selecting specific tool combinations. Although the pairing of tools in comb2 was random, the shared variants identified by specific combinations of tools can still be identified (Supplementary File). In addition to random combinations for the "comb2" category, we also calculated performance metrics such as precision, recall, and F1 scores for each individual tool. These metrics provide a clearer understanding of how each tool performs in detecting structural variants, particularly in terms of their ability to correctly identify true positives (recall) and avoid false positives (precision). By considering these performance metrics, our study not only explores random tool pairings but also allows for a more strategic selection of tool combinations based on the specific needs of the goal.

Comb6 (a combination of 6 SV calling tools) achieves a balanced performance by striking a good compromise between precision and recall, boasting a precision of 0.294 and the highest F1 score of 0.265. This indicates that Comb6 effectively minimizes false positives while still capturing a reasonable number of true positives. In contrast, Comb8 (the combination of 8 tools) delivers perfect precision, achieving a value of 1.0 due to the absence of false positives. However, this comes at the cost of a very low recall (0.081), meaning that while it detects SVs with high accuracy, it misses many true positive variants. As a result, its F1 score remains low, reflecting poor balance between detecting true variants and avoiding false positives.

| Tools | TP | FP | FN | Precision | Recall | F1 Score |
|---|---|---|---|---|---|---|
| cutesv | 28 | 9492 | 34 | 0.002941 | 0.451613 | 0.005844 |
| sniffles | 31 | 1967 | 31 | 0.015516 | 0.500000 | 0.030097 |
| delly | 43 | 3798 | 19 | 0.011195 | 0.693548 | 0.022034 |
| dysgu | 13 | 8847 | 49 | 0.001467 | 0.209677 | 0.002914 |
| debreak | 43 | 1552 | 19 | 0.026959 | 0.693548 | 0.051901 |
| nanovar | 19 | 2773 | 43 | 0.006805 | 0.306452 | 0.013315 |
| svim | 28 | 2111 | 34 | 0.013090 | 0.451613 | 0.025443 |
| severus | 57 | 1657 | 5 | 0.033256 | 0.919355 | 0.064189 |
| comb2 | 52 | 4258 | 10 | 0.012065 | 0.838710 | 0.023788 |
| comb3 | 46 | 1211 | 16 | 0.036595 | 0.741935 | 0.069750 |
| comb4 | 34 | 389 | 28 | 0.080378 | 0.548387 | 0.140206 |
| comb5 | 24 | 129 | 38 | 0.156863 | 0.387097 | 0.223256 |
| comb6 | 15 | 36 | 47 | 0.294118 | 0.241935 | 0.265487 |
| comb7 | 9 | 11 | 53 | 0.450000 | 0.145161 | 0.219512 |
| comb8 | 5 | 0 | 57 | 1000000 | 0.080645 | 0.149254 |

**Table 4**. Performance metrics for structural variant calling tools.
The performance of various structural variant (SV) calling tools based on true positives (TP), false positives (FP), and false negatives (FN). It includes precision, recall, and F1 score for each tool.

Severus, on the other hand, excels in recall with a value of 0.919, identifying most true positive SVs, making it highly sensitive to variant detection. However, this comes with a trade-off: its precision is notably low (0.033), resulting in a high number of false positives. This imbalance between detecting true positives and incorrectly identifying variants is reflected in its modest F1 score of 0.064. Collectively, these results highlight the different strengths and weaknesses of the tools, emphasizing the importance of choosing the right tool based on the specific goals of the analysis—whether it's minimizing errors or maximizing the detection of true variants.

A close examination of the variant comparison with the truthset reveals Severus as a standout tool, identifying 57 out of 62 somatic structural variants. Despite its comprehensive performance, Severus still misses some key variants, underscoring the importance of using additional tools to capture a more complete variant landscape. For instance, the variant truthset_55_1 in the truthset vcf is identified by Debreak and Nanovar, while the variant truthset_13_1 is detected solely by Delly, and the variant truthset_18_1 is captured only by Dysgu. These findings highlight the critical need to avoid relying exclusively on a single tool. Each tool brings unique strengths to variant detection, and this diversity is essential for ensuring all relevant variants are identified.

The H2009 dataset does not have an established somatic structural variant truth set; nevertheless, we performed variant calling and intersection analyses analogous to those conducted on the COLO829 dataset. Despite the lack of a truth set for comparison, we compiled a common variant list containing 284 variants detected by all tools (Supplementary File). Furthermore, we developed a list of variants identified by at least two tools (Supplementary File). This strategy, reinforced by the comprehensive analyses of the COLO829 dataset and its corresponding truth set, highlights the critical importance of identifying overlapping variants among different tools. By concentrating on these shared variants, we can enhance our understanding of the genomic landscape, paving the way for the discovery of novel variants with potential clinical implications. Moreover, this approach not only contributes to the robustness of variant detection but also aids in minimizing false positives by validating findings across multiple methodologies. Ultimately, identifying these common variants can be beneficial in constructing a new somatic structural variant truth sets of other cell lines, thereby improving the reliability and accuracy of somatic variant calls in future research. By combining tools, researchers can refine the scope of their findings and narrow down the results, focusing on strong consensus candidates that are most likely to be true variants. This strategy enhances the efficiency of downstream wetlab validation efforts by prioritizing the most reliable variant calls for further investigation.

The study highlights the significance of selecting the optimal combination of structural variant (SV) calling tools based on individual application goals, rather than making assumptions about which tools perform best. By focusing on common variations and removing biases, the study demonstrates how future studies could improve SV detection through methodical tool pairing led by comprehensive metrics for performance. Precision is emphasized in clinical genomics to reduce false positives, therefore tool combinations with high accuracy, such as least6, least7, or least8, are better suited for accurate diagnostic outcomes. In contrast, research applications that seek to uncover novel SVs may prefer techniques with higher recall, such as Severus and least2, in order to catch a broader range of possible true variants, even if it involves allowing some false positives. For general-purpose applications, combinations like least6 achieve an effective compromise between precision and recall, improving detection and accuracy for a wide range of genomic research. These combinations increase the detection of true positive variations while decreasing false positives, making them suitable for a wide range of genomic studies, including research and clinical applications.

While employing multiple sequencing Technologies (Oxford Nanopore Technologies and Pacific Biosciences can significantly improve variant resolution, it comes with a substantial financial burden, especially for researchers working in rare diseases or cancer genomics. Merging short-read and long-read data provides better accuracy but can be prohibitively expensive. In contrast, using multiple state-of-the-art tools for structural variant calling presents no additional cost, yet offers a deeper and more nuanced analysis of somatic structural variants. This approach enables researchers to hone in on true somatic structural variants with high confidence, offering a cost-effective alternative to expensive multi-sequencer strategies while maximizing detection accuracy. Ultimately,

leveraging multiple variant calling tools significantly enhances the identification of somatic structural variants, providing a robust, cost-efficient method for genomic research.

In summary, combining structural variant (SV) calls from different tools can be highly advantageous for researchers, ensuring the identification of consensus somatic SVs and reducing the risk of missing important variants. While similar approaches are common in single nucleotide variant (SNV) analyses, such a strategy has been relatively unexplored in long-read SV calling research. For rare SVs or studies focused on cancer SVs, adopting this approach could provide valuable insights and improve the accuracy and reliability of SV detection, enhancing the overall understanding of genomic alterations in various contexts.

Although this study highlights the importance of not relying on a single tool for detecting somatic structural variants, but rather combining results from several tools, there are some limitations to address. First, the lack of a truth set for many datasets prevents us from comparing the outputs of other datasets. Second, although we included the most widely used SV calling tools, there are several additional tools, such as nanomonsv[16], CAMPHOR[39], and NanoSV[40] etc., which should also be evaluated to achieve a more comprehensive assessment. Additionally, this study does not investigate the effects of coverage and tumor purity, which have been shown to influence results. Finally, this study was conducted using two tumor-normal pairs; to obtain more conclusive results, multiple datasets or simulated data should be simultaneously investigated.

The findings of this study are based on analyses conducted using two cell lines, NCI-H2009 and COLO829, the latter provides well-characterized ground truth datasets for benchmarking structural variant (SV) callers. While these datasets are invaluable for establishing a controlled framework for comparing tool performance, their applicability may be limited by the specific genomic characteristics of these cell lines. To enhance the broader relevance and generalizability of the results, future research should incorporate data from additional cancer types, including those with distinct mutational profiles, as well as primary tumour samples. Such an extension would enable the evaluation of SV caller performance across a wider spectrum of genomic and tumour contexts, capturing factors like intratumoural heterogeneity, purity variations, and diverse sequencing platforms.

This study highlights the benefits of combining multiple structural variant (SV) calling tools to improve the accuracy of somatic SV detection. By integrating various tools, researchers can leverage their respective strengths, leading to a more reliable identification of true somatic candidates and reducing the possiblity of false positives. This approach enhances the discovery of true positive somatic variants and boosts the overall reliability of somatic SV classification. Our analysis, which compared the performance of eight SV callers across different cancer datasets, underscores both the strengths and limitations of current methods. We emphasize the critical role of tool integration in advancing SV discovery and improving the precision of genomics research.

## Data availability

All scripts, pipelines and parameter settings used in the study are available on GitHub at https://github.com/skeremaydin/Long-Read-Tool-Comparison.

## Code availability

All scripts, pipelines and parameter settings used in the study are available on GitHub at https://github.com/skeremaydin/Long-Read-Tool-Comparison.

Received: 24 October 2024; Accepted: 3 March 2025
Published online: 13 March 2025

## References

1. Currall, B. B., Chiang, C., Talkowski, M. E. & Morton, C. C. Mechanisms for structural variation in the human genome. *Curr. Genet. Med. Rep.* **1**(2), 81–90. https://doi.org/10.1007/s40142-013-0012-8 (2013).
2. Alkan, C., Coe, B. P. & Eichler, E. E. Genome structural variation discovery and genotyping. *Nat. Rev. Genet.* **12**(5), 363–376. https://doi.org/10.1038/nrg2958 (2011).
3. Scott, A. J., Chiang, C. & Hall, I. M. Structural variants are a major source of gene expression differences in humans and often affect multiple nearby genes. *Genome Res.* **31**(12), 2249–2257. https://doi.org/10.1101/gr.275488.121 (2021).
4. Sudmant, P. H. et al. An integrated map of structural variation in 2,504 human genomes. *Nature* **526**(7571), 75–81. https://doi.org/10.1038/nature15394 (2015).
5. Ashby, C. et al. Structural variants shape the genomic landscape and clinical outcome of multiple myeloma. *Blood Cancer J.* **12**(5), 85. https://doi.org/10.1038/s41408-022-00673-x (2022).
6. Guan, P. & Sung, W. K. Structural variation detection using next-generation sequencing data: A comparative technical review. *Methods (San Diego, Calif.)* **102**, 36–49. https://doi.org/10.1016/j.ymeth.2016.01.020 (2016).
7. Ramroop, J. R., Gerber, M. M. & Toland, A. E. Germline variants impact somatic events during tumorigenesis. *Trends Genet.: TIG* **35**(7), 515–526. https://doi.org/10.1016/j.tig.2019.04.005 (2019).
8. Zhang, Y. et al. High-coverage whole-genome analysis of 1220 cancers reveals hundreds of genes deregulated by rearrangement-mediated cis-regulatory alterations. *Nat. Commun.* **11**(1), 736. https://doi.org/10.1038/s41467-019-13885-w (2020).
9. Sanchis-Juan, A. et al. Complex structural variants in Mendelian disorders: Identification and breakpoint resolution using short- and long-read genome sequencing. *Genome Med.* **10**, 95. https://doi.org/10.1186/s13073-018-0606-6 (2018).
10. Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: Ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **17**(6), 333–351. https://doi.org/10.1038/nrg.2016.49 (2016).
11. Sedlazeck, F. J. et al. Accurate detection of complex structural variations using single-molecule sequencing. *Nat. Methods* **15**(6), 461–468. https://doi.org/10.1038/s41592-018-0001-7 (2018).
12. Chaisson, M. J. P. et al. Multi-platform discovery of haplotype-resolved structural variation in human genomes. *Nat. Commun.* **10**(1), 1784. https://doi.org/10.1038/s41467-018-08148-z (2019).
13. Begum, G. et al. Long-read sequencing improves the detection of structural variations impacting complex non-coding elements of the genome. *Int. J. Mol. Sci.* **22**(4), 2060. https://doi.org/10.3390/ijms22042060 (2021).
14. van Dijk, E. L., Jaszczyszyn, Y., Naquin, D. & Thermes, C. The third revolution in sequencing technology. *Trends Genet.: TIG* **34**(9), 666–681. https://doi.org/10.1016/j.tig.2018.05.008 (2018).
15. Lupski, J. R. & Stankiewicz, P. T. *Genomic Disorders: The Genomic Basis of Disease* (Springer Science & Business Media, 2007).

16. Shiraishi, Y. et al. Precise characterization of somatic complex structural variations from tumor/control paired long-read sequencing data with nanomonsv. *Nucleic Acids Res.* **51**(14), e74. https://doi.org/10.1093/nar/gkad526 (2023).
17. Yang, L. A practical guide for structural variation detection in the human genome. *Curr. Protoc. Hum. Genet.* **107**(1), e103. https://doi.org/10.1002/cphg.103 (2020).
18. Bertolotti, A. C. et al. The structural variation landscape in 492 Atlantic salmon genomes. *Nat. Commun.* **11**(1), 5176. https://doi.org/10.1038/s41467-020-18972-x (2020).
19. Yi, D., Nam, J. W. & Jeong, H. Toward the functional interpretation of somatic structural variations: bulk- and single-cell approaches. *Brief. Bioinform.* **24**(5), bbad297. https://doi.org/10.1093/bib/bbad297 (2023).
20. Kim, S. Y., Jacob, L. & Speed, T. P. Combining calls from multiple somatic mutation-callers. *BMC Bioinform.* **15**, 154. https://doi.org/10.1186/1471-2105-15-154 (2014).
21. Trevarton, A. J., Chang, J. T. & Symmans, W. F. Simple combination of multiple somatic variant callers to increase accuracy. *Sci. Rep.* **13**, 8463. https://doi.org/10.1038/s41598-023-34925-y (2023).
22. Jiang, T. et al. Long-read-based human genomic structural variation detection with cuteSV. *Genome Biol.* **21**(1), 189. https://doi.org/10.1186/s13059-020-02107-y (2020).
23. Rausch, T. et al. DELLY: Structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics (Oxford, England)* **28**(18), i333–i339. https://doi.org/10.1093/bioinformatics/bts378 (2012).
24. Chen, Y. et al. Deciphering the exact breakpoints of structural variations using long sequencing reads with DeBreak. *Nat. Commun.* **14**, 283. https://doi.org/10.1038/s41467-023-35996-1 (2023).
25. Cleal, K. & Baird, D. M. Dysgu: efficient structural variant calling using short or long reads. *Nucleic Acids Res.* **50**(9), e53. https://doi.org/10.1093/nar/gkac039 (2022).
26. Tham, C. Y. et al. NanoVar: accurate characterization of patients' genomic structural variants using low-depth nanopore sequencing. *Genome Biol.* **21**(1), 56. https://doi.org/10.1186/s13059-020-01968-7 (2020).
27. Heller, D. & Vingron, M. SVIM: structural variant identification using mapped long reads. *Bioinformatics (Oxford, England)* **35**(17), 2907–2915. https://doi.org/10.1093/bioinformatics/btz041 (2019).
28. Keskus, A., Bryant, A., Ahmad, T., Yoo, B., Aganezov, S., Goretsky, A., Donmez, A., Lansdon, L. A., Rodriguez, I., Park, J., Liu, Y., Cui, X., Gardner, J., McNulty, B., Sacco, S., Shetty, J., Zhao, Y., Tran, B., Narzisi, G., Helland, A., Cook, D. E., Chang, P.-C., Kolesnikov, A., Carroll, A., Molloy, E. K., Pushel, I., Guest, E., Pastinen, T., Shafin, K., Miga, K. H., Malikic, S., Day, C.-P., Robine, N., Sahinalp, C., Dean, M., Farooqi, M. S., Paten, B. & Kolmogorov, M. Severus: Accurate detection and characterization of somatic structural variation in tumor genomes using long reads. *bioRxiv* [Preprint]. https://doi.org/10.1101/2024.03.22.24304756 (2024).
29. Espejo Valle-Inclan, J. et al. A multi-platform reference for somatic structural variation detection. *Cell Genomics* **2**(6), 100139. https://doi.org/10.1016/j.xgen.2022.100139 (2022).
30. Andrews S. *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Available online at http://www.bioinformatics.babraham.ac.uk/projects/fastqc (2010).
31. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics (Oxford, England)* **34**(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191 (2018).
32. Okonechnikov, K., Conesa, A. & García-Alcalde, F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics (Oxford, England)* **32**(2), 292–294. https://doi.org/10.1093/bioinformatics/btv566 (2016).
33. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *GigaScience* **10**(2), giab008. https://doi.org/10.1093/gigascience/giab008 (2021).
34. Jeffares, D. et al. Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. *Nat. Commun.* **8**, 14061. https://doi.org/10.1038/ncomms14061 (2017).
35. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**(16), e164. https://doi.org/10.1093/nar/gkq603 (2010).
36. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**(18), 2938–2940. https://doi.org/10.1093/bioinformatics/btx364 (2017).
37. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular visualization in R. *Bioinformatics (Oxford, England)* **30**(19), 2811–2812. https://doi.org/10.1093/bioinformatics/btu393 (2014).
38. Delahaye, C. & Nicolas, J. Sequencing DNA with nanopores: Troubles and biases. *PloS One* **16**(10), e0257521. https://doi.org/10.1371/journal.pone.0257521 (2021).
39. Fujimoto, A. et al. Whole-genome sequencing with long reads reveals complex structure and origin of structural variation in human genetic variations and somatic mutations in cancer. *Genome Med.* **13**(1), 65. https://doi.org/10.1186/s13073-021-00883-1 (2021).
40. Cretu Stancu, M. et al. Mapping and phasing of structural variation in patient genomes using nanopore sequencing. *Nat. Commun.* **8**, 1326. https://doi.org/10.1038/s41467-017-01343-4 (2017).

## Acknowledgements

Ahmet Acar would like to acknowledge Republic of Türkiye The Council of Higher Education Research Universities Support Program (Grant number: ADEP-108-2022-11202).

## Author contributions

All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by S. K. A., K. C. Y, and A. A. All authors read and agreed with the final version of the manuscript.

## Funding

This research received no funding.

## Declarations

## Competing interests

The authors declare no competing interests.

## Consent for publication

The authors declare that they consent to publication.

## Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-92750-x.

**Correspondence** and requests for materials should be addressed to A.A.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.