Exploring the RNA Gap for Improving Diagnostic Yield in Primary Immunodeficiencies

Jed J. Lye[1], Anthony Williams[1,2]* and Diana Baralle[1,3]*

[1] University of Southampton Medical School, University of Southampton, Southampton, United Kingdom, [2] Wessex Investigational Sciences Hub Laboratory (WISH Lab), Faculty of Medicine, University of Southampton, Southampton, United Kingdom, [3] Faculty of Medicine, Highfield Campus, University of Southampton, Southampton, United Kingdom

Challenges in diagnosing primary immunodeficiency are numerous and diverse, with current whole-exome and whole-genome sequencing approaches only able to reach a molecular diagnosis in 25–60% of cases. We assess these problems and discuss how RNA-focused analysis has expanded and improved in recent years and may now be utilized to gain an unparalleled insight into cellular immunology. We review how investigation into RNA biology can give information regarding the differential expression, monoallelic expression, and alternative splicing—which have important roles in immune regulation and function. We show how this information can inform bioinformatic analysis pipelines and aid in the variant filtering process, expediting the identification of causal variants—especially those affecting splicing—and enhance overall diagnostic ability. We also demonstrate the challenges, which remain in the design of this type of investigation, regarding technological limitation and biological considerations and suggest potential directions for the clinical applications.

Keywords: primary immunodeficiency disorders, clinical diagnostics, RNASeq, RNA, RNAseq analysis

INTRODUCTION

Primary immunodeficiency disorders (PID) result from altered, poor, or absent function in one or more components of the immune system, rendering the affected individuals with increased susceptibility to immune-related ailments including increased frequency and severity of infection, autoimmunity, aberrant inflammation, and malignancy (McCusker et al., 2018). The understanding of the genetic heterogeneity of PID has expanded greatly over the last decade, now encompassing a list of over 350 distinct disorders arising from at least 344 gene defects, demonstrative of the complexity of the immune system (Bousfiha et al., 2018; Picard et al., 2018). This plethora of genetic causes has brought about a need to categorize the disorders for expedited diagnosis and treatment protocols. Some broader methods simply classify the disorders into groups of innate and adaptive immunity linked to the clinical phenotype (McCusker and Warrington, 2011). The Inborn Errors of Immunity Committee (previously the International Union of Immunological Societies PID Expert Committee) has now devised a precise and useful system, which classifies disorders by the immunological pathway affected. In addition, it now has corresponding phenotypical classification systems for clinicians at the bedside to help identify the disorders. These briefly comprise nine categories: immunodeficiencies affecting cellular and humoral immunity, combined immunodeficiency disorder (CID) with associated or syndromic features, predominantly antibody deficiencies, diseases of immune dysregulation, congenital defects of phagocyte, defects in intrinsic

and innate immunity, auto-inflammatory disorders, complement deficiencies, and phenocopies of PID (Bousfiha et al., 2018).

The most common form of PID is selective immunoglobulin A deficiency, which is usually typically asymptomatic but can manifest with a variety of clinical presentations including coeliac disease, type 1 diabetes mellitus, and increased infections. With an estimated prevalence of 1 in 300–500 persons (Boyle and Buckley, 2007), and while individually rare (>1 in 2,000), the remaining disorders considered in the wider scope of PID together represent a significant burden on the health and economy of a nation. Current diagnostic levels suggest an incidence of 5.90/100,000 (Shillitoe et al., 2018); however, underdiagnsosis of PID may mean the true incidence is as high as 1:250 (Europe PIPDDfOCi.).

The importance of early diagnosis In PID cases is high, with relation to both the patient's qualitative experience and the economic cost to healthcare services. Sources vary in cost analysis of undiagnosed PID. Some say that while a diagnosed US patient costs healthcare services over US$250,000 per annum, largely due to treatment costs, an early diagnosis of the disorder can save as much as US$6,500 per patient, per annum (Abolhassani et al., 2015). An alternate source suggests an undiagnosed patient might cost the healthcare system US$102,552 annually. Once diagnosed, these costs may drop by as much as US$79,942 (Condino-Neto and Espinosa-Rosales, 2018). In a patient survey, 45% of patients reported a diagnostic wait time of between 1 and 6 years; around 1/6th reported waiting 10–20 years. Other key findings of the same survey confirmed undiagnosed patients bring about a dramatically increased burden on National Health Service (NHS) resources (UK,). Identification of the precise molecular origins for each patient's case of PID leads to improved patient care (Walter et al., 2016) and improved prognosis. The importance of correct genetic cause for a PID phenotype is demonstrated by the different treatment preferences which exist for conditions which may present with similar clinical phenotypes (Heimall et al., 2012). Precision therapeutic diagnostics can help to achieve this in part, by allowing targeted intervention to the specific molecular causes (Bonilla et al., 2015; Lenardo et al., 2016; Ramakrishnan et al., 2016).

Diagnostic Challenges in PID

Challenges in diagnosing primary immunodeficiency are numerous and diverse. Studies which correlate the phenotype and genotype have been useful in diagnostics, developing an understanding of various PID disorders (Fischer, 1993). Additionally, these correlation studies have been useful for deconvoluting the pleiotropic nature of the involved genes, through which a single variant can bring about a variety of clinical phenotypes (Meyts et al., 2016). However, the development of a universal diagnostic pipeline for PID is hindered by the heterogeneity in the presentation of disease, even among patients with what appears to be the same pathogenic genetic variant (Richardson et al., 2018). Conversely, a number of genotypes can bring about even the most well-characterized phenotype (Meyts et al., 2016). Once a clinical diagnosis of PID is suspected, mainly based upon a compatible phenotype, a family history

is usually taken and a number of subsequent laboratory tests performed to confirm the type of immune mechanism affected (Hernandez-Trujillo and Ballow, 2015). With the emergence of targeted sequencing of larger PID gene panels, clinical exomes, and complete exomes through short-read next-generation sequencing technologies, the inclusion of genetic testing within a PID diagnostic workup has become more widespread. This approach to both adult- and paediatric-onset disease has consolidated the importance of protein-based functional immune testing (cytokines, antibodies, etc.) for characterizing the nature of the phenotypic presentation, but furthermore to evaluate candidate genetic variants in such pathways that have been identified through parallel germline DNA testing.

DNA sequencing-based genetic testing is used where possible, as it provides the best diagnostic capability of existing clinically adapted methods (Condino-Neto and Espinosa-Rosales, 2018). Whole-exome sequencing (WES) has the highest success rate of the clinically adapted diagnostic methods (Gilissen et al., 2011; Boycott et al., 2017), which it achieves despite the exome comprising only ~2% of the human genome (Bamshad et al., 2011).This is in part due to 85% of currently annotated variants existing within the transcribed portion of the genome (Majewski et al., 2011). It has been hypothesized that this focus had likely led to the underestimation of the contribution to disease of non-coding variants (Kremer et al., 2018).

Due to the improvement that WES and whole-genome sequencing (WGS) bring to diagnostics, researchers are calling for universal molecular gene testing for the diagnosis of primary immune deficiencies (Heimall, 2019). Evidence from existing literature, however, suggests that even this may be inadequate; currently, WES and WGS are only able to produce reliable diagnosis in 25–60% of cases (Yang et al., 2013; Moens et al., 2014; Yang et al., 2014; Taylor et al., 2015; Meyts et al., 2016; Stray-Pedersen et al., 2017). Although many countries have undertaken whole-genome sequencing projects to evaluate this approach (Philippidis, 2018), the development of WGS as a clinically validated routine testing modality is still in its infancy. Within the UK's 100,000 Genomes Project, PID was accepted as an indication for inclusion, and plans to incorporate WGS for PID into routine clinical pathways have been approved following the transition phase of 100K Project to WGS sequencing in routine NHS care across England.

Formal confirmed genetic diagnosis of PID relies heavily on existing knowledge pertaining to consequences of the variants in the genes of relevance to the presenting phenotype and assumed mechanism of disease resulting from such variants in a dominant or recessive manner genomic sequence (Rae et al., 2018). The key to this task is the ability of bioinformatics tools to predict the significance of such variants. WES delivers around 20,000–23,000 variants per individual, and WGS produces 3–5 million per individual (Kremer et al., 2018), which makes the task of identifying a Mendelian disease variant vanishingly unlikely without a series of bioinformatics filters. Problems with the WGS/WES sequencing diagnostic methods arise when no variant, identified through a patient's genome sequencing, can be reliably linked to the clinical presentation and cytological/molecular manifestation of the disorder. Failure to identify a

definitive molecular cause occurs in about 70–75% of Mendelian conditions, according to a 2018 meta-analysis (Schwarze et al., 2018), mirrored by examples from PID (Gallo et al., 2016). The types of variants which are not always identified by current next-generation sequencing (NGS) approaches include exonic variants of unknown significance, variants in intronic and intergenic non-coding DNA (Scacheri and Scacheri, 2015), variants in the *cis*-acting regulatory elements of transcription (Bryois et al., 2014) imprinting disorders, and repeat expansions (Kremer et al., 2018).

Conventional clinical diagnostics, utilizing human phenotype ontology for integration of cases into specific diagnostic groups, and traditional genetic sequencing methods for diagnostics are still currently inadequate. While proteomic diagnostic methods are in development, they exist at a relatively early stage of development and can miss the potentially valuable RNA regulatory phenomena.

Variants Affecting Differential Expression

Identification of definitive disease-causing mutations is confounded in some cases by expression levels being modulated by variants occurring in non-coding segments and those hiding in plain sight in genes not currently understood to be linked to the disease or phenotype. Often, these can be lost during the filtering process because of a lack of integrative understanding or supporting evidence (Thormann et al., 2019).

These expression quantitative trait loci (eQTLs) elicit a powerful, sometimes synergistic effect on the expression of a large number of genes. Single-nucleotide polymorphisms (SNPs) on eQTLs affect the transcriptional level of other RNAs, modifying protein expression and causing phenotypic changes to the abilities and behaviors of cells in some immunological cases (Fairfax et al., 2014). These eQTLs individually explain a fraction of the genetic expression of specific genes. The vast majority do not exist in the coding regions of genes and are predicted to be involved in gene regulation (Casamassimi et al., 2017). It is now understood that these eQTLs have a more pronounced effect on immune regulation than the effects of age and sex, and more interestingly exclusive effects only observable during immune stimulation have been identified for some of these eQTL variants (Piasecka et al., 2018). Epigenomic studies have helped to highlight the *cis*-regulatory nature of some non-coding regions of the genome. These suggest that the enrichment of disease-risk variants in cell-specific regulatory sequences is indicative of their cell type and contextual effects (Roadmap Epigenomics et al., 2015). Large-scale investigation into the association between genetic variants and expression of genes in a tissue-specific manner (including whole blood) was carried out by the Genotype—Tissue Expression Consortium (The Genotype-Tissue (GTEX), 2015). This research did not extend to immune tissues specifically, although links between immune cell-specific gene expression levels and eQTLs have been investigated by the DICE Project (Database of Immune Cell Expression, Expression Quantitative Trait Loci and Epigenomics) (Schmiedel et al., 2018). The researchers on this project were able to positively identify a range of *cis*-eQTLs for 12,254 genes, demonstrative of the high abundance of these sites. Interestingly, many of these eQTLs had

effects which were cell type-specific. The identification of these sites and interrogation for the existence of variants will likely play a crucial role in explaining the changes in expression of key genes which lead to PID.

The Role of RNA Splicing in the Immune System and PID

Alternative splicing is the method through which the cell can produce an array of transcript isoforms derived from a single gene or multiple genes spliced together (Ward and Cooper, 2010). Introns are spliced out and exons are either ligated through transesterification reaction or, in many cases, spliced out in different combinations, leaving the remaining exons to form a mature mRNA (Ward and Cooper, 2010).

Deep surveying on alternative splicing has shown that 95% of genes which contain multiple exons undergo alternative splicing, and even when only considering moderate to high abundance events, there are reportedly 100,000 individual splicing events in major tissues (Pan et al., 2008).

Alternative splicing occurs both co-transcriptionally and post-transcriptionally, and the action of transcription factors as well as splicing factors regulates and influences splicing events in some of the most crucial mechanisms of the adaptive immune system (Heyd et al., 2006; Orvain et al., 2008; Alkhatib et al., 2012). Important examples include RNA-polymerase II as a facilitator of splicing factor recruitment (Bentley, 2014), the alternative splicing of CD45 which is necessary for the production of a range of tyrosine phosphatases, imperative for the diverse set of lineage and stage-specific receptor signal transduction thresholds in immune tissues (Zikherman and Weiss, 2008), and FOXO1-induced Ikaros splicing, essential for the recombination of immunoglobulin genes. FOXO1 is a transcription factor which, through its effects on alternative splicing, allows the immune system to produce its diverse range of antibodies/immunoglobulins (Reynaud et al., 2008).

Activation of lymphocytes is a key component of the adaptive immune response to pathogens (Bonilla and Oettgen, 2010). Part of the central activation of these cells is the degradation of IκBα and release of NF-κB, which translocates to the nucleus to initiate maturation and activation of the cell. The "CBM" complex, which brings about the degradation of IκBα, is formed by *CARMA1, BCL10*, and *MALT1* (Oeckinghaus et al., 2007). *MALT1*, a crucial component of this complex, undergoes alternative splicing of EXON 7 to produce mRNA isoforms with a differential function. The activation strength of CD4[+] T cells is mediated by the relative abundance of the alternatively spliced isoforms of MALT1, which is in part controlled by the molarity of phosphorylated splicing factor hnRNPU in the nucleus (Meininger et al., 2016). Alternative splicing, then, is a key component of the normally functioning immune system, and perturbations in canonical function can likely lead to pathology.

Variants Affecting Alternative Splicing

The impact of mutations that affect RNA processing/splicing is currently providing a diagnostic revolution. Variants which

affect splicing either occur in active splice sites, those which occur in regulatory elements, and those which occur in intronic or intergenic regions (Grodecká et al., 2017) see **Figure 1**.

Studies comparing variants affecting splicing in PID have determined that the variants which directly influence splice sites are more robustly linked to disease phenotypes than those which effect splicing regulatory elements (Grodecká et al., 2017). *Cis*-mutations in the genome can affect splicing though altering the splice site recognition or altering exon splicing enhancer or silencer sites (Ward and Cooper, 2010). Splice sites usually comprise GT and AG dinucleotides at 5′ and 3′ sites, respectively. If a variant changes this sequence, or causes another one to appear, it can affect the ability of the splicing machinery to detect the canonical splice site (Krawczak et al., 1992; Krawczak et al., 2007). Additionally, mutations in *trans*-acting splice factors—the splicing machinery of the cell—can also bring about disease by preventing these factors from performing their function of generating the required isoforms (Ward and Cooper, 2010), although these are not covered in this review.

Due to the impact of these findings, interest in the detection of splice-altering variants and activated cryptic splice site has spurred on the development of a number of *in silico* tools for prediction of splice site usage (Grodecká et al., 2017; Ohno et al., 2018). Unfortunately, these tools are often unable to discern the resulting transcripts exon use patterns (Jaganathan et al., 2019), and while their predictive ability can be enhanced by other orthogonal investigations such as mini-gene assays (Grodecká et al., 2017), the multiple facets of splicing control involve more than just the sequence of the splice site in question, as evidenced by the temporal and spatial differences in splicing patterns. Briefly, these include the activation of other splice sites within the gene, splicing quantitative trait loci, the relative abundance, phosphorylation status, and localization of different and often competing *trans*-acting factors (Wang et al., 2015).

Further complicating this process, seemingly benign, synonymous exonic variants can disrupt splicing to cause disease. Using RNA sequencing (RNASeq) to complement genomic sequencing, Cummings et al. evidenced this in the *POMGNT1* and *RYR1* genes, finding variants which were demonstrated to be causative of Mendelian diseases in muscle (Cummings et al., 2017). Part of the normal filtering process which many bioinformaticians adopt is to filter out synonymous variants very early on, but investigation using deep learning has led to the understanding that between 9% and 11% of rare genetic disorders are caused by synonymous or intronic splice-altering mutations (Jaganathan et al., 2019). Indeed, much as gene expression can be influenced by multiple loci, so too can multiple loci contribute to the occurrence of splicing events. These loci are appropriately termed splicing quantitative trait loci (sQTLs) (Jia et al., 2015). Analysis of sQTLs has been improved by RNASeq methodologies, but remains a difficult challenge as the isoform expression has to be estimated using statistical methods (Patro et al., 2017). These sQTLs are not necessarily in close proximity to the splice junction. Characterization of these sites in humans has shown

SNPs demonstrating tangible sQTL activity at 100 kb from the relative splice site (Takata et al., 2017).

Non-protein-coding genes are a significant source of disease-causing variation (Scacheri and Scacheri, 2015). Examples within the PID research and diagnosis space include a recently discovered variant occurring in coding regions for genes comprising RNA components of the minor spliceosome, which is used for the splicing of at least one exon in ~800 genes (Turunen et al., 2013). Specifically, the non-coding gene *RNU4ATAC* that produces a small nuclear RNA (snRNA) termed U4atac was discovered to cause Roifman syndrome (Merico et al., 2015; Heremans et al., 2018) by preventing canonical minor intron splicing. Compound heterozygous variants were first discovered in an affected family after traditional filtering methods had not detected viable variants; the link was confirmed by the detection of intron retention during curated splicing analysis of RNASeq data (Merico et al., 2015).

The importance of alternative splicing in the immune system has been further demonstrated in mouse models. The ImmGen Project was set up specifically to investigate gene expression and regulation in mice using microarray profiling. It found that, in mice, around 60% of genes are expressed as multiple isoforms in T or B cells, and 70% of these had an impact on the lineage differentiation (Ergun et al., 2013). Compound heterozygous mutations *in MALT1*, mentioned earlier, which is heavily implicated in the activation of T cells, have been shown to bring about profound combined immunodeficiency. One of these variants was indeed a splice site acceptor change from the consensus AG to GG, identified by whole-exome sequencing (Punwani et al., 2015).

To further complicate the already complex nexus of control mechanisms contributing to PID, a range of epigenetic mechanisms leading to primary Immunodeficiencies have been observed and reviewed (Campos-Sanchez et al., 2019). In principle, the majority of genes identified to be susceptible to variants in PID may also be subject to heritable epigenetic modifications which could lead to the same or similar symptoms, acting as a further coefficient value when calculating the potential number of disorders, including those disorders affected by splicing, which can cause PID (Zhu et al., 2018).

RNA in Diagnostics

RNA investigation technology and its literature has experienced great leaps forward in recent years in terms of technological advancement and cost reduction (Muir et al., 2016). RNA sequencing is now largely replacing microarrays as the most used quantitative method of mapping gene expression profiles (Lowe et al., 2017). The transcriptome—or RNA expression profile—of a given tissue can give unparalleled insights into the elegant inner workings of the cell. Through capture of all internal RNA species, it characterizes the cellular gene transcription architecture and can deliver an instantaneous picture of environment–cell interaction or response program (Lowe et al., 2017; Wirka et al., 2018).

A range of technologies exist for conducting RNA sequencing, each with its own strengths and limitations.

Long-read sequencing provides reliable structural information, but can have suboptimal reliability in base calling (Feng et al., 2015) or is more expensive for high-throughput analysis (Rhoads and Au, 2015). Short-read NGS RNASeq involves sonication or enzymatic degradation of RNA into smaller fragments, selection of fragments using one of a number of methods, cDNA synthesis, the construction of a library, and subsequent sequencing followed by realignment (Kukurba and Montgomery, 2015). This technology has been the currently favored approach for high-throughput analysis.

Currently, this technology generates a mixture of both quantitative and qualitative analysis opportunities of RNA species: qualitative transcriptome profiling outcomes include identification of sequence variants at the level of the genome (Neums et al., 2017), somatic cell mosaics, non-canonical splice variants, occurring either due to *cis-* or *trans-*acting factor aberrations (Ward and Cooper, 2010). Quantitative outcomes of transcriptional profiling include differentially expressed genes, alternative splicing events, and allele-specific expression quantification (Kukurba and Montgomery, 2015). Previous studies have demonstrated that when compared with large control datasets, identification of expression outliers in peripheral whole blood can contribute to the detection of disease-causing variants (Zeng et al., 2015; Zhao et al., 2016). As well as gene expression levels, perturbations in the relative abundance of specific isoforms is a driving force in the genesis of many diseases (Chen et al., 2010; Kim et al., 2018) as isoforms can have differential function (Takeda et al., 2010), or, in some cases, can be antagonistic (Eshel et al., 2008). Through RNASeq or exon junction spanning probe-based capture, changes in isoform balance can also be resolved. The sensitivity suitability of RNASeq in transcriptomic investigation and splicing was demonstrated in mouse and human models and has enabled the discovery of ~7,600 novel isoforms in mouse immune cells (Ergun et al., 2013) and detected 100,000 splicing events with at least moderate abundance (Pan et al., 2008).

Transcriptome profiling can also give insights into control mechanisms exhibited by the non-coding RNA species, such as long non-coding RNA (lncRNA) and microRNA (miRNA), the significance of which is continually being elucidated in the molecular pathology of disease (Kramer et al., 2015; DiStefano, 2018). Indeed, such examples exist in PID; miR-6891-5p accumulation is demonstrated to contribute to selective IgA deficiency, the most common form of PID (Chitnis et al., 2017). Thanks to the increasing ability of technology and steady reduction in costs, we are also able to cast a winder net. Through RNASeq-based investigation, instead of concentrating on *a priori,* system-specific gene panels that many studies target, it is possible to examine all the mRNA species destined for translation. Through these hypothesis-free methods, it is possible to create profiles of normal transcription and disease transcription in a tissue-specific manner (Gonorazky et al., 2019). Subsequent use of follow-up analysis tools can be used to generate filtering process for causal variants or for biomarker identification (Han and Jiang, 2014). It is also possible to quantify the relative expression of those genes coding for the splice factors themselves, which can directly bring about pathological processes specific to PID, such as those observed in Roifman's syndrome, mentioned earlier (Merico et al., 2015; Heremans et al., 2018).

Microfluidic technology adaptations have allowed the development of robust, single-cell transcriptomic profiling (Kimmerling et al., 2016). In combination with NGS-based technologies, the single-cell technology provides a method for profiling the transcriptomes of individual cells, giving unparalleled insights into the heterogeneity of cell populations and their transcriptional profiles (Hwang et al., 2018). Adaptations such as the SMART-seq2 or fluidigm C1 library preparation methods also now allow the production of full-length cDNAs, giving transcript isoform-level resolution. However, these methods do not yet allow multiplexing, massively increasing overall costs and labor in large cohorts (See et al., 2018). The ability to profile the entire transcriptome of a peripheral blood mononuclear cell (PBMC) culture individually would give a dramatically increased ability to understand the cell–cell interactions taking place in an immune challenge, and this approach could be utilized in those patients suspected to be genetic mosaics.

DISCUSSION

The early and accurate diagnosis of primary immunodeficiencies is important to ensure the attainment of positive patient outcomes, through minimizing the time to diagnosis, identifying molecular pathways for targeted therapy, and reducing the economic cost of ill health or inappropriate treatment options. Diagnosis of the disorders remains difficult due to clinical challenges in identifying the presence of a primary immune system disorder, stratifying the phenotype to a myriad of overlapping candidate genes and then the laborious task of variant filtering, interpretation, and lack of knowledge pertaining to variants, especially those residing in the non-coding segments of the DNA. Functional validation of a candidate variant is currently undertaken with protein-based *ex vivo* tests, which are difficult to standardize and mostly available in research laboratories. RNA profiling to identify alternative splicing, gene expression-level variation monoallelic expression may contribute a further insight into candidate variants derived from proband or family-based WES/WGS sequencing results. We propose the introduction of RNASeq-based analysis for patients who have a clinical presentation of PID, but who, despite normal baseline immune testing, cellular analysis, and having undergone WES/WGS remain undiagnosed (see **Figure 2**).

RNASeq is an emerging technology which, when combined with WES/WGS, provides unprecedented insights into differential gene expression, splicing activity, and allelic specific expression and can inform regarding other phenomena such as genetic mosaicism. However, RNASeq remains relatively novel as a diagnostic testing tool in rare diseases and the control datasets and cellular contributions to complex tissue profiles (i.e., whole blood) will require further dissection.

Utilizing candidate gene lists and large control datasets for comparison enhances the power of the transcriptional
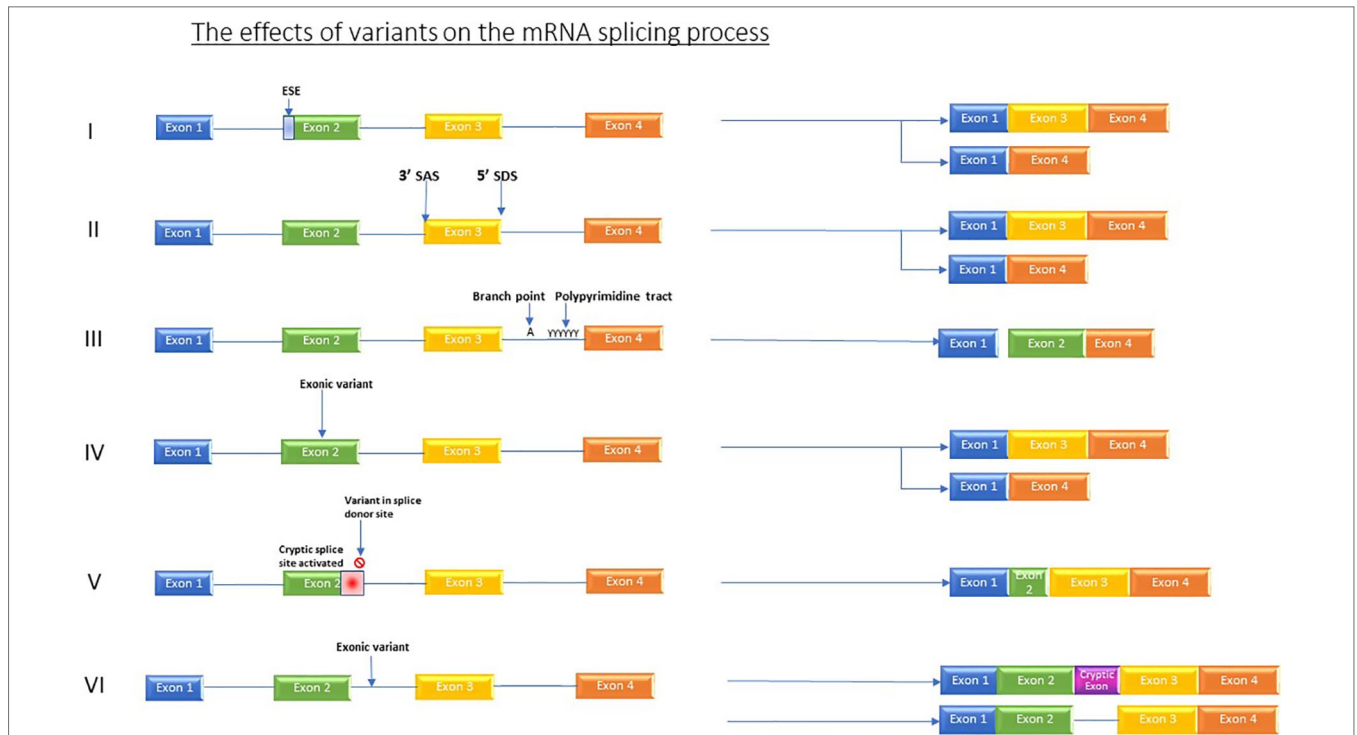
FIGURE 1 | Shows the various effects which variants can have on the splicing process. I) Variants in regulator elements such as exon splicing enhancers resulting in for example, exon skipping. II) Variant in splice acceptor OR splice donor site causes skipping of one or more exons. III) Branch point and Poly pyrimidine tract sequence variants causing exon skipping. IV) Exonic variants causing exon skipping. V) Variant in splice donor site induces activation of an alternative, cryptic splice donor site in exon. VI) Intronic variants producing new cryptic exon or retained intron.
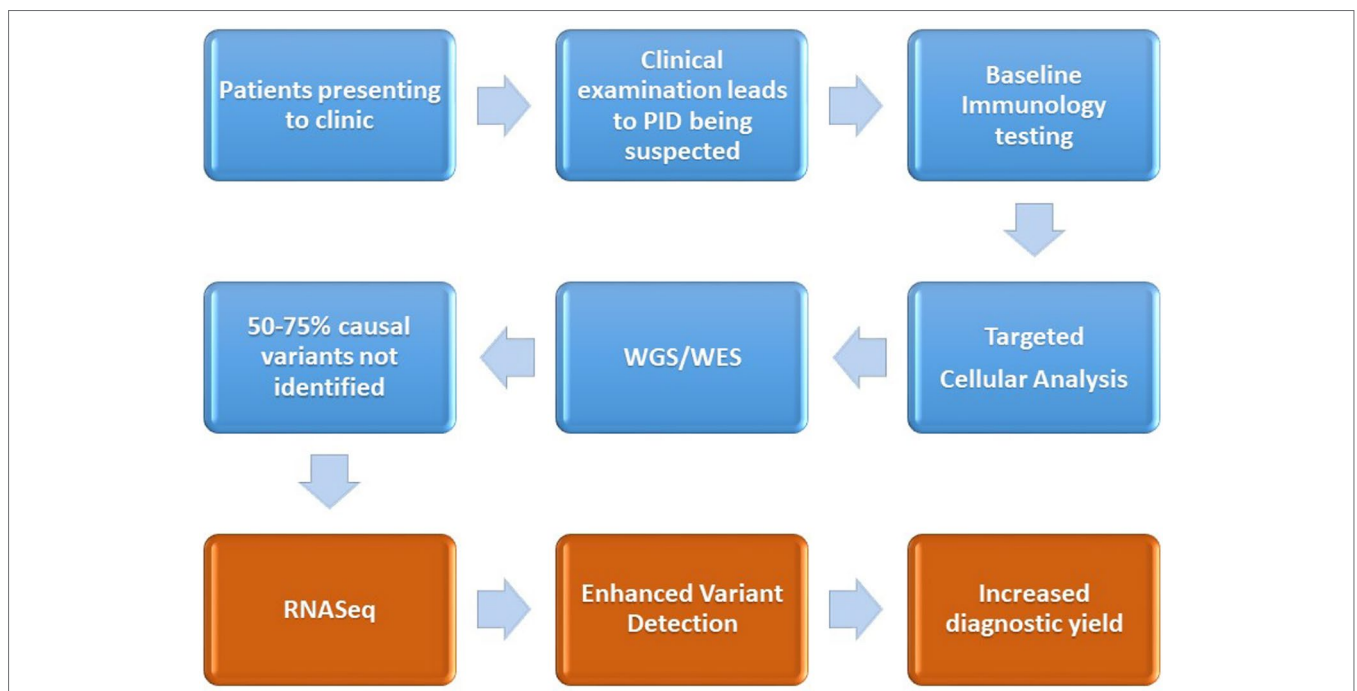
FIGURE 2 | Demonstration of the current diagnostic pathway (*blue*) for the majority of patients after a diagnosis of PID is suspected. We also outline the proposed intervention point of RNASeq and the associated enhanced variant detection (coming about through assessment of differential expression, changes to alternative splicing) and the increased diagnostic yield.

profiling through RNASeq and improves resolution for differential gene expression. Existing projects have developed these datasets for whole blood and immune cells, which provide a starting point for the interrogation of clinical samples for diagnostic research.

Immune responses to pathogenic challenges are exceptionally variable, and the variability in these responses is not easily elucidated. Environmental influences such as age, sex, seasonality, nutrition, and lifestyle all have effects on the specific response profile exhibited by individuals (Piasecka et al., 2018). These factors that influence responses can have a greater degree of significance in specific cell types. CD8[+] T cells, for example, show a high degree of heterogeneity in the context of temporal changes through the life course of the individual, and CD4[+] T cells and monocytes are heavily influenced by sex (Piasecka et al., 2018). It is therefore useful to be able to discern transcripts from different cell types within a culture. Utilizing flow cytometry to separate cell types or utilizing single-cell RNASeq is becoming an attractive option.

In order to assess the impact of genomic variation on the unstimulated immune system, the normal immune response and the immune-deficient responses, it is important to experimentally "tune out" the variations in signal arising from environmental factors. It has been established that a high degree of the cellular variation in CD8[+] cell populations can be attributed to environmental factors, which makes them a poor model for genetic variant impact. CD4[+] T cells display a large degree of heritability in these assays, and as such should provide a good level of transcriptomic heritability also. This will allow for clearer elucidation of the effects of variants on differential gene expression (Brodin et al., 2015).

The immune system's response to pathogen-based challenges is highly dynamic, and observing this response is more informative when identifying impaired response (Duffy et al., 2017). Indeed, it has been shown in innate immune system studies that the effects on differential expression of some variants can only be observed in a dynamic fashion (Fairfax et al., 2014; Lee et al., 2014). Co-culture of PBMCs provides a greater insight into activation pathways as it allows for the cell–cell communication response programs and produces similar results in terms of ranked gene expression response networks, with a few notable exceptions (Duffy et al., 2017). Studies of dynamic immune responses to challenges, in concert with machine learning, can be used to identify small groups of stimulation pathway-specific genes (Urrutia et al., 2016). Comparing the expression profiles of these genes in healthy cohorts with PID patients can potentially be utilized to identify candidate genes, which may then harbor a disease-causing variant or indicate some anomaly in the pathway for further investigation.

The transcriptomic landscape provides an excellent opportunity for advancement of diagnostic yield, and transcriptional profiling is already being utilized across a range of disorders to help build a "molecular fingerprint" of disease and better inform variant-filtering processes. The immunology community has made a case for PID diagnosis to be supported using transcriptional profiling using whole-transcriptome sequencing (Moens et al.,

2014), and these are being answered with examples in primary immunodeficiency cases such as Dock8 CID, GATA2 deficiency, and X-linked reticulate pigmentary disorder (XLPDR) (Hsu et al., 2013; Khan et al., 2016; Starokadomskyy et al., 2016). Over the coming years, an extended diagnostic approach to PID testing may develop that builds on a clinical module of phenotype, family history, and baseline immunological testing. This will be complemented by a DNA module of coding and non-coding variant analysis, utilizing sophisticated bioinformatic pipelines to prioritize candidate genetic variants of new loci that would be consistent with the clinical phenotype and family segregation. These candidate variants for monogenic disease may then be functionally interrogated *via* RNAseq for an influence within the gene itself and possibly the network within which it operates. In parallel, functional testing of candidate genes through protein-based assays may be undertaken to characterize the impact of a putative monogenic pathogenic variant within a reductionist model at the protein level. The sharing of these modular assessments across the international community will incrementally improve the standardized analysis of novel variants that will continue to grow over the next few years.

AUTHOR CONTRIBUTIONS

DB and AW devised and planned the project and co-supervised JL. JL wrote the manuscript. DB and AW reviewed, edited and contributed to the manuscript.

FUNDING

GLOSSARY OF TERMS

Term	Description
NGS (next-generation sequencing)	Next-generation sequencing describes the massively parallel, high-throughput modern evolution of sequencing.
WES (whole-exome sequencing)	Whole-exome sequencing is the process of reading and recording the nucleotide sequence of all the protein coding regions of a subject's genome (termed the exome).
WGS (whole-genome sequencing)	Whole-genome sequencing is the process of reading and recording the nucleotide sequence of the entirety of a subject's genome.
RNASeq (RNA sequencing)	RNA sequencing is the process of capturing and sequencing the available RNA from a tissue or cell. This can include all transcribed RNAs, or just the messenger RNA, depending on the capture method.
Short-read sequencing	Short-read sequencing is the process of sequencing a pre-fragmented library of nucleic acid, typically 75–300 bp in size.
Long-read sequencing	Long-read sequencing in the process of sequencing full-length nucleic acid fragments.

REFERENCES

Abolhassani, H., Naseri, A., Rezaei, N., and Aghamohammadi, A. (2015). Economic burden of common variable immunodeficiency: annual cost of disease AU - Sadeghi, Bamdad. *Expert Rev. Clin. Immunol.* 11 (5), 681–688. doi: 10.1586/1744666X.2015.1029457

Alkhatib, A., Werner, M., Hug, E., Herzog, S., Eschbach, C., and Faraidun, H. (2012). FoxO1 induces Ikaros splicing to promote immunoglobulin gene recombination. *J. Exp. Med.* 209 (2), 395–406. doi: 10.1084/jem.20110216

Anna, A., and Monika, G. (2018). Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J. Appl. Genet.* 59 (3), 253–268. doi: 10.1007/s13353-018-0444-7

Bamshad, M. J., Ng, S. B., Bigham, A. W., Tabor, H. K., Emond, M. J., Nickerson, D. A. et al. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nat. Rev. Genet.* 12, 745. doi: 10.1038/nrg3031

Bentley, D. L. (2014). Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet.* 15 (3), 163–175. doi: 10.1038/nrg3662

Bonilla, F. A., and Oettgen, H. C. (2010). Adaptive immunity. *J. Allergy Clin. Immunol.* 125 (2 Suppl 2), S33–S40. doi: 10.1016/j.jaci.2009.09.017

Bonilla, F. A., Khan, D. A., Ballas, Z. K., Chinen, J., Frank, M. M., Hsu, J. T. et al. (2015). Practice parameter for the diagnosis and management of primary immunodeficiency. *J. Allergy Clin. Immunol.* 136 (5), 1186–205.e1-78. doi: 10.1016/j.jaci.2015.04.049

Bousfiha, A., Jeddane, L., Picard, C., Ailal, F., Bobby Gaspar, H., Al-Herz, W. et al. (2018). The 2017 IUIS Phenotypic classification for primary immunodeficiencies. *J. Clin. Immunol.* 38 (1), 129–143. doi: 10.1007/s10875-017-0465-8

Boycott, K. M., Rath, A., Chong, J. X., Hartley, T., Alkuraya, F. S., Baynam, G. et al. (2017). International cooperation to enable the diagnosis of all rare genetic diseases. *Am. J. Hum. Genet.* 100 (5), 695–705. doi: 10.1016/j.ajhg.2017.04.003

Boyle, J. M., and Buckley, R. H. (2007). Population prevalence of diagnosed primary immunodeficiency diseases in the United States. *J. Clin. Immunol.* 27 (5), 497–502. doi: 10.1007/s10875-007-9103-1

Brodin, P., Jojic, V., Gao, T., Bhattacharya, S., Angel, C. J., and Furman, D. et al. (2015). Variation in the human immune system is largely driven by non-heritable influences. *Cell* 160 (1–2), 37–47. doi: 10.1016/j.cell.2014.12.020

Bryois, J., Buil, A., Evans, D. M., Kemp, J. P., Montgomery, S. B., Conrad, D. F. et al. (2014). Cis and trans effects of human genomic variants on gene expression. *PloS Genet.* 10 (7), e1004461. doi: 10.1371/journal.pgen.1004461

Campos-Sanchez, E., Martínez-Cano, J., del Pino Molina, L., López-Granados, E., and Cobaleda, C. (2019). Epigenetic Deregulation in Human Primary Immunodeficiencies. *Trends In Immunol.* 40 (1), 49–65. doi: 10.1016/j.it.2018.11.005

Casamassimi, A., Federico, A., Rienzo, M., Esposito, S., and Ciccodicola, A. (2017). Transcriptome Profiling in human diseases: new advances and perspectives. *Int. J. Mol. Sci.* 18 (8), 1652. doi: 10.3390/ijms18081652

Chen, S., Townsend, K., Goldberg, T. E., Davies, P., and Conejero-Goldberg, C. (2010). MAPT isoforms: differential transcriptional profiles related to 3R and 4R splice variants. *J. Alzheimers Dis.* 22 (4), 1313–1329. doi: 10.3233/JAD-2010-101155

Chitnis, N., Clark, P. M., Kamoun, M., Stolle, C., Brad Johnson, F., and Monos, D. S. (2017). An expanded role for HLA genes: HLA-B encodes a microRNA that regulates IGA and other immune response transcripts. *Front. Immunol.* 8 (583). doi: 10.3389/fimmu.2017.00583

Condino-Neto, A., and Espinosa-Rosales, F. J. (2018). Changing the Lives of people with primary immunodeficiencies (pi) with early testing and diagnosis. *Front. In Immunol.* 9 (1439). doi: 10.3389/fimmu.2018.01439

Cummings, B. B., Marshall, J. L., Tukiainen, T., Lek, M., Donkervoort, S., Foley, A. R. et al. (2017). Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci. Transl. Med.* 9 (386), eaal5209. doi: 10.1126/scitranslmed.aal5209

DiStefano, J. K. (2018). The Emerging Role of Long Noncoding RNAs in Human Disease. *Methods Mol. Biol. (Clifton NJ)* 1706, 91–110. doi: 10.1007/978-1-4939-7471-9_6

Duffy, D., Rouilly, V., Braudeau, C., Corbiere, V., Djebali, R., Ungeheuer, M. N. et al. (2017). Standardized whole blood stimulation improves immunomonitoring of induced immune responses in multi-center study. *Clin. Immunol. (Orlando Fla)* 183, 325–335. doi: 10.1016/j.clim.2017.09.019

Ergun, A., Doran, G., Costello, J. C., Paik, H. H., Collins, J. J., Mathis, D. et al. (2013). Differential splicing across immune system lineages. *Proc. Natl. Acad. Sci. U. S. A.* 110 (35), 14324–14329. doi: 10.1073/pnas.1311839110

Eshel, D., Toporik, A., Efrati, T., Nakav, S., Chen, A., and Douvdevani, A. (2008). Characterization of natural human antagonistic soluble CD40 isoforms produced through alternative splicing. *Mol. Immunol.* 46 (2), 250–257. doi: 10.1016/j.molimm.2008.08.280

Europe PIPDDfOCi. European Reference Paper. P1-16 worldpiweek.org.

Fairfax, B. P., Humburg, P., Makino, S., Naranbhai, V., Wong, D., Lau, E. et al. (2014). Innate Immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* 343 (6175), 1246949. doi: 10.1126/science.1246949

Feng, Y., Zhang, Y., Ying, C., Wang, D., and Du, C. (2015). Nanopore-based Fourth-generation DNA Sequencing Technology. *Genomics Proteomics Bioinf.* 13 (1), 4–16. doi: 10.1016/j.gpb.2015.01.009

Fischer, A. (1993). Primary T-cell immunodeficiencies. *Curr. Opin. In Immunol.* 5 (4), 569–578. doi: 10.1016/0952-7915(93)90040-Y

Gallo, V., Dotta, L., Giardino, G., Cirillo, E., Lougaris, V., D'Assante, R. et al. (2016). Diagnostics of primary immunodeficiencies through next-generation sequencing. *Front. In Immunol.* 7, 466. doi: 10.3389/fimmu.2016.00466

Gilissen, C., Hoischen, A., Brunner, H. G., and Veltman, J. A. (2011). Unlocking Mendelian disease using exome sequencing. *Genome Biol.* 12 (9), 228. doi: 10.1186/gb-2011-12-9-228

Gonorazky, H. D., Naumenko, S., Ramani, A. K., Nelakuditi, V., Mashouri, P., Wang, P. et al. (2019). Expanding the boundaries of rna sequencing as a diagnostic tool for rare mendelian disease. *Am. J. Hum. Genet.* 104 (3), 466–483. doi: 10.1016/j.ajhg.2019.01.012

Grodecká, L., Hujová, P., Kramárek, M., Kršjaková, T., Kováčová, T., Vondrášková, K. et al. (2017). Systematic analysis of splicing defects in selected primary immunodeficiencies-related genes. *Clin. Immunol.* 180, 33–44. doi: 10.1016/j.clim.2017.03.010

Han, H., and Jiang, X. (2014). Disease Biomarker Query from RNA-Seq Data. *Cancer Inform.* 13 (Suppl 1), 81–94. doi: 10.4137/CIN.S13876

Heimall, J., Keller, M., Saltzman, R., Bunin, N., McDonald-McGinn, D., Zakai, E. et al. (2012). Diagnosis of 22q11.2 deletion syndrome and artemis deficiency in two children with T-B-NK+ immunodeficiency. *J. Clin. Immunol.* 32 (5), 1141–1144. doi: 10.1007/s10875-012-9741-9

Heimall, J. (2019). Now is the time to use molecular gene testing for the diagnosis of primary immune deficiencies. *J. Allergy Clin. Immunol. In Pract.* 39 (1), 129–140. doi: 10.1016/j.jaip.2018.12.022

Heremans, J., Garcia-Perez, J. E., Turro, E., Schlenner, S. M., Casteels, I., Collin, R. et al. (2018). Abnormal differentiation of B cells and megakaryocytes in patients with Roifman syndrome. *J. Allergy Clin. Immunol.* 142 (2), 630–646. doi: 10.1016/j.jaci.2017.11.061

Hernandez-Trujillo, V., and Ballow, M. (2015). Diagnosing primary immunodeficiency: a practical approach for the non-immunologist AU - Lehman, Heather. *Curr. Med. Res. Opin.* 31 (4), 697–706. doi: 10.1185/03007995.2014.1001063

Heyd, F., ten Dam, G., and Moroy, T. (2006). Auxiliary splice factor U2AF26 and transcription factor Gfi1 cooperate directly in regulating CD45 alternative splicing. *Nat. Immunol.* 7 (8), 859–867. doi: 10.1038/ni1361

Hsu, A. P., Johnson, K. D., Falcone, E. L., Sanalkumar, R., Sanchez, L., Hickstein, D. D. et al. (2013). GATA2 haploinsufficiency caused by mutations in a conserved intronic element leads to MonoMAC syndrome. *Blood* 121 (19), 3830–7, s1-7. doi: 10.1182/blood-2012-08-452763

Hwang, B., Lee, J. H., and Bang, D. (2018). Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp. Mol. Med.* 50 (8), 96. doi: 10.1038/s12276-018-0071-8

Jaganathan, K., Kyriazopoulou Panagiotopoulou, S., McRae, J. F., Darbandi, S. F., Knowles, D., and Li, Y. I. et al. (2019). Predicting Splicing from Primary Sequence with Deep Learning. *Cell* 176 (3), 535–48.e24. doi: 10.1016/j.cell.2018.12.015

Jia, C., Hu, Y., Liu, Y., and Li, M. (2015). Mapping Splicing Quantitative Trait Loci in RNA-Seq. *Cancer Inform.* 14 (Suppl 1), 45–53. doi: 10.4137/CIN.S24832

Khan, S., Kuruvilla, M., Hagin, D., Wakeland, B., Liang, C., Vishwanathan, K. et al. (2016). RNA sequencing reveals the consequences of a novel insertion in dedicator of cytokinesis-8. *J. Allergy Clin. Immunol.* 138 (1), 289–92.e6. doi: 10.1016/j.jaci.2015.11.033

Kim, H. K., Pham, M. H. C., Ko, K. S., Rhee, B. D., and Han, J. (2018). Alternative splicing isoforms in health and disease. *Pflügers Arch. Eur. J. Physiol.* 470 (7), 995–1016. doi: 10.1007/s00424-018-2136-x

Kimmerling, R. J., Lee Szeto, G., Li, J. W., Genshaft, A. S., Kazer, S. W., Payer, K. R. et al. (2016). A microfluidic platform enabling single-cell RNA-seq of multigenerational lineages. *Nat. Commun.* 7 (1), 10220. doi: 10.1038/ncomms10220

Kramer, N. J., Wang, W.-L., Reyes, E. Y., Kumar, B., Chen, C.-C., Ramakrishna, C. et al. (2015). Altered lymphopoiesis and immunodeficiency in <em>miR-142</em< null mice. *Blood* 125 (24), 3720. doi: 10.1182/blood-2014-10-603951

Krawczak, M., Reiss, J., and Cooper, D. N. (1992). The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum. Genet.* 90 (1–2), 41–54. doi: 10.1007/BF00210743

Krawczak, M., Thomas, N. S., Hundrieser, B., Mort, M., Wittig, M., Hampe, J. et al. (2007). Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing. *Hum. Mutat.* 28 (2), 150–158. doi: 10.1002/humu.20400

Kremer, L. S., Wortmann, S. B., and Prokisch, H. (2018). "Transcriptomics": molecular diagnosis of inborn errors of metabolism via RNA-sequencing. *J. Inherit Metab. Dis.* 41 (3), 525–532. doi: 10.1007/s10545-017-0133-4

Kukurba, K. R., and Montgomery, S. B. (2015). RNA Sequencing and Analysis. *Cold Spring Harb. Protoc.* 2015 (11), 951–969. doi: 10.1101/pdb.top084970

Lee, M. N., Ye, C., Villani, A.-C., Raj, T., Li, W., Eisenhaure, T. M. et al. (2014). Common Genetic Variants Modulate Pathogen-Sensing Responses in Human Dendritic Cells. *Science* 343 (6175), 1246980. doi: 10.1126/science.1246980

Lenardo, M., Lo, B., and Lucas, C. L. (2016). Genomics of Immune Diseases and New Therapies. *Annu. Rev. Immunol.* 34, 121–149. doi: 10.1146/annurev-immunol-041015-055620

Lowe, R., Shirley, N., Bleackley, M., Dolan, S., and Shafee, T. (2017). Transcriptomics technologies. *PloS Comput. Biol.* 13 (5), e1005457. doi: 10.1371/journal.pcbi.1005457

Majewski, J., Schwartzentruber, J., Lalonde, E., Montpetit, A., and Jabado, N. (2011). What can exome sequencing do for you? *J. Med. Genet.* 48 (9), 580–589. doi: 10.1136/jmedgenet-2011-100223

McCusker, C., and Warrington, R. (2011). Primary immunodeficiency. *Allergy Asthma Clin. Immunol.* 7 (Suppl 1), S11–S1S. doi: 10.1186/1710-1492-7-S1-S11

McCusker, C., Upton, J., and Warrington, R. (2018). Primary immunodeficiency. *Allergy Asthma Clin. Immunol.* 14 (2), 61. doi: 10.1186/s13223-018-0290-5

Meininger, I., Griesbach, R. A., Hu, D., Gehring, T., Seeholzer, T., Bertossi, A. et al. (2016). Alternative splicing of MALT1 controls signalling and activation of CD4(+) T cells. *Nat. Commun.* 7, 11292. doi: 10.1038/ncomms11292

Merico, D., Roifman, M., Braunschweig, U., Yuen, R. K., Alexandrova, R., Bates, A. et al. (2015). Compound heterozygous mutations in the noncoding RNU4ATAC cause Roifman Syndrome by disrupting minor intron splicing. *Nat. Commun.* 6, 8718. doi: 10.1038/ncomms9718

Meyts, I., Bosch, B., Bolze, A., Boisson, B., Itan, Y., Belkadi, A. et al. (2016). Exome and genome sequencing for inborn errors of immunity. *J. Allergy Clin. Immunol.* 138 (4), 957–969. doi: 10.1016/j.jaci.2016.08.003

Moens, L. N., Falk-Sörqvist, E., Asplund, A. C., Bernatowska, E., Smith, C. I. E., and Nilsson, M. (2014). Diagnostics of primary immunodeficiency diseases: a sequencing capture approach. *PloS One* 9 (12), e114901. doi: 10.1371/journal.pone.0114901

Muir, P., Li, S., Lou, S., Wang, D., Spakowicz, D. J., and Salichos, L. et al. (2016). The real cost of sequencing: scaling computation to keep pace with data generation. *Genome Biol.* 17, 53. doi: 10.1186/s13059-016-0917-0

Neums, L., Suenaga, S., Beyerlein, P., Anders, S., Koestler, D., Mariani, A. et al. (2017). VaDiR: an integrated approach to Variant Detection in RNA. *GigaScience* 7 (2), 1–13. doi: 10.1093/gigascience/gix122

Oeckinghaus, A., Wegener, E., Welteke, V., Ferch, U., Arslan, S. C., Ruland, J. et al. (2007). Malt1 ubiquitination triggers NF-kappaB signaling upon T-cell activation. *EMBO J.* 26 (22), 4634–4645. doi: 10.1038/sj.emboj.7601897

Ohno, K., Takeda, J. I., and Masuda, A. (2018). Rules and tools to predict the splicing effects of exonic and intronic mutations. *Wiley Interdiscip. Rev. RNA* 9 (1). doi: 10.1002/wrna.1451

Orvain, C., Matre, V., and Gabrielsen, O. S. (2008). The transcription factor c-Myb affects pre-mRNA splicing. *Biochem. Biophys. Res. Commun.* 372 (2), 309–313. doi: 10.1016/j.bbrc.2008.05.054

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., and Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* 40 (12), 1413–1415. doi: 10.1038/ng.259

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14 (4), 417–419. doi: 10.1038/nmeth.4197

Philippidis, A. (2018). The 100, 000 Genomes club www.genenews.com [Available from: https://www.genengnews.com/insights/the-100000-genomes-club/.

Piasecka, B., Duffy, D., Urrutia, A., Quach, H., Patin, E., Posseme, C. et al. (2018). Distinctive roles of age, sex, and genetics in shaping transcriptional variation of human immune responses to microbial challenges. *Proc. Natl. Acad. Sci. U. S. A.* 115 (3), e488–ee97. doi: 10.1073/pnas.1714765115

Picard, C., Bobby Gaspar, H., Al-Herz, W., Bousfiha, A., Casanova, J.-L., Chatila, T. et al. (2018). International union of immunological societies: 2017 primary immunodeficiency diseases committee report on inborn errors of immunity. *J. Clin. Immunol.* 38 (1), 96–128. doi: 10.1007/s10875-017-0464-9

Punwani, D., Wang, H., Chan, A. Y., Cowan, M. J., Mallott, J., Sunderam, U. et al. (2015). Combined immunodeficiency due to MALT1 mutations, treated by hematopoietic cell transplantation. *J. Clin. Immunol.* 35 (2), 135–146. doi: 10.1007/s10875-014-0125-1

Rae, W., Ward, D., Mattocks, C., Pengelly, R. J., Eren, E., Patel, S. V. et al. (2018). Clinical efficacy of a next-generation sequencing gene panel for primary immunodeficiency diagnostics. *Clin. Genet.* 93 (3), 647–655. doi: 10.1111/cge.13163

Ramakrishnan, K. A., Pengelly, R. J., Gao, Y., Morgan, M., Patel, S. V., Davies, E. G. (2016). Precision molecular diagnosis defines specific therapy in combined immunodeficiency with megaloblastic anemia secondary to MTHFD1 deficiency. *J. Allergy Clin. Immunol. In Pract.* 4 (6), 1160–6.e10. doi: 10.1016/j.jaip.2016.07.014

Reynaud, D., Demarco, I. A., Reddy, K. L., Schjerven, H., Bertolino, E., Chen, Z. et al. (2008). Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. *Nat. Immunol.* 9 (8), 927–936. doi: 10.1038/ni.1626

Rhoads, A., and Au, K. F. (2015). PacBio Sequencing and Its Applications. *Genomics Proteomics Bioinf.* 13 (5), 278–289. doi: 10.1016/j.gpb.2015.08.002

Richardson, A. M., Moyer, A. M., Hasadsri, L., and Abraham, R. S. (2018). Diagnostic tools for inborn errors of human immunity (primary immunodeficiencies and immune dysregulatory diseases). *Curr. Allergy Asthma Rep.* 18 (3), 19. doi: 10.1007/s11882-018-0770-1

Roadmap Epigenomics, C., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A. et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* 518, 317. doi: 10.1038/nature14248

Scacheri, C. A., and Scacheri, P. C. (2015). Mutations in the noncoding genome. *Curr. Opin. Pediatr.* 27 (6), 659–664. doi: 10.1097/MOP.0000000000000283

Schmiedel, B. J., Singh, D., Madrigal, A., Valdovino-Gonzalez, A. G., White, B. M., Zapardiel-Gonzalo, J. et al. (2018). Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. *Cell* 175 (6), 1701–15.e16. doi: 10.1016/j.cell.2018.10.022

Schwarze, K., Buchanan, J., Taylor, J. C., and Wordsworth, S. (2018). Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genet. In Med.* 20 (10), 1122–1130. doi: 10.1038/gim.2017.247

See, P., Lum, J., and Chen, J. (2018). Ginhoux F. A single-cell sequencing guide for immunologists. *Front. Immunol.* 9, 2425. doi: 10.3389/fimmu.2018.02425

Shillitoe, B., Bangs, C., Guzman, D., Gennery, A. R., Longhurst, H. J., Slatter, M. et al. (2018). The United Kingdom Primary Immune Deficiency (UKPID) registry 2012 to 2017. *Clin. Exp. Immunol.* 192 (3), 284–291. doi: 10.1111/cei.13125

Starokadomskyy, P., Gemelli, T., Rios, J. J., Xing, C., Wang, R. C., and Li, H. et al. (2016).DNA polymerase-α regulates the activation of type I interferons through cytosolic RNA:DNA synthesis. *Nat. Immunol.* 17 (5), 495–504. doi: 10.1038/ni.3409

Stray-Pedersen, A., Sorte, H. S., Samarakoon, P., Gambin, T., Chinn, I. K., Coban Akdemir, Z. H. et al. (2017). Primary immunodeficiency diseases: genomic approaches delineate heterogeneous Mendelian disorders. *J. Allergy Clin. Immunol.* 139 (1), 232–245. doi: 10.1016/j.jaci.2016.05.042

Takata, A., Matsumoto, N., and Kato, T. (2017). Genome-wide identification of splicing QTLs in the human brain and their enrichment among

schizophrenia-associated loci. *Nat. Commun.* 8 (1), 14519. doi: 10.1038/ncomms14519

Takeda, N., O'Dea, E. L., Doedens, A., J-w, Kim, Weidemann, A., Stockmann, C. et al. (2010). Differential activation and antagonistic function of HIF-α isoforms in macrophages are essential for NO homeostasis. *Genes Dev.* 24 (5), 491–501. doi: 10.1101/gad.1881410

Taylor, J. C., Martin, H. C., Lise, S., Broxholme, J., Cazier, J. B., Rimmer, A. et al. (2015).Factors influencing success of clinical genome sequencing across a broad spectrum of disorders. *Nat. Genet.* 47 (7), 717–726. doi: 10.1038/ng.3304

The Genotype-Tissue Expression (GTEx) (2015) Pilot analysis: Multitissue gene regulation in humans. *Science* 348 (6235), 648. doi: 10.1126/science.1262110

Thormann, A., Halachev, M., McLaren, W., Moore, D. J., Svinti, V., Campbell, A. et al. (2019). Flexible and scalable diagnostic filtering of genomic variants using G2P with Ensembl VEP. *Nat. Commun.* 10 (1), 2373. doi: 10.1038/s41467-019-10016-3

Turunen, J. J., Niemela, E. H., Verma, B., and Frilander, M. J. (2013). The significant other: splicing by the minor spliceosome. *Wiley Interdiscip. Rev. RNA* 4 (1), 61–76. doi: 10.1002/wrna.1141

UK, P. Patients' experience survey of Primary Immunodeficiency Disorders services. www.piduk.org; 2016 September 2016.

Urrutia, A., Duffy, D., Rouilly, V., Posseme, C., Djebali, R., Illanes, G. et al. (2016). Standardized Whole-Blood Transcriptional Profiling Enables the Deconvolution of Complex Induced Immune Responses. *Cell Rep.* 16 (10), 2777–2791. doi: 10.1016/j.celrep.2016.08.011

Walter, J. E., Farmer, J. R., Foldvari, Z., Torgerson, T. R., and Cooper, M. A. (2016).Mechanism-Based strategies for the management of autoimmunity and immune dysregulation in primary immunodeficiencies. *J. Allergy Clin. Immunol. In Pract.* 4 (6), 1089–1100. doi: 10.1016/j.jaip.2016.08.004

Wang, Y., Liu, J., Huang, B. O., Xu, Y.-M., Li, J., Huang, L.-F. et al. (2015). Mechanism of alternative splicing and its regulation. *BioMed. Rep.* 3 (2), 152–158. doi: 10.3892/br.2014.407

Ward, A. J., and Cooper, T. A. (2010). The pathobiology of splicing. *J. Pathol.* 220 (2), 152–163. doi: 10.1002/path.2649

Wirka, R. C., Pjanic, M., and Quertermous, T. (2018). Advances in Transcriptomics: Investigating Cardiovascular Disease at Unprecedented Resolution. *Circ. Res.* 122 (9), 1200–1220. doi: 10.1161/CIRCRESAHA.117.310910

Yang, Y., Muzny, D. M., Reid, J. G., Bainbridge, M. N., Willis, A., Ward, P. A. et al. (2013). Clinical Whole-exome sequencing for the diagnosis of mendelian disorders. *New Engl. J. Med.* 369 (16), 1502–1511. doi: 10.1056/NEJMoa1306555

Yang, Y., Muzny, D. M., Xia, F., Niu, Z., Person, R., Ding, Y. et al. (2014). Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* 312 (18), 1870–1879. doi: 10.1001/jama.2014.14601

Zeng, Y., Wang, G., Yang, E., Ji, G., Brinkmeyer-Langford, C. L., Cai, J. J. (2015). Aberrant Gene Expression in Humans. *PloS Genet.* 11 (1), e1004942. doi: 10.1371/journal.pgen.1004942

Zhao, J., Akinsanmi, I., Arafat, D., Cradick, T. J., Lee, C. M., Banskota, S. et al. (2016). A Burden of Rare Variants Associated with Extremes of Gene Expression in Human Peripheral Blood. *Am. J. Hum. Genet.* 98 (2), 299–309. doi: 10.1016/j.ajhg.2015.12.023

Zhu, L.-Y., Zhu, Y.-R., Dai, D.-J., Wang, X., and Jin, H.-C. (2018). Epigenetic regulation of alternative splicing. *Am. J. Cancer Res.* 8 (12), 2346–2358.

Zikherman, J., and Weiss, A. (2008).Alternative splicing of CD45: the tip of the iceberg. *Immunity* 29 (6), 839–841. doi: 10.1016/j.immuni.2008.12.005