Genetics in Medicine Open (2024) 2, 101816

ELSEVIER

Genetics in Medicine OPEN
An Official Journal of the ACMG

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

ARTICLE

Next-generation variant exon screening: Moving forward in routine genetic disease investigations

Conghui Wang[1], Panlai Shi[1], Hongbin Liang[2], David S. Cram[2], Donald A. Leigh[2,*], Xiangdong Kong[1,*]

[1]*Genetic and Prenatal Diagnosis Center, Department of Obstetrics and Gynecology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China;* [2]*Genetics and Precision Medicine Centre, The First People's Hospital of Kunming, Kunming, Yunnan, China*

ARTICLE INFO

Article history:
Received 14 February 2023
Received in revised form
17 January 2024
Accepted 19 January 2024
Available online 29 January 2024

Keywords:
Copy-number variants
Nucleotide variants
Skin diseases
Variant exon screening

ABSTRACT

Purpose: Patients with genetic diseases often seek testing to reach a firm diagnosis. Based on clinical phenotypes, exome sequencing for small-nucleotide variations or array-based methods for copy-number variations (CNVs) are commonly offered to identify the underlying causative genetic variants. In this study, we investigated whether data from a standard ES test could be used to additionally identify pathogenic CNVs and increase diagnostic yield.
Methods: Prospectively, 134 patients presenting with a skin condition suspected of being genetic in origin were offered the next-generation variant exon screening (ngVES) test. Sequencing data were analyzed for both single-nucleotide variants and CNVs using established algorithms.
Results: The positive detection rate for skin diseases using ngVES was 66% (88/134) with the most common diagnoses being neurofibromatosis type1 (n = 48) and tuberous sclerosis type2 (n = 12). The diagnostic increased yield from 58% to 66% was the result of additional detection of pathogenic CNVs. Each of the 9 CNVs were verified by independent genetic tests.
Conclusion: The advances in the ngVES bioinformatics pipeline are proofs of concept, which improved identification of genetic variants associated with skin disease. Simultaneous single-nucleotide variants/INDEL and CNV detection by this approach demonstrates ngVES potential as a first-tier screen for any suspected genetic disease.

Introduction

Worldwide, genetic disease syndromes are relatively common with a birth incidence of single gene and chromosomal disorders estimated at 5% and 1.5%, respectively.[1] These genetic syndromes are generally caused by the inheritance of familial pathogenic variants or by de novo variants. The majority of identified pathogenic variants comprise

The Article Publishing Charge (APC) for this article was paid by Conghui Wang.
Conghui Wang, Panlai Shi, and Hongbin Liang have contributed equally to this work.
*Correspondence and requests for materials should be addressed to Xiangdong Kong, No.1, Jianshe East Rd, Erqi District, Zhengzhou, Henan Province 450000, China. *Email address:* kongxd@263.net OR Donald Arthur Leigh, No. 1228, Beijing Rd, Panlong District, Kunming, Yunnan Province 650000, China. *Email address:* don.leigh09@gmail.com

doi: https://doi.org/10.1016/j.gimo.2024.101816

single-nucleotide variants (SNVs), small insertions and deletions (INDELs), and copy-number variations (CNVs). Other possibilities include translocation disruptions and uniparental disomy. In both pediatric and adult patients, a combination of clinical examination with a series of specific confirmatory molecular tests can often identify the underlying genetic basis for many, but not all, syndromes. However, reaching a final diagnosis can be important for individual patients and allow clinicians to implement any available effective treatments to alleviate disease symptoms.[2,3] For many patients with complex or slightly atypical phenotypes, the clinical diagnosis can sometimes be confounding and may require even further genome studies to assist in making a firm diagnosis.[4]

A number of molecular techniques are currently used for genetic diagnosis. If the phenotype is commonly associated with a known disease, then a simple variant screen or a direct gene(s) exon sequence (singular or clinical panel) is typically used. Extending this may involve gene sequencing if clinicians are confident of the gene/phenotype relationship. For patients for whom a simple pathogenic variation is not identifiable in any suspected gene or with phenotypes that are of less obvious gene association, the approach may also involve CNV assessments by chromosome microarray analysis (CMA) using either a high-density oligonucleotide microarray or a single-nucleotide polymorphism (SNP) array.[5] When targeted appropriately, both array platforms have the capacity to identify CNVs as small as 20 to 50 kb in size with high reliability and accuracy, although this may vary in the different chromosome regions especially near telomeres and centromeres. A newer option, next-generation sequencing (NGS) of random chromosome regions (CNV-seq) can also be used[6,7]; however, CNV resolution is only around 100 kb. Both the CMA platforms and CNV-seq cover the entire chromosome, and analysis will often involve discriminating between polymorphic CNVs, CNVs of unknown significance, and known pathogenic/likely pathogenic CNVs. In many cases, this assessment may involve identification of any genes within the CNV interval. If no likely pathogenic CNVs are found, further analysis by exome sequencing (ES) is often used to search for other possible genes or pathogenic SNVs and small INDELs[8] that may have been missed in the first approaches. In cases that no pathogenic variants are found, genome sequencing (GS) may be used as a last resort to identify rare non-exon pathogenic variants.[9] As testing progresses or the complexity of tests increases, the time required and overall costs similarly increase.

A recent study of patients with genetic diseases demonstrated that CMA for CNVs combined with ES for SNVs/INDELS resulted in a higher diagnostic yield of reportable pathogenic variants.[10-12] In a recent study of fetuses with structural anomalies, use of simultaneous ES and CNV-seq tests resulted in a higher diagnostic yield after including CNV detection.[13] These findings suggest that early, more complete investigation of patients may potentially benefit from a single test that could simultaneously screen for a greater variety of pathogenic variant genetic types. The sequential approach in disease analysis often involves first screening for common variants (either SNVs or CNVs), to sequencing selected gene regions or whole gene analysis, to gene panels or even exome analysis, and potentially right through to genome sequencing. This is often a cascade system that takes time, resources, and money. An early test that is more comprehensive could increase the likelihood of an early positive diagnosis, thus reducing time to diagnosis and assist in identifying any possible treatment options, as well as ultimately decreasing testing and personal costs.

In investigating common CNV linked diseases, options often revolve around specific gene centered associations[14] because these are the most likely affected elements. Recent advances in bioinformatics have resulted in the development of new tools and algorithms for CNV detection by assessing the relative probe/gene copies from the available SNP information in ES data. In comparative studies, the majority of these tools[15,16] showed encouraging results for detection of known pathogenic CNVs previously identifiable by CMA. However, with typical ES probe sets largely limited to gene exonic regions randomly spaced across the genome, among individual gene probe sets coverage and depth can be variable within a subset of genes. Several issues remain, including false-positive and false-negative CNV calls and for the larger pathogenic CNVs (>100 kb), correct interval length assessments can sometimes be problematic. Accurate measurement of any chromosome mosaicism present may also be important.

For complex diseases or phenotypes with variable or unknown common genetic causes, putative exome screening is often utilized. Different manufacturers' exome probe sets have different effective gene coverage with many exons or even whole genes subject to poor overall analysis and therefore rarely offer true exome coverage. Rather than use the more common term exome sequencing, it is more appropriate to refer to it as ES and consider how the data are best utilized in any investigation. To improve variant detection utilizing standard ES data, we developed an advanced bioinformatics pipeline, termed next-generation variant exon screening (ngVES). This approach using the standard ES data, in addition to routine SNV/INDEL analysis, also enables detection of other sequence variation, such as loss of heterozygosity, as well as segmental copy-number changes. In a proof-of-concept study, using a large cohort of patients with skin diseases suspected of being genetic in origin, we demonstrated that the ngVES pipeline is not only highly reliable and accurate for detection of pathogenic SNVs/INDELs but also for pathogenic CNVs.

Materials and Methods

Study samples

For the clinical ngVES study, we prospectively analyzed 135 genomic DNA samples extracted from peripheral blood

cells from patients presenting with clinical features of skin disease suspected to potentially be genetic in origin (Supplemental Table 1).

ngVES

ngVES uses the additional CNV analysis tool exome hidden Markov model (XHMM)[17] to perform secondary data mining, and it includes the detection of CNVs utilizing standard ES data. ES was performed by standard procedures using the IGT-T192V1 Plus panel (Integrated DNA Technologies) to capture selected targets. This panel covers a 41.97 Mb target region of the human genome, including the CDS regions of 20,716 nuclear coding genes and their immediate intronic sequences, as well as the whole mitochondrial genome. A total of 50 ng of genomic DNA purified from peripheral blood (Qiagen DNA Blood Midi/Mini kit, Qiagen GmbH) was fragmented in Tagmentation Buffer 1 to an average size of between 200 to 300 bp and then enriched with Bead-Linked Transposomes (Illumina). Fragments were then end-repaired, and 1 adenosine base added at the 3' end. The modified fragments were then ligated with barcoded adapters (Fast DNA Library Prep Set CW3045M, CWBIO Inc). After purification with XP beads and PCR amplification with sequencing primers (6-12 cycles), libraries were hybridized with the IDT exon baits. Following elution, targets were then re-amplified by PCR (12 cycles) and purified with XP beads. Target quantity was evaluated using Qubit 4.0 (Thermo Fisher Scientific Inc). Finally, sequence ready target libraries were analyzed on the Novaseq6000 platform (Illumina) using the 150 bp pair-end sequencing mode. Raw image files were processed using CASAVA v1.82 for base calling and generating the raw sequencing data. For analysis of SNVs/indels and CNVs, FASTQ files were first aligned to the human reference genome (hg19) by the Burrows-Wheeler Aligner (bwa-0.7.15), and PCR duplications that cause bias in copy-number analysis were filtered out using Sambamba (v0.6.6) (https://github.com/biod/sambamba). In QC analysis, the average read number per sample was 36.8M (26-52M), the average sequencing depth was 114× (90-150×, with 20× coverage above 95%) and the sequencing Q30 accuracy was ≥90%.

SNVs/indel variants and SNPs were called using version 3.8 of the Genome Analysis Toolkit (GATK). Annotation databases utilized included (1) human population databases, such as gnomAD (http://gnomad.broadinstitute.org/), the 1000 Genome Project (http://browser.1000genomes.org), and dbSNP (http://www.ncbi.nlm.nih.gov/snp), (2) in silico prediction algorithms, such as SIFT (http://sift.jcvi.org), FATHMM (http://fathmm.biocompute.org.uk), Mutation Assessor (http://mutationassessor.org), CADD (http://cadd.gs.washington.edu), and SPIDEX (http://tools.genes.toronto.edu/), and (3) disease and phenotype databases, such as OMIM (http://www.omim.org), ClinVar (http://www.ncbi.nlm.nih.gov/clinvar), HGMD (http://www.hgmd.org), and HPO (https://hpo.jax.org/app/). Variants were classified into 5 categories, "pathogenic," "likely pathogenic," "uncertain significance," "likely benign," and "benign," according to the American College of Medical Genetics and Genomics (ACMG) guidelines for interpretation of genetic variants.[18]

The CNV analysis tool exome hidden Markov model (XHMM) was used for exon CNV calling. First, at least 20 libraries analyzed in the same flow cell were used to build a baseline reference set with a minimum correlation coefficient of >0.94 for copy number. The copy-number ratio of exon CNVs was calculated by dividing the mapped exon reads per kilobase per million (RPKM) value of the target sample by the average RPKM value of reference libraries in the same sequencing flow cell. Allowing a 20% standard deviation, exon copy numbers of 0.8 to 1.2 and 2.8 to 3.2 against reference (1.8-2.2) was used to define deletions and duplications, respectively.

Confirmatory molecular tests for CNVs

Pathogenic variants detected by ngVES were independently confirmed by further molecular tests. SNVs and INDELs were verified by either interactive genomics viewer plot sequence alignments of the allelic molecules and/or by Sanger sequencing. Small CNVs (<100 kb) were confirmed by multiplex ligation-dependent probe amplification.

Results

Clinical ngVES

A prospective group of 134 patients with a variety of skin disease conditions (Supplemental Table 1) were referred for ngVES for confirmatory genetic diagnosis. Based on best-match phenotype-genotype associations of genes known to be linked to skin diseases, 88 (66%) of patients had a positive gene finding (Figure 1), which was then reported to the clinician to aid clinical management of the condition. Of the 88 positive patients, 79 (90%) carried pathogenic SNVs/INDELs and 8 (9.5%) carried pathogenic CNVs and 1 carried both a SNV and a CNV (0.5%). Overall, the diagnostic rate of ES analysis (SNVs/INDELs) improved from 58% to 66% with ngVES analysis (SNVs/INDELs plus CNVs). This was considered a small, but clinically significant, increased diagnostic yield (Figure 1). Among the 88 positive patients, 75 (85%) were classified with an autosomal dominant (AD) disorder, 8 (9%) with an autosomal recessive (AR) disorder, 2 (2.5%) with an X-linked recessive (XR) disorder, 2 (2.5%) with a X-linked dominant disorder, and 1 (1%) with a dual AD and XR disorder (Figure 1).

Molecular diagnosis

Of the 88 positive patients, 76 were diagnosed with an AD skin condition. The most common causative AD variants

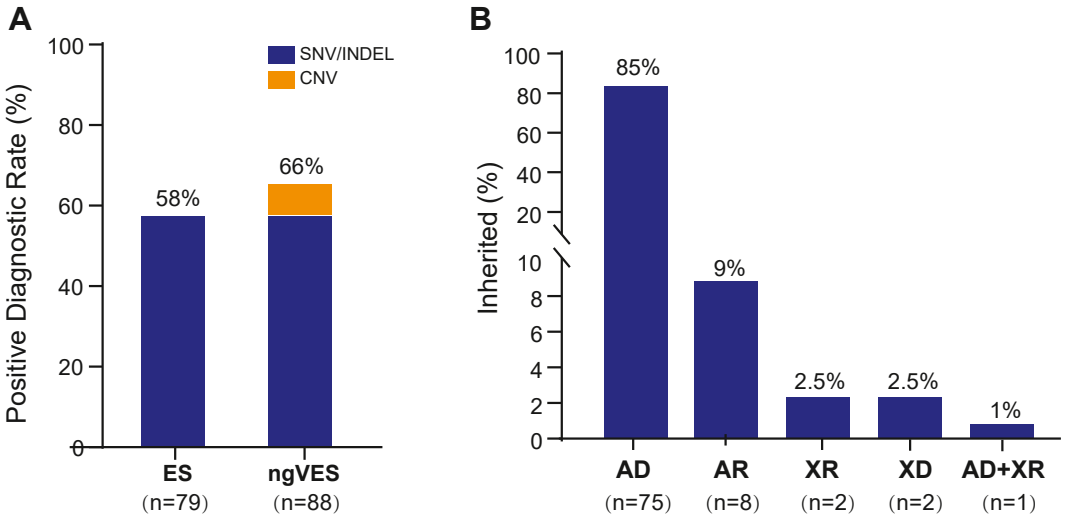

Figure 1 Performance of ngVES for diagnosis of skin conditions. A. Diagnostic yield of ES compared with ngVES. B. Disease inheritance pattern. AD, autosomal dominant; AR, autosomal recessive; INDEL, insertions/deletions; SNVs, single-nucleotide variations; XD, X-linked dominant; XR, X-linked recessive.

were identified within *NF1* (48; 63%), followed by *TSC2* (12; 16%), *FLG* (4; 5%), and *TSC1* (2; 3%) genes (Figure 2). One male patient with a *FLG* gene variant also had a variant in the *STS* gene suggesting an AD plus XR skin condition(s). In the remaining 9 patients (12%), there were individual cases identified with variants in *CARD14*, *FLCN1*, *KRT5*, *KRT6a*, *KRT6b*, *KRT9*, *KRT14*, *PITX2*, and *PTPN1* genes. For the 8 positive patients diagnosed with an AR skin condition, the 2 causative allelic variants were identified in the *CTSC*, *COL17A1*, *LAMB3*, *LIPA*, *LYST*, *NTRK1*, *SERPINB7*, and *TGM1* genes. Two cases of XR skin disease in male patients were caused by variants in the *STS* gene and another case with a variant in the *ATP7A* gene. There were 2 cases of an X-linked dominant skin disease in female patients associated with variants in the *IKBKG* gene.

The 75 AD gene positive patients all presented with a variety of skin features with café-au-lait spots being the most common (Figure 2, Supplemental Table 1). In 45 patients presenting with café-au-lait spots and 3 patients presenting with freckles in the armpit and groin regions, a pathogenic variant in the *NF1* gene was identified, confirming a diagnosis of neurofibromatosis type 1. In 14 patients that presented with hypomelanotic macules and/or angiofibromas, 12 had *TSC2* variants (1 mosaic, Supplemental Table 1), and 2 had *TSC1* variants, confirming a diagnosis of tuberous sclerosis type 2 and tuberous sclerosis type 1, respectively. For the remaining 16 patients with other types of abnormal skin features, a positive gene diagnosis consistent with a known syndrome or skin disease was identified.

Confirmation of pathogenic CNV variants detected by ngVES

The 9 pathogenic CNVs identified by the ngVES pipeline were independently assessed by other molecular methods (Supplemental Table 1). Five pathogenic CNVs were exonic gene deletions. There were 3 cases for the *TSC2* gene in which exons 1-9, exons 31-42 and exons 1-2 were deleted, 1 case for *NF1* in which exons 35-36 were deleted, and 1 case for the X-dominant gene *IKBKG* in which exon 4 was deleted. In all 5 cases, independent MLPA confirmed the CNVs called by ngVES (Figure 3). The remaining 4 cases were large whole gene deletions, including 1 case of a 1.4 Mb deletion involving *NF1* in which 10 neighboring OMIM genes were also deleted (Figure 4), 2 male cases of a 135 kb deletion involving the XR gene *STS* and 3 neighboring OMIM genes (Figure 5), and 1 case of a 5.8 kb deletion exclusively involving the entire *PITX2* gene. Independent MLPA and CNV-seq testing confirmed these 3 deletions (Figures 4 and 5).

Discussion

We developed and applied a new pipeline called ngVES for identifying pathogenic and likely pathogenic SNVs/INDELs and CNVs from standard ES data. In a prospective clinical study involving a large cohort of patients with a variety of skin diseases, the combined SNVs/INDELs/CNVs identified with this single test approach provided a higher positive diagnostic rate compared with ES or CNV-seq alone. For the skin diseases analyzed in this study, SNVs/indels were the major causative variants and small CNVs minor causative variants- both classes being identifiable by the ngVES pipeline. On this basis, we suggest that ngVES potentially offers advantages over either a single CNV-seq or ES investigation for other genetic conditions. This could include, cardiac[19] or neurodevelopmental disorders[20] in which CNVs are the more common pathogenic variants, but SNVs and or INDELS may also be involved.

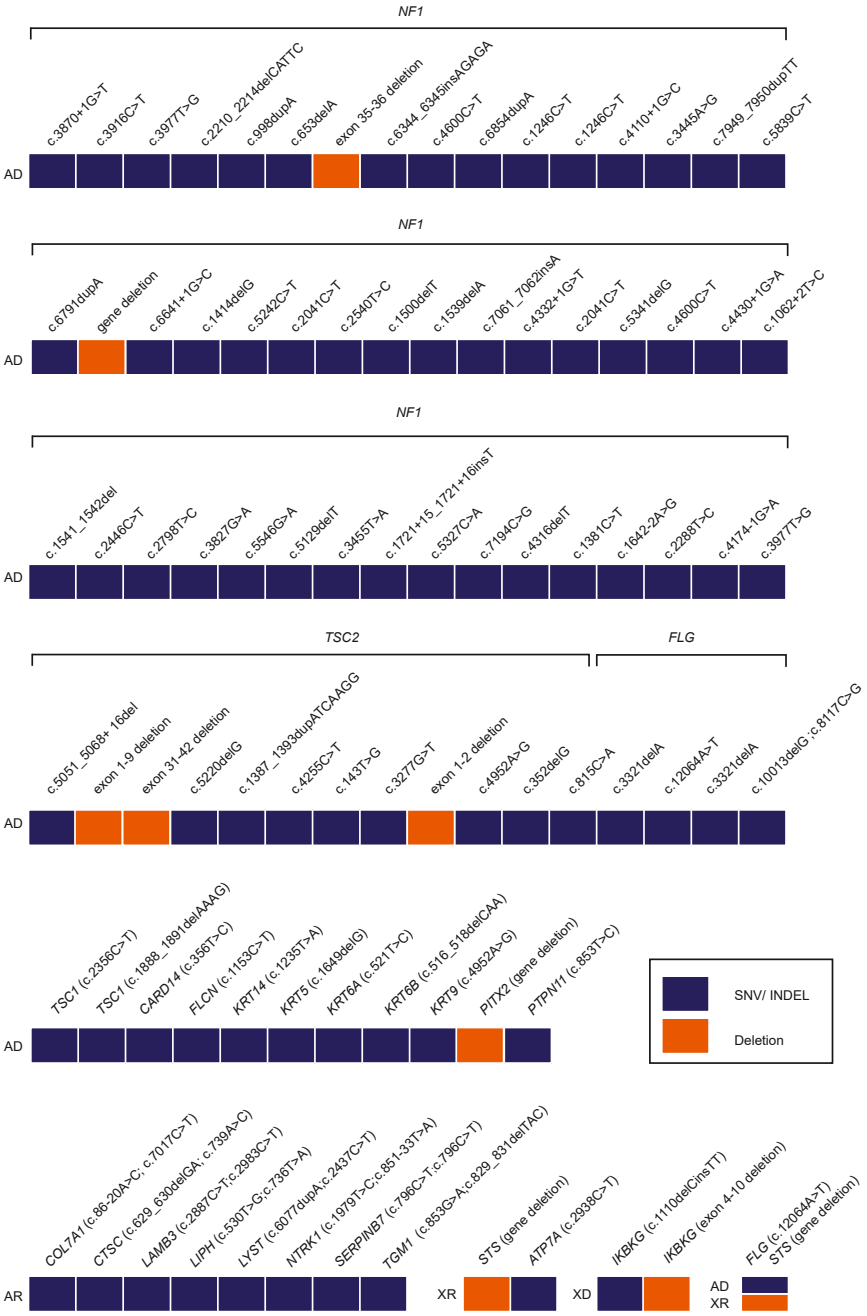

Figure 2 Positive molecular diagnosis for 88 patients with skin diseases. The genes and the genetic variants (SNVs/INDELS or CNVs) identified by ngVES for each positive sample are indicated for the different inheritance modes. AD, autosomal dominant; AR, autosomal recessive; XD, X-linked dominant; XR, X-linked recessive.

Although skin diseases with features of café-au-lait spots have a strong association with variants in the *NF1* gene[21] and hypomelanotic macules and angiofibromas are typical features of patients with *TSC1* or *TSC2* gene variants,[22] it can be argued that a directed screen would be the most logical initial approach. However, ngVES offers the more comprehensive initial screen with its ability to identify both SNVs and INDELs, as well as CNVs. In the model example, identification of *NF1* and *TSC1/2* gene variants is particularly important beyond initial skin disease presentation because it can also point to underlying systemic disease and tumorigenesis enabling more rapid implementation of emerging treatments,[22,23] as well as having implications for the immediate and extended family. Some genetic skin diseases are also quite difficult to diagnose; therefore, a firm molecular diagnosis can assist the dermatologist to determine the best treatment regimen.

In general, it is acknowledged in the literature that current ES approaches or panel approaches are not capable of identifying pathogenic variants outside the immediate exon gene regions, and important variants can potentially be missed in the analysis because of the variable depth of

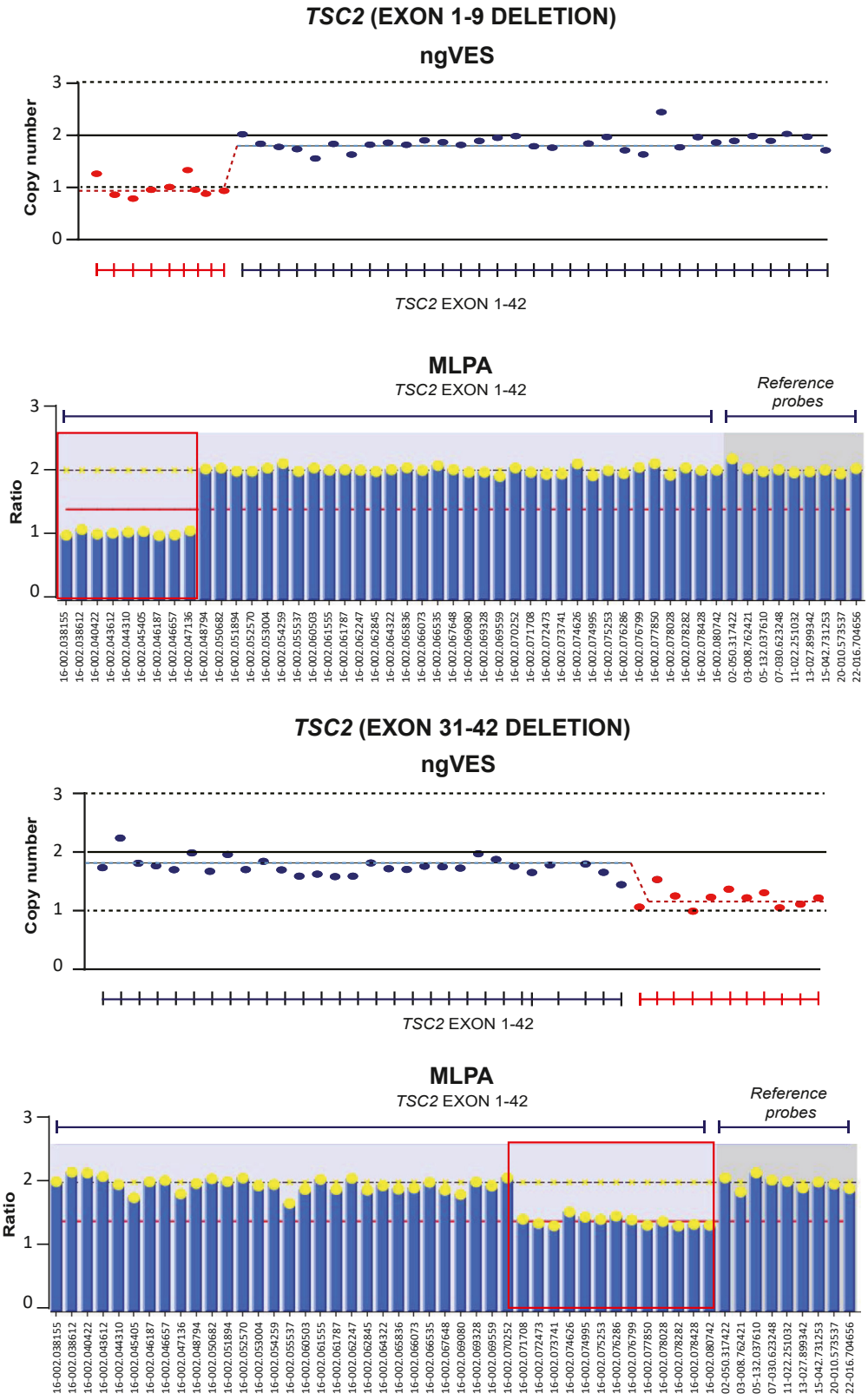

Figure 3 **Detection of small exon deletions by ngVES.** ngVES CNV plots and confirmatory MLPA plots are shown for 2 patients with *TSC2* exonic deletions.

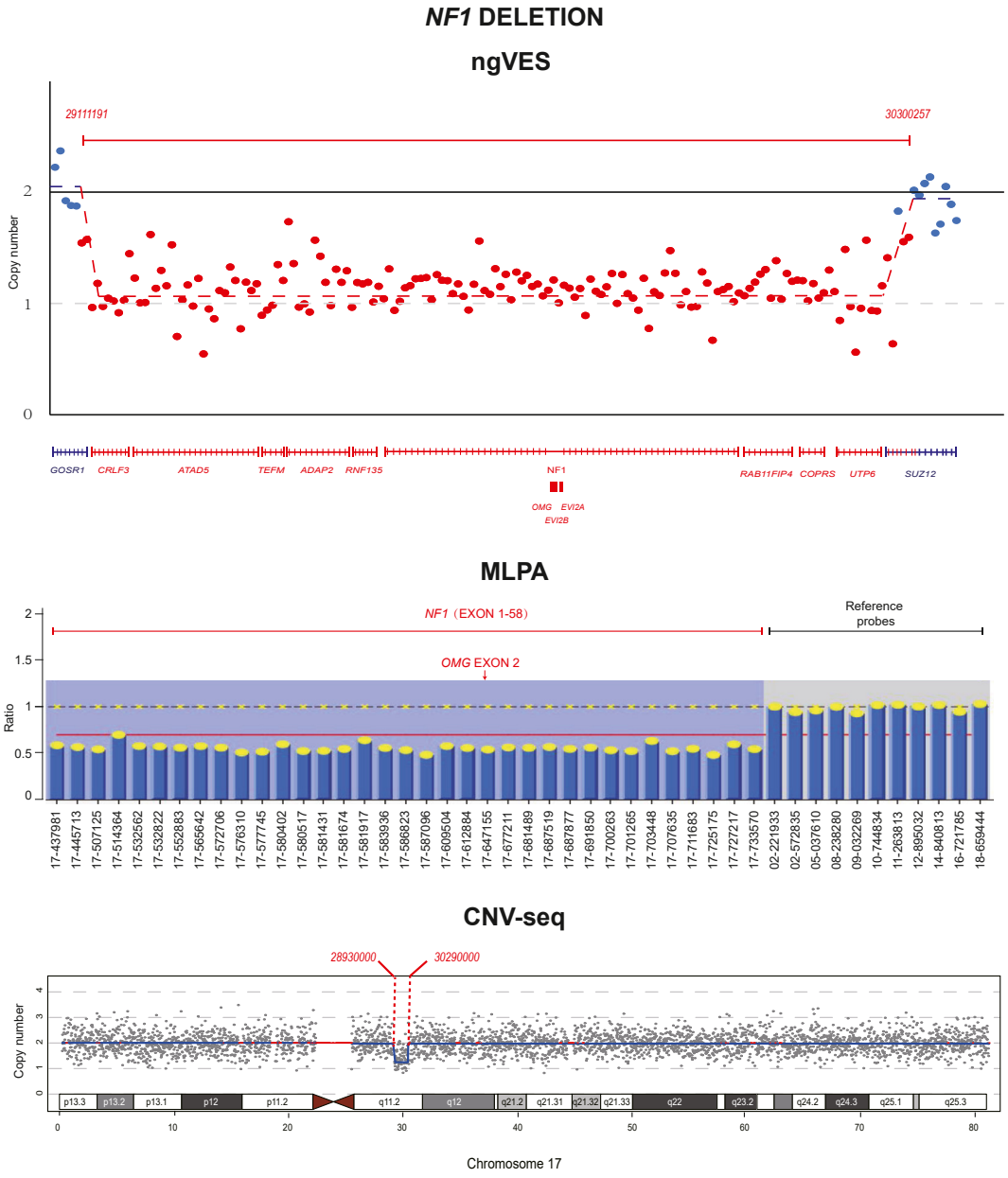

Figure 4 **Detection of a 1.4 Mb deletion involving the *NF1* gene.** The ngVES copy-number plot and the confirmatory MLPA and CNV-seq plots are shown. The deletion interval (hg19 coordinates) encompasses the OMIM genes *CRLF3, ATAD5, TEFM, ADAP2, RNF135, NF1, OMG, EVI2B, EVI2A, RAB11FIP4*, and *SUZ12*.

sequencing reads across some exonic regions. In similar studies using the same basic ES methodology it is possible that some relevant variants may not have been detectable with the specific ES panel selected. An alternative strategy to improve diagnostic rates would be GS, but this is a more expensive procedure because of the larger amount of sequencing involved and may not be particularly suitable for some CNV detection because of lower coverage (average sequencing depth of around 30) - especially if mosaicism is present. Another possibly more fruitful approach that could address some of these problems is gene sequencing. This approach would use a defined skin disease panel but expanded to include probes for intronic and 5 and 3 prime gene regions and then balanced for probe performance (equal depth all gene target regions). Applying a ngVES bioinformatics approach to whole gene data from an expanded and optimized skin disease panel could improve diagnostic gains for both SNV/INDEL and CNV pathogenic variants. Such an expanded, directed gene panel approach may logically also be useful for patients with other types of genetic conditions. However, any selected panel approach, whether for skin diseases or any other genetic condition, immediately limits discovery to the selected targets and may miss other contributory genetics. Any bias at the start will limit discovery at the analysis point. ES reduces any such initial biases by looking in a more global manner at the outset.

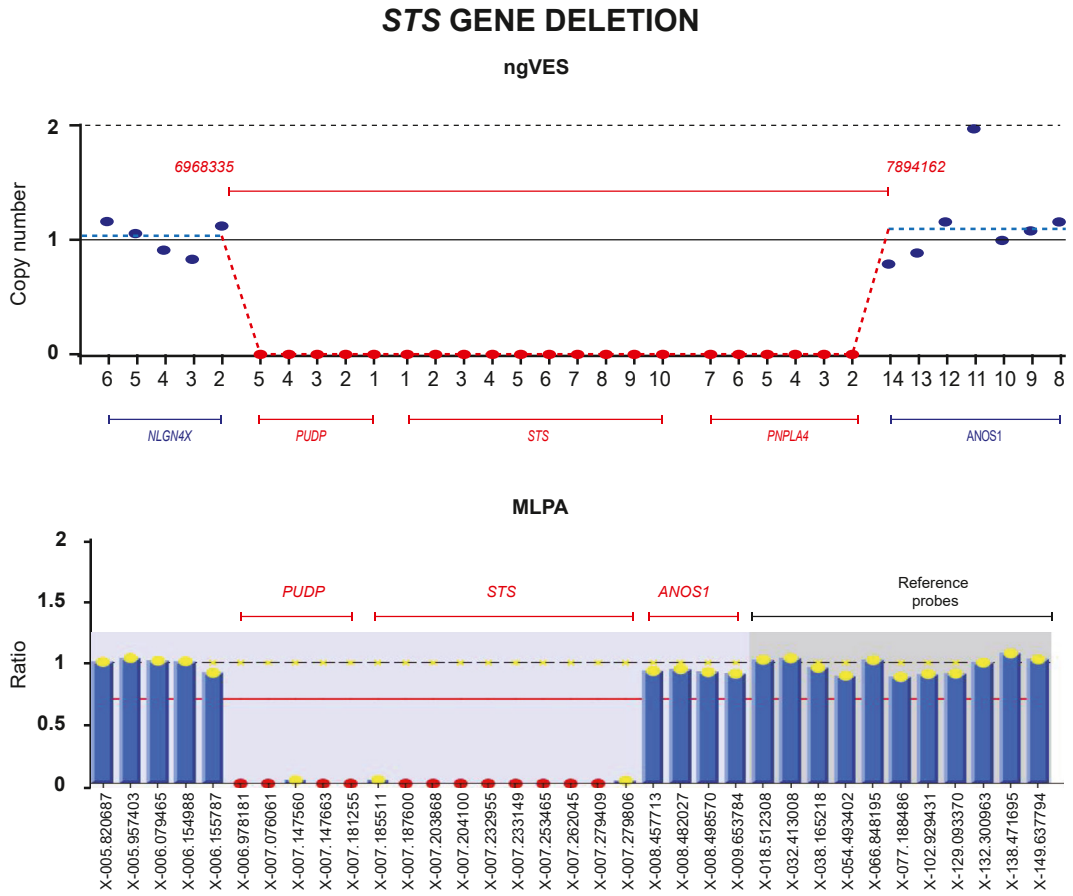

Figure 5 Detection of a 135 kb deletion involving the *STS* gene. The ngVES plot and the confirmatory MLPA plot are shown. The deletion interval (hg19 coordinates) encompasses the OMIM genes *PUDP*, *STS*, and *PNPLA4*.

In current clinical practice, when investigating a potential genetic disease, clinicians need to decide on which diagnostic approach to use. When a phenotype is apparently clear, a single gene test or a panel test may be logical and their first choice.[24] If the standard panels fail to identify a possible genetic cause, then further testing is often initiated. With no initial positive panel result or with less defined or unusual phenotypes, a CNV test (often by CMA or CNV-seq) may be the next choice, followed by an ES screen for the majority of these samples that show no obvious deletions or duplications events.[4] Nearly half of all samples will not have variations identifiable by CNV or sequence analysis; therefore, further testing will probably involve genome sequencing, although diagnostic gains here are often quite low and may sometimes reflect prior poor execution of either the current ES protocols or CNV interpretations.[25] As shown in this study, a single selected approach will sometimes miss important pathogenic variations that could have further clinical considerations and may also require interventions. The extended benefits of the ngVES pipeline are therefore in its broad approach to genetic screening in covering suspected targets for both sequence variation plus CNV changes, potentially obviating the need for separate tests for SNP/indels and CNVs.

A significant advantage of ngVES demonstrated here is the ability to identify very small CNVs involving only a single exon or a single gene in which such changes are not currently detectable using CNV-seq or by many current CMA platforms in which probe designs and regional gene coverage can vary significantly. ngVES provides a non-biased approach to investigating the genetic basis of any disease and presents a more comprehensive option that we suggest may be suitable for first-tier screening. We further suggest that using a novel bioinformatics approach, CNVs called by ngVES can be independently confirmed using the existing ES information provided in the sequencing data and benchmarking against reference data from the same CNV interval region in other samples analyzed in the same sequencing flow cell. Thus, for small exon CNVs or larger CNVs involving the gene of interest, it will be possible to determine with a high degree of confidence whether the CNV called is a true- or false-positive result utilizing inherent SNPs. If any CNV still falls into the gray zone, then other confirmatory molecular tests can be used independently to validate the CNV called by ngVES before results are reported to the patient. It is reasonable to also suggest that analysis can be extended, using this SNP assessment approach, to simply and conveniently investigate and

exclude rare, but potentially important, uniparental disomy occurrences with no further separate testing required.

The important improvement of ES demonstrated here is based purely on bioinformatics; therefore, ngVES can be translated to laboratories everywhere and thus offers wider applications in many areas of disease screening and diagnosis. The main limitation in such an application will revolve around expertise in ES application and related technology, including target enrichment processes. This will undoubtedly include choice of PCR sets or hybridization enrichment probe sets. For maximum gains to be achieved, the laboratory must have appropriate skills in all areas of NGS with rigor in NGS and analysis/reporting. This may be especially important when the ngVES technology is ready for routine prenatal applications because the ES technology is already in transition to this area.[26,27]

Despite some potential limitations, ngVES is moving toward a more universal screening approach with the potential for increased detection rates and reduced follow-up requirements. The cost of sequencing has dropped dramatically over the last decade and the availability of commercial probe sets has increased substantially. Therefore, we can propose that ngVES also has obvious potential in any carrier screening program as a first-tier approach covering SNVs, INDELs, and CNVs over much of the genome. This composite analysis approach demonstrated herein also has important implications for more effective carrier screening in cases which the use of ngVES may offer efficient means compared with the use of multiple selected gene variant panels. This would be a better utilization of genetic data and is suitable for comprehensive carrier testing in any population background.

In conclusion, being a composite bioinformatics analysis, ngVES requires no changes in laboratory technology or workflow. Sequencing costs are low and are the same as a singular ES investigation, but the patient receives the benefit of a more comprehensive hybrid test. In any disease investigations, the simultaneous examination of other genes/chromosomes may, in some cases, refine clinical diagnosis and can be especially important to resolve the genetic basis of mixed phenotypes. From the laboratory perspective, efficiencies are improved by rationalization of methodologies with fewer wet lab requirements compared with cascade screening, with the added benefit of potential cost reductions for both laboratory and patient.

Data Availability

Requests for data can be directed to Xiangdong Kong: kongxd@263.net.

Acknowledgments

The authors thank the National Key R&D Program (2018YFC1002203); the Major Science and Technology Projects in Yunnan Province (202302AA310018) and Spring City Plan: the High-level Talent Promotion and Training Project of Kunming (2022SCP002) for research support.

Funding

The study was supported by grants from National Key R&D Program (2018YFC1002203).

Author Information

Conceptualization: X.K., D.S.C., D.A.L.; Data Curation: C.W., P.S., H.L.; Funding Acquisition: X.K. Writing-original draft: D.S.C., D.A.L.; Writing-review and editing: X.K., D.S.C., D.A.L.

ORCIDs

Conghui Wang: http://orcid.org/0000-0001-8617-6223
Panlai Shi: http://orcid.org/0000-0001-5826-1340
Hongbin Liang: http://orcid.org/0000-0002-1308-1335
David S. Cram: http://orcid.org/0000-0002-2587-2476
Donald A. Leigh: http://orcid.org/0000-0002-9025-4244
Xiangdong Kong: http://orcid.org/0000-0001-6014-221X

Ethics Declaration

The study was approved by the Medical Ethics Committee of the First Affiliated Hospital of Zhengzhou University (KS-20189-KY-36). All patients who participated in the study provided written informed consent for undertaking genetic diagnostic testing.

Conflict of Interest

The authors declare no conflicts of interest.

Additional Information

The online version of this article (https://doi.org/10.1016/j.gimo.2024.101816) contains supplementary material, which is available to authorized users.

References

1. Verma IC, Puri RD. Global burden of genetic disease and the role of genetic screening. *Semin Fetal Neonatal Med*. 2015;20(5):354-363. http://doi.org/10.1016/j.siny.2015.07.002

2. Rehm HL. Evolving health care through personal genomics. *Nat Rev Genet*. 2017;18(4):259-267. http://doi.org/10.1038/nrg.2016.162
3. Pogue RE, Cavalcanti DP, Shanker S, et al. Rare genetic diseases: update on diagnosis, treatment and online resources. *Drug Discov Today*. 2018;23(1):187-195. http://doi.org/10.1016/j.drudis.2017.11.002
4. Aylsworth AS. Defining Disease Phenotypes (Chapter 2). In: Haines JL, Pericak-Vance MA, eds. *Genetic Analysis of Complex Diseases*. 2nd ed. John Wiley & Sons; 2005:51-82.
5. Levy B, Wapner R. Prenatal diagnosis by chromosomal microarray analysis. *Fertil Steril*. 2018;109(2):201-212. http://doi.org/10.1016/j.fertnstert.2018.01.005
6. Liang D, Peng Y, Lv W, et al. Copy number variation sequencing for comprehensive diagnosis of chromosome disease syndromes. *J Mol Diagn*. 2014;16(5):519-526. http://doi.org/10.1016/j.jmoldx.2014.05.002
7. 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
8. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of Mendelian disorders. *N Engl J Med*. 2013;369(16):1502-1511. http://doi.org/10.1056/NEJMoa1306555
9. Hayeems RZ, Dimmock D, Bick D, et al. Clinical utility of genomic sequencing: a measurement toolkit. *npj Genom Med*. 2020;5(1):56. http://doi.org/10.1038/s41525-020-00164-7
10. Sun Y, Ye X, Fan Y, et al. High detection rate of copy number variations using capture sequencing data: a retrospective study. *Clin Chem*. 2020;66(3):455-462. http://doi.org/10.1093/clinchem/hvz033
11. Yuan B, Wang L, Liu P, et al. CNVs cause autosomal recessive genetic diseases with or without involvement of SNV/indels. *Genet Med*. 2020;22(10):1633-1641. http://doi.org/10.1038/s41436-020-0864-8
12. Yang L, You C, Qiu S, et al. Novel and de novo point and large microdeletion mutation in PRRT2-related epilepsy. *Brain Behav*. 2020;10(5):e01597. http://doi.org/10.1002/brb3.1597
13. Chen X, Jiang Y, Chen R, et al. Clinical efficiency of simultaneous CNV-seq and whole-exome sequencing for testing fetal structural anomalies. *J Transl Med*. 2022;20(1):10. http://doi.org/10.1186/s12967-021-03202-9
14. Hansen AW, Murugan M, Li H, et al. A genocentric approach to discovery of Mendelian disorders. *Am J Hum Genet*. 2019;105(5):974-986. http://doi.org/10.1016/j.ajhg.2019.09.027
15. 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
16. Zhao L, Liu H, Yuan X, Gao K, Duan J. Comparative study of whole exome sequencing-based copy number variation detection tools. *BMC Bioinformatics*. 2020;21(1):97. http://doi.org/10.1186/s12859-020-3421-1
17. Fromer M, Purcell SM. Using XHMM software to detect copy number variation in whole-exome sequencing data. *Curr Protoc Hum Genet*. 2014;81:7.23.1-7.23.21. http://doi.org/10.1002/0471142905.hg0723s81
18. 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
19. Zhu X, Li J, Ru T, et al. Identification of copy number variations associated with congenital heart disease by chromosomal microarray analysis and next-generation sequencing. *Prenat Diagn*. 2016;36(4):321-327. http://doi.org/10.1002/pd.4782
20. Zhai Y, Zhang Z, Shi P, Martin DM, Kong X. Incorporation of exome-based CNV analysis makes trio-WES a more powerful tool for clinical diagnosis in neurodevelopmental disorders: a retrospective study. *Hum Mutat*. 2021;42(8):990-1004. http://doi.org/10.1002/humu.24222
21. Lalor L, Davies OMT, Basel D, Siegel DH. Café au lait spots: when and how to pursue their genetic origins. *Clin Dermatol*. 2020;38(4):421-431. http://doi.org/10.1016/j.clindermatol.2020.03.005
22. Salussolia CL, Klonowska K, Kwiatkowski DJ, Sahin M. Genetic etiologies, diagnosis, and treatment of tuberous sclerosis complex. *Annu Rev Genomics Hum Genet*. 2019;20:217-240. http://doi.org/10.1146/annurev-genom-083118-015354
23. Wilson BN, John AM, Handler MZ, Schwartz RA. Neurofibromatosis type 1: new developments in genetics and treatment. *J Am Acad Dermatol*. 2021;84(6):1667-1676. http://doi.org/10.1016/j.jaad.2020.07.105
24. Segal MM. Genome interpretation: clinical correlation is recommended. *Appl Transl Genom*. 2015;6:26-27. http://doi.org/10.1016/j.atg.2015.07.002
25. Schwarze K, Buchanan J, Taylor JC, Wordsworth S. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genet Med*. 2018;20(10):1122-1130. http://doi.org/10.1038/gim.2017.247
26. Lord J, McMullan DJ, Eberhardt RY, et al. Prenatal exome sequencing analysis in fetal structural anomalies detected by ultrasonography (PAGE): a cohort study. *Lancet*. 2019;393(10173):747-757. http://doi.org/10.1016/S0140-6736(18)31940-8
27. Petrovski S, Aggarwal V, Giordano JL, et al. Whole-exome sequencing in the evaluation of fetal structural anomalies: a prospective cohort study. *Lancet*. 2019;393(10173):758-767. http://doi.org/10.1016/S0140-6736(18)32042-7