Clinical Genetics and Genomics for the Immunologist: A Primer

Maxwell H. Bannister, BS[a], Xiao P. Peng, MD, PhD[b,c,d,*]

KEYWORDS

- Molecular diagnostic testing • Next-generation sequencing
- Inborn error of immunity • Primary immunodeficiency
- Variant of uncertain significance • Somatic mosaicism • Genetic analysis
- Immunogenetics

KEY POINTS

- An increasing arsenal of available tools is helping to elicit the pathophysiologies of GIDs, though current discovery still focuses on Mendelian, single-locus mechanisms.
- Up-front genetic testing adds valuable molecular information to clinical and immune data, but genomics is one amongst many tools for guiding diagnosis and care.
- Formulating a diagnostic strategy requires understanding the genetic principles associated with a patient's differential and awareness of the strengths and limitations of available assays.
- The authors provide a practical guide to inform such decision-making and address issues related to suboptimal diagnostic yields, patient counseling, and results reporting.

INTRODUCTION

Human immune-mediated disease involves complex interactions between the environment and biologic processes of immunity that exist in all living cells, not just "professional" immune cells. Our field continues to identify Mendelian forms of genetically driven immune disease (GID), as well as to uncover the genetic architectures and paradigms driving these disorders.

[a] Medical Scientist Training Program, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [b] Genetics of Blood and Immunity, Montefiore Einstein; [c] New York Center for Rare Diseases; [d] Division of Pediatric Genetic Medicine, Department of Pediatrics, The Children's Hospital at Montefiore, The University Hospital for Albert Einstein College of Medicine, 3411 Wayne Avenue, 9th Floor, Bronx, NY 10467, USA
* Corresponding author. Division of Pediatric Genetic Medicine, Department of Pediatrics, Children's Hospital at Montefiore, Montefiore Medical Center, Albert Einstein College of Medicine, 3411 Wayne Ave, Bronx, NY 10467.
E-mail address: xpeng@montefiore.org

Immunol Allergy Clin N Am 45 (2025) 153–171
https://doi.org/10.1016/j.iac.2025.01.002 **immunology.theclinics.com**

Abbreviations	
ACMG	American College of Medical Genetics
ALPS-FAS	autoimmune lymphoproliferative syndrome caused by mutations in the FAS gene, which encodes the first apoptosis signal (receptor)
AFs	allele frequencies
CFH-CFHR	the complement factor H and complement factor H-related gene locus
CMA	chromosomal microarray
CNV	copy number variation
DNA	deoxyribonucleic acid
DRAGEN	dynamic read analysis for GENomics
F8	coagulation factor VIII
GATK	genome analysis toolkit
IKBKG	inhibitor of nuclear factor kappa-B kinase subunit gamma
IUIS	International Union of Immunological Sciences
GID	genetically driven immune disorder
LP	likely pathogenic
LRS	long-read sequencing
MAVE	multiplexed assays of variant effect
NGS	next-generation sequencing
NCF1	neutrophil cytosolic factor 1
NLRP3	NOD-, LRR-, and pyrin domain-containing protein 3
NOD2	nucleotide-binding oligomerization domain containing 2
OGM	optical genome mapping
P	pathogenic
PCR	polymerase chain reaction
RNA	ribonucleic acid
SBS	sequencing-by-synthesis
SAMD9/9L	sterile alpha motif domain containing 9/9-like
SNP	single nucleotide polymorphism
SNV	single nucleotide variant
SRS	short-read sequencing
SV	structural variant
TGP	targeted gene panel
UBA1	ubiquitin-like modifier activating enzyme 1
UNC13D	Unc-13 homolog D
VEXAS	vacuoles, E1 enzyme, X-linked, autoinflammatory, somatic
VUSes	variants of uncertain significance
WAS	the Wiskott-Aldrich syndrome gene
WES	whole exome sequencing
WGS	whole genome sequencing

Historically referred to as primary immunodeficiencies, immune disorders resulting from high-impact genetic lesions are now known to present with a wide range of immune-mediated manifestations beyond infection susceptibility—including immune dysregulation, autoinflammation, atopy, hematologic complications, and malignancy risk. This distinction has been recognized with a shift in the field toward increasingly widespread use of the term inborn errors of immunity. However, as this term fails to accommodate conditions driven by somatic changes, the authors adopt the term GID in this review.

The advent of next-generation sequencing (NGS) and increasing availability of diagnostic genetic testing has enabled us to add a valuable genotypic layer of information to the clinical and immunophenotypic spectrum of immune disease, expanding our understanding of pathobiology to inform more personalized management. Though genomics is only one of many tools for establishing molecular diagnosis, clinical genetic testing has become increasingly cost-effective and widely available. It has proven

particularly crucial for the diagnosis of GIDs for which we currently have few to no diagnostic immune studies.

A common misconception about the clinical practice of genetics is that it is tantamount to the ordering of genomic studies, rather than strategizing the best way to achieve patient-specific molecular diagnose(s) and treatment. Sometimes, the best strategies do not involve genomics at all—these are discussed in subsequent chapters. Herein, we provide a practical guide to understanding, choosing, and interpreting currently available genetic testing for patients with GID. A brief review of some current genetic landscapes and paradigms provides the context for testing strategy and assay choice, followed by a primer on genetic analysis, variant interpretation, and what to do with reported results.

DISCUSSION

Genetic Paradigms

The International Union of Immunological Sciences (IUIS) most recently released a GID nosology in 2022, encompassing 10 subcategories grouped on the basis of shared pathobiologies and phenotypic features.[1] Just as every aspect of immunity is represented in this catalog of immune disease, so do the pathogenic mechanisms span every facet of cellular physiology. However, significant clinical and mechanistic overlap is also recognized across these categories.

While heterogeneity in the expression of human genetic disorders is not a new concept, GIDs demonstrate this to an incredible extent and diversity, making molecular diagnosis and genetic interpretation challenging. On one level, there is a vastly heterogeneous genetic landscape with most clinical presentations or phenotypes associated with multiple genes, some yet to be discovered. On a deeper level, these individual genes are often associated with multiple disease-causing mechanisms and inheritance patterns. Even for conditions caused by recessively inherited enzymatic loss-of-function such as deficiency of adenosine deaminase 2 (DADA2), genotype–phenotype relationships may not be clear or straightforward.[2,3] Additionally, modifiers on the DNA, RNA, protein, environmental, and other levels further contribute to the variable expressivity and incomplete penetrance typically seen with GIDs.[4] Thus, the same mutation in close kindred may present very differently, while different mutations in different genes can present very similarly for different individuals. As a consequence, molecular diagnosis in and of itself is important but not always sufficient for informing management. One may still need to more precisely characterize each individual's specific expression of disease and potential for additional complications. At the end of the day, treatment must be guided by the synthesis of many forms of data, of which genetic information is simply one.

Spotlight on somatic variation

Post-zygotic somatic mutations can phenocopy or rescue germline GIDs, adding additional layers of complexity to genetic detection and analysis, and harboring important implications for reproductive counseling.

Mosaicism refers to the presence of 2 or more genetically distinct cell populations within an individual, all derived from a single fertilized egg.[5] This can be caused by de novo variation arising at any stage of post-zygotic development, with the extent of mosaic involvement dependent on the specific timing of the mutational event (**Fig. 1**). There is no germline involvement with somatic mosaicism, while there is exclusively germline involvement with gonadal mosaicism, and gonosomal mosaicism refers to involvement of both somatic and germline tissues.

It is usually difficult to distinguish among these forms of involvement without a family history indicating disease recurrence in other members or sampling of multiple tissues

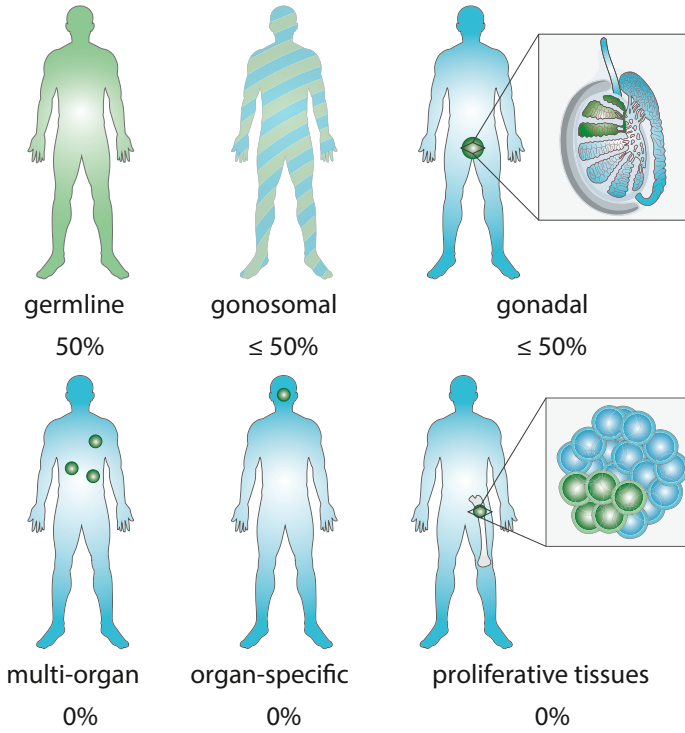

Fig. 1. The tissue distribution of mosaic variants has important implications for disease presentation, diagnostic approach, and inheritance counseling. *Green, affected, blue, unaffected; percentages indicate risk of passing on a dominantly inherited trait.*

from the proband. Thus, for both gonosomal and gonadal mosaicism, routine genetic testing on blood or buccal samples may fail to detect a disease-causing variant in a clinically unaffected parent who harbors ongoing potential for transmission to offspring. Even if trio NGS-based testing suggests that the variant found in the affected proband arose *de novo*, a recurrence risk of 1% is typically quoted during counseling, given the possibility of undetected parental mosaicism.[6]

Detection and sampling strategy will be informed by disease-relevant allele frequencies (AFs) and tissue(s). Some conditions are obligately driven by somatic mutations, while others may involve germline disease where mosaicism may develop. In the latter situation, mosaic variation may cause disease or ameliorate/rescue clinical and cellular phenotypes. The latter is of particular concern for disorders of genome maintenance and related conditions such as ribosomopathies, where selection pressure may exist *against* disease-causing mutations in hematopoietic cells. Indeed, this is so commonly seen for the *SAMD9/9L* syndromes that the accumulation of additional somatic changes in peripheral blood and/or bone marrow is sometimes used as a diagnostic hallmark of the condition.[7] In these cases, seemingly clinically unaffected family members may still be at risk for disease manifestations, so additional sequencing and/or tissue sampling may be needed to determine the true extent of mosaicism.

In some situations involving concern for low variant allele frequencies (VAFs), *a priori* enrichment for a relevant cell population that expresses the mutation of interest may enhance detection—this is currently clinically available for ALPS-FAS.[8] However, this

may not be always necessary—for example, even though the *UBA1* mutations causing VEXAS are found exclusively in myeloid cells, clinical disease tends to correlate with mutational burden.[9] In contrast, very low VAFs may still be disease relevant for mosaic variations of typically germline GIDs such as Blau syndrome due to *NOD2* gain-of-function (GOF) mutations or *NLRP3*-related inflammasomopathy.[10]

Genetic Testing for Immune Disease

Diagnostic testing choice is guided by many considerations, but the basic aim is to choose the assay(s) and analysis most appropriate for the paradigms relevant to each patient's genetic differential (Fig. 2).

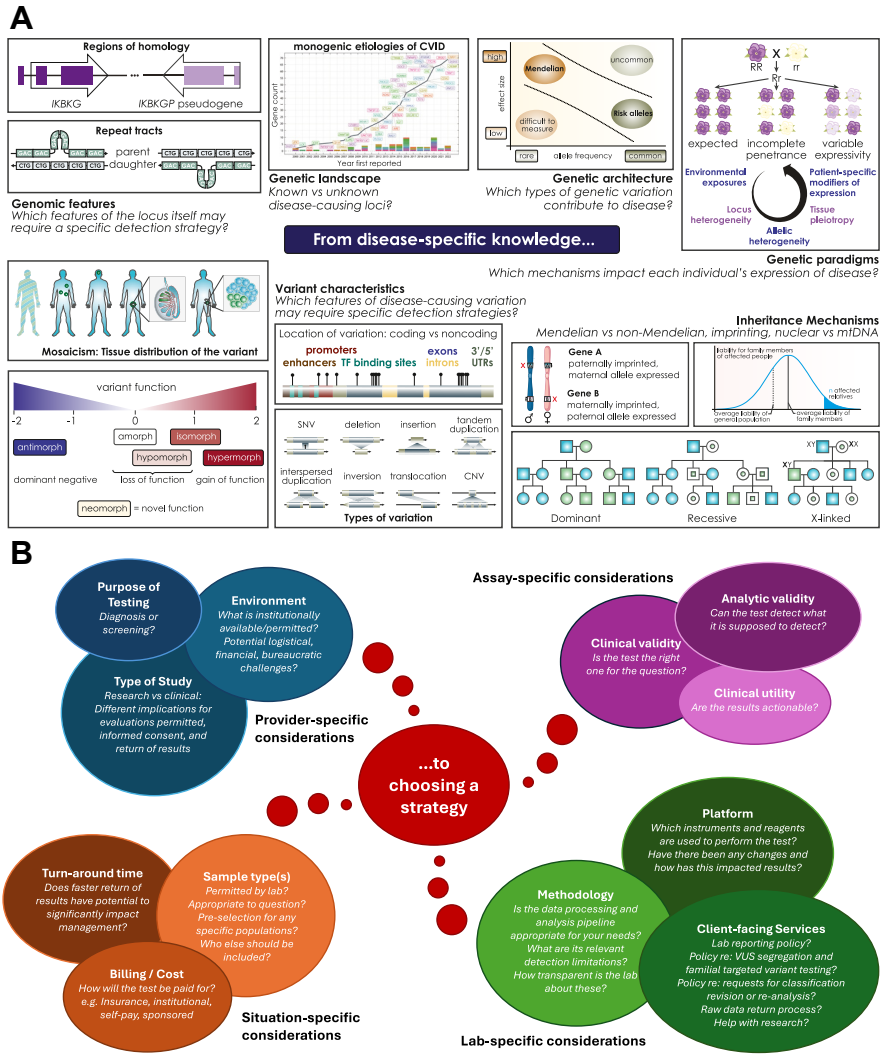

Fig. 2. Genetic testing choice is predicated on understanding and aligning (*A*) the genetic paradigms and pitfalls of a patient's differential with (*B*) the goals of testing and features of available diagnostic tests.

In this section, we review current clinically available DNA-based assays through emerging clinical diagnostic tools (**Tables 1** and **2**). We refer the reader to recent reviews for discussion of other molecular diagnostic studies, such as those based on DNA methylation, RNA sequencing, proteomics, or metabolic profiling.[11]

Single gene testing

There are many approaches to single gene testing depending on the type(s) of lesion(s) being queried, but gene-specific Sanger sequencing or NGS is still commonly offered by many commercial laboratories. This remains a reasonable first-line approach for patients with highly specific clinical presentations due to germline or high-frequency somatic mutations (eg, *UBA1*) as traditional Sanger sequencing can detect AFs as low as approximately 15% to 20%.[12,13] Other applications may include molecular confirmation in patients with demonstrated protein defects by flow cytometry (eg, WAS)[14] or enzyme activity (eg, ADA2),[15] or evaluation for mutations in hotspots or founder populations. In the case of loci with extensive sequence homology (eg, *IKBKG, NCF1, CFH-CFHR*) or known recurrent structural variation (eg, *UNC13D, F8*), this may be combined with other orthogonal approaches for mutation detection such as multiplex ligation-dependent probe amplification[16] or PCR-based strategies.[17–20]

Chromosomal microarray

The earliest chromosomal microarrays (CMAs) relied on comparative genome hybridization to draw relative dosage-based conclusions about copy number variation (CNV) on the order of magnitude of 10^5 to 10^6 base pairs (bp). Current commercial microarrays use allele-specific DNA probes targeting regions in which there is single nucleotide polymorphism (SNP) variation between individuals, with greater probe density in known gene-rich regions. Thus, detection may vary from locus to locus depending on the distance spanned by individual probes, but reliable detection is generally afforded at approximately 10^4 bp order of magnitude and higher. SNP arrays can also identify regions of homozygosity including regions of uniparental isodisomy that may harbor recessive disease-causing variants. Larger CNVs and structural variants (SVs) including balanced translocations and more complex rearrangements may be better revealed by karyotype, while smaller changes may be imputed from NGS data using tailored bioinformatics strategies. However, the latter also relies on having sufficient depth and breadth of coverage across the region(s) of interest. In general, CNVs on the order of 10^2 to 10^3 bp remain a detection challenge for current clinical platforms, though this may soon change with the advent of optical genome mapping (OGM) and long-read sequencing (LRS).[21]

Next-generation sequencing (NGS)

NGS—also known as second generation, massively parallel or short-read sequencing (SRS)—broadly relies on sequencing-by-synthesis (SBS) or sequencing-by-hybridization strategies. Illumina's high-throughput SBS instruments dominate the current market and are widely used for research-grade and clinical-grade targeted gene panels (TGPs), whole exome sequencing (WES) or whole genome sequencing (WGS). However, even more accurate and cost-efficient technologies continue to emerge. On the horizon are "third generation" LRS platforms that perform true single-molecule sequencing on much longer DNA strands, holding promise for improved detection of disease-causing lesions that have thus far eluded NGS (see **Table 1**).[22]

While the specific parameters underlying genetic tests may vary widely, each can be conceived of as optimizing tradeoffs between breadth and depth of coverage, in addition to turnaround time, storage space, and costs (see **Table 2**). Breadth of coverage

Table 1
Available genomic sequencing platforms and their clinical applications

	Generations of Available Sequencing Platforms		
	Sanger (1st gen)	**Short-Read NGS (2nd gen)**	**Long-Read Single Molecule (3rd gen)**
Technology	Size-separation of end-labeled DNA fragments	Sequencing by synthesis or by hybridization	Single molecule real-time detection
Throughput	Low (1 piece of DNA per run)	High (10^6–10^{10} reads per run)	Varies but can be high
Sample input	Scales with number of loci interrogated	Low	Higher
Read lengths	500–1000 bp	50–300 bp	10^4–10^5 bp
Accuracy	High	High	Lower, though improving
Runtimes	Minutes to hours	1–10 d	1–2 d
Data file sizes	Small	Large	Largest
Data analysis	Straightforward and standardized	More complex but largely standardized	Complex and still evolving
Cost	High per base, low per run	Low per base, moderate per run	Low per base, high per run
Pros	Rapid, low cost, high accuracy interrogation of specific loci	Relatively fast and cost-effective interrogation of many loci with high accuracy	Fewer reads required for detection confidence, ability to capture modified DNA states and difficult-to-detect regions
Cons	Poorly scalable, poorly amenable to SV detection	Difficulty with SV detection and mapping of nonunique sequence	Lower SNV calling accuracy, high data storage, and computing requirements

Table 2
The most commonly used clinical-grade DNA diagnostic assays based on the technologies in Table 2

Comparison of Current Commercially Available Clinical Assays					
			NGS-Based		
	Sanger	**SNP-Based CMA**	**TGP (Often)**	**WES**	**WGS**
Target	Single gene	Whole genome*	2–650+ genes	1.5% of gene	Whole genome*
Coverage	High for locus of interest, absent elsewhere	Variable (higher in disease gene-enriched regions)	High for loci of interest, absent elsewhere (depends in backbone)	High for coding regions of most genes (depends on specific capture system used)	Most uniform across loci
Read depths	N/A		>300x	>50–100x	>30–60x
Variants detected	$<10^2$	10^1–10^2	10^2–10^4 (depends on panel size and content)	10^4–10^5	10^6
Order price	$10s–$100	~$500	$250–3500 (depends on panel size, depth, and content)	$750–$6000 (lower for proband-only)	$1500–$9000 (lower for proband-only)
TAT	2–3 wk	3–4 wk	Routine: 308 wk Rapid: Within 1 wk	Routine: 6–8 wk Rush: 3–4 wk Rapid: 7–10 d (for preliminary result)	
Pros	Fast, low cost, accurate, lowest probability of VUSes	Detects CNVs and ROHs including UPD across genome (~10^4–10^6 bp in size)	Faster, lower cost, higher depth and accuracy for chosen loci than WES/WGS May be customized with addition of probes or orthogonal assays	Captures significantly more loci than TGPs Greater average read depth for detection of potential mosaicism than WGS Can impute CNVs bioinformatically	Less sequencing and amplication bias than WES Better detection of noncoding regions and SVs than WES

Cons	Detection limited to specific loci chosen Poor detection of SVs/CNVs	Low resolution Limited population data Cannot detect smaller CNCs or copy number neutral SVs	Detection limited to specific loci chosen May share detection pitfalls with WES/WGS Proband-only testing	Higher cost and data storage burden More complex computational and analytical requirements Higher probability of VUS detection More elements of consent, including incidental findings Can include family members for comparison Can reanalyze raw data without re-sequencing	
Clinical scenario	*WAS* sequencing after abnormal WAS protein flow cytometry result	Detection of 22q11.2 deletion given clinical features of DiGeorge syndrome	CGD panel testing that includes *NCF1* exon 2-specific detection after abnormal DHR result	Complex, nonspecific phenotype with many loci on differential Multiple candidates have potential for mosaicism, that is, *SAMD9/SAMD9L, TLR8, NLRC4, FAS*	Potential for disease-causing noncoding genes and/or variation, that is, *GATA2, UNC13D, RNU4ATAC, and RMRP*

* "Whole" is somewhat misleading in this case - more precise would be "most of the genome detectable by this technology".

essentially represents the range of genomic loci included in detection, while depth of coverage equals the number of times any given nucleotide is sequenced. As the total data generated approximates the product of these 2 values, maximizing one parameter usually occurs at the expense of the other, unless one specifically funds the additional data generation. Commercial laboratories rarely provide exact numbers but usually guarantee an average coverage level at a certain minimum depth across the genome, for example, 100% coverage at min 10x depth. For coverages less than 100%, additional clarification may be required to understand whether focal or more diffusely distributed in nature.

Depth is a particularly important consideration where concern for mosaicism exists. For example, depths of coverage around 10^2 to 10^3 orders of magnitude are usually achievable with most TGPs or WES but would be prohibitive for WGS given its much greater breadth. VAFs as low as 5% to 10% may be detectable by WES, though the genetic testing laboratory may not specifically look for these variants unless the ordering provider specifies their loci of concern up front. In contrast, use of a focused TGP, particularly if designed explicitly for somatic variant detection, may be the best strategy for detecting lower abundance disease-causing variants. However, even with sufficient coverage depth, additional variant calling and analysis strategies may still be needed for true detection—these will be discussed in the following section.

Targeted gene panels. One cannot assume that all similarly named TGPs offer the same content and detection parameters. Some laboratories provide all TGPs on an exome or genome backbone, with only the indicated set of genes uncovered for analysis. Others may include additional NGS capture probes for known intronic variants or design orthogonal strategies for identification of known difficult-to-detect variants for which NGS may be suboptimal. Moreover, TGP gene content varies widely and continues to evolve. Currently reported diagnostic yields of even large GID-related TGPs rarely exceed 30%, though may be higher in selected cohorts or those enriched for endogamy.[23,24]

TGPs remain useful first-line tests for presentations with a limited genetic differential or requiring faster turn-around times than routine WES/WGS. Other advantages over broader testing include the possibility of fewer variants of uncertain significance (VUSes) or incidental findings. However, one disadvantage of TGPs in the context of rapidly expanding genetic landscapes is the need for frequent content updates to accommodate new gene-disease relationships. At a certain point, this becomes unwieldy and as challenging to obtain insurance authorization for as broader testing.

Whole exome and genome sequencing. Broad NGS is increasingly used up front to try to shorten diagnostic odysseys for patients with suspected GIDs, especially if poorly differentiated, multisystem involvement and/or multiple facets of immune pathology are present. First-line WES/WGS has been shown to improve diagnostic yields and reduce costs over other approaches such as TGPs, particularly if trios are available. However, because of heterogeneity in patient cohorts and selection criteria, published diagnostic yields range from 10% to 70%, with larger cohorts generally reporting lower yields than smaller specific cohorts.[25,26] Moreover, socioeconomic factors may constrain the availability of such testing for all patients.

WES typically encompasses the bulk of coding exons with variable lengths of flanking noncoding sequence in the nuclear genome, while WGS uses unenriched and unselected sequence. Both WES and WGS have proven helpful for identifying novel GID-causing variation and/or gene-disease relationships. However, WGS does not require additional steps that may introduce bias, guarantees more uniformly distributed

genome-wide coverage for CNV and single nucleotide variant (SNV) detection, and can also capture disease-relevant noncoding variation. Thus, WGS may be preferable to WES if the need for periodic reanalysis is anticipated, though periodic WES reanalysis can also increase diagnostic yields without additional sequencing.[27,28]

Widespread clinical WGS use is currently limited by acceptable sample types, cost, and insurance reimbursement. Of note, genome-wide deep sequencing for detection of somatic GID lesions is currently clinically unavailable. Thus, in situations where potential for mosaicism is relevant, TGPs and WES may be preferred over WGS. Finally, it is important to remember that short-read NGS in general may be suboptimal for detecting variation in some GID genes, so other approaches would need to be considered.[29]

Overview of Next-Generation Sequencing Workflow

NGS raw data generation and processing

After clinical evaluation and pretest informed consent counseling, sample choice is one of the first important decisions a provider needs to make (**Fig. 3**). Buccal swab or saliva samples are often noninvasive and easy-to-access alternatives to blood when germline disease is suspected. However, buccal and saliva samples are also not free from leukocyte contamination in the context of allogeneic transplant or clonal malignancy, so skin biopsy for a keratinocyte and/or fibroblast sample may need to be considered.[30,31]

In a typical NGS workflow, the extracted genomic DNA undergoes fragmentation and size selection for DNA library preparation. Unlike WES, WGS does not include the additional intervening steps of PCR amplification and target enrichment using exon capture kits. Raw data are generated as a text-based FASTQ file, which stores raw nucleotide sequence information with corresponding quality scores.

Each read then needs to be mapped or aligned to a reference genome. Though most research sequencing currently uses the hg38 assembly (released in 2013) as reference, most commercial laboratories continue to align to the older hg19 reference genome, even though hg38 offers fewer gaps and errors. Some laboratories are just now beginning the process of lifting over to hg38 but remain concerned about the potential for information loss from their large proprietary variant databases. Lifting over

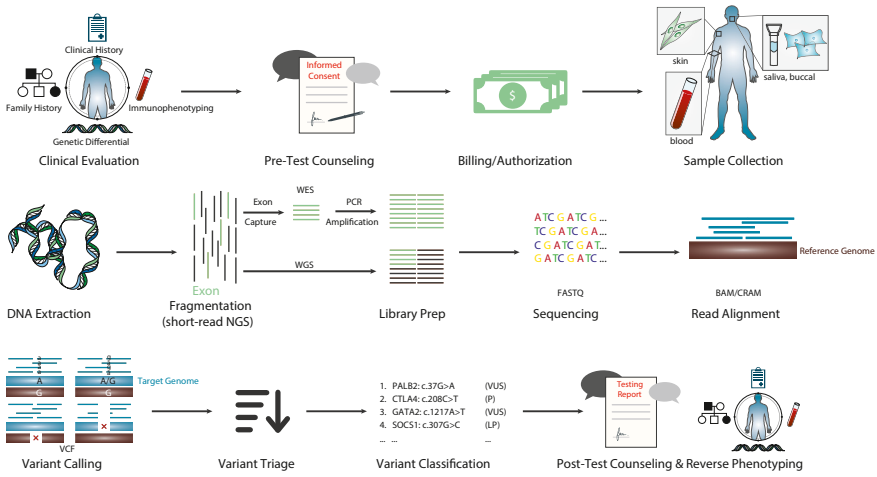

Fig. 3. A schematic depicting end-to-end clinical evaluation and diagnostic NGS workflow.

directly to the more diversely representative and complete telomere-to-telomere CHM13 reference sequence is another option, but its annotation is still ongoing.[32]

During this process, nonunique reads that map to multiple locations in a reference sequence are typically thrown out. This means that up to 50% of generated reads may not be used, even if containing potentially disease-relevant information. Mutations in regions featuring increased sequence homology may still be detectable by NGS, but the detection may be suboptimal.

Successful read alignments are stored in a text-based sequence Alignment/Mapping (SAM) file. Binary alignment map (BAM) and compressed reference-oriented alignment map (CRAM) files are both compressed versions of SAM files, encoded in different ways. BAM files use a lossless compression strategy that serializes all the data together, while CRAM enables greater compressibility by storing only the base calls that diverge from a designated reference sequence file, which is stored separately.

Then, a variant call file is generated listing all identified variation from the reference and associated AFs. Different variant calling algorithms may be needed to identify SNVs versus CNVs and SVs, or to distinguish germline versus somatic variants. In general, the Broad Institute-developed genome analysis toolkit (GATK) has been the industry standard for identifying germline SNPs and indels. Subsequent improvements on this foundation include the DRAGEN-GATK system developed in conjunction with Illumina, which provides greater speed and accuracy. Variant callers have been extensively benchmarked and reviewed so we refer you to that literature.[33,34]

Variant analysis and classification

Currently, the major bottleneck to clinical reporting times remains variant triage and classification. Of the approximately 3 million variants called by WGS or 20,000 by WES, 100s to 1000s may be clinically relevant, presenting an intrinsic signal-to-noise challenge. Laboratories continue to rely on a combination of computational tools and manual review to annotate, filter, and narrow down the variants ultimately chosen for clinical reporting.

The goal of clinical-grade genetic analysis is to identify any variants that may contribute to the clinical presentation of the patient, as reported by the ordering providers. *Laboratories will not know to look for variants related to unreported patient phenotypes, nor can they report variants in unknown or unproven gene-disease relationships.* Novel genotype–phenotype discovery and/or use of non-clinically proven tools for such discovery requires research-grade protocols and informed consent from the patient.

In 2015, the American College of Medical Genetics (ACMG) introduced a now widely used framework for assessing the potential impact of a variant on the function of its relevant gene (and often, protein product).[35] These criteria encompass various categories of data (eg, population or cohort genetics, experimental studies, in silico tools, model organisms) with graded strengths of evidence. At present, it is best suited for the assessment of small variants in protein-coding genes. It is also largely based on genetic assumptions about Mendelian disease derived from the disorders generally seen by clinical geneticists, so some criteria may be less relevant to GID paradigms. These criteria are then integrated into a standard terminology, with variants assigned a classification as Benign, Likely Benign, VUS, Likely Pathogenic (LP), or Pathogenic (P) to provide a “shared gestalt” about potential molecular impact.

Two analytical approaches are often employed in parallel—a phenotype-driven analysis focused on candidate genes of relevance to the patient’s clinical presentation, and a more agnostic filtering process looking for variants highly likely to be deleterious that

may play a role in the patient's phenotype. Rarely are all potential clinical presentations of a GID described in the literature, so a key principle is not to base phenotype-driven analysis on specific word-matching of clinical signs or symptoms (as is often done for dysmorphology syndromes), but to endeavor to match pathobiology instead—in other words, to determine whether the clinical presentation provided could be plausibly seen in the context of the known immunobiological mechanism(s). The availability of family member sequencing data may aid further variant prioritization, particularly if a specific inheritance pattern is suspected. However, given the widespread incomplete penetrance seen with GIDs, familial segregation data may not always be informative.

Results interpretation and follow-up

A clinical genetic testing report offers a wealth of information beyond the test results themselves, so it is often helpful to read the fine print (**Fig. 4**). Laboratories may have slightly different reporting policies, but most will report P/LP variants with a known

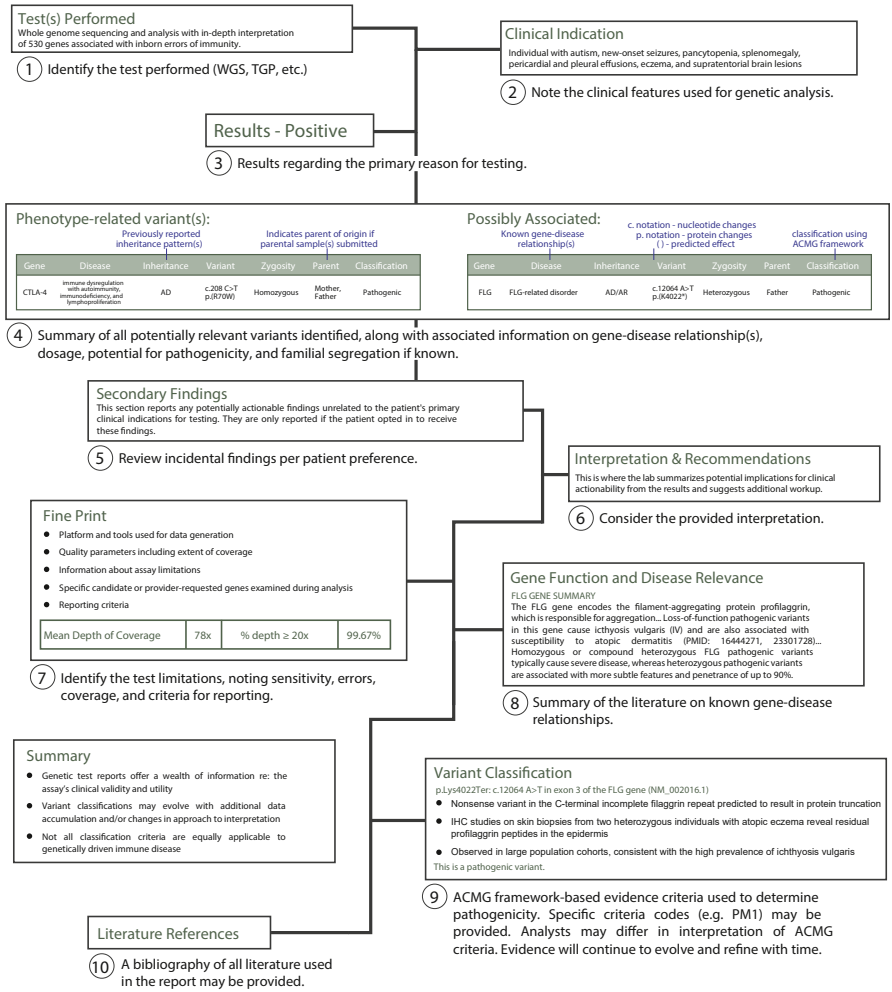

Fig. 4. Flow chart showing how to navigate the wealth of information provided in a clinical genetic testing report.

association to the patient's reported clinical presentation, but differ on whether or not to report heterozygous VUSes associated with only autosomal recessive conditions. During pretest counseling for WES/WGS, patients are usually asked whether they would like to know about any potential secondary or incidental findings identified in genes unrelated to the primary indication for testing. However, these may also differ across laboratories—with some choosing only to report the secondary findings officially recommended by the ACMG and others offering additional "perks" including pharmacogenetic data.

It is important to keep in mind that *variant classification does not take responsibility for determining a variant's relevance to a patient's specific clinical findings*—ultimately, this remains the onus of the patient's clinical providers. Thus, post-test counseling is important for helping the patient and their providers achieve a shared understanding of what the test results mean for the patient's own care as well as for family members. If a definitive diagnosis is made, reverse phenotyping may be needed to better characterize the patient's specific expression of disease for informing treatment. If a conclusive diagnosis is not made, additional phenotyping may still be needed to help establish the relevance of a potential VUS or to suggest specific candidate pathways to examine more deeply.

Some variants classified as P/LP may have little or no clinical impact, while others classified as a VUS may indeed be the patient's disease-causing lesion. In these latter situations, a VUS classification does not negate a patient's clinical diagnosis, nor should it hinder management guided by the patient's clinical and immune phenotype. However, there may be additional strategies a clinician can use to help upgrade the VUS into P/LP territory.

VUS resolution remains a major bottleneck for the achievement of definitive molecular diagnoses. The term itself implies anywhere between 10% and 90% likelihood of pathogenicity based on available data. Many laboratories offer free VUS resolution testing for additional first-degree family members to generate segregation data if it can support a variant's pathogenicity, but this may be less helpful for GIDs as previously noted.

Ideally, one would be able to identify a research group who can functionally assess the impact of the variant in an *in vitro* context, independent of the patient's other potential genetic contributions. Being able to demonstrate a qualitative and/or quantitative impact on RNA and/or protein production, as well as on downstream cellular consequences, helps to fulfill the functional classification criteria often most difficult to supply. Being able to provide complementation rescue data for loss-of-function phenotypes is often even more satisfactory. In recent years, there has been a collaborative effort to develop multiplexed assays of variant effect (MAVEs) in an attempt to rapidly and systematically augment the body of functional evidence available to support variant classification for diverse genes.[36] However, this is still very much a work in progress, with very few MAVEs currently validated for immune genes. Moreover, GID genes often play diverse roles, so a single variant may exert diverse downstream outcomes that arise with different thresholds of dysfunction, or only be elicited in specific tissues or under specific conditions.[37–40] These nuances may not be sufficiently reflected in one gene-specific assay.

However, even for genes where these assays may not be available, we are lucky in the GID field to have access to many clinical-grade cellular immune studies that may help provide additional supportive information about the impact of a lesion on a specific protein, or on a pathway or cell type for which the gene may be relevant.[41] Whether this patient-derived data can only be considered additional phenotypic evidence or can also serve as functional molecular evidence remains subject

to interpretation and may need to be decided on a gene-by-gene basis by ClinGen groups.

Future Perspectives

Even with extensive genetic testing, more than half of patients with GID currently fail to receive a molecular diagnosis from clinical genetic testing. This reflects limitations in both our technology and knowledge. Standard SRS platforms and computational algorithms may be insufficient for detecting disease-relevant somatic, noncoding, or structural variation. Furthermore, the genetic landscape of GIDs remains incompletely understood, as evidenced by the ever-expanding IUIS nosology. In addition, not all strong genetic contributions are necessarily monogenic, and we have barely begun to dissect how more complex combinations of variants may interact to cause immune disease. Thus, iterative genetic analysis coupled with ongoing phenotypic refinement to identify the relevant pathways and processes driving a patient's disease are essential for maximizing diagnostic yields with the current tools at our disposal. As academic laboratories continue to develop valuable immune studies to improve our ability to interrogate perturbations in diverse immune pathways, genomics technologies also continue to evolve in parallel.

Unlike SRS, OGM and LRS perform true single-molecule detection on native DNA molecules, eliminating potential for capture or amplification bias and circumventing the read-mapping concerns that plague loci such as *IKBKG* and *NCF1*.[42] OGM is emerging as an efficient and cost-effective approach for detecting genome-wide lesions above sizes of approximately 300 to 500 bp that elude standard cytogenetic techniques, but it lacks the resolution for pinpointing exact sequence context.[43] LRS provides such resolution in addition to phasing and DNA methylation information but remains associated with less accurate SNV detection and much higher computing and storage costs.[44] However, both technologies and their associated bioinformatic tools have matured to such an extent that they are now increasingly used for research-based diagnosis of unsolved patients and entering clinical development.[22,45]

SUMMARY

Twenty-one years ago, the first draft of the human genome was published. This effort took 13 years, 20 sequencing centers, and thousands of sequencing instruments to complete. Around this time, our field was aware of around 100 GIDs. However, with the widespread application of NGS to novel gene-disease discovery since then, we now know of over 600. In the past 5 to 10 years, we have also begun to routinely order clinical genetic testing on our patients to look for changes at the molecular level that help inform their care. All this progress is worth marveling at.

However, despite exponential growth in our understanding of GIDs over the past few decades, much more remains to be learned. With the combination of new functional studies, -omics modalities, and expanding pharmacologic space, the achievement of molecular diagnosis is poised to have greater impact than ever on clinical management targeted to each patient's specific pathophysiology. As such, it remains imperative for clinicians caring for these patients to understand how human genetics may contribute to clinical immune disease and how to use the assays available for interrogating this contribution. Finally, as our field continues to make advances in the practice of molecular medicine, we must also ensure that we do not inadvertently exacerbate already existing health care disparities. Unless we ensure that there is greater inclusiveness and equality of representation in our databases, unless we can improve access to

clinical services and counseling resources for underserved communities, we have not begun to make real progress.

CLINICS CARE POINTS

- No single assay, molecular or otherwise, is sufficient to address the diverse pathobiology of GIDs.
  - A multimodal diagnostic approach with synthesis across different data types is essential.
- Evaluation is often iterative, even with a positive molecular diagnosis.
  - Genetic contributions are only one piece of information about an individual's expression of disease but can inform the collection of additionally valuable information.
- GID paradigms may differ from those of other Mendelian disease landscapes.
  - This is important to be aware of when developing testing and analysis strategies or applying genetic evidence frameworks for variant classification.
- Despite the expanding richness of genomic tools, the diagnostic yield of testing will only ever be as good as the information provided.
  - Clinical evaluation and genetic analysis should focus on identifying relevant underlying pathobiologies, rather than simply matching phenotypes.
- Variant identification, triage, and classification remain an ongoing bottleneck in the molecular diagnostic process.
  - Both patient cell-based and *in vitro* functional studies should be used to help inform VUS resolution.
- Laboratories classify variants in the context of gene–disease relationship knowledge, but providers are responsible for interpreting the relevance and actionability of a genetic test report in the context of their patient.
  - Providers have a responsibility to provide appropriate pre- and post-test counseling and to help minimize inequalities of care that may inadvertently arise from genomics advances.

DISCLOSURES

Mr M.H. Bannister received funding from NIH training grant T32 GM136577. Dr X.P. Peng is a consultant for Genesis Therapeutics, MedZown, and Pharming.

REFERENCES

1. Bousfiha A, Moundir A, Tangye SG, et al. The 2022 update of IUIS phenotypical classification for human inborn errors of immunity. J Clin Immunol 2022;42(7):1508–20.
2. Chen L, Mamutova A, Kozlova A, et al. Comparison of disease phenotypes and mechanistic insight on causal variants in patients with DADA2. J Allergy Clin Immunol 2023;152(3):771–82.
3. Wouters M, Ehlers L, Van Eynde W, et al. Dominant negative *ADA2* mutations cause ADA2 deficiency in heterozygous carriers. medRxiv [Preprint] 2024. 2024.12.09.24317629.
4. Akalu YT, Bogunovic D. Inborn errors of immunity: an expanding universe of disease and genetic architecture. Nat Rev Genet 2024;25(3):184–95.
5. Thorpe J, Osei-Owusu IA, Avigdor BE, et al. Mosaicism in human health and disease. Annu Rev Genet 2020;54(1):487–510.
6. Lee M, Lui ACY, Chan JCK, et al. Revealing parental mosaicism: the hidden answer to the recurrence of apparent de novo variants. Hum Genom 2023;17(1):91.

7. Sahoo SS, Pastor VB, Goodings C, et al. Clinical evolution, genetic landscape and trajectories of clonal hematopoiesis in SAMD9/SAMD9L syndromes. Nat Med 2021;27(10):1806–17.
8. Bleesing JJ, Nagaraj CB, Zhang K. Autoimmune lymphoproliferative syndrome. In: Adam MP, Feldman J, Mirzaa GM, et al, editors. GeneReviews®. Seattle (WA): University of Washington; 1993. Available at: http://www.ncbi.nlm.nih.gov/books/NBK1108/. Accessed August 22, 2024.
9. Beck DB, Ferrada MA, Sikora KA, et al. Somatic mutations in UBA1 and severe adult-onset autoinflammatory disease. N Engl J Med 2020;383(27):2628–38.
10. Aluri J, Cooper MA. Somatic mosaicism in inborn errors of immunity: current knowledge, challenges, and future perspectives. Semin Immunol 2023;67:101761.
11. Kernohan KD, Boycott KM. The expanding diagnostic toolbox for rare genetic diseases. Nat Rev Genet 2024;25(6):401–15.
12. Deininger MW, Hodgson JG, Shah NP, et al. Compound mutations in BCR-ABL1 are not major drivers of primary or secondary resistance to ponatinib in CP-CML patients. Blood 2016;127(6):703–12.
13. Rohlin A, Wernersson J, Engwall Y, et al. Parallel sequencing used in detection of mosaic mutations: comparison with four diagnostic DNA screening techniques. Hum Mutat 2009;30(6):1012–20.
14. Chiang SCC, Vergamini SM, Husami A, et al. Screening for Wiskott-Aldrich syndrome by flow cytometry. J Allergy Clin Immunol 2018;142(1):333–5.e8.
15. Aksentijevich I, Sampaio Moura N, Barron K. Adenosine deaminase 2 deficiency. In: Adam MP, Feldman J, Mirzaa GM, et al, editors. GeneReviews®. University of Washington, Seattle; 1993. Available at: http://www.ncbi.nlm.nih.gov/books/NBK544951/. Accessed August 22, 2024.
16. Abarrategui-Garrido C, Martínez-Barricarte R, López-Trascasa M, et al. Characterization of complement factor H–related (CFHR) proteins in plasma reveals novel genetic variations of CFHR1 associated with atypical hemolytic uremic syndrome. Blood 2009;114(19):4261–71.
17. Roos D, de Boer M, Köker MY, et al. Chronic granulomatous disease caused by mutations other than the common GT deletion in NCF1, the gene encoding the p47phox component of the phagocyte NADPH oxidase. Hum Mutat 2006;27(12):1218–29.
18. Zhang K, Astigarraga I, Bryceson Y, et al. Familial hemophagocytic lymphohistiocytosis. In: Adam MP, Feldman J, Mirzaa GM, et al, editors. GeneReviews®. Seattle (WA): University of Washington; 1993. Available at: http://www.ncbi.nlm.nih.gov/books/NBK1444/. Accessed August 22, 2024.
19. Konkle BA, Nakaya Fletcher S. Hemophilia A. In: Adam MP, Feldman J, Mirzaa GM, et al, editors. GeneReviews®. Seattle (WA): University of Washington; 1993. Available at: http://www.ncbi.nlm.nih.gov/books/NBK1404/. Accessed August 22, 2024.
20. Wright JT, Grange DK, Fete M. Hypohidrotic ectodermal dysplasia. In: Adam MP, Feldman J, Mirzaa GM, et al, editors. GeneReviews®. Seattle (WA): University of Washington; 1993. Available at: http://www.ncbi.nlm.nih.gov/books/NBK1112/. Accessed August 22, 2024.
21. Lenahan AL, Squire AE, Miller DE. Panels, exomes, genomes, and more—finding the best path through the diagnostic odyssey. Pediatr Clin North Am 2023;70(5):905–16.
22. Mastrorosa FK, Miller DE, Eichler EE. Applications of long-read sequencing to Mendelian genetics. Genome Med 2023;15(1):42.
23. Peng X, Kaviany S. Approach to diagnosing inborn errors of immunity. Rheum Dis Clin N Am 2023;49(4):731–9.

24. Peng XP, Al-Ddafari MS, Caballero-Oteyza A, et al. Next generation sequencing (NGS)-based approach to diagnosing Algerian patients with suspected inborn errors of immunity (IEIs). Clin Immunol 2023;256:109758.
25. von Hardenberg S, Klefenz I, Steinemann D, et al. Current genetic diagnostics in inborn errors of immunity. Front Pediatr 2024;12. https://doi.org/10.3389/fped.2024.1279112.
26. Vorsteveld EE, Hoischen A, van der Made CI. Next-generation sequencing in the field of primary immunodeficiencies: current yield, challenges, and future perspectives. Clin Rev Allergy Immunol 2021;61(2):212–25.
27. Similuk MN, Yan J, Ghosh R, et al. Clinical exome sequencing of 1000 families with complex immune phenotypes: toward comprehensive genomic evaluations. J Allergy Clin Immunol 2022;150(4):947–54.
28. Vorsteveld EE, Van der Made CI, Smeekens SP, et al. Clinical exome sequencing data from patients with inborn errors of immunity: cohort level meta-analysis and the benefit of systematic reanalysis. Clin Immunol 2024. https://doi.org/10.1101/2024.06.14.24308832.
29. Lee K, Abraham RS. Next-generation sequencing for inborn errors of immunity. Hum Immunol 2021;82(11):871–82.
30. Theda C, Hwang SH, Czajko A, et al. Quantitation of the cellular content of saliva and buccal swab samples. Sci Rep 2018;8(1):6944.
31. Thiede C, Prange-Krex G, Freiberg-Richter J, et al. Buccal swabs but not mouthwash samples can be used to obtain pretransplant DNA fingerprints from recipients of allogeneic bone marrow transplants. Bone Marrow Transplant 2000;25(5):575–7.
32. Nurk S, Koren S, Rhie A, et al. The complete sequence of a human genome. Science 2022;376(6588):44–53.
33. Xu C. A review of somatic single nucleotide variant calling algorithms for next-generation sequencing data. Comput Struct Biotechnol J 2018;16:15–24.
34. Barbitoff YA, Abasov R, Tvorogova VE, et al. Systematic benchmark of state-of-the-art variant calling pipelines identifies major factors affecting accuracy of coding sequence variant discovery. BMC Genom 2022;23(1):155.
35. 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):405–24.
36. Arbesfeld JA, Da EY, Stevenson JS, et al. Mapping MAVE data for use in human genomics applications. bioRxiv 2023. https://doi.org/10.1101/2023.06.20.545702.
37. Caballero-Oteyza A, Crisponi L, Peng XP, et al. *OTULIN*-related conditions: report of a new case and review of the literature using GenIA. Clin Immunol 2024;265:110292.
38. Dabbah-Krancher G, Snow AL. Mistuned NF-κB signaling in lymphocytes: lessons from relevant inborn errors of immunity. Clin Exp Immunol 2023;212(2):117–28.
39. Baysac K, Sun G, Nakano H, et al. PLCG2-associated immune dysregulation (PLAID) comprises broad and distinct clinical presentations related to functional classes of genetic variants. J Allergy Clin Immunol 2024;153(1):230–42.
40. Cosson C, Riou R, Patoli D, et al. Functional diversity of NLRP3 gain-of-function mutants associated with CAPS autoinflammation. J Exp Med 2024;221(5):e20231200.
41. Abraham RS. How to evaluate for immunodeficiency in patients with autoimmune cytopenias: laboratory evaluation for the diagnosis of inborn errors of immunity associated with immune dysregulation. Hematology 2020;2020(1):661–72.

42. Conlin LK, Aref-Eshghi E, McEldrew DA, et al. Long-read sequencing for molecular diagnostics in constitutional genetic disorders. Hum Mutat 2022;43(11):1531–44.
43. Smith AC, Hoischen A, Raca G. Cytogenetics is a science, not a technique! Why optical genome mapping is so important to clinical genetic laboratories. Cancers 2023;15(22):5470.
44. Hook PW, Timp W. Beyond assembly: the increasing flexibility of single-molecule sequencing technology. Nat Rev Genet 2023;24(9):627–41.
45. Barseghyan H, Eisenreich D, Lindt E, et al. Optical genome mapping as a potential routine clinical diagnostic method. Genes 2024;15(3):342.