Exploring the RNA gap for improving diagnostic yield in Primary Immunodeficiencies.’

Jed J. Lye1, Anthony Williams1\* and Diana Baralle1.

*1 University of Southampton Medical School, University of Southampton, UK*

**\*Corresponding authors**

*Professor. Diana Baralle*

*Faculty of Medicine*

*University of Southampton,*

*Duthie Building,*

*Life Sciences Building*

*Highfield Campus,*

*Southampton,*

*Hampshire,*

*UK*

*SO17 1BJ*

*D.Baralle@soton.ac.uk*

*Professor Anthony Williams*

*Wessex Investigational Sciences Hub laboratory (WISH Lab)*

*Mailpoint 850*

*Faculty of Medicine*

*University of Southampton*

*Tremona Road*

*Southampton*

*Hampshire*

*UK*

*S016 6YD*

*apw2@soton.ac.uk*

**KEYWORDS**: Primary Immunodeficiency Disorders, Clinical Diagnostics, RNASeq, RNA

**Abstract**

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 give be utilised to gain an unparalleled insight into cellular immunology. We review how investigation into RNA biology can give information regarding the differential expression, mono-allelic 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.

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 RNA, 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 75-300bp in size. |
| Long read sequencing | Long read sequencing in the process of sequencing full length nucleic acid fragments. |

*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 (1). 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 (2, 3). This plethora of genetic causes has brought about a need to categorise 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 (4). 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 of nine categories; Immunodeficiencies affecting cellular and humoral immunity, 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 (2).

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 (5), and whilst individually rare (>1 in 2000) , 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 (6), however underdiagnsosis of PID may mean the true incidence is as high as 1:250 (7).

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 whilst a diagnosed US patient costs healthcare services over US$250,000 PA, largely due to treatment costs, an early diagnosis of the disorder can save as much as US$6500 per patient, per annum (8). 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 $79,942 (9). In a patient survey, 45% of patients reported a diagnostic wait time of between 1-6 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 NHS resources (10). Identification of the precise molecular origins for each patient’s case of PID leads to improved patient care (11), 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 (12). Precision therapeutic diagnostics can help to achieve this in part, by allowing targeted intervention to the specific molecular causes (13-15) .

*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 (16). Additionally, these correlation studies have been useful for de-convoluting the pleiotropic nature of the involved genes, through which a single variant can bring about a variety of clinical phenotypes (17). However the development of a universal diagnostic pipeline for PIDD is hindered by the heterogeneity in presentation of disease, even among patients with what appears to be the same pathogenic genetic variant (18). Conversely, a number of genotypes can bring about even the most well characterised phenotype (17). Once a clinical diagnosis of PID is suspected, mainly based upon a compatible phenotype, a family history is usually taken and number of subsequent laboratory tests performed to confirm the type of immune mechanism affected can be performed (19). 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 work up 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 characterising 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 adopted methods (9). Whole exome sequencing (WES) has the highest success rate of clinically adopted diagnostic methods, (20, 21), which it achieves despite the exome comprising only ~2% of the human genome (22).This is in part due to 85% of currently annotated variants existing within the transcribed portion of the genome (23). It has been hypothesised that this focus had likely led to the underestimation of the contribution to disease of non-coding variants (24).

Due to the improvement that WES and WGS brings to diagnostics researchers are calling for universal molecular gene testing for the diagnosis of primary immune deficiencies (25). Evidence from existing literature, however, suggests that even this may be inadequate; currently WES and whole genome sequencing (WGS) is only able to produce reliable diagnosis in 25-60% of cases (26-31). Although many countries have undertaken whole genome sequencing projects to evaluate this approach (32), the development of whole genome sequencing (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 (33). 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 (24), 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 patient’s genome sequencing, can be reliably linked to the clinical presentation and cytological/molecular manifestation of the disorder. Failure to identify definitive molecular cause occurs in about 70-75% of mendelian conditions according to a 2018 meta-analysis (34) mirrored by examples from PID (35). The types of variants which are not always identified by current NGS approaches include exonic variants of unknown significance, variants in intronic and intergenic non-coding DNA (36), variants in the *cis-*acting regulatory elements of transcription (37) imprinting disorders and repeat expansions (24).

Conventional clinical diagnostics, utilising human phenotype ontology for integration of cases into specific diagnostic groups, and traditional genetic sequencing methods for diagnostics are still currently inadequate. Whilst 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 filtering process because of a lack of integrative understanding or supporting evidence (38).

These expression quantitative loci (eQTL’s), elicit a powerful, sometimes synergistic effect on the expression of a large number of genes. Single nucleotide polymorphisms (SNP’s) on eQTL loci affect the transcriptional level of other RNA’s, modifying protein expression and causing phenotypic changes to the abilities and behaviours of cells in some immunological cases (39). These expression quantitative trait loci (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 (40). 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 (41). 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 (42). Large scale investigation into 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 (43). This research did not extend to immune tissues specifically, although links between immune cell specific gene expression levels and eQTLs has been investigated by the DICE project (database of immune cell expression, expression quantitative trait loci and epigenomics) (44). 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 (45). 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 (45).

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 (46).

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 (47-49). Important examples include RNA-polymerase II as a facilitator of splicing factor recruitment (50), 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 (51), and FOX01 induced Ikaros splicing, essential for the recombination of immunoglobulin genes. FOXO1 is a transcription factor, which through it’s effects on alternative spilcing allows the immune system to produce its diverse range of antibodies/immunoglobulins (52).

Activation of lymphocytes is a key component of the adaptive immune response to pathogens (53). 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 (54). 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 (55). 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, those which occur in intronic or intergenic regions. (56).

Figure 2 Variants affecting Splicing



Figure 2 - 1) Variants which occur in known active splice sites (or branchpoint/polypyrimidine tract) leading to whole exon skipping 2) Deep intronic variants which bring about activation of cryptic splice sites and inclusion of cryptic exons, 3) cryptic splice site activation in existing exons – causing fragment loss 4) canonical splice site variant, leading to the inclusion of a cryptic exon, 5) mutations in regulatory elements such as exon splicing enhancers, resulting in for example; exon skipping. Adapted from (Abramowicz and Monika, 2018) (57)

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 (56). *Cis*-mutations in the genome can affect splicing though altering the splice site recognition or altering exon splicing enhancer or silencer sites (45). 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 (58, 59). 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 (45), 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 (56, 60). Unfortunately, these tools are often unable to discern the resulting transcripts exon use patterns (61), and whilst their predictive ability can be enhanced by other orthogonal investigations such as mini-gene assays (56), 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 localisation of different and often competing *trans*-acting factors (62).

Further complicating this process, seemingly benign, synonymous exonic variants can disrupt splicing to cause disease. Using RNASeq to compliment 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 (63). 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%-11% of rare genetic disorders are caused by synonymous or intronic splice altering mutations (61). Indeed, much as gene expression can be influenced by multiple loci, so too can multiple loci contribute to the occurence of splicing events. These loci are appropriately termed splicing quantitative trait loci (sQTLs) (64). 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 (65). These sQTL’s are not necessarily in close proximity to the splice junction, characterization of these sites in humans have shown SNP’S demonstrating tangible sQTL activity at 100kb from the relative splice site (66).

Non protein-coding genes, are a significant source of disease causing variation (36). 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 (67). Specifically, the noncoding gene RNU4ATAC that produces a small nuclear RNA (snRNA) termed U4atac was discovered to cause Roifman syndrome (68, 69), 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 (69).

The importance of alternative splicing in the immune system, has 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 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 (70). Compound heterozygous mutations *in MALT1,* mentioned earlier which is heavily implicated in 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 (71) .

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 (72). 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 (73).

*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 (74). RNA Sequencing is now largely replacing microarrays as the most used quantitative methods of mapping gene expression profiles (75). The transcriptome – or RNA expression profile of a given tissue can give unparalleled insight into the elegant inner workings of the cell. Through capture of all internal RNA species it characterises the cellular gene transcription architecture and can deliver an instantaneous picture of environment–cell interaction or response programme (75, 76) .

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 sub-optimal reliability in base calling (77), or is more expensive for high throughput analysis (78). 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 (79). This technology has been the currently favoured 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 (80), somatic cell mosaics, non-canonical splice variants, occurring either due to *cis* or *trans-*acting factor aberrations (45). Quantitative outcomes of transcriptional profiling include differentially expressed genes, alternative splicing events and allele specific expression quantification (79). 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 (81, 82). As well as gene expression levels, perturbations in the relative abundance of specific isoforms is a driving force in the genesis of many diseases (83, 84) as isoforms can have differential function (85), or in some cases, can be antagonistic (86). 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 ~7600 novel isoforms in mouse Immune Cells (70) and detected 100,000 splicing events with at least moderate abundance (46).

Transcriptome profiling can also give insight into control mechanisms exhibited by the non-coding RNA species, such as lncRNA, and miRNA, the significance of which is continually being elucidated in the molecular pathology of disease (87, 88). Indeed such examples exist in PID; miR-6891-5p accumulation is demonstrated to contribute to selective IgA deficiency, the most common form of PID (89). 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 of 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 (90). Subsequent use of follow up analysis tools can be used to generate filtering process for causal variants, or for biomarker identification (91). It is also possible to quantify the relative expression of those genes coding for the splice factors themselves, which can directly bring about specific pathological processes specific to PID such as those observed in Roifman’s syndrome, mentioned earlier (68, 69).

Micro-fluidic technology adaptations have allowed the development of robust, single-cell transcriptomic profiling (92). In combination with NGS based technologies the single cell technology provides a method for profiling the transcriptomes of individual cells, giving unparalleled insight into the heterogeneity of cell populations and their transcriptional profiles (93). Adaptations such as the SMART-seq2 or fluidigm C1 library preparation methods also now allow the production of full length cDNA’s, giving transcript isoform level resolution. However, these methods do not yet allow multiplexing, massively increasing overall costs and labour in large cohorts (94). The ability to profile the entire transcriptome of a 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 utilised in those patients suspected to be genetic mosaics.

*Discussion*

The early and accurate diagnosis of Primary Immuno-deficiencies is important to ensure the attainment of positive patient outcomes, through minimising 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 standardise and mostly available in research laboratories. RNA profiling to identify alternative splicing, gene expression level variation monoallelic expression may contribute a further insight into candidates 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 fig 2.).

Figure 2.



Figure 2. Demonstrates 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.

RNASeq is an emerging technology which, when combined with WES/WGS provides unprecedented insight into differential gene expression, splicing activity, allelic specific expression and can inform regarding other phenomenon’s 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.

Utilising candidate gene lists and large control datasets for comparison enhances the power of the transcriptional 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 have all have effects on the specific response profile exhibited by individuals (41). 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 (41). It is therefore useful to be able to discern transcripts from different cell types within a culture. Utilising flow cytometry to separate cell types or utilising 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 (95).

The immune systems response to pathogen-based challenges is highly dynamic, and observing this response is more informative when identifying impaired response (96). Indeed is 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 (39, 97). Co-culture of PBMC’s 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 (96). Studies of dynamic immune responses to challenges, in concert with machine learning can be used to identify small groups of stimulation pathway-specific genes (98). Comparing the expression profiles of these genes in healthy cohorts with PID patients can potentially be utilised to identify candidate genes, which may then harbour 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 utilised 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 (31), and these are being answered with examples in primary immunodeficiency cases such as Dock8 CID, GATA2 deficiency, X-linked reticulate pigmentary disorder (XLPDR) (99-101). 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 complimented by a DNA module of coding and non-coding variant analysis, utilising sophisticated bioinformatic pipelines to prioritise candidate genetic variants of new loci that would be consistent 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 characterise 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 standardised analysis of novel variants that will continue to grow over the next few years .

**References**

1. McCusker C, Upton J, Warrington R. Primary immunodeficiency. Allergy, Asthma & Clinical Immunology. 2018;14(2):61.

2. Bousfiha A, Jeddane L, Picard C, Ailal F, Bobby Gaspar H, Al-Herz W, et al. The 2017 IUIS Phenotypic Classification for Primary Immunodeficiencies. Journal of clinical immunology. 2018;38(1):129-43.

3. Picard C, Bobby Gaspar H, Al-Herz W, Bousfiha A, Casanova J-L, Chatila T, et al. International Union of Immunological Societies: 2017 Primary Immunodeficiency Diseases Committee Report on Inborn Errors of Immunity. Journal of clinical immunology. 2018;38(1):96-128.

4. McCusker C, Warrington R. Primary immunodeficiency. Allergy Asthma Clin Immunol. 2011;7 Suppl 1(Suppl 1):S11-S.

5. Boyle JM, Buckley RH. Population Prevalence of Diagnosed Primary Immunodeficiency Diseases in the United States. J Clin Immunol. 2007;27(5):497-502.

6. Shillitoe B, Bangs C, Guzman D, Gennery AR, Longhurst HJ, Slatter M, et al. The United Kingdom Primary Immune Deficiency (UKPID) registry 2012 to 2017. Clinical & Experimental Immunology. 2018;192(3):284-91.

7. Europe PIPDDfOCi. European Reference Paper. worldpiweek.org.

8. Abolhassani H, Naseri A, Rezaei N, Aghamohammadi A. Economic burden of common variable immunodeficiency: annual cost of disease AU - Sadeghi, Bamdad. Expert Review of Clinical Immunology. 2015;11(5):681-8.

9. Condino-Neto A, Espinosa-Rosales FJ. Changing the Lives of People With Primary Immunodeficiencies (PI) With Early Testing and Diagnosis. Frontiers in Immunology. 2018;9(1439).

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

11. Walter JE, Farmer JR, Foldvari Z, Torgerson TR, Cooper MA. Mechanism-Based Strategies for the Management of Autoimmunity and Immune Dysregulation in Primary Immunodeficiencies. The journal of allergy and clinical immunology In practice. 2016;4(6):1089-100.

12. Heimall J, Keller M, Saltzman R, Bunin N, McDonald-McGinn D, Zakai E, et al. Diagnosis of 22q11.2 deletion syndrome and artemis deficiency in two children with T-B-NK+ immunodeficiency. J Clin Immunol. 2012;32(5):1141-4.

13. Bonilla FA, Khan DA, Ballas ZK, Chinen J, Frank MM, Hsu JT, et al. Practice parameter for the diagnosis and management of primary immunodeficiency. The Journal of allergy and clinical immunology. 2015;136(5):1186-205.e1-78.

14. Lenardo M, Lo B, Lucas CL. Genomics of Immune Diseases and New Therapies. Annual review of immunology. 2016;34:121-49.

15. Ramakrishnan KA, Pengelly RJ, Gao Y, Morgan M, Patel SV, Davies EG, et al. Precision Molecular Diagnosis Defines Specific Therapy in Combined Immunodeficiency with Megaloblastic Anemia Secondary to MTHFD1 Deficiency. The journal of allergy and clinical immunology In practice. 2016;4(6):1160-6.e10.

16. Fischer A. Primary T-cell immunodeficiencies. Current Opinion in Immunology. 1993;5(4):569-78.

17. Meyts I, Bosch B, Bolze A, Boisson B, Itan Y, Belkadi A, et al. Exome and genome sequencing for inborn errors of immunity. The Journal of allergy and clinical immunology. 2016;138(4):957-69.

18. Richardson AM, Moyer AM, Hasadsri L, Abraham RS. Diagnostic Tools for Inborn Errors of Human Immunity (Primary Immunodeficiencies and Immune Dysregulatory Diseases). Current Allergy and Asthma Reports. 2018;18(3):19.

19. Hernandez-Trujillo V, Ballow M. Diagnosing primary immunodeficiency: a practical approach for the non-immunologist AU - Lehman, Heather. Current Medical Research and Opinion. 2015;31(4):697-706.

20. Boycott KM, Rath A, Chong JX, Hartley T, Alkuraya FS, Baynam G, et al. International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases. The American Journal of Human Genetics. 2017;100(5):695-705.

21. Gilissen C, Hoischen A, Brunner HG, Veltman JA. Unlocking Mendelian disease using exome sequencing. Genome biology. 2011;12(9):228.

22. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, et al. Exome sequencing as a tool for Mendelian disease gene discovery. Nature Reviews Genetics. 2011;12:745.

23. Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? Journal of medical genetics. 2011;48(9):580-9.

24. Kremer LS, Wortmann SB, Prokisch H. "Transcriptomics": molecular diagnosis of inborn errors of metabolism via RNA-sequencing. J Inherit Metab Dis. 2018;41(3):525-32.

25. Heimall J. Now Is the Time to Use Molecular Gene Testing for the Diagnosis of Primary Immune Deficiencies. The journal of allergy and clinical immunology In practice. 2019.

26. Meyts I, Bosch B, Bolze A, Boisson B, Itan Y, Belkadi A, et al. Exome and genome sequencing for inborn errors of immunity. The Journal of allergy and clinical immunology. 2016;138(4):957-69.

27. Taylor JC, Martin HC, Lise S, Broxholme J, Cazier JB, Rimmer A, et al. Factors influencing success of clinical genome sequencing across a broad spectrum of disorders. Nature genetics. 2015;47(7):717-26.

28. Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA, et al. Clinical Whole-Exome Sequencing for the Diagnosis of Mendelian Disorders. New England Journal of Medicine. 2013;369(16):1502-11.

29. Yang Y, Muzny DM, Xia F, Niu Z, Person R, Ding Y, et al. Molecular findings among patients referred for clinical whole-exome sequencing. JAMA. 2014;312(18):1870-9.

30. Stray-Pedersen A, Sorte HS, Samarakoon P, Gambin T, Chinn IK, Coban Akdemir ZH, et al. Primary immunodeficiency diseases: Genomic approaches delineate heterogeneous Mendelian disorders. The Journal of allergy and clinical immunology. 2017;139(1):232-45.

31. Moens LN, Falk-Sörqvist E, Asplund AC, Bernatowska E, Smith CIE, Nilsson M. Diagnostics of primary immunodeficiency diseases: a sequencing capture approach. PloS one. 2014;9(12):e114901-e.

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

33. Rae W, Ward D, Mattocks C, Pengelly RJ, Eren E, Patel SV, et al. Clinical efficacy of a next-generation sequencing gene panel for primary immunodeficiency diagnostics. Clinical Genetics. 2018;93(3):647-55.

34. Schwarze K, Buchanan J, Taylor JC, Wordsworth S. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. Genetics in Medicine. 2018;20(10):1122-30.

35. Gallo V, Dotta L, Giardino G, Cirillo E, Lougaris V, D'Assante R, et al. Diagnostics of Primary Immunodeficiencies through Next-Generation Sequencing. Frontiers in immunology. 2016;7:466-.

36. Scacheri CA, Scacheri PC. Mutations in the noncoding genome. Curr Opin Pediatr. 2015;27(6):659-64.

37. Bryois J, Buil A, Evans DM, Kemp JP, Montgomery SB, Conrad DF, et al. Cis and Trans Effects of Human Genomic Variants on Gene Expression. PLOS Genetics. 2014;10(7):e1004461.

38. Thormann A, Halachev M, McLaren W, Moore DJ, Svinti V, Campbell A, et al. Flexible and scalable diagnostic filtering of genomic variants using G2P with Ensembl VEP. Nature communications. 2019;10(1):2373-.

39. Fairfax BP, Humburg P, Makino S, Naranbhai V, Wong D, Lau E, et al. Innate Immune Activity Conditions the Effect of Regulatory Variants upon Monocyte Gene Expression. Science. 2014;343(6175):1246949.

40. Casamassimi A, Federico A, Rienzo M, Esposito S, Ciccodicola A. Transcriptome Profiling in Human Diseases: New Advances and Perspectives. International journal of molecular sciences. 2017;18(8):1652.

41. Piasecka B, Duffy D, Urrutia A, Quach H, Patin E, Posseme C, et al. Distinctive roles of age, sex, and genetics in shaping transcriptional variation of human immune responses to microbial challenges. Proceedings of the National Academy of Sciences of the United States of America. 2018;115(3):E488-e97.

42. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518:317.

43. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science. 2015;348(6235):648.

44. Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, et al. Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. Cell. 2018;175(6):1701-15.e16.

45. Ward AJ, Cooper TA. The pathobiology of splicing. J Pathol. 2010;220(2):152-63.

46. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nature genetics. 2008;40(12):1413-5.

47. Orvain C, Matre V, Gabrielsen OS. The transcription factor c-Myb affects pre-mRNA splicing. Biochemical and biophysical research communications. 2008;372(2):309-13.

48. Heyd F, ten Dam G, Moroy T. Auxiliary splice factor U2AF26 and transcription factor Gfi1 cooperate directly in regulating CD45 alternative splicing. Nature immunology. 2006;7(8):859-67.

49. Alkhatib A, Werner M, Hug E, Herzog S, Eschbach C, Faraidun H, et al. FoxO1 induces Ikaros splicing to promote immunoglobulin gene recombination. The Journal of experimental medicine. 2012;209(2):395-406.

50. Bentley DL. Coupling mRNA processing with transcription in time and space. Nature reviews Genetics. 2014;15(3):163-75.

51. Zikherman J, Weiss A. Alternative splicing of CD45: the tip of the iceberg. Immunity. 2008;29(6):839-41.

52. Reynaud D, Demarco IA, Reddy KL, Schjerven H, Bertolino E, Chen Z, et al. Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. Nature immunology. 2008;9(8):927-36.

53. Bonilla FA, Oettgen HC. Adaptive immunity. The Journal of allergy and clinical immunology. 2010;125(2 Suppl 2):S33-40.

54. Oeckinghaus A, Wegener E, Welteke V, Ferch U, Arslan SC, Ruland J, et al. Malt1 ubiquitination triggers NF-kappaB signaling upon T-cell activation. The EMBO journal. 2007;26(22):4634-45.

55. Meininger I, Griesbach RA, Hu D, Gehring T, Seeholzer T, Bertossi A, et al. Alternative splicing of MALT1 controls signalling and activation of CD4(+) T cells. Nat Commun. 2016;7:11292-.

56. Grodecká L, Hujová P, Kramárek M, Kršjaková T, Kováčová T, Vondrášková K, et al. Systematic analysis of splicing defects in selected primary immunodeficiencies-related genes. Clinical Immunology. 2017;180:33-44.

57. Anna A, Monika G. Splicing mutations in human genetic disorders: examples, detection, and confirmation. J Appl Genet. 2018;59(3):253-68.

58. Krawczak M, Thomas NS, Hundrieser B, Mort M, Wittig M, Hampe J, et al. Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing. Human mutation. 2007;28(2):150-8.

59. Krawczak M, Reiss J, Cooper DN. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. Human genetics. 1992;90(1-2):41-54.

60. Ohno K, Takeda JI, Masuda A. Rules and tools to predict the splicing effects of exonic and intronic mutations. Wiley interdisciplinary reviews RNA. 2018;9(1).

61. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting Splicing from Primary Sequence with Deep Learning. Cell. 2019;176(3):535-48.e24.

62. Wang Y, Liu J, Huang BO, Xu Y-M, Li J, Huang L-F, et al. Mechanism of alternative splicing and its regulation. Biomed Rep. 2015;3(2):152-8.

63. Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. Sci Transl Med. 2017;9(386):eaal5209.

64. Jia C, Hu Y, Liu Y, Li M. Mapping Splicing Quantitative Trait Loci in RNA-Seq. Cancer Inform. 2015;14(Suppl 1):45-53.

65. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14(4):417-9.

66. Takata A, Matsumoto N, Kato T. Genome-wide identification of splicing QTLs in the human brain and their enrichment among schizophrenia-associated loci. Nature Communications. 2017;8(1):14519.

67. Turunen JJ, Niemela EH, Verma B, Frilander MJ. The significant other: splicing by the minor spliceosome. Wiley interdisciplinary reviews RNA. 2013;4(1):61-76.

68. Heremans J, Garcia-Perez JE, Turro E, Schlenner SM, Casteels I, Collin R, et al. Abnormal differentiation of B cells and megakaryocytes in patients with Roifman syndrome. The Journal of allergy and clinical immunology. 2018;142(2):630-46.

69. Merico D, Roifman M, Braunschweig U, Yuen RK, Alexandrova R, Bates A, et al. Compound heterozygous mutations in the noncoding RNU4ATAC cause Roifman Syndrome by disrupting minor intron splicing. Nature communications. 2015;6:8718.

70. Ergun A, Doran G, Costello JC, Paik HH, Collins JJ, Mathis D, et al. Differential splicing across immune system lineages. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(35):14324-9.

71. Punwani D, Wang H, Chan AY, Cowan MJ, Mallott J, Sunderam U, et al. Combined immunodeficiency due to MALT1 mutations, treated by hematopoietic cell transplantation. J Clin Immunol. 2015;35(2):135-46.

72. Campos-Sanchez E, Martínez-Cano J, del Pino Molina L, López-Granados E, Cobaleda C. Epigenetic Deregulation in Human Primary Immunodeficiencies. Trends in Immunology. 2019;40(1):49-65.

73. Zhu L-Y, Zhu Y-R, Dai D-J, Wang X, Jin H-C. Epigenetic regulation of alternative splicing. Am J Cancer Res. 2018;8(12):2346-58.

74. Muir P, Li S, Lou S, Wang D, Spakowicz DJ, Salichos L, et al. The real cost of sequencing: scaling computation to keep pace with data generation. Genome Biol. 2016;17:53-.

75. Lowe R, Shirley N, Bleackley M, Dolan S, Shafee T. Transcriptomics technologies. PLoS Comput Biol. 2017;13(5):e1005457-e.

76. Wirka RC, Pjanic M, Quertermous T. Advances in Transcriptomics: Investigating Cardiovascular Disease at Unprecedented Resolution. Circulation research. 2018;122(9):1200-20.

77. Feng Y, Zhang Y, Ying C, Wang D, Du C. Nanopore-based Fourth-generation DNA Sequencing Technology. Genomics, Proteomics & Bioinformatics. 2015;13(1):4-16.

78. Rhoads A, Au KF. PacBio Sequencing and Its Applications. Genomics, Proteomics & Bioinformatics. 2015;13(5):278-89.

79. Kukurba KR, Montgomery SB. RNA Sequencing and Analysis. Cold Spring Harb Protoc. 2015;2015(11):951-69.

80. Neums L, Suenaga S, Beyerlein P, Anders S, Koestler D, Mariani A, et al. VaDiR: an integrated approach to Variant Detection in RNA. GigaScience. 2017;7(2).

81. Zeng Y, Wang G, Yang E, Ji G, Brinkmeyer-Langford CL, Cai JJ. Aberrant Gene Expression in Humans. PLOS Genetics. 2015;11(1):e1004942.

82. Zhao J, Akinsanmi I, Arafat D, Cradick TJ, Lee CM, Banskota S, et al. A Burden of Rare Variants Associated with Extremes of Gene Expression in Human Peripheral Blood. American journal of human genetics. 2016;98(2):299-309.

83. Chen S, Townsend K, Goldberg TE, Davies P, Conejero-Goldberg C. MAPT isoforms: differential transcriptional profiles related to 3R and 4R splice variants. J Alzheimers Dis. 2010;22(4):1313-29.

84. Kim HK, Pham MHC, Ko KS, Rhee BD, Han J. Alternative splicing isoforms in health and disease. Pflügers Archiv - European Journal of Physiology. 2018;470(7):995-1016.

85. Takeda N, O'Dea EL, Doedens A, Kim J-w, Weidemann A, Stockmann C, et al. Differential activation and antagonistic function of HIF-α isoforms in macrophages are essential for NO homeostasis. Genes & Development. 2010;24(5):491-501.

86. Eshel D, Toporik A, Efrati T, Nakav S, Chen A, Douvdevani A. Characterization of natural human antagonistic soluble CD40 isoforms produced through alternative splicing. Molecular immunology. 2008;46(2):250-7.

87. DiStefano JK. The Emerging Role of Long Noncoding RNAs in Human Disease. Methods in molecular biology (Clifton, NJ). 2018;1706:91-110.

88. Kramer NJ, Wang W-L, Reyes EY, Kumar B, Chen C-C, Ramakrishna C, et al. Altered lymphopoiesis and immunodeficiency in &lt;em&gt;miR-142&lt;/em&gt; null mice. Blood. 2015;125(24):3720.

89. Chitnis N, Clark PM, Kamoun M, Stolle C, Brad Johnson F, Monos DS. An Expanded Role for HLA Genes: HLA-B Encodes a microRNA that Regulates IgA and Other Immune Response Transcripts. Front Immunol. 2017;8(583).

90. Gonorazky HD, Naumenko S, Ramani AK, Nelakuditi V, Mashouri P, Wang P, et al. Expanding the Boundaries of RNA Sequencing as a Diagnostic Tool for Rare Mendelian Disease. American journal of human genetics. 2019;104(3):466-83.

91. Han H, Jiang X. Disease Biomarker Query from RNA-Seq Data. Cancer Inform. 2014;13(Suppl 1):81-94.

92. Kimmerling RJ, Lee Szeto G, Li JW, Genshaft AS, Kazer SW, Payer KR, et al. A microfluidic platform enabling single-cell RNA-seq of multigenerational lineages. Nature Communications. 2016;7(1):10220.

93. Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med. 2018;50(8):96-.

94. See P, Lum J, Chen J, Ginhoux F. A Single-Cell Sequencing Guide for Immunologists. Front Immunol. 2018;9:2425-.

95. Brodin P, Jojic V, Gao T, Bhattacharya S, Angel CJ, Furman D, et al. Variation in the human immune system is largely driven by non-heritable influences. Cell. 2015;160(1-2):37-47.

96. Duffy D, Rouilly V, Braudeau C, Corbiere V, Djebali R, Ungeheuer MN, et al. Standardized whole blood stimulation improves immunomonitoring of induced immune responses in multi-center study. Clinical immunology (Orlando, Fla). 2017;183:325-35.

97. Lee MN, Ye C, Villani A-C, Raj T, Li W, Eisenhaure TM, et al. Common Genetic Variants Modulate Pathogen-Sensing Responses in Human Dendritic Cells. Science. 2014;343(6175):1246980.

98. Urrutia A, Duffy D, Rouilly V, Posseme C, Djebali R, Illanes G, et al. Standardized Whole-Blood Transcriptional Profiling Enables the Deconvolution of Complex Induced Immune Responses. Cell reports. 2016;16(10):2777-91.

99. Khan S, Kuruvilla M, Hagin D, Wakeland B, Liang C, Vishwanathan K, et al. RNA sequencing reveals the consequences of a novel insertion in dedicator of cytokinesis-8. Journal of Allergy and Clinical Immunology. 2016;138(1):289-92.e6.

100. Hsu AP, Johnson KD, Falcone EL, Sanalkumar R, Sanchez L, Hickstein DD, et al. GATA2 haploinsufficiency caused by mutations in a conserved intronic element leads to MonoMAC syndrome. Blood. 2013;121(19):3830-7, s1-7.

101. Starokadomskyy P, Gemelli T, Rios JJ, Xing C, Wang RC, Li H, et al. DNA polymerase-α regulates the activation of type I interferons through cytosolic RNA:DNA synthesis. Nature immunology. 2016;17(5):495-504.