Genetics in Medicine (2023) 25, 100316

ELSEVIER

Genetics in Medicine
An Official Journal of the ACMG

www.journals.elsevier.com/genetics-in-medicine

ACMG STATEMENT

Points to consider in the detection of germline structural variants using next-generation sequencing: A statement of the American College of Medical Genetics and Genomics (ACMG)

ARTICLE INFO

Article history:
Received 26 September 2022
Received in revised form
29 September 2022
Accepted 30 September 2022
Available online 12 December 2022

Keywords
Balanced rearrangements
Copy number variants
Exome sequencing
Genome sequencing
Structural variants

Introduction

Clinical laboratories are increasingly using next-generation sequencing (NGS) to detect germline structural variants (SVs), including both copy number variants (CNVs) and balanced chromosomal rearrangements (BCRs). Initially, the evaluation of intragenic deletions and duplications was performed in conjunction with gene-panel testing; however, clinical laboratories now also use exome sequencing (ES) and genome sequencing (GS) for genome-wide assessment of SVs.

Although technically feasible, simultaneously testing for a wide range of genomic alterations (including potentially single-nucleotide variants [SNVs], indels, CNVs, BCRs, regions of homozygosity [ROHs], mitochondrial variants, and repeat expansions [REs]) is a complex task, which may present clinical laboratories with unanticipated technical and interpretive challenges. In addition, the availability of a single test for a variety of variant types introduces a new dilemma for clinicians regarding how to best incorporate this type of comprehensive testing into the diagnostic workup of a patient. Should comprehensive genomic sequencing only be considered upon the failure of more targeted testing to provide a definitive answer, or should it be used as the first-line test to shorten the diagnostic odyssey, increase cost-effectiveness, reduce use of health care resources, and obtain a diagnosis through less invasive means? Are there genetic conditions and variant types

The Board of Directors of the American College of Medical Genetics and Genomics approved this statement on September 19, 2022.
*Correspondence: ACMG. *E-mail address:* documents@acmg.net
A full list of authors and affiliations appears at the end of the paper.

doi: https://doi.org/10.1016/j.gim.2022.09.017

for which comprehensive sequencing used as the first-line test might not reach the sensitivity and accuracy of the specific targeted assays?

Here we address some of these questions as they relate to NGS-based clinical evaluation for unbalanced germline variants (copy number gains and losses ranging in size from a single exon to a whole-chromosome), which is offered as a routine test by an increasing number of diagnostic laboratories. We describe important considerations for both clinical laboratories and clinical providers for comprehensive germline CNV detection using short-read sequencing (SRS) assays (including gene panels, ES, focused exomes/exome slices/virtual panels, and low-pass and ≥30X GS) as well as other sequencing technologies such as long-read sequencing (LRS) and complementary nonsequencing methods such as optical mapping (OM).

Evaluation for balanced germline variants (translocations, inversions, and balanced insertions) using sequencing data and other novel technologies is briefly discussed. Although it is still not readily available as a clinically validated test, such analysis is likely to become routine in the near future, thanks to increasing use of GS and rapid advancement of computational tools and novel methodologies like LRS and OM. This document also mentions evaluation for ROH and disease-causing REs, but does not include detailed discussion of these variants. Although clinically significant SVs in mitochondrial DNA do exist, mitochondrial SV detection is outside of the scope of this document. Somatic SV detection is also outside of the scope of this document and will be addressed through a separate joint project between the American College of Medical Genetics and Genomics (ACMG) and the Cancer Genomics Consortium (CGC).

In the following sections, we first describe the types of variants that are the focus of this document and list traditional and emerging approaches for their detection in the clinical setting. This is followed by a brief description of SRS-based and other novel methodologies used for the detection of SVs, including a discussion of their key advantages and limitations (Figure 1). Next, we lay out the best practices and points to be considered by clinical laboratories that already offer testing for SVs using NGS or are considering introducing this type of analysis. Finally, we elaborate on the important points to be considered by providers when deciding whether to prioritize comprehensive genome analysis over more targeted testing for SV detection.

Abbreviations

AF – allele frequency
BCR – balanced chromosomal rearrangement
CMA – chromosomal microarray
CNV – copy number variant
ES – exome sequencing
GS – genome sequencing
ID/DD – intellectual disability/developmental delay
LRS – long read sequencing
MCA – multiple congenital anomalies
MLPA – multiplex ligation-dependent probe amplification
MV – mitochondrial variant
NGS – next-generation sequencing
NIPS – noninvasive prenatal screening
OM – optical mapping
POC – products of conception
QC – quality control
qPCR – quantitative polymerase chain reaction (PCR)
RE – repeat expansion
ROH – regions of homozygosity
SNV – single nucleotide variant
SOP – standard operating procedure
SRS – short read sequencing
ssDNA – single-stranded DNA
STR – short tandem repeat
SV – structural variant; includes both CNVs and balanced rearrangements

Relevant Variant Types

CNVs

In addition to established assays for copy number gains and losses, including chromosomal microarrays (CMAs), karyotype analysis, fluorescence in situ hybridization (FISH), multiplex ligation-dependent probe amplification (MLPA), and quantitative polymerase chain reaction (qPCR), CNV detection from NGS has also entered routine clinical practice. Gene-panel testing often includes detection of intragenic deletions and duplications, and GS can be used for genome-wide CNV assessment. The broad and relatively uniform sequencing coverage gives ≥30X GS a considerable potential to reliably identify CNVs ranging in size from single exon to whole-chromosome aneuploidy. Nevertheless, the accurate detection of CNVs using methods that incorporate target enrichment (such as panels and ES), as well as the detection of small (exon-level) CNVs, remain challenging. Sensitivity, specificity, and resolution for CNV detection are variable and assay dependent.[1,2] Resolution for CNV detection may be difficult to define and may vary between genomic regions.[1,3]

Balanced SVs

Conventional cytogenetic analysis remains the gold standard for the detection of balanced germline rearrangements in the clinical setting; however, owing to its low resolution, it does not enable the identification of potentially disrupted genes at the breakpoints, and as a result, it does not allow laboratories to decipher the molecular mechanisms through which a specific rearrangement may be causing a disease. SRS has the potential to detect balanced SVs with base-level resolution with the use of appropriate assays and analytical tools.[3] Special short-read library preparation methods (eg, mate-pair/long-insert libraries) have been designed to detect BCRs throughout the genome.[4,5] However, most of these approaches have only been used on a research basis, and clinical testing for balanced SVs by SRS remains limited. An increasing number of novel technologies, including GS

Assay Type	Variant Type			
	Copy Number Variants (CNVs)	Balanced Structural Variants (SVs)	Regions of Homozygosity (ROH)	Repeat Expansions (REs)
Karyotype	✓	✓		
Chromosomal Microarray	✓		✓	
Gene Panels	✓			
Exome Sequencing	✓		✓	
Virtual Panels	✓			
Low-Pass Genome Sequencing	✓		✓	
Genome Sequencing	✓	✓	✓	✓
Mate-Pair Sequencing	✓	✓		
Optical Mapping	✓	✓	✓	✓
Long-Read Sequencing	✓	✓	✓	✓

Figure 1 Comparison of common methods for detection of structural variants. Checkmarks indicate variants that can potentially be detected by each assay. Assay limitations are highlighted by the absence of checkmarks. Details and conditions relevant for detection of individual variant types are described in the text.

(short and long read) and OM, appear promising for the accurate and robust detection of balanced SVs.

ROH

Extended ROH (typically defined by clinical laboratories as being >3-5 Mb) may be identified during a diagnostic evaluation by CMA platforms that use SNV probes. Their presence alone is not diagnostic of a genetic disorder, but, depending on the genomic content involved, it may indicate an increased risk for recessive disorders and/or imprinting disorders (caused by uniparental disomy).[6,7] Thus, there is clinical utility in detecting ROH during a genome-wide assessment for sequence variants and SVs. Furthermore, SNV analysis may be necessary for the detection of polyploidy, which is important in the diagnostic evaluation of products of conception (POCs).[8] Using specialized analytical tools, ROH can be detected from GS and to a lesser extent from ES.[9,10] Resolution for ROH detection from sequencing is variable and assay dependent; it will be higher from GS but lower from ES and low-pass GS.[11,12]

REs

RE disorders represent a class of genetic conditions caused by expansions of short tandem repeats (STRs). They are traditionally diagnosed using assays based on PCR amplification and fragment size analysis or Southern blot analysis.[13,14] Reliable detection of large STR expansions by SRS has been difficult in the past, but it is becoming more attainable through the development of advanced analytical tools.[15-18] RE disorders (with confirmation by a gold-standard clinical assay) may soon become an integral component of comprehensive genomic sequencing in many clinical laboratories. Novel technologies such as LRS and OM offer unique advantages for evaluation of REs.[19-21] Advancement of LRS methods, with their potential to directly sequence and accurately assess the size of REs and detect additional diagnostic or prognostic markers such as repeat interruptions and DNA methylation state, may in the near future completely transform clinical diagnostics for RE disorders.[21]

Relevant Testing Methodologies

Gene panels

Gene panels are developed and optimized for specific clinical indications; target enrichment during library preparation and subsequent sequencing only includes the set of genes associated with the phenotype of interest (tens or hundreds of genes as opposed to ~20,000 genes included in ES).[22] Gene panels typically provide complete, high depth coverage for the coding regions of targeted genes, which tends to improve the performance for detection of full or partial gene deletions and duplications. In addition to screening for CNVs using sequencing data, laboratories often complement gene panels with custom or commercial exon-focused arrays or confirmatory qPCR assays, allowing them to provide clinical testing for not only sequence variants but also partial and full deletions and duplications involving the genes of interest.

ES

Owing to the biases introduced by the target-enrichment process and lack of coverage of noncoding regions, ES may not provide optimal performance for the detection of all CNVs. Larger CNVs (affecting an entire gene or multiple contiguous genes) can be detected, but intragenic and intergenic CNVs may not always be accurately identified.[2,23] In the clinical setting, analysis of ES data can be a useful screen for large CNVs if a case is negative for disease-causing SNVs and indels.

Virtual panels

Virtual panels, also referred to as custom panels or exome slices, represent the practice of sequencing the exome, genome, or a large set of genes and computationally parsing this single data set into multiple sets of genes (offered as separate tests for different indications). While sequencing information is analyzed for target genes, data outside the regions of interest are masked during analysis. From the perspective of CNV/SV analysis, virtual panels have all the limitations of ES but are being used by providers in the same manner as panels, namely for individuals with specific phenotypes and a relatively high prior probability of an abnormality in one of the analyzed genes. Laboratories typically attempt to complement any SNV/indel analysis with CNV assessment, especially if the analyzed genes have deletions and duplications as a common mutational mechanism.

Low-pass GS

Low-pass GS is performed at a reduced depth of coverage (eg, <15×) that does not allow for the reliable detection of sequence variants but supports affordable detection of genome-wide CNVs. Low-pass GS shows similar diagnostic yield and possible enhanced resolution for CNV detection compared with CMA testing in postnatal and prenatal clinical settings.[12,24,25] Depending on the coverage depth, testing for extended ROH may not be included or may have lower resolution than CMA.[25,26]

GS (≥30×)

GS is the most comprehensive genetic test currently available that can detect different types of SVs in addition to sequence variants. Relatively uniform coverage (as compared with targeted assays) allows for a higher resolution and better sensitivity and specificity for the detection of CNVs.[27] By using specialized analytical tools, extended ROH, REs, and balanced rearrangements can also be identified.[28-30] However, variant identification is largely limited to nonrepetitive regions of the genome where short reads can be reliably aligned.

Additional Established and Emerging Methods for the Detection of SVs

Mate-pair sequencing

Mate-pair/long-insert sequencing is a variation of standard paired-end sequencing in which DNA is fragmented to 2 to 5 kilobase (kb) inserts, circularized, and refragmented to standard paired-end sequencing length. This protocol enables the detection of balanced structural rearrangements[5,31-33] and CNVs (≥5 kb).[34] Research has demonstrated that mate-pair sequencing identifies >90% of cytogenetically visible BCRs, revises the breakpoint location by at least 1 sub-band in 93% of cases, and detects additional complexity below the resolution of karyotyping in 26% of cases.[35,36] Clinical studies showed that mate-pair sequencing accurately identified SVs previously detected by karyotyping, FISH, and CMA; in some patients, it allowed recognition of the causal genes by identifying breakpoint junctions with greater resolution.[32,33,37] The vast majority of BCRs that are missed by mate-pair sequencing are located in repetitive regions of the genome that are not accessible by SRS.

LRS

Also known as third-generation sequencing, LRS has several advantages over sequencing technologies that use short reads. SRS creates products up to 600 base pairs (bp) in size and has limited ability to evaluate repetitive regions and SVs. LRS can generate products larger than 10 kb in size and serves as a high throughput platform for characterization of genomes through highly contiguous assemblies.[38] Several studies showed technical feasibility of using LRS for the analysis of trinucleotide RE and human disease-

associated chromosomal SVs.[39-42] However, clinical use is limited by several factors including higher cost, lower throughput, and for some platforms, higher error rate relative to SRS. SRS appears to be as sensitive as LRS to detect SVs in gene dense regions of the genome with unique sequence, whereas LRS may perform better in repetitive regions.[43]

OM

Although OM does not entail NGS, it is briefly discussed because of its broad research use for detection of SVs, as well as a growing interest in its implementation in clinical testing. OM uses long DNA molecules (up to megabases in length) that are fluorescently stained and enzymatically labeled with fluorophore tags of an alternate color, targeting specific 6 bp recognition sites throughout the genome. The DNA is linearized within nanochannels and then imaged, with the color patterns producing digital barcodes that can be assembled into contigs and compared with a reference genome in which the locations of recognition sites are known,[44] allowing for detection of SVs in the test genome. OM can detect both balanced and unbalanced SVs with high sensitivity, specificity, and resolution[44-46] and can also characterize variants not accessible to SRS such as kb-long tandem repeats.[47] OM has revealed pathogenic SVs in postnatal and prenatal samples of patients with genetic disorders;[19,47-49] it does not sequence DNA, but may be used as a complementary assay to short-read NGS to perform comprehensive evaluation for SVs. Typically, the CNV resolution is limited by the amount of genomic space between fluorescent labels (~0.5-1.5 kb).[44]

Algorithms for the Detection of SVs From SRS

A large number of analytical tools for SV detection from sequencing data are currently available, ranging from open source to commercial solutions and software packages. They differ by the type of sequencing data (eg, targeted vs genome) and have different strengths and limitations. Laboratories should understand the performance characteristics of the tools they employ. Technical considerations for alignment of short reads can be different for the evaluation of SVs compared with that of SNVs and indels. Because aberrantly mapped reads are one of the main sources of evidence for SV detection, laboratories should use aligners and settings that allow retention, and support partial alignment, of reads that do not map completely to the reference genome. Regarding the algorithms that perform analysis of read alignments to detect SVs, they can be broken down into 3 categories according to the employed evidence type, including those based on: (1) depth of coverage, (2) discordant read pairs (read pairs from the same fragment whose alignments to the reference genome are incorrectly oriented or span a larger or smaller genomic distance than the average insert size of the library), and (3) reads that span a breakpoint junction (known also as split reads) (Supplemental Figure 1 and Figure 2). Algorithms that apply de novo assembly of human genomes for SV detection also exist;[50] however, they are primarily used for research purposes. Depth-based algorithms are essential for detecting aneuploidies, large CNVs, terminal deletions, microdeletions/duplications of recurrent genomic-disorder regions mediated by low-copy repeats, and other repeat-mediated CNVs; these abnormalities have important clinical implications, but may be missed by algorithms that only evaluate discordant read pairs and/or split reads. In contrast, algorithms that evaluate discordant read pairs and/or split reads are needed for the detection of BCRs and small CNVs (<5 kb), which can be missed by read-depth only algorithms. The combination of multiple algorithms (often referred to as an ensemble method) typically improves sensitivity to detect SVs[3,51-54] but is more computationally expensive than using a single algorithm, has its own mechanisms of producing false positive or false negative results, and also requires additional strategies for merging SVs identified by different methods. A good understanding of performance characteristics of individual analytical tools is required for their incorporation into the ensemble approach. The choice of the SV detection tool (or tools) should match desired performance and application. CNV detection at the resolution of CMA can be achieved with a single tool based on depth of coverage, but higher resolution as well as the detection of multiple types of SVs may require combining multiple tools. The performance of different combinations of open-source algorithms for SV detection from GS data has previously been described.[3,54,55] In addition to performance (sensitivity and specificity) in the detection of SVs, other considerations when choosing analytical tools include cost (in particular, for commercial software), time and computing power required to run the analysis, and ease of use.

Points to Consider for Laboratories—Method Selection

Introducing SV detection from genomic sequencing as a new assay

- The frequencies and types of SVs to be detected (CNVs, balanced SVs, ROH, REs) for a particular patient population and/or condition should be considered when choosing a specific assay to implement. For example, comprehensive SV assessment from SRS may be desirable if the laboratory predominantly tests pediatric and prenatal cases with complex phenotypes and has a high ES and/or GS volume. Laboratories that mostly test prenatal samples and POCs, in which the focus is on the detection of large-scale chromosomal abnormalities (aneuploidy and polyploidy), may

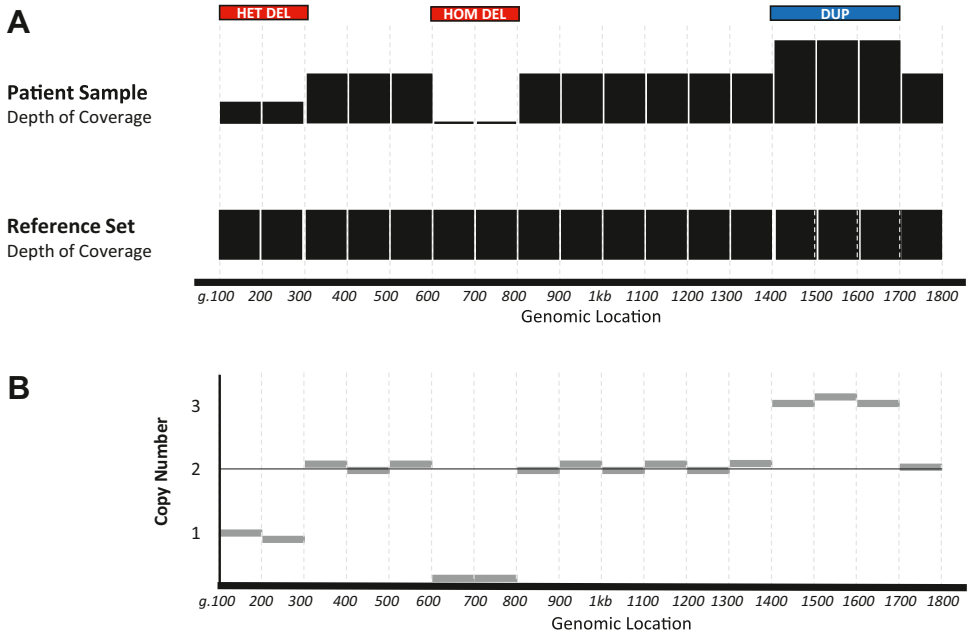


Figure 2 Copy number variant detection using depth of coverage of next-generation sequencing short reads. (A) A bin is a genomic region whose boundaries are specified by the laboratory (gray dotted lines; eg, 100 base pairs). After sequencing, map alignment, read filtering, and normalization, the number of reads in each bin of the patient sample (top) is compared with the number of reads in the corresponding bin of the reference set (pool of normal samples; bottom). Read-depth comparison coupled with complex copy number variant detection algorithms are applied to the data to identify deletions (red) and duplications (blue). Relative copy number information is calculated and can be translated into a user interface (B) for analysis and interpretation. DEL, deletion; DUP, duplication; HET, heterozygous; HOM, homozygous.

instead opt for CNV detection from low-pass GS or CMA testing.[25]

- The strengths and limitations of each assay for SV detection should be well understood when choosing a specific assay to implement. For example, low-pass GS typically does not allow for the detection of triploidy, which requires robust SNV data, and thus may be inferior to CMA for a laboratory with a high volume of POC cases.
- If switching from a traditional technology within the laboratory, the performance of the sequencing-based assay for SV detection should be at least comparable (although preferably superior) to the current method. The laboratory should consider whether the new method offers advantages in performance (sensitivity, resolution), price, etc that would justify investing resources into validation and implementation of the new technology.
- Sequencing data from genomes and larger capture-based libraries with backbone probes typically show more stability and less batch-to-batch variation than the libraries that target a smaller total percentage of the genome.[27] For sequencing assays developed to include CNV detection, laboratories, which have adequate resources, may consider GS with PCR-free library preparation, which avoids introduction of PCR and enrichment bias in the depth of coverage.[3] Alternatively, increasing the number of target areas and including backbone coverage may be used to increase the number of normal copy number anchors used to calculate relative copy number changes across the genome.[56]
- Laboratories may want to consider reimbursement in their decision to replace traditional SV assays with new technologies because SV assessment using comprehensive GS may not be covered by all payers.
- When introducing genomic SV detection as a new assay, the laboratory will need to decide whether the established assay will be discontinued or offered in parallel with the new assay. Although maintaining both assays may increase the cost, the established assay may have utility for confirmation of findings or testing a subset of samples for which the performance of the new assay is suboptimal.

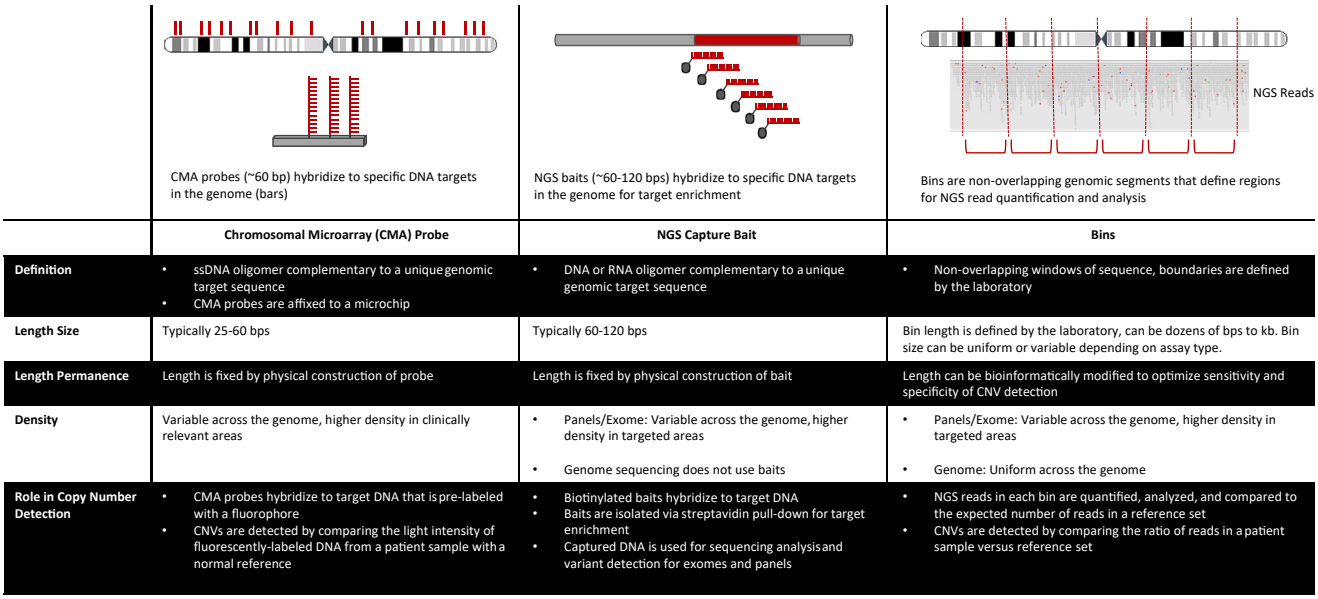
Choosing a solution for data analysis

- The decision regarding whether to develop a custom analytical pipeline or to use a commercial software package will likely depend on the bioinformatics expertise available in-house. Developing a custom pipeline for data analysis and other necessary infrastructure (a graphical interface for variant review and visualization, local or cloud-based data storage, network administration, etc) requires extensive bioinformatic expertise and continued support and will likely be a viable option only for laboratories that have access to an in-house bioinformatics team.

- Laboratories should choose algorithms (or a combination of algorithms) with the desired performance for the SVs being tested. A familiarity with the limitations and strengths of each algorithm is important, and a combination of algorithms with complementary strengths may be necessary for the detection of CNVs with high resolution or for the detection of BCRs with desired sensitivity and specificity.

Points to Consider for Laboratories—Validation and Protocol Development

- A validation should be performed that encompasses every type of SV targeted by the assay, and the size range should be defined for each type. If the analysis of a new type of SV is added at a later date, that variant type should be validated as well.
- Although it is not feasible to verify every genomic region, to ensure adequate assessment of the test performance, validation should include a sufficient number of unique samples that provide representation of different variant classes. Reference materials with robust genome-wide characterization of SVs have been developed and can be used for optimization, quality control, and to assess analytical performance of the assay.[57] For clinical validation, using at least 30 unique samples has been recommended for NGS assays and new CMA platforms.[58,59] It has also been suggested that evaluation of ≥59 variants of each type allows determination of performance characteristics of an assay with 95% confidence and 95% reliability.[60,61] For CNV validation, specifically, a selection of at least 30 samples with well-characterized CNVs of different sizes, copy number states, and contexts (segmental-duplications-mediated vs unique-breakpoint CNVs, high/low GC content, complex regions, etc) should be tested.[62] This should include specific validation samples for any recurrent microdeletion/duplication syndrome regions. Using samples with multiple previously detected abnormalities (for example with CNVs of various sizes and copy number states) may be an economical way to reach the desired number of variants for validation. This may be accomplished by running samples known from previous CMA testing to harbor those recurrent CNVs.
- Detected CNVs, their sizes, and boundaries may differ significantly between assays, depending on their resolution and other performance characteristics; establishing concordance between CNVs detected during validation and the ones identified by gold-standard methods may therefore be challenging. Laboratories are advised to define their concordance metrics in the validation plan and focus primarily on achieving high (>95%) concordance in the detection of clinically significant CNVs. The desired percentage of overlap (reciprocal overlap threshold) between CNV boundaries reported by 2 different methods should also be defined and may vary depending on how technically similar the methods are. For example, a 75% overlap may be applied when comparing CNVs between CMA and GS, and 90% when comparing GS with GS. When comparing GS with ES, the reciprocal overlap threshold may be defined based on affected genes and exons rather than the genomic coordinates of detected CNVs. If additional CNVs are detected by the method under validation (owing for example to its higher resolution), a subset of those (including at least the ones classified as pathogenic or likely pathogenic) should be confirmed by a third, independent method.
- Detection of CNVs by depth of coverage analysis using target enrichment–based assays (ES and gene panels) typically involves a comparison of the test sample to a reference.[63] The required reference set (also known as reference pool) should be defined as part of the validation. Reference samples should be processed in the same way as clinical samples and should be from the same specimen types. Laboratories may choose between using a dynamic reference (samples derived from the same batch) vs a static reference (same sample data are used as the reference across batches)[64,65] (see "Reference set quality for CNV detection from ES and gene panels"). A procedure should also be developed for when and how the reference set will be updated. For example, the laboratory may want to update the reference set when there are changes made in reagents, library preparation process, sequencing processes, or bioinformatics pipelines. CNV detection from GS data often uses intrasample normalization and may not require establishing a reference pool,[54] but data from normal controls and validation samples can be used to identify challenging genomic regions and technical artifacts.
- Appropriate quality control parameters should be determined during the validation (see "Quality assessment of CNVs").
- Laboratories should establish an SV confirmation policy that outlines which findings require confirmation and which can be reported solely based on the primary analysis. The confirmation policy may evolve over time as the laboratory gains sufficient experience with assay performance for the various types and sizes of SVs. The confirmation strategy may depend on the type of SVs being interrogated. Confirmation of CNVs may require a separate cytogenetic or molecular method such as karyotyping, CMA, FISH, or qPCR (which should also be a validated method in the laboratory). Some balanced abnormalities (like large inversions or translocations) may be confirmed by karyotype or FISH analysis, whereas submicroscopic changes may require amplification across the breakpoint junctions followed by Sanger sequencing. Increased confidence may be obtained through independent analyses that support the same SV (eg, if a deletion is detected by a depth of coverage algorithm



	Chromosomal Microarray (CMA) Probe	NGS Capture Bait	Bins
Definition	• ssDNA oligomer complementary to a unique genomic target sequence • CMA probes are affixed to a microchip	• DNA or RNA oligomer complementary to a unique genomic target sequence	• Non-overlapping windows of sequence, boundaries are defined by the laboratory
Length Size	Typically 25-60 bps	Typically 60-120 bps	Bin length is defined by the laboratory, can be dozens of bps to kb. Bin size can be uniform or variable depending on assay type.
Length Permanence	Length is fixed by physical construction of probe	Length is fixed by physical construction of bait	Length can be bioinformatically modified to optimize sensitivity and specificity of CNV detection
Density	Variable across the genome, higher density in clinically relevant areas	• Panels/Exome: Variable across the genome, higher density in targeted areas • Genome sequencing does not use baits	• Panels/Exome: Variable across the genome, higher density in targeted areas • Genome: Uniform across the genome
Role in Copy Number Detection	• CMA probes hybridize to target DNA that is pre-labeled with a fluorophore • CNVs are detected by comparing the light intensity of fluorescently-labeled DNA from a patient sample with a normal reference	• Biotinylated baits hybridize to target DNA • Baits are isolated via streptavidin pull-down for target enrichment • Captured DNA is used for sequencing analysis and variant detection for exomes and panels	• NGS reads in each bin are quantified, analyzed, and compared to the expected number of reads in a reference set • CNVs are detected by comparing the ratio of reads in a patient sample versus reference set

Figure 3 Difference between CMA probes, NGS capture baits, and bins for CNV analysis from NGS data. bps, base pairs; CNV, copy number variant; kb, kilobase; NGS, next-generation sequencing; ssDNA, single-stranded DNA.

and split reads containing nondeleted sequences from both sides of the deletion are also observed). However, a completely orthogonal method is generally recommended, at least until the laboratory gains confidence in the assay performance.

- The resolution for CNV detection should be defined during validation.
  - Important concepts for defining resolution are baits or capture probes for target enrichment–based assays (ES and gene panels) and analytical bins or windows for GS. Baits are physical DNA probes used during library preparation to select targeted regions. Bins are nonoverlapping regions of the genome within which the read depth is calculated and compared with the reference[66,67] (Figure 3). CNV detection strategies may define bin boundaries either by specific size (eg, 1 kb, 10 kb) and chromosomal coordinates (eg, an exon and/or a location corresponding to a bait target) or by the expected/desired number of reads within a region during reference set creation (eg, 1 bin = the genomic area containing a predefined number of reads in a set of normal samples). CNV detection typically requires that several consecutive bins show deviation from the reference. Size of analytical bins will modify resolution, sensitivity, and specificity of CNV detection. Smaller bin sizes allow higher resolution but increase probability of false positive calls (Figure 4). For capture-based sequencing and subsequent CNV analysis of both on- and off-target regions, bin sizes will be variable. High coverage regions will generate bins with shorter genomic segments, which will typically be more reliable for CNV detection than bins that span larger genomic segments but have lower coverage.[66,68]
- Thresholds for calling individual gains and losses should be set during validation to balance sensitivity and specificity; they may be defined according to the minimum size of detectable deletion or duplication or minimum number of consecutive affected bins; allowable level of variability between bins within a suspected CNV may also be defined (Figure 5).
- Because the most popular CNV calling algorithms are coverage dependent, considerations should be made for coverage variability when establishing assay resolution. This is particularly true for target-enrichment methods such as ES in which bait density and bin sizes may be uneven. Furthermore, variability of sequence coverage will further depend on GC content and degree of homology with other regions in the genome.[66] There is more variability in performance (or complete inability to make CNV calls) in highly homologous and GC rich regions. Owing to these considerations, it may not be feasible to give a single precise definition of resolution for CNV detection across an exome or genome. Consequently, a laboratory may quote higher resolution and sensitivity for CNV detection in high coverage, nonhomologous regions and lower sensitivity and resolution in low-coverage regions.
- Laboratories should provide the resolution for CNV detection in the test description; as mentioned earlier, the resolution may be expressed by the minimal number of affected exons, capture baits, or analytical

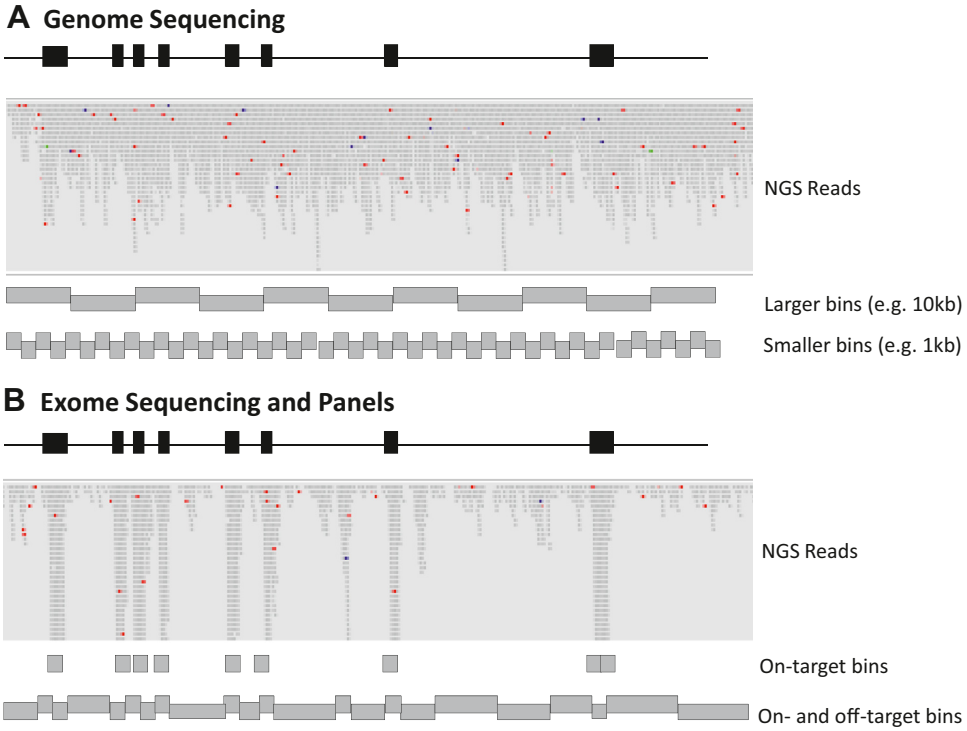


Figure 4 Coverage uniformity and bin size affect copy number variant detection sensitivity and specificity. In the example sequencing data for a gene with normal copy number dosage (black squares denote exons), genome sequencing (A) captures reads in both exon and intronic sequences, allowing for more coverage uniformity and higher breakpoint resolution than exome or panel sequencing (B), which do not routinely sequence noncoding regions. Bin size can be modified to balance sensitivity and specificity. Smaller bins (small gray bars) will typically yield greater sensitivity and higher breakpoint resolution but are more prone to false positives. Larger bins (large gray bars) give greater specificity, lower resolution, and may be at risk for higher rates of false negatives for smaller copy number variants (CNVs). For capture-based sequencing (B), bin boundaries may be fixed and located around on-target areas that focus interrogation on coding regions. Some CNV detection methods investigate both on- and off-target regions in which bin boundaries are defined by a user-specified number of reads. This methodology produces bins that are smaller in size in high coverage regions and larger in low-coverage regions. NGS, next-generation sequencing.

bins (used typically for target-enrichment assays such as targeted panels or ES) or based on size (typically used for GS).

- It may be necessary to independently define resolution for copy number gains and copy number losses, because these may vary significantly. Lower sensitivity and resolution may be determined for duplications, which are generally harder to detect than deletions.[69]
- When offering virtual panels or when testing patients whose phenotype suggests specific genetic conditions, it may be beneficial for the laboratory to provide resolution estimates for genes or regions of particular clinical interest with a disclaimer for the rest of the genome.
- Some clinically relevant areas of the genome may require combinatorial strategies to yield reliable CNV data. Genes with pseudogenes (eg, *CYP21A2*, *GBA*), high homology to other areas of the genome, or repetitive sequence may necessitate long-range PCR before sequencing and copy number analysis to specifically amplify the active gene and distinguish it from the pseudogene.[70]
- Challenging genomic regions and recurrent technical artifacts should be documented and cataloged during the validation process for each SV type and each workflow in the laboratory. For capture-based enrichment sequencing assays, bait performance is expected to be variable depending on hybridization efficiency and subsequent PCR bias. Baits that show high inconsistency in depth of coverage and poor performance for CNV analysis may still have utility for detecting sequence variants. Individual baits with suboptimal performance are accounted for by CNV callers looking at groups of neighboring bins (corresponding to multiple enrichment baits) to find trends that may indicate a copy number change (Figure 4). Increasing the number of samples analyzed during the validation increases the likelihood of detecting possible recurrent technical artifacts.
- Laboratories should choose strategies to address regions that are identified during the validation to have

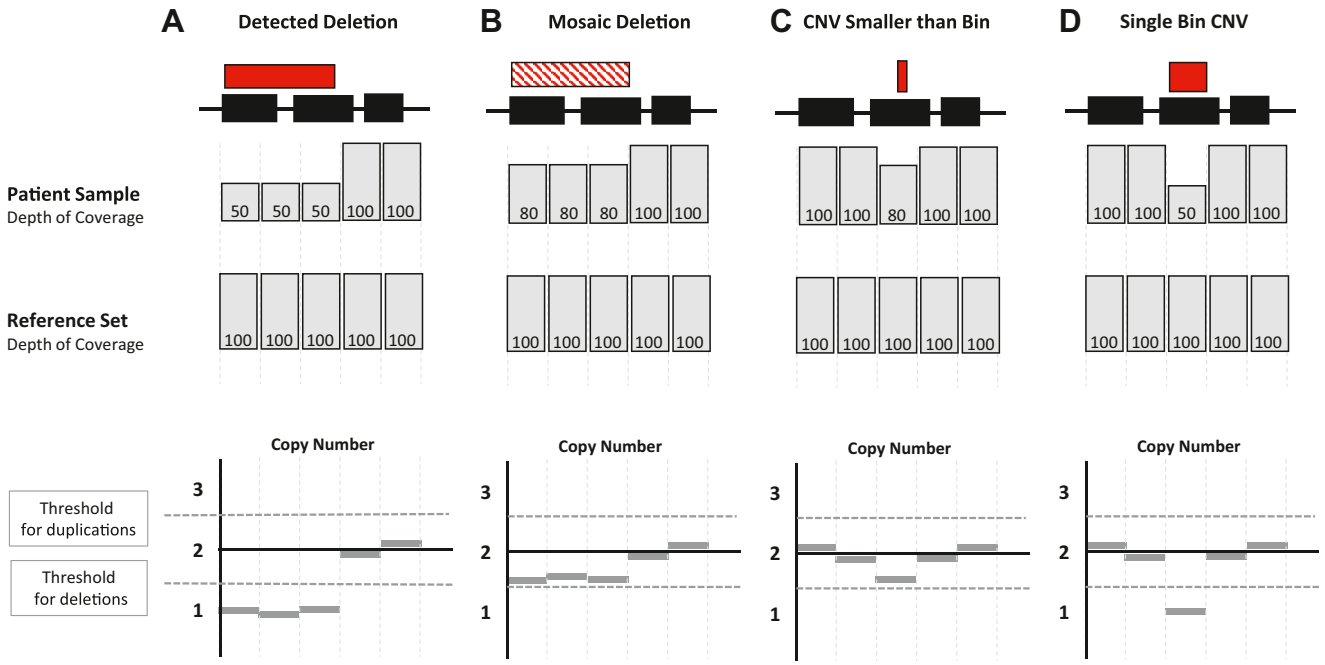


Figure 5 Examples of CNVs that may escape read depth–based detection methods. A detected deletion (A, red bar) results in significantly decreased read depth in multiple consecutive bins within the patient sample compared with the reference set. However, a mosaic deletion (B, striped bar) or deletion smaller than the target bin (C) may not result in significant enough changes in read depth to cross the threshold for detection. Depending on the sensitivity of the calling algorithm and coverage in the target region, decreased read depth in a single bin (D) may not be sufficient for CNV detection. The depth of coverage values used in this example are hypothetical. Actual sensitivity and specificity is dependent on multiple factors including the NGS library, depth of coverage, window sizes, and algorithm(s) used for CNV detection. CNVs that escape read depth–based detection methods may be identified through other CNV identification methods using NGS data (split reads, long-read sequencing, etc). CNV, copy number variant.

poor performance for CNV detection. The options include masking known regions of unreliable CNV detection (eg, high sequence homology and high variability in the read depth) upstream of analysis or identifying these recurrent regions as artifacts during analysis. Laboratories are advised to disclose the regions where copy number state cannot be calculated accurately, especially if they contain known dosage sensitive genes or loci.

Points to Consider for Laboratories—Quality Control and Proficiency Testing

General quality control

- Quality metric values should be established throughout all steps involved in specimen processing, including DNA extraction, library preparation, sequencing, reference set creation and implementation, bioinformatic data pipeline analysis, and SV data assessment.[62] The quality control (QC) metrics and standard cutoffs for SV detection should be established during the validation and defined in the laboratory standard operating procedures. Laboratory quality assurance programs should include monitoring of QC metrics on a regular basis to ensure clinical-grade SV detection performance and reliability.
- Metrics and cutoff thresholds established for SV analysis are separate from those applied to sequencing variant detection. Laboratories should have protocols in place for cases in which sequence variants meet QC requirements but SV results do not or vice versa. As a general principle, CNV detection is highly dependent on the depth of sequence coverage. Read depths as low as 15X to 30X are widely quoted in the literature as sufficient to accurately identify SNVs and small indels in reads from capture-based methods.[71-74] Specific read depth requirements for CNV detection will depend on the sequencing platform, assay type, enrichment method, and bioinformatics tools used. For CNV detection using GS, particularly when using PCR-free library preparation methods, most analytical tools show good performance with the standard read depths recommended for SNV detection (eg, 30X).[3] Low-pass GS with low depth of coverage (approximately 5X) has been used to accurately identify CNVs >25 kb in size but it is not reliable for the detection of CNVs of 25 kb or smaller and SNVs.[12] Assays based on target enrichment (ES and gene panels) have shown

inferior CNV detection at read depths of 20X when compared with read depths of more than 50X.[75,76] Higher coverage appears to allow for increased sensitivity; average read depths of 1000X to 2000X, achievable in gene-panel testing, have been shown to increase performance when identifying single-exon events.[77,78] Although sequence coverage may be sufficient for SNV detection, coverage that falls below the thresholds for CNV calling, or sequence coverage variability that is discordant with the reference set, may not yield reliable CNV data.

- Inclusion of a control sample with defined SVs in each sequencing batch/run may be considered to ensure that sufficient coverage and sequence quality were achieved to detect variants at the validated sensitivity and specificity of the assay. Inclusion of a positive control sample may also be considered if clinically significant recurrent CNVs are expected to be encountered based on the genes being interrogated.

Factors that affect sample data quality

The quality of sample-level data affects the sensitivity and resolution of CNV detection. Quality metrics for each sample should be evaluated before reporting patient results.

- DNA quality and specimen type: these factors greatly influence QC metrics particularly for CNV data derived from capture-based libraries. Samples with decreased DNA quality (eg, POCs, prenatal samples, postmortem samples) or specimen types with increased protein or chemical impurities (eg, saliva vs blood) may yield lower quality SV results owing to DNA degradation and/or contaminants that inhibit optimal DNA extraction and/or subsequent DNA processing and amplification. QC thresholds may vary across specimen types, DNA extraction methods (manual vs automated, different chemistries, etc), and NGS libraries.
- Depth of coverage: in general, sensitivity and resolution of copy number detection increases with deeper coverage and greater uniformity across the target regions. Laboratories should be aware of any library-specific regions of low coverage owing to technical difficulty capturing or mapping reads; these areas are prone to technical artifacts, in particular in samples with suboptimal quality and/or decreased coverage.[62,79] If the average sequence coverage of a sample is outside the minimum thresholds defined for the reference set, laboratories may design protocols to repeat sequencing or request additional samples as needed.
- Dosage uniformity score: the metric that evaluates how variable the coverage is within a sample is a useful QC parameter to identify poor-quality samples. Dosage uniformity outliers can have false positive and false negative calls; a repeat library preparation may be attempted if the dosage uniformity is outside of the acceptable range determined during the validation.[51] A new sample may have to be requested if data quality remains poor upon repeated sequencing attempts.

Reference set quality for CNV detection from ES and gene panels

- Samples used to construct each reference set should be of the same sample type and be processed in the same manner as the clinical samples to which the set will be applied. To avoid skewing of the expected number of reads at the reference baseline, reference samples should be devoid of aneuploidies and the number of CNVs within clinically relevant areas should be limited. The number of samples required to establish a reliable reference set should be determined based on the requirements of the selected analytical tools or by testing how change in the number of reference samples affects detection accuracy for known CNVs. In general, increasing the reference set size will decrease variability and improve CNV calling.
- Static reference sets (those that are established at one time point and applied to subsequent sequencing runs) are highly sensitive to changes that alter depth of coverage, including extraction method, library preparation and processing, and/or bioinformatic pipeline processing.[64,65] To ensure optimal usability of a static reference set, clinical samples should be processed as consistently as possible with minimal or no changes in the protocol. Changes in methodology may lead to recurrent false calls, inability to analyze CNVs, and/or need to construct new reference sets. There are currently no specific recommendations regarding the lifespan of static reference sets, and laboratories should develop quality metrics to monitor static reference set applicability over time and determine when a new reference set should be established.
- Dynamic reference sets are created for each sample batch after sequencing.[63] QC assessment for dynamic reference may include average and standard deviation of the number of CNVs detected per batch, incidence of CNV type and/or location, and quality metrics of specific samples used to create the reference sets. Known positive CNV controls may be included in each batch to check the sensitivity and resolution of dynamic reference sets.

Quality assessment of CNVs

Indicators and parameters to assess the quality of CNV determination and distinguish between technical artifacts and true copy number changes depend on the sequencing platform and analytical tools. However, some parameters are generally applicable and are similar to the strategies used to assess quality of CNVs in microarray analysis:

1. Ratio of sample to reference reads: the metric that compares the relative ratio of patient sample reads to the reference for each bin within the CNV should correspond to the expected copy number (heterozygous vs homozygous deletion, duplication vs triplication). Deviations from the expected values may indicate low quality CNV calling or potential mosaicism (Figure 5).
2. Number of affected bins: an increase in the number of consecutive affected bins correlates with the level of confidence of a call. The test to reference ratio (copy number) determined by individual bins within a suspected CNV should be consistent, with little variability (Figure 5).
3. Intersample bin variation: bin characteristics (sample to reference ratios) can be compared between samples of the same type that have been processed within the same batch to determine consistency.
4. Evidence of a breakpoint: sudden drops/elevations in depth of coverage (referred to as cliff-edges) along with the presence of split reads may be indicative of a CNV breakpoint (with discordant read pairs additionally supporting the presence of a CNV). Split reads that contain sequence from both sides of an expected deletion may be confirmatory. Split reads can also be used to determine the location and orientation of duplications or to screen for transposable elements or other insertions that disrupt a gene.
5. Variant allele fraction (VAF): VAF within the affected area should correlate with the expected copy number. Within heterozygous deletions, the majority of sequence variants should have approximately 100% fraction. For regions of suspected duplication, the VAF should be approximately 33% or 66%; however, it is important to be aware that even in the absence of copy number changes, VAFs for heterozygous variants may deviate from the expected value. If included in the user interface, SNV allele fraction tracks can investigate ROH as well as help to confirm deletions, duplications, and CNV mosaicism.
6. Manual review: laboratories should decide whether to include visualization and manual review in the process of quality assessment for SVs and whether manual review will be implemented for all reportable SVs or a subset that does not meet the predefined quality metrics. In addition to the Integrated Genome Viewer, which now supports multilocus view for SVs (https://pypi.org/project/igv-reports/), multiple open source (Collins RL, Stone MR, Brand H, Glessner JT, Talkowski ME. CNView: a visualization and annotation tool for copy number variation from whole-genome sequencing. 2016. https://doi.org/10.1101/049536),[80] and commercial solutions are available to support visualization and review of SVs detected from NGS data. Laboratories should be aware that manual review of SV calls relying on split read evidence may require specific settings to ensure that those reads are not filtered out of the display and can be visualized.

Filtering strategies for detected SVs

Analysis of ES and GS data may detect a large number of SVs in each sample,[3,43] depending on the sequencing technology (long-read vs short-read sequencing) and platform, library preparation method (standard vs mate pair), depth of coverage (low-pass vs 30X), number and type of applied detection algorithms (with more algorithms typically increasing the number of SV calls), stringency of QC metrics for SV calls, genetic ancestry, and other factors. To efficiently identify variants that may be related to the patient's phenotype, laboratories should employ clinically appropriate filtering strategies. The following characteristics may be used to filter out, annotate, or prioritize SVs:

1. Size and gene content: CNVs whose size is below the target resolution of the assay and that do not overlap known protein coding genes or regulatory elements can potentially be eliminated from further consideration.
2. QC metrics: SVs that do not meet predetermined QC metrics may be artifacts and can potentially be eliminated from further consideration.
3. Allele frequency in publicly available and internal data sets: SVs observed with allele frequency of ≥1% in internal and public data sets are generally benign and can potentially be eliminated from further consideration. For all other SVs, annotation with allele frequency can be helpful in the interpretation of their clinical significance. Ideally, allele frequency should be determined using both internal laboratory data sets as well as external, publicly available data sets (including the Genome Aggregation Database-SV derived from short-read GS,[52] 1000 Genomes Project derived from short-read and long-read GS,[81] Icelandic data set derived from long-read GS,[82] Database of Genomic Variants derived from CMA data,[83] and others). The advantage of using internal data sets is that the samples are typically uniformly processed and analyzed, controlling for technical variables; in addition, some laboratories have access to unique populations that are not publicly available. However, many laboratories have too few samples to accurately determine allele frequency of individual SVs. Determination of allele frequency necessitates defining the required overlap between CNVs detected in different samples (reciprocal overlap threshold) to consider them the same, which can be challenging. Thresholds ranging from 50% to 90% have been used in different studies, with lower thresholds having to be used when comparing across different technologies. To maximize accuracy in CNV comparison, clinical laboratories are encouraged to use higher reciprocal overlap thresholds

to be familiar with the thresholds used in the data sets they apply for annotation and to annotate their samples using a data set that identified SVs by a similar platform and detection method (eg, a data set derived from GS data should be used to annotate SVs detected by GS).

4. Association with the patient's phenotype: SVs that involve genes associated with the patient's reported clinical phenotypes can be prioritized for evaluation. Phenotype driven analysis is an established approach in evaluation of SNVs and indels detected using ES and GS and can also be applied to SVs.

Proficiency testing

Proficiency testing for detection of SVs can be performed either in-house using de-identified and previously analyzed samples or by exchanging samples with external laboratories. A platform-neutral proficiency testing challenge for genome-wide detection of germline CNVs is also expected to become available in the near future through the College of American Pathologists. The gold-standard technology for the confirmation of detected SVs is dependent on the target SV type and size. Larger CNVs are commonly analyzed using CMA, whereas intragenic CNVs may be best captured using high density arrays, MLPA, or breakpoint amplification followed by Sanger sequencing. Caution should be exercised if sending isolated DNA to an outside facility for capture-based CNV detection through NGS, because different DNA extraction methodologies may affect performance for CNV detection. Sending original specimens, such as cell pellets if available, may be optimal. For all test samples, the capabilities and limitations of the chosen methodology for proficiency testing (eg, deletion vs duplication size resolution, bait coverage, or bait density in the region of interest) should be well understood before the application of the challenge. The College of American Pathologists does not currently have any specific guidance regarding proficiency testing for SV assessment from genomic sequencing, and laboratories should follow the general biannual proficiency testing requirement.

Points to Consider for Laboratories—Interpretation and Reporting

- As with all assays in the clinical laboratory, the interpretation of SVs should be performed by individuals with experience in evaluating SVs detected by the specific assay used. The acquisition of relevant experience, and the training and competency associated with it, will be laboratory dependent.
- Detected CNVs and balanced genomic rearrangements should be reported using the International System for Human Cytogenomic Nomenclature or Human Genome Variation Society nomenclature. When exact breakpoints are determined, sequence-based nomenclature for description of chromosome rearrangements is available in the 2020 edition of the International System for Human Cytogenomic Nomenclature.[84]
- The interpretation and classification of CNVs should follow established guidelines irrespective of the assay used to detect them. Riggs et al[85] described a quantitative, evidence-based scoring system for large CNVs designed to assist laboratories with the evaluation of germline CNVs and to enable consistency of interpretation and reporting across laboratories. In addition, Abou Tayoun et al[86] provided recommendations for interpreting different types of intragenic deletions and duplications, building on the 2015 ACMG/Association for Molecular Pathology sequence variant interpretation guideline[87] to also promote consistency of interpretation across laboratories.
- As with all clinical assays, the laboratory should define their reporting criteria and clearly state these in the clinical report. The limitations of the assay, the sensitivity, specificity, and resolution of the assay should also be stated in the report.
- For laboratories with separate protocols for CNV vs sequence variant interpretation, protocols should be standardized so that classification of intragenic exon-level deletions and duplications is consistent with classification of sequence variants.
- The detection of certain rearrangements, such as trisomy 21, requires cytogenetic follow up to determine whether the abnormality is due to a chromosomal rearrangement not detectable from SRS, such as a Robertsonian translocation. This information is essential because it may have reproductive consequences. The requirement for follow up for certain findings will be the same irrespective of the molecular methodology used to detect the rearrangement. Similarly, certain unbalanced rearrangements should be investigated to determine their chromosomal mechanism (eg, to establish if a duplication is in tandem or due to an insertion or whether copy number changes are associated with a derivative or a recombinant chromosome from a balanced rearrangement in a parent). In addition, parental testing using an appropriate methodology should be recommended to determine the inheritance and recurrence risk of these rearrangements.
- If the specific boundaries of a rearrangement cannot be determined by a targeted assay, then a genome CNV test could be recommended when appropriate.
  - Example report comment: *The breakpoints of this deletion/duplication are approximate. If more accurate breakpoint resolution is clinically necessary, copy number detection using GS/high resolution microarray/alternate method is recommended.*
- Recommendations for further genetic testing and opportunities for SV reanalysis should be included when appropriate.

- Laboratories should have a policy regarding pretest counseling, obtaining consent, and reporting of established pathogenic and likely pathogenic SVs that involve recommended ACMG-secondary finding genes. It is advised to counsel patients that it may not always be possible to avoid detection or reporting of these SVs in the context of a multigenic CNV.
- Reports should include descriptions of test performance, listing limitations of detection capabilities for each variant type tested, including SVs. If certain genomic regions are not evaluated for SVs, then that should also be mentioned in the report. For virtual panels, it needs to be clear whether CNVs are evaluated and with what performance/limitations.
  - Example disclaimer for reportable CNVs detected by capture-based methods: *Copy number analysis and variant interpretation is restricted to the genomic location of the gene(s) in this panel. For CNVs that include either the first or last exon of a gene, actual breakpoints may extend beyond the gene.*

Points to Consider for Providers

In a recent evidence-based guideline, ACMG recommended ES and GS as a first-tier or second-tier test for patients with one or more congenital anomalies before the age 1 year or for patients with intellectual disability/developmental delay with onset before the age 18 years.[88] This takes into consideration that the most appropriate diagnostic testing strategy may differ between patients and should be chosen using clinical judgment and shared decision making. In general, providers should consider the following when selecting the optimal genetic test: patient presentation, financial effect of testing on the patient, turnaround time, likely diagnostic yield, and assay limitations,[88] as described in the following.

Patient presentation

- Test selection should be based on the clinical phenotype, medical and family history, results of ancillary testing, and scope of the differential diagnosis.[89] If the clinician suspects a particular disorder or if a patient presents with a phenotype known to be commonly caused by a specific set of genes, a targeted test may be more appropriate than a genome-wide assay. Nonspecific or overlapping presentations may call for broader testing strategies (CMA, ES/GS).
- Mutational spectrum and molecular mechanism for the phenotype need to be considered; particular attention should be placed on CNV burden or known structural abnormalities associated with the suspected gene(s)/disorder(s). In addition, if haploinsufficiency and/or triplosensitivity are known causes of the disorder(s) suspected, providers should ensure that CNVs can be adequately evaluated by the methodology selected. Clinicians are encouraged to consult with the laboratory regarding appropriate coverage when evaluating for a specific genetic disorder or abnormality. Note that conventional cytogenetics may be superior to other testing methodologies in specific circumstances; for instance, determining whether a trisomy was caused by a Robertsonian translocation can presently only be achieved by karyotype analysis.
- Regardless of whether or not a particular disorder is suspected, providers should always provide laboratories with as much detailed phenotypic and family history information as possible; this information is critical in the variant filtering, annotation, and classification process.[90] Failure to provide complete information could result in the true causative variant being removed by filtering.[91] Providers should make every effort to provide this information at the time of test ordering or when additional information is requested from the testing laboratory. Standardized phenotype forms provided by the laboratory may assist the clinician in providing a thorough appraisal of relevant features.
- Accurate inheritance/phasing information is also often necessary to correctly classify variants. Parental samples should be pursued whenever feasible at the direction of the clinical laboratory; note that some laboratories/assays require parental samples to be sent in conjunction with the proband samples and others will make requests on a case-by-case basis at the time of resulting. Diagnostic yield has been shown to be higher when analyzing a trio as compared to proband-only ES.[92]

Assay cost

- Assays may be covered at different levels (or not at all) by different third-party payers.[93] In addition, there may be varying criteria for eligibility (eg, only specific diagnosis codes are covered, coverage is restricted to certain age groups, etc); this may also be the case within publicly funded health care systems. The reimbursement of GS is evolving, with inconsistent coverage by payors. As the cost of GS decreases, it may become comparable in price (or even more affordable) than running multiple assays to assess different variant types. However, if insurance covers all or part of any of those other assays and not GS, it may still be more financially prudent to order multiple tests sequentially instead of a single GS test. Providers should be aware of the potential costs associated with their chosen testing approach, and, when possible, select an approach that maximizes diagnostic yield while attempting to minimize costs to the patient and/or health care system. The genomics community should continue to collect the data necessary to

demonstrate clinical utility and cost-effectiveness to encourage more robust payer coverage in the future.[94-96]

Turnaround time

- The amount of time needed to obtain final results from an assay should be considered when determining an appropriate testing strategy in time-sensitive contexts, such as those of prenatal testing or evaluation of a critically ill newborn. In these scenarios, waiting for the results of multiple, sequential assays may not be feasible. Although it is technically possible to achieve rapid turnaround for GS testing,[97] providers should consult with their laboratory regarding their specific capabilities if time is a concern. Cost (as described earlier) and available sample type may also influence provider ordering decisions in these time-sensitive situations.

Assay limitations

- Genomic sequencing adoption into clinical practice and accompanying technological evolution has occurred very rapidly. In the setting of an assay whose analysis is complex and computationally-dependent, this has contributed to variability and inconsistency between laboratories. Providers should be familiar with assay limitations and aware that the limitations may differ between laboratories because of different platforms and/or bioinformatics pipelines.[98,99] The technical limitations of the specific test being ordered should be discussed with patients and families as part of pretest counseling (and/or consenting).[100-102] The provider should try to align patient expectations with possible results of the testing. This should be done in the context of contracting with the family in a balanced, client-centered approach that balances the values of autonomy, beneficence, and nonmaleficence.
- Inappropriate test selection can lead to misdiagnosis, patient frustration, increased health care costs, and delays in care.[103,104] When in doubt, nongenetic providers should seek guidance from genetics professionals regarding the best testing approach. Where available, institutions should take advantage of genetic test use management services; preanalytical review by these services has been recognized to result in order-error correction and cost savings.[105,106] There are many considerations to take into account when selecting the suitable genetic testing and laboratory, including gene-panel size and content, method of variant detection, experience with variant interpretation, and policies regarding variant reclassification and recontact.[107] In addition to these considerations, providers should become familiar with specific laboratory practices and limitations when arranging institutional preferred test status to certain send-out laboratories/tests. Many laboratories provide information online regarding cost, turnaround time, and approximate diagnostic yield; however, this information may not be able to clarify a provider's specific question or address a specific clinical scenario. When clarification is needed (pretest or post-test), direct consultation with a laboratory director or representative is recommended.[58]

Special considerations may affect testing decisions in the prenatal setting:

- For a fetus with imaging abnormalities and/or abnormal noninvasive prenatal screening, standard CMA and karyotyping should be considered. If negative, fetal ES or GS may be considered. At present, there are no data supporting the clinical use of ES/GS for other reproductive indications, such as the identification of sonographic markers suggestive of aneuploidy or a history of recurrent unexplained pregnancy loss.[108]
- Clinicians should be aware of reporting limitations regarding ES/GS methodologies for fetal samples. Laboratories may differ in genes interrogated (known vs candidate genes) and type of variants reported (SNVs vs CNVs) or may limit reporting to only known pathogenic and likely pathogenic variants. Postnatal reanalysis should be considered if the initial prenatal test was nondiagnostic.

Conclusion

Comprehensive testing for germline CNVs from sequencing is already being offered by clinical laboratories, and its use is likely to grow over time. With increased use of GS, LRS, and novel SV assays such as OM, high resolution testing for BCRs is also likely to move into routine clinical practice. As the number of variant types being assessed from sequencing data increases, laboratories will need to continue to perform appropriate validations and implement appropriate quality control practices to ensure the accuracy of all reported results. Given the complexity of genomic testing, laboratories should expect to continue to help guide the selection of appropriate methodologies for specific clinical indications.

Acknowledgments

E.K. was a director in the Medical Genetics Laboratory at Sanford Imagenetics during the preparation of this document.

Conflict of Interest

A.B. is a salaried employee and holds stock options of Invitae Corporation. B.A.S. is a salaried employee of Medical Science & Computing, LLC. All other authors declare no conflicts of interest.

Additional Information

The online version of this article (https://doi.org/10.1016/j.gim.2022.09.017) contains supplementary material, which is available to authorized users.

Authors

Gordana Raca[1], Caroline Astbury[2], Andrea Behlmann[3], Mauricio J. De Castro[4,5], Scott E. Hickey[6,7], Ender Karaca[8], Chelsea Lowther[9], Erin Rooney Riggs[10], Bryce A. Seifert[11], Erik C. Thorland[12], Joshua L. Deignan[13]; on behalf of the ACMG Laboratory Quality Assurance Committee[14,*]

Affiliations

[1]Department of Pathology and Laboratory Medicine, Children's Hospital Los Angeles, University of Southern California, Los Angeles, CA; [2]Robert J. Tomsich Pathology & Laboratory Medicine Institute, Cleveland Clinic, Cleveland, OH; [3]Invitae, San Francisco, CA; [4]Molecular Genetics Laboratory, Air Force Medical Genetics Center, Keesler Air Force Base, Biloxi, MS; [5]The University of Mississippi Medical Center, Jackson, MS; [6]Division of Genetic and Genomic Medicine, Nationwide Children's Hospital, Columbus, OH; [7]Department of Pediatrics, College of Medicine, The Ohio State University, Columbus, OH; [8]Department of Pathology, Baylor University Medical Center, Dallas, TX; [9]Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA; [10]Autism & Developmental Medicine Institute, Geisinger, Danville, PA; [11]Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD; [12]Genomics Laboratory, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN; [13]Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, UCLA, Los Angeles, CA; [14]American College of Medical Genetics and Genomics, Bethesda, MD

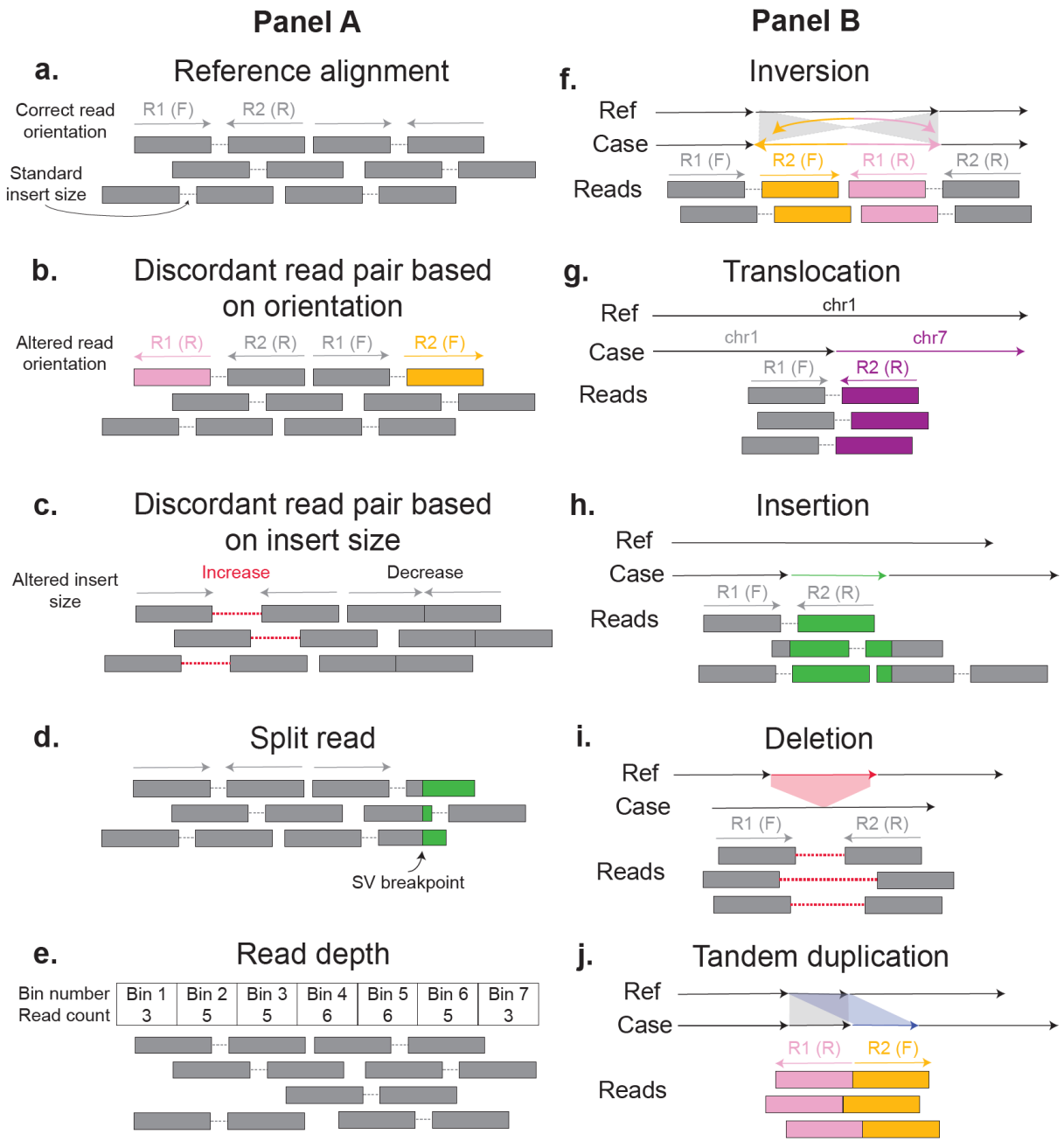
References

1. Retterer K, Scuffins J, Schmidt D, et al. Assessing copy number from exome sequencing and exome array CGH based on CNV spectrum in a large clinical cohort. *Genet Med*. 2015;17(8):623-629. http://doi.org/10.1038/gim.2014.160
2. Pfundt R, Del Rosario M, Vissers LELM, et al. Detection of clinically relevant copy-number variants by exome sequencing in a large cohort of genetic disorders. *Genet Med*. 2017;19(6):667-675. http://doi.org/10.1038/gim.2016.163
3. Chaisson MJP, Sanders AD, Zhao X, et al. Multi-platform discovery of haplotype-resolved structural variation in human genomes. *Nat Commun*. 2019;10(1):1784. http://doi.org/10.1038/s41467-018-08148-z
4. Chen W, Ullmann R, Langnick C, et al. Breakpoint analysis of balanced chromosome rearrangements by next-generation paired-end sequencing. *Eur J Hum Genet*. 2010;18(5):539-543. http://doi.org/10.1038/ejhg.2009.211
5. Talkowski ME, Ernst C, Heilbut A, et al. Next-generation sequencing strategies enable routine detection of balanced chromosome rearrangements for clinical diagnostics and genetic research. *Am J Hum Genet*. 2011;88(4):469-481. http://doi.org/10.1016/j.ajhg.2011.03.013
6. Del Gaudio D, Shinawi M, Astbury C, et al. Diagnostic testing for uniparental disomy: a points to consider statement from the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2020;22(7):1133-1141. http://doi.org/10.1038/s41436-020-0782-9
7. Gonzales PR, Andersen EF, Brown TR, et al. Interpretation and reporting of large regions of homozygosity and suspected consanguinity/uniparental disomy, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2022;24(2):255-261. http://doi.org/10.1016/j.gim.2021.10.004
8. Wang Y, Cheng Q, Meng L, et al. Clinical application of SNP array analysis in first-trimester pregnancy loss: a prospective study. *Clin Genet*. 2017;91(6):849-858. http://doi.org/10.1111/cge.12926
9. Carr IM, Bhaskar S, O'Sullivan J, et al. Autozygosity mapping with exome sequence data. *Hum Mutat*. 2013;34(1):50-56. http://doi.org/10.1002/humu.22220
10. Quinodoz M, Peter VG, Bedoni N, et al. AutoMap is a high performance homozygosity mapping tool using next-generation sequencing data. *Nat Commun*. 2021;12(1):518. http://doi.org/10.1038/s41467-020-20584-4
11. Ceballos FC, Hazelhurst S, Ramsay M. Assessing runs of homozygosity: a comparison of SNP array and whole genome sequence low coverage data. *BMC Genomics*. 2018;19(1):106. http://doi.org/10.1186/s12864-018-4489-0
12. Chaubey A, Shenoy S, Mathur A, et al. Low-pass genome sequencing: validation and diagnostic utility from 409 clinical cases of low-pass genome sequencing for the detection of copy number variants to replace constitutional microarray. *J Mol Diagn*. 2020;22(6):823-840. http://doi.org/10.1016/j.jmoldx.2020.03.008
13. Bean L, Bayrak-Toydemir P. American College of Medical Genetics and Genomics standards and guidelines for clinical genetics laboratories, 2014 edition: technical standards and guidelines for Huntington disease. *Genet Med*. 2014;16(12):e2. http://doi.org/10.1038/gim.2014.146
14. Spector E, Behlmann A, Kronquist K, et al. Laboratory testing for fragile X, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021;23(5):799-812. http://doi.org/10.1038/s41436-021-01115-y
15. Tang H, Kirkness EF, Lippert C, et al. Profiling of short-tandem-repeat disease alleles in 12,632 human whole genomes. *Am J Hum Genet*. 2017;101(5):700-715. http://doi.org/10.1016/j.ajhg.2017.09.013
16. Dashnow H, Lek M, Phipson B, et al. STRetch: detecting and discovering pathogenic short tandem repeat expansions. *Genome Biol*. 2018;19(1):121. http://doi.org/10.1186/s13059-018-1505-2
17. Tankard RM, Bennett MF, Degorski P, Delatycki MB, Lockhart PJ, Bahlo M. Detecting expansions of tandem repeats in cohorts sequenced with short-read sequencing data. *Am J Hum Genet*. 2018;103(6):858-873. http://doi.org/10.1016/j.ajhg.2018.10.015
18. van der Sanden BPGH, Corominas J, de Groot M, et al. Systematic analysis of short tandem repeats in 38,095 exomes provides an additional diagnostic yield. *Genet Med*. 2021;23(8):1569-1573. http://doi.org/10.1038/s41436-021-01174-1
19. Dai Y, Li P, Wang Z, et al. Single-molecule optical mapping enables quantitative measurement of D4Z4 repeats in facioscapulohumeral muscular dystrophy (FSHD). *J Med Genet*. 2020;57(2):109-120. http://doi.org/10.1136/jmedgenet-2019-106078
20. Stence AA, Thomason JG, Pruessner JA, et al. Validation of optical genome mapping for the molecular diagnosis of facioscapulohumeral

muscular dystrophy. *J Mol Diagn*. 2021;23(11):1506-1514. http://doi.org/10.1016/j.jmoldx.2021.07.021
21. Chintalaphani SR, Pineda SS, Deveson IW, Kumar KR. An update on the neurological short tandem repeat expansion disorders and the emergence of long-read sequencing diagnostics. *Acta Neuropathol Commun*. 2021;9(1):98. http://doi.org/10.1186/s40478-021-01201-x
22. Bean LJH, Funke B, Carlston CM, et al. Diagnostic gene sequencing panels: from design to report-a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2020;22(3):453-461. http://doi.org/10.1038/s41436-019-0666-z
23. Tan R, Wang Y, Kleinstein SE, et al. An evaluation of copy number variation detection tools from whole-exome sequencing data. *Hum Mutat*. 2014;35(7):899-907. http://doi.org/10.1002/humu.22537
24. Chau MHK, Wang H, Lai Y, et al. Low-pass genome sequencing: a validated method in clinical cytogenetics. *Hum Genet*. 2020;139(11):1403-1415. http://doi.org/10.1007/s00439-020-02185-9
25. Wang H, Dong Z, Zhang R, et al. Low-pass genome sequencing versus chromosomal microarray analysis: implementation in prenatal diagnosis. *Genet Med*. 2020;22(3):500-510. http://doi.org/10.1038/s41436-019-0634-7
26. Dong Z, Chau MHK, Zhang Y, et al. Low-pass genome sequencing-based detection of absence of heterozygosity: validation in clinical cytogenetics. *Genet Med*. 2021;23(7):1225-1233. http://doi.org/10.1038/s41436-021-01128-7
27. Hehir-Kwa JY, Pfundt R, Veltman JA. Exome sequencing and whole genome sequencing for the detection of copy number variation. *Expert Rev Mol Diagn*. 2015;15(8):1023-1032. http://doi.org/10.1586/14737159.2015.1053467
28. Dolzhenko E, van Vugt JJFA, Shaw RJ, et al. Detection of long repeat expansions from PCR-free whole-genome sequence data. *Genome Res*. 2017;27(11):1895-1903. http://doi.org/10.1101/gr.225672.117
29. Nilsson D, Pettersson M, Gustavsson P, et al. Whole-genome sequencing of cytogenetically balanced chromosome translocations identifies potentially pathological gene disruptions and highlights the importance of microhomology in the mechanism of formation. *Hum Mutat*. 2017;38(2):180-192. http://doi.org/10.1002/humu.23146
30. Shen W, Szankasi P, Durtschi J, Kelley TW, Xu X. Genome-wide copy number variation detection using NGS: data analysis and interpretation. *Methods Mol Biol*. 2019;1908:113-124. http://doi.org/10.1007/978-1-4939-9004-7_8
31. Drucker TM, Johnson SH, Murphy SJ, Cradic KW, Therneau TM, Vasmatzis G. BIMA V3: an aligner customized for mate pair library sequencing. *Bioinformatics*. 2014;30(11):1627-1629. http://doi.org/10.1093/bioinformatics/btu078
32. Johnson SH, Smadbeck JB, Smoley SA, et al. SVAtools for junction detection of genome-wide chromosomal rearrangements by mate-pair sequencing (MPseq). *Cancer Genet*. 2018;221:1-18. http://doi.org/10.1016/j.cancergen.2017.11.009
33. Smadbeck JB, Johnson SH, Smoley SA, et al. Copy number variant analysis using genome-wide mate-pair sequencing. *Genes Chromosomes Cancer*. 2018;57(9):459-470. http://doi.org/10.1002/gcc.5
34. Collins RL, Brand H, Redin CE, et al. Defining the diverse spectrum of inversions, complex structural variation, and chromothripsis in the morbid human genome. *Genome Biol*. 2017;18(1):36. http://doi.org/10.1186/s13059-017-1158-6
35. Chiang C, Jacobsen JC, Ernst C, et al. Complex reorganization and predominant non-homologous repair following chromosomal breakage in karyotypically balanced germline rearrangements and transgenic integration. *Nat Genet*. 2012;44(4):390-397. http://doi.org/10.1038/ng.2202
36. Redin C, Brand H, Collins RL, et al. The genomic landscape of balanced cytogenetic abnormalities associated with human congenital anomalies. *Nat Genet*. 2017;49(1):36-45. http://doi.org/10.1038/ng.3720
37. Vergult S, Van Binsbergen E, Sante T, et al. Mate pair sequencing for the detection of chromosomal aberrations in patients with intellectual disability and congenital malformations. *Eur J Hum Genet*. 2014;22(5):652-659. http://doi.org/10.1038/ejhg.2013.220
38. Pollard MO, Gurdasani D, Mentzer AJ, Porter T, Sandhu MS. Long reads: their purpose and place. *Hum Mol Genet*. 2018;27(R2):R234-R241. http://doi.org/10.1093/hmg/ddy177
39. Loomis EW, Eid JS, Peluso P, et al. Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene. *Genome Res*. 2013;23(1):121-128. http://doi.org/10.1101/gr.141705.112
40. Chaisson MJP, Huddleston J, Dennis MY, et al. Resolving the complexity of the human genome using single-molecule sequencing. *Nature*. 2015;517(7536):608-611. http://doi.org/10.1038/nature13907
41. Merker JD, Wenger AM, Sneddon T, et al. Long-read genome sequencing identifies causal structural variation in a Mendelian disease. *Genet Med*. 2018;20(1):159-163. http://doi.org/10.1038/gim.2017.86
42. Liu Q, Zhang P, Wang D, Gu W, Wang K. Interrogating the "unsequenceable" genomic trinucleotide repeat disorders by long-read sequencing. *Genome Med*. 2017;9(1):65. http://doi.org/10.1186/s13073-017-0456-7
43. Zhao X, Collins RL, Lee WP, et al. Expectations and blind spots for structural variation detection from long-read assemblies and short-read genome sequencing technologies. *Am J Hum Genet*. 2021;108(5):919-928. http://doi.org/10.1016/j.ajhg.2021.03.014
44. Jeffet J, Margalit S, Michaeli Y, Ebenstein Y. Single-molecule optical genome mapping in nanochannels: multidisciplinarity at the nanoscale. *Essays Biochem*. 2021;65(1):51-66. http://doi.org/10.1042/EBC20200021
45. Chan S, Lam E, Saghbini M, et al. Structural variation detection and analysis using Bionano optical mapping. *Methods Mol Biol*. 2018;1833:193-203. http://doi.org/10.1007/978-1-4939-8666-8_16
46. Yuan Y, Chung CYL, Chan TF. Advances in optical mapping for genomic research. *Comput Struct Biotechnol J*. 2020;18:2051-2062. http://doi.org/10.1016/j.csbj.2020.07.018
47. Barseghyan H, Tang W, Wang RT, et al. Next-generation mapping: a novel approach for detection of pathogenic structural variants with a potential utility in clinical diagnosis. *Genome Med*. 2017;9(1):90. http://doi.org/10.1186/s13073-017-0479-0
48. Chen M, Zhang M, Qian Y, et al. Identification of a likely pathogenic structural variation in the LAMA1 gene by Bionano optical mapping. *NPJ Genom Med*. 2020;5:31. http://doi.org/10.1038/s41525-020-0138-z
49. Sahajpal NS, Barseghyan H, Kolhe R, Hastie A, Chaubey A. Optical genome mapping as a next-generation cytogenomic tool for detection of structural and copy number variations for prenatal genomic analyses. *Genes (Basel)*. 2021;12(3):398. http://doi.org/10.3390/genes12030398
50. Chaisson MJP, Wilson RK, Eichler EE. Genetic variation and the de novo assembly of human genomes. *Nat Rev Genet*. 2015;16(11):627-640. http://doi.org/10.1038/nrg3933
51. Werling DM, Brand H, An JY, et al. An analytical framework for whole-genome sequence association studies and its implications for autism spectrum disorder. *Nat Genet*. 2018;50(5):727-736. http://doi.org/10.1038/s41588-018-0107-y
52. Collins RL, Brand H, Karczewski KJ, et al. A structural variation reference for medical and population genetics. *Nature*. 2020;581(7809):444-451. https://doi.org/10.1038/s41586-020-2287-8
53. Trost B, Walker S, Wang Z, et al. A comprehensive workflow for read depth-based identification of copy-number variation from whole-genome sequence data. *Am J Hum Genet*. 2018;102(1):142-155. http://doi.org/10.1016/j.ajhg.2017.12.007
54. Abel HJ, Larson DE, Regier AA, et al. Mapping and characterization of structural variation in 17,795 human genomes. *Nature*. 2020;583(7814):83-89. http://doi.org/10.1038/s41586-020-2371-0
55. Ho SS, Urban AE, Mills RE. Structural variation in the sequencing era. *Nat Rev Genet*. 2020;21(3):171-189. http://doi.org/10.1038/s41576-019-0180-9
56. Vetro A, Goidin D, Lesende I, et al. Diagnostic application of a capture based NGS test for the concurrent detection of variants in sequence and copy number as well as LOH. *Clin Genet*. 2018;93(3):545-556. http://doi.org/10.1111/cge.13060

57. Zook JM, Hansen NF, Olson ND, et al. A robust benchmark for detection of germline large deletions and insertions. *Nat Biotechnol*. 2020;38(11):1347-1355. http://doi.org/10.1038/s41587-020-0538-8
58. Rehder C, Bean LJH, Bick D, et al. Next-generation sequencing for constitutional variants in the clinical laboratory, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021;23(8):1399-1415. http://doi.org/10.1038/s41436-021-01139-4
59. Shao L, Akkari Y, Cooley LD, et al. Chromosomal microarray analysis, including constitutional and neoplastic disease applications, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021;23(10):1818-1829. http://doi.org/10.1038/s41436-021-01214-w
60. Jennings LJ, Arcila ME, Corless C, et al. Guidelines for validation of next-generation sequencing-based oncology panels: a joint consensus recommendation of the Association for Molecular Pathology and College of American Pathologists. *J Mol Diagn*. 2017;19(3):341-365. http://doi.org/10.1016/j.jmoldx.2017.01.011
61. Jennings L, Van Deerlin VM, Gulley ML. Recommended principles and practices for validating clinical molecular pathology tests. *Arch Pathol Lab Med*. 2009;133(5):743-755. http://doi.org/10.5858/133.5.743
62. Marshall CR, Chowdhury S, Taft RJ, et al. Best practices for the analytical validation of clinical whole-genome sequencing intended for the diagnosis of germline disease. *NPJ Genom Med*. 2020;5:47. http://doi.org/10.1038/s41525-020-00154-9
63. Zhao M, Wang Q, Wang Q, Jia P, Zhao Z. Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. *BMC Bioinformatics*. 2013;14(Suppl 11):S1. http://doi.org/10.1186/1471-2105-14-S11-S1
64. Roca I, González-Castro L, Fernández H, Couce ML, Fernández-Marmiesse A. Free-access copy-number variant detection tools for targeted next-generation sequencing data. *Mutat Res Rev Mutat Res*. 2019;779:114-125. http://doi.org/10.1016/j.mrrev.2019.02.005
65. Singh AK, Olsen MF, Lavik LAS, Vold T, Drabløs F, Sjursen W. Detecting copy number variation in next generation sequencing data from diagnostic gene panels. *BMC Med Genomics*. 2021;14(1):214. http://doi.org/10.1186/s12920-021-01059-x
66. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput Biol*. 2016;12(4):e1004873. http://doi.org/10.1371/journal.pcbi.1004873
67. Magi A, Tattini L, Pippucci T, Torricelli F, Benelli M. Read count approach for DNA copy number variants detection. *Bioinformatics*. 2012;28(4):470-478. http://doi.org/10.1093/bioinformatics/btr707
68. Bellos E, Coin LJM. cnvOffSeq: detecting intergenic copy number variation using off-target exome sequencing data. *Bioinformatics*. 2014;30(17):i639-i645. http://doi.org/10.1093/bioinformatics/btu475
69. Teo SM, Pawitan Y, Ku CS, Chia KS, Salim A. Statistical challenges associated with detecting copy number variations with next-generation sequencing. *Bioinformatics*. 2012;28(21):2711-2718. http://doi.org/10.1093/bioinformatics/bts535
70. Kerkhof J, Schenkel LC, Reilly J, et al. Clinical validation of copy number variant detection from targeted next-generation sequencing panels. *J Mol Diagn*. 2017;19(6):905-920. http://doi.org/10.1016/j.jmoldx.2017.07.004
71. Jones MA, Bhide S, Chin E, et al. Targeted polymerase chain reaction-based enrichment and next generation sequencing for diagnostic testing of congenital disorders of glycosylation. *Genet Med*. 2011;13(11):921-932. http://doi.org/10.1097/GIM.0b013e318226fbf2
72. Gowrisankar S, Lerner-Ellis JP, Cox S, et al. Evaluation of second-generation sequencing of 19 dilated cardiomyopathy genes for clinical applications. *J Mol Diagn*. 2010;12(6):818-827. http://doi.org/10.2353/jmoldx.2010.100014
73. Lohmann K, Klein C. Next generation sequencing and the future of genetic diagnosis. *Neurotherapeutics*. 2014;11(4):699-707. http://doi.org/10.1007/s13311-014-0288-8
74. Gargis AS, Kalman L, Berry MW, et al. Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nat Biotechnol*. 2012;30(11):1033-1036. http://doi.org/10.1038/nbt.2403
75. Yao R, Yu T, Qing Y, Wang J, Shen Y. Evaluation of copy number variant detection from panel-based next-generation sequencing data. *Mol Genet Genomic Med*. 2019;7(1):e00513. http://doi.org/10.1002/mgg3.513
76. Yao R, Zhang C, Yu T, et al. Evaluation of three read-depth based CNV detection tools using whole-exome sequencing data. *Mol Cytogenet*. 2017;10:30. http://doi.org/10.1186/s13039-017-0333-5
77. Walsh T, Lee MK, Casadei S, et al. Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing. *Proc Natl Acad Sci U S A*. 2010;107(28):12629-12633. http://doi.org/10.1073/pnas.1007983107
78. Schenkel LC, Kerkhof J, Stuart A, et al. Clinical next-generation sequencing pipeline outperforms a combined approach using Sanger sequencing and multiplex ligation-dependent probe amplification in targeted gene panel analysis. *J Mol Diagn*. 2016;18(5):657-667. http://doi.org/10.1016/j.jmoldx.2016.04.002
79. Zare F, Hosny A, Nabavi S. Noise cancellation using total variation for copy number variation detection. *BMC Bioinformatics*. 2018;19(Suppl 11):361. http://doi.org/10.1186/s12859-018-2332-x
80. Dolzhenko E, Weisburd B, Ibañez K, et al. REViewer: haplotype-resolved visualization of read alignments in and around tandem repeats. *Genome Med*. 2022;14(1):84. http://doi.org/10.1186/s13073-022-01085-z
81. Sudmant PH, Rausch T, Gardner EJ, et al. An integrated map of structural variation in 2,504 human genomes. *Nature*. 2015;526(7571):75-81. http://doi.org/10.1038/nature15394
82. Beyter D, Ingimundardottir H, Oddsson A, et al. Long-read sequencing of 3,622 Icelanders provides insight into the role of structural variants in human diseases and other traits. *Nat Genet*. 2021;53(6):779-786. http://doi.org/10.1038/s41588-021-00865-4
83. MacDonald JR, Ziman R, Yuen RKC, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res*. 2014;42(Database issue):D986-D992. http://doi.org/10.1093/nar/gkt958
84. McGowan-Jordan J, Hastings RJ, Moore S, eds. *ISCN 2020: An International System for Human Cytogenomic Nomenclature (2020)*. Karger Publishers; 2020.
85. Riggs ER, Andersen EF, Cherry AM, et al. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet Med*. 2020;22(2):245-257. https://doi.org/10.1038/s41436-019-0686-8
86. Abou Tayoun AN, Pesaran T, DiStefano MT, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Hum Mutat*. 2018;39(11):1517-1524. http://doi.org/10.1002/humu.23626
87. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424. http://doi.org/10.1038/gim.2015.30
88. Manickam K, McClain MR, Demmer LA, et al. Exome and genome sequencing for pediatric patients with congenital anomalies or intellectual disability: an evidence-based clinical guideline of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021;23(11):2029-2037. http://doi.org/10.1038/s41436-021-01242-6
89. McCarthy JJ, Mendelsohn BA. Appendix 5. Selecting and ordering a lab test. In: *Precision Medicine: A Guide to Genomics in Clinical Practice*. McGraw-Hill Education; 2017.

90. Conway ME, Kalejta CD, Sternen DL, Singh IR. The importance of genetics experts in optimizing genetic test orders through prospective and retrospective reviews. *Am J Clin Pathol*. 2020;153(4):537-547. http://doi.org/10.1093/ajcp/aqz188
91. Bush LW, Beck AE, Biesecker LG, et al. Professional responsibilities regarding the provision, publication, and dissemination of patient phenotypes in the context of clinical genetic and genomic testing: points to consider-a statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2018;20(2):169-171. http://doi.org/10.1038/gim.2017.242
92. Brockman DG, Austin-Tse CA, Pelletier RC, et al. Randomized prospective evaluation of genome sequencing versus standard-of-care as a first molecular diagnostic test. *Genet Med*. 2021;23(9):1689-1696. http://doi.org/10.1038/s41436-021-01193-y
93. Retterer K, Juusola J, Cho MT, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med*. 2016;18(7):696-704. http://doi.org/10.1038/gim.2015.148
94. Dimmock D, Caylor S, Waldman B, et al. Project Baby Bear: rapid precision care incorporating rWGS in 5 California children's hospitals demonstrates improved clinical outcomes and reduced costs of care. *Am J Hum Genet*. 2021;108(7):1231-1238. http://doi.org/10.1016/j.ajhg.2021.05.008
95. Stark Z, Schofield D, Martyn M, et al. Does genomic sequencing early in the diagnostic trajectory make a difference? A follow-up study of clinical outcomes and cost-effectiveness. *Genet Med*. 2019;21(1):173-180. https://doi.org/10.1038/s41436-018-0006-8
96. Farnaes L, Hildreth A, Sweeney NM, et al. Rapid whole-genome sequencing decreases infant morbidity and cost of hospitalization. *NPJ Genom Med*. 2018;3:10. http://doi.org/10.1038/s41525-018-0049-4
97. Phillips KA, Douglas MP, Marshall DA. Expanding use of clinical genome sequencing and the need for more data on implementation. *JAMA*. 2020;324(20):2029-2030. http://doi.org/10.1001/jama.2020.19933
98. Clark MM, Hildreth A, Batalov S, et al. Diagnosis of genetic diseases in seriously ill children by rapid whole-genome sequencing and automated phenotyping and interpretation. *Sci Transl Med*. 2019;11(489):eaat6177. http://doi.org/10.1126/scitranslmed.aat6177
99. Roy S, Coldren C, Karunamurthy A, et al. Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: a joint recommendation of the Association for Molecular Pathology and the College of American Pathologists. *J Mol Diagn*. 2018;20(1):4-27. http://doi.org/10.1016/j.jmoldx.2017.11.003
100. Zhang K, Lin G, Han D, et al. An initial survey of the performances of exome variant analysis and clinical reporting among diagnostic laboratories in China. *Front Genet*. 2020;11:582637. http://doi.org/10.3389/fgene.2020.582637
101. Brett GR, Wilkins EJ, Creed ET, et al. Genetic counseling in the era of genomics: what's all the fuss about? *J Genet Couns*. 2018;27(5):1010-1021. http://doi.org/10.1007/s10897-018-0216-x
102. Sanderson SC, Hill M, Patch C, Searle B, Lewis C, Chitty LS. Delivering genome sequencing in clinical practice: an interview study with healthcare professionals involved in the 100 000 Genomes Project. *BMJ Open*. 2019;9(11):e029699. http://doi.org/10.1136/bmjopen-2019-029699
103. Jamal L, Schupmann W, Berkman BE. An ethical framework for genetic counseling in the genomic era. *J Genet Couns*. 2020;29(5):718-727. http://doi.org/10.1002/jgc4.1207
104. Farmer MB, Bonadies DC, Mahon SM, et al. Adverse events in genetic testing: the fourth case series. *Cancer J*. 2019;25(4):231-236. http://doi.org/10.1097/PPO.0000000000000391
105. Montanez K, Berninger T, Willis M, Harding A, Lutgendorf MA. Genetic testing costs and compliance with clinical best practices. *J Genet Couns*. 2020;29(6):1186-1191. http://doi.org/10.1002/jgc4.1285
106. Londre GK, Zaleski CA, Conta JH. Adding value to genetic testing through utilization management: commercial laboratory's experience. *Am J Med Genet A*. 2017;173(5):1433-1435. http://doi.org/10.1002/ajmg.a.38147
107. Riley JD, Procop GW, Kottke-Marchant K, Wyllie R, Lacbawan FL. Improving molecular genetic test utilization through order restriction, test review, and guidance. *J Mol Diagn*. 2015;17(3):225-229. http://doi.org/10.1016/j.jmoldx.2015.01.003
108. Monaghan KG, Leach NT, Pekarek D, Prasad P, Rose NC. The use of fetal exome sequencing in prenatal diagnosis: a points to consider document of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2020;22(4):675-680. http://doi.org/10.1038/s41436-019-0731-7



Supplementary Figure 1. Identification of structural variants from genome sequencing data.

Panel A. Evidence used by algorithms to identify structural variants (SVs) from paired-end genome sequencing (GS) short read data. (a) Reference genome alignment: Read 1 (R1) and Read 2 (R2) are in the correct forward-reverse (FR) pattern. Discordant read pair alignment: when a pair of reads from the same fragment have an (b) orientation or (c) insert size that differs from the typical reference alignment. (d) Split read: when a read overlaps the breakpoint of an SV. (e) Read depth: when the number of reads overlapping a fixed or user-defined bin varies across a region, as depicted in Figures 2-4. Assembly based methods (not shown) have also been developed but are computationally expensive.

Panel B. Signatures of balanced and unbalanced SVs based on discordant read pairs. Split reads for one or more breakpoint(s) and read depth evidence may also be available to support the presence of an SV. Genomic sequence impacted by an SV is depicted in green. (f) Inversion supported by discordant read pairs at both breakpoints based on altered orientation. (g) Translocation and (h) insertion supported by one read in read-pair mapping to alternative location in the genome. (i) Deletion supported by increased insert size when reads are mapped back to the reference genome. (j) Tandem duplication supported by decreased insert size when mapped back to the reference genome and altered orientation of reads flanking the duplication breakpoint. Complex SVs are comprised of two or more distinct SV signatures.

Ref, reference; R1, read one of read-pair; R2, read two of read-pair; R, reverse read orientation; F, forward read orientation; Inv, inversion; Tloc, translocation; Ins, insertion; Del, deletion; Dup, duplication.