

Transcriptome analysis provides critical answers to the “variants of uncertain significance” conundrum

Mackenzie Postel¹, Julie O. Culver², Charité Ricker², and David Craig¹

¹University of Southern California Keck School of Medicine

²University of Southern California Norris Comprehensive Cancer Center

January 11, 2022

Abstract

The vast volume of data that has been generated as a result of the next-generation sequencing revolution is overwhelming to sift through and interpret. Parsing functional vs. non-functional and benign vs. pathogenic variants continues to be a challenge. Out of three billion bases, the genomes of two given individuals will only differ by about 3 million variants (0.1%). Furthermore, only a small fraction of these are biologically-relevant and, of those that are functional, only a handful actually drive disease pathology. While whole genome and exome sequencing have transformed our collective understanding of the role that genetics plays in disease pathogenesis, there are certain conditions and populations for whom DNA-level data has failed to produce a molecular diagnosis. Patients of non-White race/non-European ancestry are disproportionately affected by “variants of unknown/uncertain significance” (VUS). This limits the scope of precision medicine for minority patients and perpetuates health disparities. VUS often include deep intronic and splicing variants which are difficult to interpret in DNA alone. RNA analysis is capable of illuminating the consequences of VUS thereby allowing for their reclassification as pathogenic vs. benign. Here we review the critical role, going forward, of transcriptome analysis for clarifying VUS in both neoplastic and non-neoplastic diseases.

Title Page

Transcriptome analysis provides critical answers to the “variants of uncertain significance” conundrum

Authors

Mackenzie Postel^{1,2}, Julie O. Culver², Charité Ricker², David W. Craig^{1,2,*}

Affiliations

¹. Department of Translational Genomics, University of Southern California, Los Angeles, CA

². USC Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA 90033, USA;

Correspondence

*Correspondence: davidwcr@usc.edu

Abstract

The vast volume of data that has been generated as a result of the next-generation sequencing revolution is overwhelming to sift through and interpret. Parsing functional vs. non-functional and benign vs. pathogenic variants continues to be a challenge. Out of three billion bases, the genomes of two given individuals will only differ by about 3 million variants (0.1%). Furthermore, only a small fraction of these are biologically-relevant and, of those that are functional, only a handful actually drive disease pathology. While whole genome and exome sequencing have transformed our collective understanding of the role that genetics plays in disease pathogenesis, there are certain conditions and populations for whom DNA-level data has failed to produce a molecular diagnosis. Patients of non-White race/non-European ancestry are disproportionately affected by “variants of unknown/uncertain significance” (VUS). This limits the scope of precision medicine for minority patients and perpetuates health disparities. VUS often include deep intronic and splicing variants which are difficult to interpret in DNA alone. RNA analysis is capable of illuminating the consequences of VUS thereby allowing for their reclassification as pathogenic vs. benign. Here we review the critical role, going forward, of transcriptome analysis for clarifying VUS in both neoplastic and non-neoplastic diseases.

Keywords: splicing variants, deep intronic variants, variants of unknown significance, genetic ancestry

Main Body

Introduction

Next-generation sequencing (**NGS**) technology has fundamentally altered how scientists identify the genetic basis of disease. Rather than sequencing and interrogating one gene at a time, clinicians and researchers now routinely sequence at a genome-wide level and, in turn, face a new problem: isolating the critical disease-causing, pathogenic variants from millions of others. While most variants are easily excluded upon the basis of frequency and gene biology, there are many variants of unknown/uncertain significance (**VUS**) for which it is not evident whether a variant is pathogenic or benign.

The term VUS is an exceptionally loaded with definitions and guidelines coming from the American College of Medical Genetics (**ACMG**) (Richards et al., 2018), Association for Molecular Pathology (**AMP**) (M. M. Li et al., 2017), and International Agency for Research on Cancer (**IARC**) (Plon et al., 2008). VUS have been the subject of entire special issues of this journal, *Human Mutation*, such as in 2008. The precise meaning of VUS depends upon context. For example, in clinical testing entire variant classes (such as synonymous variants) are presumed likely benign at baseline, ensuring VUS are less frequently reported. Still, VUS can take a significant psychological and emotional toll on patients and their families when reported and may lead to confusion and clinical mismanagement. Importantly, as we will discuss, VUS affect minority populations disproportionately (Culver et al., 2013; Gelfman et al., 2017; Maurano et al., 2012; Park et al., 2018; Qian et al., 2021; Vaz-Drago et al., 2017). Therefore, VUS are an issue with far-reaching impact, both in research and clinical reporting. At a simple level, the sheer numbers of intronic, synonymous, and other deep variants contribute to their overall predictive value as a variant class. Additionally, it is well-recognized that many important disease-causing variants fall into these neglected “opaque” areas of variant interpretation.

There is an overwhelming need for better tools capable of elucidating the functional consequences of rare molecular events and, thereby, illuminating how genetic variation impacts biology and disease pathogenesis. Analysis of DNA-level information alone in the context of clinical phenotypes is unlikely to clarify this information given rare events’ limited sample size and lack of statistical power. Compared to genomic analysis – which outlines the repertoire of functions encoded in DNA that a cell *could* perform – transcriptomic analysis enables a deeper understanding of how a cell is *actually* functioning. In other words, RNA expression profiling provides critical insight into how genotype translates into phenotype (Mantione et al., 2014).

Here we will highlight how gene expression profiling can be used to reveal the presence of deep intronic

and splicing variants. We will review the emerging role of RNA sequencing as a tool to differentiate VUS as pathogenic vs. benign, especially in populations of diverse races, ethnicities, and genetic ancestries as individuals of non-European ancestry have significantly higher rates of VUS (Petrovski & Goldstein, 2016). Finally, we will explore the ways in which transcriptomic analysis can be translated into clinical diagnostics for cancer (i.e., detection of causative variants, both germline and somatic) and how it can be applied to address cancer disparities.

Leveraging the transcriptome

The term “RNA-seq” was coined in 2008 and, around this same time, several landmark studies hinted toward the potential diagnostic utility of RNA sequencing data using short-read NGS. One major effort was the ENCODE Project Consortium, which stated: “*to understand the human genome... and the ways in which its defects can give rise to disease, we need a more transparent view of the information it encodes*” (Birney et al., 2007). An understanding of human disease necessitates an understanding of the link between genomic variation and transcription; this was the goal of the Genotype-Tissue Expression (**GTEX**) project started in 2010 (Consortium, 2013). Analogous to the 1000 Genomes Project, which aimed to provide a catalogue of common variation, GTEx aimed to catalogue genotype and gene-expression correlations across dozens of tissues and hundreds (eventually thousands) of individuals.

An improved ability to predict the consequences of genetic variation on transcription looms as one of the largest opportunities and biggest challenges for addressing VUS. RNA analysis allows for the appreciation of phenomena like differential expression, allele-specific expression, alternative splicing, isoform switching, and many other events that would not be readily apparent in DNA data. Current approaches are largely limited to annotation of variant classes that are both rare and predictable for transcript disruption, such as nonsense mutations that lead to loss of transcript expression. More common variant classes are confounded by their higher prevalence and lower accuracy in predicting deleterious impact to gene expression.

Germline variants and somatic mutations can impact RNA species in a variety of ways, and they are often grouped into categories according to functional impact (i.e. High, Moderate, Low/Modifier) using annotation ontology software such as *ANNOVAR* (Wang et al., 2010; Yang & Wang, 2015) and *SnpEff* (Cingolani et al., 2012). Variants falling into the *High* functional impact category include nonsense variants introducing premature stop-codons, stop-loss, 5' untranslated region (**UTR**) premature start codons, and start-loss variants. Many of these lead to lowered transcription, but also full removal of transcripts through nonsense-mediated decay (**NMD**) (Chakravorty & Hegde, 2018). Another relatively rare variant class includes splice-site disrupting variants within the +/-2 base pair (**bp**) region of the exon-intron boundaries, which can impact splicing and result in exon skipping, exon usage, mutually exclusive/inclusive exons, and/or intron retention (Woolfe et al., 2010). Variants within the *Moderate* category typically are predicted to impact amino acid composition, including missense amino changes from non-synonymous variants and in-frame insertions/deletions.

The remainder of ontology variants are more frequently found per genome and are often pragmatically filtered out due to their low predictive value. These include synonymous variants, regulatory region variants, 3-prime (3') UTR variants, 5-prime (5') UTR variants, and intronic variants. It is essential to highlight splice region variants proximal to exon boundaries. Indeed, as shown in Figure 1A, variant alleles and exonic position well beyond the 2bp exon donor/acceptor window impact function. Variants annotated as missense and synonymous can alter splicing, exposing so-called cryptic splice-sites and splice enhancing or silencing elements (Woolfe et al., 2010). Moreover, deeper within the intron, variants as far out as 18 bp can impact spliceosome binding at a key motif referred to as the lariat junction.

Many of these low impact variants are found in higher numbers per genome and any given variant has a low probability of being functional. Still, there are numerous examples whereby variants drive disease via cryptic splice-sites (ex: *HBB*, IVS2+705T>G for β-thalasemia7), UTRs (ex: *GJB1* in Charcot-Marie-Tooth neuropathy), and other deep intronic variants (ex: *COL4A5*IVS6+1873G>A in Alport Syndrome) (Dobkin

et al., 1983; King et al., 2002). Frequently, the mechanism of impact is evident in transcriptome data. Indeed, novel RNA functions and phenomena have been elucidated that extend far beyond the traditional protein-coding role described by the “central dogma.” We are also beginning to better understand the ways in which messenger RNA (**mRNA**) can be processed – chiefly by alternative splicing, RNA editing, and crosstalk between the two processes – to create the great variety that characterizes the human transcriptome (Tang et al., 2020). Examples of unique phenomena into which gene expression data provides insight, and which have significant disease implications, include splicing mutations, deep intronic truncating mutations leading to NMD and loss of mRNA, and allele-specific expression (**ASE**).

Aberrant Splicing Mutations Beyond 2bp

Variants disrupting RNA splicing occur throughout the gene and, without RNA-based functional evidence, can be annotated into less obvious categories including as synonymous, intronic, and UTR variants. Functionally, splicing occurs via a complex of proteins and small nuclear RNAs called the spliceosome that form complementary RNA-RNA complexes with target RNAs (Anna & Monika, 2018). The spliceosome catalyzes the splicing of pre-mRNA into mRNA and circular lariat RNAs, the latter of which are destroyed (Talhouarne & Gall, 2018). Variants at the canonical boundary motifs between exons and introns are *GT*(at the 5' end) and *AG* (at the 3' end; Figure 1A) are generally prioritized but, as shown in this figure, variants upstream and downstream can disrupt function at a lower certainty/frequency (Anna & Monika, 2018). Adding to the mechanistic complexity is the existence of a variety of spliceosomes (“major” and “minor”) as well as noncanonical — aka “cryptic” or “pseudo” — splice site sequences (e.g. *GC-AG* and *AT-AC*). Which splice site is used largely depends upon the presence of cis-acting regulatory elements (splicing enhancers vs. suppressors), as well as the physical structures of a branch site and polypyrimidine tract (these sequences bind spliceosome proteins) (Anna & Monika, 2018).

In order to minimize false positive rate, candidate variant lists derived from DNA sequencing typically only take into consideration *canonical* changes to boundary motifs (as mentioned earlier: 5'-*GT* and 3'-*AG*) or variants that lie within 100 base pairs of canonical splice sites (that is to say *not* deep intronic variants). This approach is pragmatic, as these variants are both highly penetrant and rare leading to overall high precision (or positive predictive value). Without additional functional data, the large number of variants further down- or up-stream lead to overall low precision.

RNA-seq provides a direct functional measure of splicing, allowing for fewer variants, with higher inherent accuracy and precision, to be considered. Furthermore, it provides a direct measure of spliced events. As shown in Figure 1B , there are at least 7 classes of splicing events including: (1) exon skipping, (2) mutually-exclusive exons (i.e. a coordinated set of splicing events where, as a result, only one of two exon is retained), (3) exon scrambling, (4) intron retention, (5) alternative promoters and terminators, (6) alternative donor sites and (7) alternative acceptor sites (aka alternative 5' and 3' splice sites). It also allows for consideration of variants that are typically filtered out including synonymous variants, UTR variants, and so by supplementation with transcriptomic data. (Chen & Weiss, 2015; Pohl et al., 2013; Shi et al., 2018)

The emergence of RNA-seq in conjunction with germline DNA sequencing has led to optimization, development, and training of new bioinformatic tools for splice-site prediction and splice-site detection. Splice-site prediction focuses on using advanced algorithms and artificial intelligence (**AI**) methods to predict from DNA variants alone whether a variant impact splicing. Pragmatically, these approaches suffer from the sheer number of variants and have higher utility when RNA-seq data is also readily available, reducing the overall search space. A useful metric when it comes to assessing alternative splicing events is “**PSI** ” or percent spliced in; also known as the exon-inclusion ratio, PSI indicates how often a given exon occurs in all isoforms of the gene that contains said exon (Tanner et al., 2021). Splice variant prediction tools typically employ one of two strategies: adaptive models or random forest analysis. The essential difference between these two methods is whether one uses large databases from which non-pathogenic vs. pathogenic splice variants can be trained. Methods like those employed by Liu et al (Liu et al., 2016) utilize random forests to generate

probability scores that can be used to estimate the potential of a 3-12bp window within the lariat RNA junction. *DeepSplice* is an example of a tool that utilizes deep convolutional neural networks, as well as paired events to reduce false positives (Zhang et al., 2018). *SpliceAI* uses a deep neural network to model mRNA splicing from noncoding sites, yielding a 10% rate of pathogenic variant discovery in neurodevelopmental disorders (Jaganathan et al., 2019; Sanders et al., 2020). These are just a few emerging algorithms and, in practice, databases serve as vehicles for querying calculated predictive scores, such as with *dbNSFP* (Liu et al., 2020).

Other informatic algorithms directly utilize RNA-seq data in their analysis, such as *LeafCutterMD* (Jenkinson et al., 2020) and *FRASER* (Find RAre Splicing Events in RNA-seq (Mertes et al., 2021). *SpliceSeq* is an example of a tool that utilizes “splice graph” models (Ryan et al., 2012). Tools like *SpliceV* facilitate discovery and visualization of splicing events (Ungerleider & Flemington, 2019). Underlying these approaches lies layers of additional nuance with variation in alignment strategies (such as variant-aware aligners) and assembly (Hong et al., 2018). As methods rapidly evolve, implemented approaches vary substantially across and within labs, which can adversely impact database resources. Still, it is worth noting that such splicing analysis not only has diagnostic but also therapeutic utility; for instance, tumors with identified splicing variants could be targeted with therapies that inhibit the spliceosome, splicing regulatory proteins, or aberrant splicing products (Lee & Abdel-Wahab, 2016; Scotti & Swanson, 2016).

Indirect Insights From Allele-Specific Expression & Nonsense-Mediated Decay

Another indirect phenomenon into which RNA-seq can provide insight (and which DNA-only analysis cannot appreciate) is allele-specific expression (**ASE**). The initial definition of ASE was a purely germline one: preferential expression of one parental allele in a heterozygous individual (Shao et al., 2019). Fundamentally with RNA-seq, ASE requires heterozygous proxy variants, often single nucleotide polymorphism (**SNP**s), that allow phasing (or at least quantification) of the relative expression between maternally- and paternally-inherited chromosomes (Figure 2). These proxy-SNPs include common and rare synonymous, UTR, and missense variants. Importantly, ASE is the end-result and, in practice, one is frequently inferring loss of a pathogenic allele by monogenic expression of the wild-type allele. Processes such as NMD will actively remove transcripts containing pathogenic mutations.

Interestingly, detection of ASE can lead to discoveries of VUS and cancer-related genes. One study in individuals with hereditary pancreatic cancer identified a heterozygous SNP in *BRCA2*(rs144848) displaying ASE in RNA; the variant was a VUS and the mechanism behind it was determined to be a truncating mutation leading to NMD (Tan et al., 2008). NMD is the process whereby a trimeric complex of proteins degrades mRNA that would otherwise result in potentially-pathogenic truncated proteins; it has been posited that NMD also plays a role in the regulation of *normal* mRNA expression, as well as regulation of alternative splicing via degradation of splice variants with premature termination codons (Brogna & Wen, 2009). One common mechanism in the pathogenesis of cancer is a nonsense, somatic mutation in a tumor suppressor that leads to NMD and, by means of haploinsufficiency or dominant negative effects, loss of function.

One key historical example leading to ASE involves measuring X-chromosome inactivation or “skewing” (a physiological process normal in females) and genomic imprinting to gain insight into the preferential selection of variants along the X-chromosome (Shao et al., 2019). In genetic females, allele-specific expression also reflects the loss of transcript expression from one allele due to X-inactivation, an important methylation-driven “dosage compensation mechanism” (Shvetsova et al., 2019). Historically, X-skewing has been detected using the *HUMARA* assay, which takes advantage of differential methylation at the androgen receptor locus. However, RNA-seq along expressed X-chromosome genes provides additional insights. One can directly observe skewing directionality (ex: 95% skewing in X and preferential expression of wild-type alleles) (Szeleniger et al., 2014).

Recognizing that there is natural variability in skewing in healthy females, information regarding X-inactivation is less discriminatory but still useful in unexpected ways. For example, in assessing a candidate

variant within the X-linked *TAF1* in a boy, our group leveraged RNA-seq data for the boy's healthy mother which showed extreme X-skewing towards the protective wild-type allele and away from our candidate VUS (Hurst et al., 2018). In other words, RNA-seq was used to evaluate a model where X-skewing protected the mother. Further sequencing showed that the *de novo* event founded with her and was likely selected, given a history of unsuccessful pregnancies by the boy's parents. This example has generalizability and is relevant to other X-linked disorders, including Alport, Charcot–Marie–Tooth, Fabry, Fanconi, Fragile X, Hunter, Rett, and Wiskott–Aldrich syndromes (Migeon, 2020).

Outlier Analysis as a Solution to Power Problems

Case-control studies are feasible for *common* variants but are far more difficult to accomplish for *rare* VUS where the controls overwhelmingly outnumber the cases. A number of solutions have been put forth to solve this problem. These include likelihood ratio tests, burden tests with genetic scores, adaptive burden tests with data-adaptive weights/thresholds, variance-component tests, exponential combination tests, normal transformation with trait “winsorization” (to find a balance between Type I error and statistical power), or a combination thereof (Auer et al., 2016; Li et al., 2021).

Alternatively, one key solution to address the inherent lack of statistical power when it comes to analyzing rare variants is to use *gene expression outlier* analysis. Rare variants are often associated with extremes of expression, whether over- or under-expression (X. Li et al., 2017). Variants can be interpreted through careful integration of DNA data with RNA-seq data from other patients or from public resources. Cummings et al. (2017) were one of the first to demonstrate the fact that consideration of such “outlier” variants in RNA-seq data results in an improved diagnostic rate (35% in that study, specifically) in patients for whom a molecular diagnosis cannot be made from DNA-data alone.(Cummings et al., 2017) Gene expression outlier analysis has often been employed in the field of cancer genomics to identify cancer drivers for a specific subset of cancer types or cancer outliers (Alshalalfa et al., 2012; Mori et al., 2013). Fundamentally, outlier analysis does not actually identify a mechanism but rather a gene that is fundamentally different from others expressed in a cohort. As will be discussed later, there is a fundamental presumption that the comparison cohort is relevant.

THE Clinical UTILITY and diagnostic yield of rna

While *analytical* validity refers to the sensitivity, specificity, and accuracy of a diagnostic test in terms of its ability to measure a biomarker in a lab setting, *clinical* validity refers to the accuracy and predictive value of that test when it comes predicting clinical diagnosis. Both terms are distinct from *clinical utility*, which refers to a test's ability to make a difference – that is its potential to impact patient quality of care/life by guiding clinical decision-making (Byron et al., 2016). Here we will outline the clinical utility of transcriptome analysis for the diagnosis of both neoplastic and non-neoplastic diseases.

Mendelian Disorders

Transcriptome analysis is a boon to the diagnosis of rare Mendelian diseases. Historically, genetic counseling has relied upon whole exome sequencing to identify causative disease variants; however, this DNA-only approach has left up to 75% of patients without genetic diagnoses (Stenton & Prokisch, 2020). When integrated with genome sequencing – and, especially, in situations when said genome sequencing encompasses both exons and introns – gene expression profiling has been shown to significantly boost molecular diagnostic rates; yields have been shown to increase by 10-35% (Lee et al., 2020; Maddirevula et al., 2020; Stenton & Prokisch, 2020). This is because RNA data both: (1) puts any variants identified in DNA into context by revealing their transcript-level consequences (ex: allele-specific expression due to nonsense-mediated decay, imprinting, and/or expression of splice variants), and (2) illuminates phenomena (like gene expression

outliers) that may not pass the threshold of detection in DNA data alone (but that are crucial to the pathogenesis of a given disease) (Lee et al., 2020).

Gene expression profiling has also improved clinicians' ability to diagnose, stratify, and subtype autoimmune diseases, like systemic lupus erythematosus, as well as degenerative diseases like Age-Related Macular Degeneration (Alarcón-Riquelme, 2019). Additionally, consideration of the transcriptomic landscape has shed light on the fact that many of these diseases are heterogenous with a spectrum of causative molecular events (Morello et al., 2019). RNA-seq is also capable of overcoming the “bottleneck of variant interpretation” in patients with inborn errors of metabolism, mitochondrialopathies, and/or unsolved muscle disorders, leading to significantly increased diagnostic yields (Kremer et al., 2018; Thompson et al., 2020).

It is important to note that recent studies have shown, particularly in the case of monogenetic neuromuscular disorders, that blood-based RNA-seq is not sufficient for diagnosis; however, RNA-seq performed on myotubes generated by trans-differentiation of patient fibroblasts was capable of identifying a molecular culprit (predominantly splicing variants) in 36% of patients for whom DNA-only analysis had failed to do so (Gonorazky et al., 2019). This highlights the fact that several methodological improvements must be made to hasten the progress of translating transcriptome analysis from the benchtop to the bedside, and to enhance diagnostic sensitivity. These include refinement of *ex vivo* trans-differentiation of accessible cells to more disease-relevant cell types (Lee et al., 2020).

Hereditary Cancer

Cancer genomic analysis involves the identification of inherited (“germline”) risk variants and acquired (“somatic”) mutations in DNA and RNA (Koeppel et al., 2018). Transcriptome analysis has been shown to be capable of identifying rare, causative variants by revealing changes in splicing and gene expression that were undetected by DNA sequencing (Yuan et al., 2020). Since examples of RNA-seq analysis in the conjunction of cancer risk prediction is more recent, we will dissect these papers in greater detail. Before we do, two key distinctions should be made regarding hereditary cancer studies. First, there is a general bias towards using RNA-sequencing in conjunction with panel-based clinical sequencing, to reduce the genomic search space considerably. If one focuses on all oncogenic or tumor-suppressor genes it changes the prevalence of background events, and ultimately the precision and/or diagnostic yield. Second, the context of reporting is distinct from Mendelian studies, a search or diagnostic odyssey. Typically, hereditary cancer VUS create unique stress, and there is some implied interpretation of negative findings.

A series of papers from 2019 through 2021 illustrate and give further insights into these distinctions. First, Conner et al. (2019) found that, by supplementing DNA genetic testing with RNA, heterozygous duplication events in *MSH2* – which were previously classified as VUS in five individuals with Lynch Syndrome – were able to be reclassified as pathogenic or likely pathogenic (Conner et al., 2019). Similarly, Karam et al. (2019) showed that, by supplementing DNA with RNA genetic testing in cases suspicious for hereditary cancer in which the variant in question involved a potential splice site alteration, (1) inconclusive DNA-based results were resolved in 49 of 56 inconclusive cases (88%) studied, with 26 (47%) being reclassified as clinically-actionable and 23 (41%) being clarified as benign; (2) the study estimated that 2% of patients receiving paired DNA/RNA testing would benefit by the addition of RNA by further characterization of splice-site VUS (Karam et al., 2019). Two other studies found that the addition of transcriptomic analysis to hereditary cancer testing enabled 60% and 20%, respectively, of splicing VUS to be reclassified as (likely) pathogenic (Agiannitopoulos et al., 2021; Rofes et al., 2020). Landrith et al. (2020) performed germline RNA-seq to profile 18 genes (i.e. *APC*, *ATM*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *MLH1*, *MSH2*, *MSH6*, *MUTYH*, *NF1*, *PALB2*, *PMS2*, *PTEN*, *RAD51C*, *RAD51D*, and *TP53*) in patients with suspected hereditary cancer syndromes. The investigators demonstrated a 9.1% relative increase in the detection of pathogenic variants afforded by augmenting DNA data with RNA analysis (Landrith et al., 2020). Deep intronic variants have also been identified in *BRCA1/2*, by virtue of RNA analysis, in patients with familial breast and ovarian cancers (Anczuków et al., 2012; Montalban et al., 2019).

As is evident from the studies mentioned above, RNA deep intronic mutations and splicing aberrations are unique mechanisms of carcinogenesis which, based upon DNA data alone, are still often classified as VUS (Urbanski et al., 2018). Splicing mutations (which can be present in both pre-mRNA exons and introns (the latter of which has historically been harder to detect using traditional DNA analyses) lead to abnormal mRNA phenomena (e.g. exon skipping, intron inclusion, cryptic splice site activation) and the production of abnormal proteins with diagnostic value (Shi et al., 2018). Expression changes in splicing regulators can be used as biomarkers for cancer diagnosis (ex:*hnRNPA2/B1*, an RNA-binding protein involved in mRNA splicing, is a sensitive and specific early-diagnostic marker of lung neoplasms) (Zhang et al., 2021). RNA-seq has shown utility in diagnosing germline splicing variants in hereditary cancer genes that were not evident in DNA analysis (Urbanski et al., 2018). While splicing variants make up 11% of hereditary cancer gene VUS, they make up 55% of those VUS that are “likely pathogenic”(Parsons et al., 2019).

Larger-scale reports have been published by clinical genetic companies where RNA-seq was used in conjunction with panel-based studies across thousands of individuals. Ambry recently released a series of “RNA Case Studies” that demonstrate the clinical diagnostic utility of transcriptomic data, particularly for identifying intronic variants (AmbryGenetics, 2019). One such scenario was the case of a 33-year-old male, with a personal and family history of colon polyps, for whom no clinically-significant variants could be detected via DNA-only analysis. When genetic analysis was supplemented with transcriptomic analysis (i.e. Ambry’s *+RNAinsight* ® panel), however, abnormal *APC* transcripts were detected prompting further investigation via targeted Sanger DNA sequencing. This resulted in the confirmation of a deep intronic, likely pathogenic variant. Transcriptomic data enabled the patient’s provider to make a genetic diagnosis of familial adenomatous polyposis.(AmbryGenetics, 2019) Other examples include a likely pathogenic intronic variant that was identified outside of DNA analytical range in the gene *ATM*(c.497-2661A>G), and exon skipping variants in *MSH6*leading to Lynch Syndrome. Ambry’s *+RNAinsight* ®panel, mentioned in the 2 cases above, analyzes 91 cancer driver genes, and can be paired with most DNA panels; it has shown to be capable of reclassifying >70% of VUS (AmbryGenetics, 2021).

Similarly, a recent study by Invitae aimed to exemplify the utility of RNA analysis for reclassifying splicing VUS (Truty et al., 2021). The investigators analyzed a significantly large sample consisting of nearly 700k patients from a clinical cohort plus individuals from two large public datasets (i.e. *ClinVar* and Genome Aggregation Database/gnomAD) (Truty et al., 2021). In their clinical cohort, Invitae found that 5.4% of individuals had at least one splicing VUS (most of which were identified outside of essential splice sites), and that splicing variants represented 13% of all variants classified as (likely) pathogenic or VUS. They estimated that, in the clinical cohort, RNA analysis would be capable of clarifying/reclassifying splicing VUSs in 1.7% of cases. In comparison to the clinical cohort, in *ClinVar* and *gnomAD* , Invitae observed that splicing VUS comprised nearly 5% and 9% of reported variants, respectively. Invitae concluded that, in all 3 cohorts, individuals would have a tangible, clinical-diagnostic benefit from RNA testing (Truty et al., 2021).

Not only can transcriptome characterization classify VUS as (likely) pathogenic, but it can also clarify variants as *benign* . For example, RNA data supported a variant *downgrade* of a likely pathogenic splice site variant at a canonical splice site (Shamseldin et al., 2021). In the case of *CDH1* c.387+1G>A, various clinical laboratories initially reported the variant in multiple Hispanic/Latino patients as “likely pathogenic” on the basis of the “+1” position of the variant. This led to the diagnosis of hereditary diffuse gastric cancer syndrome, a condition requiring complex management because of its association with a very high risk of early onset gastric cancer and lobular breast cancer. However, the variant was studied in more detail because the patients with this variant lacked the associated phenotype of the condition. The variant was experimentally demonstrated to result in the activation of a cryptic in-frame donor splice site, leading to the recommendation by ACMG and AMP that variants at this position not be considered as likely pathogenic (Maoz et al., 2016).

In large part, we have limited this review to germline-inherited variation due to space and scope. However, clearly, RNA-sequencing has utility in the context of somatic variation, and, in fact, this can be the basis of treatment decisions. It is worth highlighting that a 2021 study in *Oncogene* examined somatic variation

across over 1,000 pan-cancer, paired whole genomes and transcriptomes to understand the role of splicing mutations in tumorigenesis. The investigators identified about 700 somatic intronic mutations; nearly half were within deep intronic regions and, of those, 38% activated cryptic splice sites. A subset of the deep intronic mutations resulted in splicing enhancers or silencers alterations. They found that intronic mutations often affected tumor suppressor genes, and those hematological malignancies, particularly, harbor many deep intronic mutations. Taken as a whole, this paper suggests considerable insights can be gained well beyond germline analysis of VUS (Jung et al., 2021).

Limitations & Future Directions

The progress of RNA-based diagnostics is encouraging, especially as new and translational gene expression profiling techniques emerge (Wang et al., 2020). Gene expression profiling allows for, not only, the identification of fusion transcripts, but also the detection of phenomena like differential expression, ASE, alternative splicing, and the presence of non-coding RNAs (Conner et al., 2019). Both targeted RNA microarrays and RNA-seq have shown analytical validity when it comes to diagnostics for pediatric, adolescent/young adult, and adult patients (Vaske et al., 2019).

Conflicting Lines of Evidence

One fallacy of reasoning – commonly and erroneously applied to the analysis of variant lists such as variant call format (**VCF**) files – is the assumption that the *absence* of a transcript variant means that the variant is absent from the specimen. This common misconception lead to the development of genomic VCFs (**gVCFs**) which call every position – both variants and wild type/reference.

The only way to move forward with statistical power and confidence is through collaborative efforts and the creation of diverse and devoted databases. *ClinVar* (Rehm et al., 2017) and *gnomAD*(Karczewski et al., 2020) are under-appreciated summary-level datasets. *gnomAD*’s focus on categorizing rare events was foundational. At the RNA-level, this approach has not yet been adopted outside of isolated cases; burgeoning examples are *RNAcentral* (a database of non-coding RNAs) (Petrov et al., 2015) and *SpliceDB* (a database of canonical and non-canonical mammalian splice sites) (Burset et al., 2001).

With the clinical implementation of any new “translational” technology, one must approach variant curation and interpretation of functional evidence with caution. Interpretation can be more complex than anticipated; there are many potential pitfalls. For example, Nix et al. once posited that a partial exon-skipping mutation identified in *BRCA2* was pathogenic; it was later found to occur in many healthy controls (Mundt et al., 2017).

Differences in RNA-seq Library Preparation & Analysis Methods

Unlike genomic sequencing of DNA, differences in collection methods, library preparation, tissue sources, etc. massively impact RNA-seq analysis and interpretation. The first and most apparent variable is the tissue source for RNA and its relevance to the disease or phenotype. For example, how well can RNA from whole-blood provide insights into neurological disorders? GTEx provides an initial framework to evaluate this question showing typically >40% of genes expressed at reasonably high levels, and experiences reviewed in previous sections frequently faced a similar question (Consortium, 2013). Likely, customized assays leveraging enrichment may increase this dynamic range of RNA species, recognizing many genes will not have the expression needed for interpretation via RNA-seq. Nonetheless, many of the studies highlighted showed >10% improvement in diagnostic yield despite such changes.

Without question, the ability to look across rare DNA variation across thousands of individuals, such as through resources like gnomAD, has profoundly influenced the interpretation of genomic variants. Aggregation of RNA - even within the same lab will face significant and un-ignorable challenges. As has been experienced by consortiums and labs, aggregation of RNA-seq across samples, studies, and library preps

typically recapitulates multiple technical variables to drive the largest proportion. Efforts to normalize or adjust to these technical differences are an active area of research beyond the scope of this review.

Even still, when examining consortiums such as PsychENCODE (Psych et al., 2015) and AMP-AD (Hodes & Buckholtz, 2016), among others, eliminating technical variation from RNA-seq experiments is challenging, particularly if one is interested in rare events. To illustrate this point, we consider the recent release of 4,871 longitudinally-collected samples from 1,570 clinically-phenotyped individuals from the Parkinson’s Progression Marker Initiative (PPMI), conducted using random priming for PaxGene collected whole-blood with paired whole-genome sequencing (Craig et al., 2021). Forthcoming efforts from TopMED will utilize the same PaxGene whole-blood protocols but will differ in using mRNA-seq from poly-A priming. These two methods lead to different species with random priming, showing pre-spliced RNA and non-polyA-tailed transcripts. Algorithms trained on these methods will fundamentally differ in their core measures, such as PSI. Even within the same dataset, we have observed significant differences in gene/exon usage that depended on read lengths of paired 100bp vs. a 125bp subset.

While daunting, solutions are emerging for aggregating RNA such as through the ARCHS⁴ aggregation across mouse and human RNA-seq studies (Lachmann et al., 2018). Other examples include in-house solutions or those specific to a given group; it becomes a question of sensitivity. Our group successfully employed outlier analysis to identify causative variants in a cohort collected over 5 years that was sequenced by different labs using different methods.

Fragmentation of RNA-seq Databases and Standards

Though the RNA-based diagnostics described here have potential, there are still obstacles that must be overcome before they will be incorporated into routine clinical practice. These challenges include the need for scientific rigor, reproducibility, accuracy, precision, clinical validity, and clinical utility. Standards must be created for test thresholds and normalized reporting, and databases must be established (Tahiliani et al., 2020; Wang et al., 2020). These databases must be designed so as to not fall prey to any logical fallacies (ex: the “marker-positive fallacy”).

Issues of database size, diversity, and representation (both in the sense of race/ethnicity and cases/controls), population structure, and cryptic relatedness must be considered (Update., 1996). We must also acknowledge, and attempt to address, limitations (ex: the half-life/stability of RNA) and potential confounders (e.g. temporal changes in RNA expression, differences in RNA capture from fresh frozen vs. formalin fixed paraffin embedded samples, and phenomena like clonal hematopoiesis of indeterminate potential in liquid biopsies) (Wang et al., 2020).

Investigators must carefully consider the tissue from which they are isolating RNA given the fact that expression patterns differ across tissues (and, on the circadian-level, RNA expression can even differ in the same tissue at different time points) (Maddirevula et al., 2020). It is important to balance preference for minimally-invasive techniques with considerations of differential tissue expression. One recent study found that, when comparing brain vs. blood vs. human B-lymphoblastoid cell lines (**LCL**), LCLs possessed isoform diversity for neurodevelopmental genes similar to that of brain tissue; LCLs also expressed these genes more highly compared to blood (Rentas et al., 2020). The authors of this paper described an RNA-seq pipeline with 90% sensitivity and claimed that findings in LCLs outperformed those in blood and had implications for the molecular diagnosis of >1000 genetic syndromes (Rentas et al., 2020).

Another limitation is the fact that expression quantitative trait loci (**eQTL**) databases – like GTEx Portal – are limited to common variants (i.e. variants with a minor allele frequency >1%). This means that such datasets are not applicable toward understanding VUS which, although rare in the general/overall population, disproportionately impact Non-White/European groups. RNA analysis is also limited by the fact that most tools utilize transcripts defined by a Gene Transfer Format (GTF) file and find it difficult to annotate the 3' untranslated region (3' UTR) (Shenker et al., 2015). Therefore, there is a critical need for more rigorous, reproducible, and representative RNA databases and tools.

VUS as a Manifestation of Cancer Disparities

One anecdotal trend that we have noticed within our own group and across collaborative efforts is that RNA data allows for the identification of previously missed variation particularly in individuals of non-European ancestry. For example, in *Human Mutation* we reported a variant within 3bp of the exon boundary using an outlier approach in individuals of African ancestry. The molecular consequences of this variant included exon skipping, altered isoform usage, and loss of canonical isoform expression – events not evident in DNA data alone (McCullough et al., 2020). Patients who self-identify as Hispanic/Latinx, Black/African, and Asian/Pacific Islander experience more advanced stage disease at time of screening, significantly lower diagnostic yields, and higher rates of VUS and variant reclassification compared to their European/Caucasian counterparts (Dutil et al., 2019; Kinney et al., 2018; Kowalski et al., 2019; Marco-Puche et al., 2019; Ndugga-Kabuye & Issaka, 2019; Roberts et al., 2020; Slavin et al., 2018; Urbina-Jara et al., 2019). Individuals from non-European populations will have more private variation for one of three reasons: (1) they are poorly represented in reference datasets, (2) they have greater African ancestry, or (3) they come from a population that has undergone recent expansions (ex: Bangladesh) (Halperin et al., 2017).

A recent study reported by Ambry Genetics found that their *BRCAplus*, *BreastNext*, and *CancerNext* panels yielded [?]2-3x fewer VUS for Non-Hispanic whites than for minority populations (AmbryGenetics, 2017). Another study reports VUS frequencies in the tumor suppressor genes *BRCA1/2* to be 4.4% in Caucasians, 8.9% in African Americans, and 8.0% in Hispanic/Latinos; for larger hereditary cancer panels, this study reported VUS frequencies of 22.1% in Caucasians, 30.3% in African Americans, and 24.9% in Hispanics/Latinos (Appelbaum et al., 2020).

One important distinction to make here is the difference between race/ethnicity and genetic ancestry. While race and ethnicity are social constructs, ancestry is a biological/genetic construct resulting from human migrations throughout history resulting in biogeographical genetic variation (Batai et al., 2021). An example of how genetic ancestry can further clarify race/ethnicity-based disparities is the fact that higher African ancestry in Hispanic/Latinos (who are typically “admixed” with genetic contributions from African, European, and American Indian aka Native/Indigenous American ancestries) is associated with more aggressive breast cancer subtypes and a greater likelihood of receiving inconclusive VUS during genetic testing (Chapman-Davis et al., 2021; Dutil et al., 2019; Kinney et al., 2018; Kowalski et al., 2019; Marco-Puche et al., 2019; Ndugga-Kabuye & Issaka, 2019; Roberts et al., 2020; Slavin et al., 2018; Urbina-Jara et al., 2019; Virlogeux et al., 2015). Gene expression profiling may be able to help shed light on and alleviate these inequities (Fresard et al., 2019; Wai et al., 2020).

Conclusions

VUS cause significant psychological distress to patients and disproportionately limit the promise of precision medicine for minority patients (Landry et al., 2018). RNA data provides critical answers to the question of VUS, particularly in terms of clarifying deep intronic and splicing variants as pathogenic vs. benign. This necessitates the development of more rigorous, reproducible, and representative RNA databases and analytical tools.

conflict of interest Statement

The authors have no conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

REFERENCES

- Agiannitopoulos, K., Pepe, G., Papadopoulou, E., Tsaousis, G. N., Kampouri, S., Maravelaki, S., Fassas, A., Christodoulou, C., Iosifidou, R., Karageorgopoulou, S., Markopoulos, C., Natsiopoulos, I., Papazisis, K., Vasilaki-Antonatou, M., Venizelos, V., Ozmen, V., Tansan, S., Kaban, K., Eniu, D. T., Chiorean, A., & Nasioulas, G. (2021, 2021 May-Jun). Clinical Utility of Functional RNA Analysis for the Reclassification of Splicing Gene Variants in Hereditary Cancer. *Cancer Genomics Proteomics*, 18 (3), 285-294. <https://doi.org/10.21873/cgp.20259>
- Alarcon-Riquelme, M. E. (2019, 02 26). New Attempts to Define and Clarify Lupus. *Curr Rheumatol Rep*, 21 (4), 11. <https://doi.org/10.1007/s11926-019-0810-4>
- Alshalalfa, M., Bismar, T. A., & Alhajj, R. (2012). Detecting cancer outlier genes with potential rearrangement using gene expression data and biological networks. *Adv Bioinformatics*, 2012 , 373506. <https://doi.org/10.1155/2012/373506>
- Anczukow, O., Buisson, M., Leone, M., Coutanson, C., Lasset, C., Calender, A., Similnikova, O. M., & Mazoyer, S. (2012, Sep 15). BRCA2 deep intronic mutation causing activation of a cryptic exon: opening toward a new preventive therapeutic strategy. *Clin Cancer Res*, 18 (18), 4903-4909. <https://doi.org/10.1158/1078-0432.CCR-12-1100>
- Anna, A., & Monika, G. (2018, Aug). Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J Appl Genet*, 59 (3), 253-268. <https://doi.org/10.1007/s13353-018-0444-7>
- Appelbaum, P. S., Paren, E., Berger, S. M., Chung, W. K., & Burke, W. (2020, 03). Is there a duty to reinterpret genetic data? The ethical dimensions. *Genet Med*, 22 (3), 633-639. <https://doi.org/10.1038/s41436-019-0679-7>
- Auer, P. L., Reiner, A. P., & Leal, S. M. (2016, 08). The effect of phenotypic outliers and non-normality on rare-variant association testing. *Eur J Hum Genet*, 24 (8), 1188-1194. <https://doi.org/10.1038/ejhg.2015.270>
- Batai, K., Hooker, S., & Kittles, R. A. (2021, 06). Leveraging genetic ancestry to study health disparities. *Am J Phys Anthropol*, 175 (2), 363-375. <https://doi.org/10.1002/ajpa.24144>
- Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigo, R., Gingeras, T. R., Margulies, E. H., Weng, Z., Snyder, M., Dermitzakis, E. T., Thurman, R. E., Kuehn, M. S., Taylor, C. M., Neph, S., Koch, C. M., Asthana, S., Malhotra, A., Adzhubei, I., Greenbaum, J. A., Andrews, R. M., Flieck, P., Boyle, P. J., Cao, H., Carter, N. P., Clelland, G. K., Davis, S., Day, N., Dhami, P., Dillon, S. C., Dorschner, M. O., Fiegler, H., Giresi, P. G., Goldy, J., Hawrylycz, M., Haydock, A., Humbert, R., James, K. D., Johnson, B. E., Johnson, E. M., Frum, T. T., Rosenzweig, E. R., Karnani, N., Lee, K., Lefebvre, G. C., Navas, P. A., Neri, F., Parker, S. C., Sabo, P. J., Sandstrom, R., Shafer, A., Verte, D., Weaver, M., Wilcox, S., Yu, M., Collins, F. S., Dekker, J., Lieb, J. D., Tullius, T. D., Crawford, G. E., Sunyaev, S., Noble, W. S., Dunham, I., Denoeud, F., Reymond, A., Kapranov, P., Rozowsky, J., Zheng, D., Castelo, R., Frankish, A., Harrow, J., Ghosh, S., Sandelin, A., Hofacker, I. L., Baertsch, R., Keefe, D., Dike, S., Cheng, J., Hirsch, H. A., Sekinger, E. A., Lagarde, J., Abril, J. F., Shahab, A., Flamm, C., Fried, C., Hackermuller, J., Hertel, J., Lindemeyer, M., Missal, K., Tanzer, A., Washietl, S., Korbel, J., Emanuelsson, O., Pedersen, J. S., Holroyd, N., Taylor, R., Swarbreck, D., Matthews, N., Dickson, M. C., Thomas, D. J., Weirauch, M. T., Gilbert, J., Drenkow, J., Bell, I., Zhao, X., Srinivasan, K. G., Sung, W. K., Ooi, H. S., Chiu, K. P., Foissac, S., Alioto, T., Brent, M., Pachter, L., Tress, M. L., Valencia, A., Choo, S. W., Choo, C. Y., Ucla, C., Manzano, C., Wyss, C., Cheung, E., Clark, T. G., Brown, J. B., Ganesh, M., Patel, S., Tammana, H., Chrast, J., Henrichsen, C. N., Kai, C., Kawai, J., Nagalakshmi, U., Wu, J., Lian, Z., Lian, J., Newburger, P., Zhang, X., Bickel, P., Mattick, J. S., Carninci, P., Hayashizaki, Y., Weissman, S., Hubbard, T., Myers, R. M., Rogers, J., Stadler, P. F., Lowe, T. M., Wei, C. L., Ruan, Y., Struhl, K., Gerstein, M., Antonarakis, S. E., Fu, Y., Green, E. D., Karaoz, U., Siepel, A., Taylor, J., Liefer, L. A., Wetterstrand, K. A., Good, P. J., Feingold, E. A., Guyer, M. S., Cooper,

G. M., Asimenos, G., Dewey, C. N., Hou, M., Nikolaev, S., Montoya-Burgos, J. I., Loytynoja, A., Whelan, S., Pardi, F., Massingham, T., Huang, H., Zhang, N. R., Holmes, I., Mullikin, J. C., Ureta-Vidal, A., Paten, B., Seringhaus, M., Church, D., Rosenbloom, K., Kent, W. J., Stone, E. A., Batzoglou, S., Goldman, N., Hardison, R. C., Haussler, D., Miller, W., Sidow, A., Trinklein, N. D., Zhang, Z. D., Barrera, L., Stuart, R., King, D. C., Ameur, A., Enroth, S., Bieda, M. C., Kim, J., Bhinge, A. A., Jiang, N., Liu, J., Yao, F., Vega, V. B., Lee, C. W., Ng, P., Yang, A., Moqtaderi, Z., Zhu, Z., Xu, X., Squazzo, S., Oberley, M. J., Inman, D., Singer, M. A., Richmond, T. A., Munn, K. J., Rada-Iglesias, A., Wallerman, O., Komorowski, J., Fowler, J. C., Couttet, P., Bruce, A. W., Dovey, O. M., Ellis, P. D., Langford, C. F., Nix, D. A., Euskirchen, G., Hartman, S., Urban, A. E., Kraus, P., Van Calcar, S., Heintzman, N., Kim, T. H., Wang, K., Qu, C., Hon, G., Luna, R., Glass, C. K., Rosenfeld, M. G., Aldred, S. F., Cooper, S. J., Halees, A., Lin, J. M., Shulha, H. P., Xu, M., Haidar, J. N., Yu, Y., Iyer, V. R., Green, R. D., Wadelius, C., Farnham, P. J., Ren, B., Harte, R. A., Hinrichs, A. S., Trumbower, H., Clawson, H., Hillman-Jackson, J., Zweig, A. S., Smith, K., Thakkapallayil, A., Barber, G., Kuhn, R. M., Karolchik, D., Armengol, L., Bird, C. P., de Bakker, P. I., Kern, A. D., Lopez-Bigas, N., Martin, J. D., Stranger, B. E., Woodroffe, A., Davydov, E., Dimas, A., Eyras, E., Hallgrimsdottir, I. B., Huppert, J., Zody, M. C., Abecasis, G. R., Estivill, X., Bouffard, G. G., Guan, X., Hansen, N. F., Idol, J. R., Maduro, V. V., Maskeri, B., McDowell, J. C., Park, M., Thomas, P. J., Young, A. C., Blakesley, R. W., Muzny, D. M., Sodergren, E., Wheeler, D. A., Worley, K. C., Jiang, H., Weinstock, G. M., Gibbs, R. A., Graves, T., Fulton, R., Mardis, E. R., Wilson, R. K., Clamp, M., Cuff, J., Gnerre, S., Jaffe, D. B., Chang, J. L., Lindblad-Toh, K., Lander, E. S., Koriabine, M., Nefedov, M., Osoegawa, K., Yoshinaga, Y., Zhu, B., de Jong, P. J., Consortium, E. P., Program, N. C. S., Center, B. C. o. M. H. G. S., Center, W. U. G. S., Institute, B., & Institute, C. s. H. O. R. (2007, Jun 14). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, *447* (7146), 799-816. <https://doi.org/10.1038/nature05874>

Brogna, S., & Wen, J. (2009, Feb). Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat Struct Mol Biol*, *16* (2), 107-113. <https://doi.org/10.1038/nsmb.1550>

Burset, M., Seledtsov, I. A., & Solovyev, V. V. (2001, Jan 01). SpliceDB: database of canonical and non-canonical mammalian splice sites. *Nucleic Acids Res*, *29* (1), 255-259. <https://doi.org/10.1093/nar/29.1.255>

Byron, S. A., Van Keuren-Jensen, K. R., Engelthaler, D. M., Carpten, J. D., & Craig, D. W. (2016, May). Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nat Rev Genet*, *17* (5), 257-271. <https://doi.org/10.1038/nrg.2016.10>

Chakravorty, S., & Hegde, M. (2018, Jun). Clinical Utility of Transcriptome Sequencing: Toward a Better Diagnosis for Mendelian Disorders. *Clin Chem*, *64* (6), 882-884. <https://doi.org/10.1373/clinchem.2017.276980>

Chapman-Davis, E., Zhou, Z. N., Fields, J. C., Frey, M. K., Jordan, B., Sapra, K. J., Chatterjee-Paer, S., Carlson, A. D., & Holcomb, K. M. (2021, 01). Racial and Ethnic Disparities in Genetic Testing at a Hereditary Breast and Ovarian Cancer Center. *J Gen Intern Med*, *36* (1), 35-42. <https://doi.org/10.1007/s11606-020-06064-x>

Chen, J., & Weiss, W. A. (2015, Jan 02). Alternative splicing in cancer: implications for biology and therapy. *Oncogene*, *34* (1), 1-14. <https://doi.org/10.1038/onc.2013.570>

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012, Apr-Jun). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnPEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*, *6* (2), 80-92. <https://doi.org/10.4161/fly.19695>

Conner, B. R., Hernandez, F., Souders, B., Landrith, T., Boland, C. R., & Karam, R. (2019, 05). RNA Analysis Identifies Pathogenic Duplications in MSH2 in Patients With Lynch Syndrome. *Gastroenterology*, *156* (6), 1924-1925.e1924. <https://doi.org/10.1053/j.gastro.2019.01.248>

Consortium, G. T. (2013, Jun). The Genotype-Tissue Expression (GTEx) project. *Nat Genet*, *45* (6), 580-585. <https://doi.org/10.1038/ng.2653>

- Craig, D. W., Keuren-Jensen, K. V., & Initiative, P. P. M. (2021). RNA sequencing of whole blood reveals early alterations in immune cells and gene expression in Parkinson's disease. *Nature Aging*, 1, 734–747.
- Culver, J. O., Brinkerhoff, C. D., Clague, J., Yang, K., Singh, K. E., Sand, S. R., & Weitzel, J. N. (2013, Nov). Variants of uncertain significance in BRCA testing: evaluation of surgical decisions, risk perception, and cancer distress. *Clin Genet*, 84 (5), 464-472. <https://doi.org/10.1111/cge.12097>
- Cummings, B. B., Marshall, J. L., Tukiainen, T., Lek, M., Donkervoort, S., Foley, A. R., Bolduc, V., Waddell, L. B., Sandaradura, S. A., O'Grady, G. L., Estrella, E., Reddy, H. M., Zhao, F., Weisburd, B., Karczewski, K. J., O'Donnell-Luria, A. H., Birnbaum, D., Sarkozy, A., Hu, Y., Gonorazky, H., Claeys, K., Joshi, H., Bournazos, A., Oates, E. C., Ghaoui, R., Davis, M. R., Laing, N. G., Topf, A., Genotype-Tissue Expression, C., Kang, P. B., Beggs, A. H., North, K. N., Straub, V., Dowling, J. J., Muntoni, F., Clarke, N. F., Cooper, S. T., Bonnemann, C. G., & MacArthur, D. G. (2017, Apr 19). Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci Transl Med*, 9 (386). <https://doi.org/10.1126/scitranslmed.aal5209>
- Dobkin, C., Pergolizzi, R. G., Bahre, P., & Bank, A. (1983, Mar). Abnormal splice in a mutant human beta-globin gene not at the site of a mutation. *Proc Natl Acad Sci U S A*, 80 (5), 1184-1188. <https://doi.org/10.1073/pnas.80.5.1184>
- Dutil, J., Teer, J. K., Golubeva, V., Yoder, S., Tong, W. L., Arroyo, N., Karam, R., Echenique, M., Matta, J. L., & Monteiro, A. N. (2019, 11 28). Germline variants in cancer genes in high-risk non-BRCA patients from Puerto Rico. *Sci Rep*, 9 (1), 17769. <https://doi.org/10.1038/s41598-019-54170-6>
- Fresard, L., Smail, C., Ferraro, N. M., Teran, N. A., Li, X., Smith, K. S., Bonner, D., Kernohan, K. D., Marwaha, S., Zappala, Z., Balliu, B., Davis, J. R., Liu, B., Prybol, C. J., Kohler, J. N., Zastrow, D. B., Reuter, C. M., Fisk, D. G., Grove, M. E., Davidson, J. M., Hartley, T., Joshi, R., Strober, B. J., Utiramerur, S., Lind, L., Ingelsson, E., Battle, A., Bejerano, G., Bernstein, J. A., Ashley, E. A., Boycott, K. M., Merker, J. D., Wheeler, M. T., Montgomery, S. B., Network, U. D., & Consortium, C. R. C. (2019, 06). Identification of rare-disease genes using blood transcriptome sequencing and large control cohorts. *Nat Med*, 25 (6), 911-919. <https://doi.org/10.1038/s41591-019-0457-8>
- Gelfman, S., Wang, Q., McSweeney, K. M., Ren, Z., La Carpio, F., Halvorsen, M., Schoch, K., Ratzon, F., Heinzen, E. L., Boland, M. J., Petrovski, S., & Goldstein, D. B. (2017, Aug 9). Annotating pathogenic non-coding variants in genic regions. *Nat Commun*, 8 (1), 236. <https://doi.org/10.1038/s41467-017-00141-2>
- Gonorazky, H. D., Naumenko, S., Ramani, A. K., Nelakuditi, V., Mashouri, P., Wang, P., Kao, D., Ohri, K., Viththiyapaskaran, S., Tarnopolsky, M. A., Mathews, K. D., Moore, S. A., Osorio, A. N., Villanova, D., Kemaladewi, D. U., Cohn, R. D., Brudno, M., & Dowling, J. J. (2019, 03 07). Expanding the Boundaries of RNA Sequencing as a Diagnostic Tool for Rare Mendelian Disease. *Am J Hum Genet*, 104 (3), 466-483. <https://doi.org/10.1016/j.ajhg.2019.01.012>
- Halperin, R. F., Carpten, J. D., Manojlovic, Z., Aldrich, J., Keats, J., Byron, S., Liang, W. S., Russell, M., Enriquez, D., Claasen, A., Cherni, I., Awuah, B., Oppong, J., Wicha, M. S., Newman, L. A., Jaigge, E., Kim, S., & Craig, D. W. (2017, 10 19). A method to reduce ancestry related germline false positives in tumor only somatic variant calling. *BMC Med Genomics*, 10 (1), 61. <https://doi.org/10.1186/s12920-017-0296-8>
- Hodes, R. J., & Buckholtz, N. (2016). Accelerating Medicines Partnership: Alzheimer's Disease (AMP-AD) Knowledge Portal Aids Alzheimer's Drug Discovery through Open Data Sharing. *Expert Opin Ther Targets*, 20 (4), 389-391. <https://doi.org/10.1517/14728222.2016.1135132>
- Hong, J. H., Ko, Y. H., & Kang, K. (2018). RNA variant identification discrepancy among splice-aware alignment algorithms. *PLoS One*, 13 (8), e0201822. <https://doi.org/10.1371/journal.pone.0201822>
- Hurst, S. E., Liktor-Busa, E., Moutal, A., Parker, S., Rice, S., Szelinger, S., Senner, G., Hammer, M. F., Johnstone, L., Ramsey, K., Narayanan, V., Perez-Miller, S., Khanna, M., Dahlin, H., Lewis, K., Craig, D., Wang, E. H., Khanna, R., & Nelson, M. A. (2018, Sep). A novel variant in. *Neuronal Signal*, 2 (3), NS20180141. <https://doi.org/10.1042/NS20180141>

- Jaganathan, K., Kyriazopoulou Panagiotopoulou, S., McRae, J. F., Darbandi, S. F., Knowles, D., Li, Y. I., Kosmicki, J. A., Arbelaez, J., Cui, W., Schwartz, G. B., Chow, E. D., Kanterakis, E., Gao, H., Kia, A., Batzoglou, S., Sanders, S. J., & Farh, K. K. (2019, Jan 24). Predicting Splicing from Primary Sequence with Deep Learning. *Cell*, 176 (3), 535-548 e524. <https://doi.org/10.1016/j.cell.2018.12.015>
- Jenkinson, G., Li, Y. I., Basu, S., Cousin, M. A., Oliver, G. R., & Klee, E. W. (2020, Nov 1). LeafCutterMD: an algorithm for outlier splicing detection in rare diseases. *Bioinformatics*, 36 (17), 4609-4615. <https://doi.org/10.1093/bioinformatics/btaa259>
- Jung, H., Lee, K. S., & Choi, J. K. (2021, 02). Comprehensive characterisation of intronic mis-splicing mutations in human cancers. *Oncogene*, 40 (7), 1347-1361. <https://doi.org/10.1038/s41388-020-01614-3>
- Karam, R., Conner, B., LaDuca, H., McGoldrick, K., Krempely, K., Richardson, M. E., Zimmermann, H., Gutierrez, S., Reineke, P., Hoang, L., Allen, K., Yussuf, A., Farber-Katz, S., Rana, H. Q., Culver, S., Lee, J., Nashed, S., Toppmeyer, D., Collins, D., Haynes, G., Pesaran, T., Dolinsky, J. S., Tippin Davis, B., Elliott, A., & Chao, E. (2019, Oct 2). Assessment of Diagnostic Outcomes of RNA Genetic Testing for Hereditary Cancer. *JAMA Netw Open*, 2 (10), e1913900. <https://doi.org/10.1001/jamanetworkopen.2019.13900>
- Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alfoldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., Walters, R. K., Tashman, K., Farjoun, Y., Banks, E., Poterba, T., Wang, A., Seed, C., Whiffin, N., Chong, J. X., Samocha, K. E., Pierce-Hoffman, E., Zappala, Z., O'Donnell-Luria, A. H., Minikel, E. V., Weisburd, B., Lek, M., Ware, J. S., Vittal, C., Armean, I. M., Bergelson, L., Cibulskis, K., Connolly, K. M., Covarrubias, M., Donnelly, S., Ferriera, S., Gabriel, S., Gentry, J., Gupta, N., Jeandet, T., Kaplan, D., Llanwarne, C., Munshi, R., Novod, S., Petrillo, N., Roazen, D., Ruano-Rubio, V., Saltzman, A., Schleicher, M., Soto, J., Tibbetts, K., Tolonen, C., Wade, G., Talkowski, M. E., Neale, B. M., Daly, M. J., MacArthur, D. G., & Consortium, G. A. D. (2020, 05). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, 581 (7809), 434-443. <https://doi.org/10.1038/s41586-020-2308-7>
- King, K., Flinter, F. A., Nihalani, V., & Green, P. M. (2002, Dec). Unusual deep intronic mutations in the COL4A5 gene cause X linked Alport syndrome. *Hum Genet*, 111 (6), 548-554. <https://doi.org/10.1007/s00439-002-0830-3>
- Kinney, A. Y., Howell, R., Ruckman, R., McDougall, J. A., Boyce, T. W., Vicuna, B., Lee, J. H., Guest, D. D., Rycroft, R., Valverde, P. A., Gallegos, K. M., Meisner, A., Wiggins, C. L., Stroup, A., Paddock, L. E., & Walters, S. T. (2018, 10). Promoting guideline-based cancer genetic risk assessment for hereditary breast and ovarian cancer in ethnically and geographically diverse cancer survivors: Rationale and design of a 3-arm randomized controlled trial. *Contemp Clin Trials*, 73 , 123-135. <https://doi.org/10.1016/j.cct.2018.09.005>
- Koeppel, F., Bobard, A., Lefebvre, C., Pedrero, M., Deloger, M., Bourdin, Y., Richon, C., Chen-Min-Tao, R., Robert, G., Meurice, G., Rouleau, E., Michiels, S., Massard, C., Scoazec, J. Y., Solaro, E., Soria, J. C., Andre, F., & Lacroix, L. (2018, 2018 Jul/Aug). Added Value of Whole-Exome and Transcriptome Sequencing for Clinical Molecular Screenings of Advanced Cancer Patients With Solid Tumors. *Cancer J*, 24 (4), 153-162. <https://doi.org/10.1097/PPO.0000000000000322>
- Kowalski, M. H., Qian, H., Hou, Z., Rosen, J. D., Tapia, A. L., Shan, Y., Jain, D., Argos, M., Arnett, D. K., Avery, C., Barnes, K. C., Becker, L. C., Bien, S. A., Bis, J. C., Blangero, J., Boerwinkle, E., Bowden, D. W., Buyske, S., Cai, J., Cho, M. H., Choi, S. H., Choquet, H., Cupples, L. A., Cushman, M., Daya, M., de Vries, P. S., Ellinor, P. T., Faraday, N., Fornage, M., Gabriel, S., Ganesh, S. K., Graff, M., Gupta, N., He, J., Heckbert, S. R., Hidalgo, B., Hodonsky, C. J., Irvin, M. R., Johnson, A. D., Jorgenson, E., Kaplan, R., Kardia, S. L. R., Kelly, T. N., Kooperberg, C., Lasky-Su, J. A., Loos, R. J. F., Lubitz, S. A., Mathias, R. A., McHugh, C. P., Montgomery, C., Moon, J. Y., Morrison, A. C., Palmer, N. D., Pankratz, N., Papanicolaou, G. J., Peralta, J. M., Peyser, P. A., Rich, S. S., Rotter, J. I., Silverman, E. K., Smith, J. A., Smith, N. L., Taylor, K. D., Thornton, T. A., Tiwari, H. K., Tracy, R. P., Wang, T., Weiss, S. T., Weng, L. C., Wiggins, K.

- L., Wilson, J. G., Yanek, L. R., Zollner, S., North, K. E., Auer, P. L., Raffield, L. M., Reiner, A. P., Li, Y., Consortium, N. T.-O. f. P. M. T., & Group, T. H. H. W. (2019, 12). Use of >100,000 NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium whole genome sequences improves imputation quality and detection of rare variant associations in admixed African and Hispanic/Latino populations. *PLoS Genet*, 15 (12), e1008500. <https://doi.org/10.1371/journal.pgen.1008500>
- Kremer, L. S., Wortmann, S. B., & Prokisch, H. (2018, 05). "Transcriptomics": molecular diagnosis of inborn errors of metabolism via RNA-sequencing. *J Inherit Metab Dis*, 41 (3), 525-532. <https://doi.org/10.1007/s10545-017-0133-4>
- Lachmann, A., Torre, D., Keenan, A. B., Jagodnik, K. M., Lee, H. J., Wang, L., Silverstein, M. C., & Ma'ayan, A. (2018, Apr 10). Massive mining of publicly available RNA-seq data from human and mouse. *Nat Commun*, 9 (1), 1366. <https://doi.org/10.1038/s41467-018-03751-6>
- Landrith, T., Li, B., Cass, A. A., Conner, B. R., LaDuca, H., McKenna, D. B., Maxwell, K. N., Domchek, S., Morman, N. A., Heinlen, C., Wham, D., Koptiuch, C., Vagher, J., Rivera, R., Bunnell, A., Patel, G., Geurts, J. L., Depas, M. M., Gaonkar, S., Pirzadeh-Miller, S., Krukenberg, R., Seidel, M., Pilarski, R., Farmer, M., Pyrtel, K., Milliron, K., Lee, J., Hoodfar, E., Nathan, D., Ganzak, A. C., Wu, S., Vuong, H., Xu, D., Arulmoli, A., Parra, M., Hoang, L., Molparia, B., Fennessy, M., Fox, S., Charpentier, S., Burdette, J., Pesaran, T., Profato, J., Smith, B., Haynes, G., Dalton, E., Crandall, J. R., Baxter, R., Lu, H. M., Tippin-Davis, B., Elliott, A., Chao, E., & Karam, R. (2020). Splicing profile by capture RNA-seq identifies pathogenic germline variants in tumor suppressor genes. *NPJ Precis Oncol*, 4 , 4. <https://doi.org/10.1038/s41698-020-0109-y>
- Landry, L. G., Ali, N., Williams, D. R., Rehm, H. L., & Bonham, V. L. (2018, 05). Lack Of Diversity In Genomic Databases Is A Barrier To Translating Precision Medicine Research Into Practice. *Health Aff (Millwood)*, 37 (5), 780-785. <https://doi.org/10.1377/hlthaff.2017.1595>
- Lee, H., Huang, A. Y., Wang, L. K., Yoon, A. J., Renteria, G., Eskin, A., Signer, R. H., Dorrani, N., Nieves-Rodriguez, S., Wan, J., Douine, E. D., Woods, J. D., Dell'Angelica, E. C., Fogel, B. L., Martin, M. G., Butte, M. J., Parker, N. H., Wang, R. T., Shieh, P. B., Wong, D. A., Gallant, N., Singh, K. E., Tavyev Asher, Y. J., Sinsheimer, J. S., Krakow, D., Loo, S. K., Allard, P., Papp, J. C., Palmer, C. G. S., Martinez-Agosto, J. A., Nelson, S. F., & Network, U. D. (2020, 03). Diagnostic utility of transcriptome sequencing for rare Mendelian diseases. *Genet Med*, 22 (3), 490-499. <https://doi.org/10.1038/s41436-019-0672-1>
- Lee, S. C., & Abdel-Wahab, O. (2016, 09 07). Therapeutic targeting of splicing in cancer. *Nat Med*, 22 (9), 976-986. <https://doi.org/10.1038/nm.4165>
- Li, J., Kong, N., Han, B., & Sul, J. H. (2021, 06). Rare variants regulate expression of nearby individual genes in multiple tissues. *PLoS Genet*, 17 (6), e1009596. <https://doi.org/10.1371/journal.pgen.1009596>
- Li, M. M., Datto, M., Duncavage, E. J., Kulkarni, S., Lindeman, N. I., Roy, S., Tsimerman, A. M., Vnencak-Jones, C. L., Wolff, D. J., Younes, A., & Nikiforova, M. N. (2017, Jan). Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn*, 19 (1), 4-23. <https://doi.org/10.1016/j.jmoldx.2016.10.002>
- Li, X., Kim, Y., Tsang, E. K., Davis, J. R., Damani, F. N., Chiang, C., Hess, G. T., Zappala, Z., Strober, B. J., Scott, A. J., Li, A., Ganna, A., Bassik, M. C., Merker, J. D., Hall, I. M., Battle, A., Montgomery, S. B., Consortium, G., Laboratory, D. t. A. C. C. L. A. W. G., Group, S. M. g. A. W., groups, E. G. e., Fund, N. C., NIH/NCI, NIH/NHGRI, NIH/NIMH, NIH/NIDA, Site—NDRI, B. C. S., Site—RPCI, B. C. S., Resource—VARI, B. C., Bank, B. B. R. U. o. M. B. E., Management, L. B. P., Study, E., & Visualization—EBI, G. B. D. I., & Genome Browser Data Integration & Visualization—UCSC Genomics Institute, U. i. o. C. S. C. (2017, 10 11). The impact of rare variation on gene expression across tissues. *Nature*, 550 (7675), 239-243. <https://doi.org/10.1038/nature24267>
- Liu, X., Li, C., Mou, C., Dong, Y., & Tu, Y. (2020, Dec 2). dbNSFP v4: a comprehensive database

- of transcript-specific functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Genome Med*, 12 (1), 103. <https://doi.org/10.1186/s13073-020-00803-9>
- Liu, X., Wu, C., Li, C., & Boerwinkle, E. (2016, Mar). dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Hum Mutat*, 37 (3), 235-241. <https://doi.org/10.1002/humu.22932>
- Maddirevula, S., Kuwahara, H., Ewida, N., Shamseldin, H. E., Patel, N., Alzahrani, F., AlSheddi, T., AlObeid, E., Alenazi, M., Alsaif, H. S., Alqahtani, M., AlAli, M., Al Ali, H., Helaby, R., Ibrahim, N., Abdulwahab, F., Hashem, M., Hanna, N., Monies, D., Derar, N., Alsagheir, A., Alhashem, A., Alsaleem, B., Alhebbi, H., Wali, S., Umarov, R., Gao, X., & Alkuraya, F. S. (2020, Jun 17). Analysis of transcript-deleterious variants in Mendelian disorders: implications for RNA-based diagnostics. *Genome Biol*, 21 (1), 145. <https://doi.org/10.1186/s13059-020-02053-9>
- Mantione, K. J., Kream, R. M., Kuzelova, H., Ptacek, R., Raboch, J., Samuel, J. M., & Stefano, G. B. (2014, Aug 23). Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq. *Med Sci Monit Basic Res*, 20 , 138-142. <https://doi.org/10.12659/MSMBR.892101>
- Maoz, C., Culver, J., & al., e. (2016). *Genetic, clinical and molecular characterization of a suspected CDH1 mutation* Collaborative Group of the Americas on Inherited Gastrointestinal Cancer (CGA-IGC) annual conference poster session.,
- Marco-Puche, G., Lois, S., Benitez, J., & Trivino, J. C. (2019). RNA-Seq Perspectives to Improve Clinical Diagnosis. *Front Genet*, 10 , 1152. <https://doi.org/10.3389/fgene.2019.01152>
- Maurano, M. T., Humbert, R., Rynes, E., Thurman, R. E., Haugen, E., Wang, H., Reynolds, A. P., Sandstrom, R., Qu, H., Brody, J., Shafer, A., Neri, F., Lee, K., Kutyavin, T., Stehling-Sun, S., Johnson, A. K., Canfield, T. K., Giste, E., Diegel, M., Bates, D., Hansen, R. S., Neph, S., Sabo, P. J., Heimfeld, S., Raubitschek, A., Ziegler, S., Cotsapas, C., Sotoodehnia, N., Glass, I., Sunyaev, S. R., Kaul, R., & Stamatoyannopoulos, J. A. (2012, Sep 7). Systematic localization of common disease-associated variation in regulatory DNA. *Science*, 337 (6099), 1190-1195. <https://doi.org/10.1126/science.1222794>
- McCullough, C. G., Szelinger, S., Belnap, N., Ramsey, K., Schrauwen, I., Claasen, A. M., Burke, L. W., Siniard, A. L., Huentelman, M. J., Narayanan, V., & Craig, D. W. (2020, 02). Utilizing RNA and outlier analysis to identify an intronic splice-altering variant in AP4S1 in a sibling pair with progressive spastic paraparesia. *Hum Mutat*, 41 (2), 412-419. <https://doi.org/10.1002/humu.23939>
- Mertes, C., Scheller, I. F., Yepez, V. A., Celik, M. H., Liang, Y., Kremer, L. S., Gusic, M., Prokisch, H., & Gagneur, J. (2021, Jan 22). Detection of aberrant splicing events in RNA-seq data using FRASER. *Nat Commun*, 12 (1), 529. <https://doi.org/10.1038/s41467-020-20573-7>
- Migeon, B. R. (2020, 07). X-linked diseases: susceptible females. *Genet Med*, 22 (7), 1156-1174. <https://doi.org/10.1038/s41436-020-0779-4>
- Montalban, G., Bonache, S., Moles-Fernandez, A., Gisbert-Beamud, A., Tenes, A., Bach, V., Carrasco, E., Lopez-Fernandez, A., Stjepanovic, N., Balmana, J., Diez, O., & Gutierrez-Enriquez, S. (2019, 02). Screening of. *J Med Genet*, 56 (2), 63-74. <https://doi.org/10.1136/jmedgenet-2018-105606>
- Morello, G., Guarnaccia, M., Spampinato, A. G., Salomone, S., D'Agata, V., Conforti, F. L., Aronica, E., & Cavallaro, S. (2019, 07 10). Integrative multi-omic analysis identifies new drivers and pathways in molecularly distinct subtypes of ALS. *Sci Rep*, 9 (1), 9968. <https://doi.org/10.1038/s41598-019-46355-w>
- Mori, K., Oura, T., Noma, H., & Matsui, S. (2013). Cancer outlier analysis based on mixture modeling of gene expression data. *Comput Math Methods Med*, 2013 , 693901. <https://doi.org/10.1155/2013/693901>
- Mundt, E., Nix, P., Brown, K., Bowles, K. R., & Manley, S. (2017, Dec 1). Complexities of Variant Classification in Clinical Hereditary Cancer Genetic Testing. *J Clin Oncol*, 35 (34), 3796-3799. <https://doi.org/10.1200/JCO.2017.74.5182>

Ndugga-Kabuye, M. K., & Issaka, R. B. (2019, 10). Inequities in multi-gene hereditary cancer testing: lower diagnostic yield and higher VUS rate in individuals who identify as Hispanic, African or Asian and Pacific Islander as compared to European. *Fam Cancer*, 18 (4), 465-469. <https://doi.org/10.1007/s10689-019-00144-6>

Park, E., Pan, Z., Zhang, Z., Lin, L., & Xing, Y. (2018, Jan 4). The Expanding Landscape of Alternative Splicing Variation in Human Populations. *Am J Hum Genet*, 102 (1), 11-26. <https://doi.org/10.1016/j.ajhg.2017.11.002>

Parsons, M. T., Tudini, E., Li, H., Hahnen, E., Wappenschmidt, B., Feliubadalo, L., Aalfs, C. M., Agata, S., Aittomaki, K., Alducci, E., Alonso-Cerezo, M. C., Arnold, N., Auber, B., Austin, R., Azzollini, J., Balmana, J., Barbieri, E., Bartram, C. R., Blanco, A., Blumcke, B., Bonache, S., Bonanni, B., Borg, A., Bortesi, B., Brunet, J., Bruzzone, C., Bucksch, K., Cagnoli, G., Caldes, T., Caliebe, A., Caligo, M. A., Calvello, M., Capone, G. L., Caputo, S. M., Carnevali, I., Carrasco, E., Caux-Moncoutier, V., Cavalli, P., Cimi, G., Clarke, E. M., Concolino, P., Cops, E. J., Cortesi, L., Couch, F. J., Darder, E., de la Hoya, M., Dean, M., Debatin, I., Del Valle, J., Delnatte, C., Derive, N., Diez, O., Ditsch, N., Domchek, S. M., Dutrannois, V., Eccles, D. M., Ehrencrona, H., Enders, U., Evans, D. G., Farra, C., Faust, U., Felbor, U., Feroce, I., Fine, M., Foulkes, W. D., Galvao, H. C. R., Gambino, G., Gehrig, A., Gensini, F., Gerdes, A. M., Germani, A., Giesecke, J., Gismondi, V., Gomez, C., Gomez Garcia, E. B., Gonzalez, S., Grau, E., Grill, S., Gross, E., Guerrieri-Gonzaga, A., Guillaud-Bataille, M., Gutierrez-Enriquez, S., Haaf, T., Hackmann, K., Hansen, T. V. O., Harris, M., Hauke, J., Heinrich, T., Hellebrand, H., Herold, K. N., Honisch, E., Horvath, J., Houdayer, C., Hubbel, V., Iglesias, S., Izquierdo, A., James, P. A., Janssen, L. A. M., Jeschke, U., Kaulfuss, S., Keupp, K., Kiechle, M., Kolbl, A., Krieger, S., Kruse, T. A., Kvist, A., Laloo, F., Larsen, M., Lattimore, V. L., Lautrup, C., Ledig, S., Leinert, E., Lewis, A. L., Lim, J., Loeffler, M., Lopez-Fernandez, A., Lucci-Cordisco, E., Maass, N., Manoukian, S., Marabelli, M., Matricardi, L., Meindl, A., Michelli, R. D., Moghadasi, S., Moles-Fernandez, A., Montagna, M., Montalban, G., Monteiro, A. N., Montes, E., Mori, L., Moserle, L., Muller, C. R., Mundhenke, C., Naldi, N., Nathanson, K. L., Navarro, M., Nevanlinna, H., Nichols, C. B., Niederacher, D., Nielsen, H. R., Ong, K. R., Pachter, N., Palmero, E. I., Papi, L., Pedersen, I. S., Peissel, B., Perez-Segura, P., Pfeifer, K., Pineda, M., Pohl-Rescigno, E., Poplawski, N. K., Porfirio, B., Quante, A. S., Ramser, J., Reis, R. M., Revillion, F., Rhiem, K., Riboli, B., Ritter, J., Rivera, D., Rofes, P., Rump, A., Salinas, M., Sanchez de Abajo, A. M., Schmidt, G., Schoenwiese, U., Seggewiss, J., Solanes, A., Steinemann, D., Stiller, M., Stoppa-Lyonnet, D., Sullivan, K. J., Susman, R., Sutter, C., Tavtigian, S. V., Teo, S. H., Teule, A., Thomassen, M., Tibiletti, M. G., Tischkowitz, M., Tognazzo, S., Toland, A. E., Tornero, E., Torngren, T., Torres-Esquius, S., Toss, A., Trainer, A. H., Tucker, K. M., van Asperen, C. J., van Mackelenbergh, M. T., Varesco, L., Vargas-Parra, G., Varon, R., Vega, A., Velasco, A., Vesper, A. S., Viel, A., Vreeswijk, M. P. G., Wagner, S. A., Waha, A., Walker, L. C., Walters, R. J., Wang-Gohrke, S., Weber, B. H. F., Weichert, W., Wieland, K., Wiesmuller, L., Witzel, I., Wockel, A., Woodward, E. R., Zachariae, S., Zampiga, V., Zeder-Goss, C., Lazaro, C., De Nicolo, A., Radice, P., Engel, C., Schmutzler, R. K., Goldgar, D. E., Spurdle, A. B., & Investigators, K. (2019, 09). Large scale multifactorial likelihood quantitative analysis of BRCA1 and BRCA2 variants: An ENIGMA resource to support clinical variant classification. *Hum Mutat*, 40 (9), 1557-1578. <https://doi.org/10.1002/humu.23818>

Petrov, A. I., Kay, S. J. E., Gibson, R., Kulesha, E., Staines, D., Bruford, E. A., Wright, M. W., Burge, S., Finn, R. D., Kersey, P. J., Cochrane, G., Bateman, A., Griffiths-Jones, S., Harrow, J., Chan, P. P., Lowe, T. M., Zwieb, C. W., Wower, J., Williams, K. P., Hudson, C. M., Gutell, R., Clark, M. B., Dinger, M., Quek, X. C., Bujnicki, J. M., Chua, N. H., Liu, J., Wang, H., Skogerbo, G., Zhao, Y., Chen, R., Zhu, W., Cole, J. R., Chai, B., Huang, H. D., Huang, H. Y., Cherry, J. M., Hatzigeorgiou, A., Pruitt, K. D., & RNACentral Consortium. (2015, 01). RNACentral: an international database of ncRNA sequences. *Nucleic Acids Res*, 43 (Database issue), D123-129. <https://doi.org/10.1093/nar/gku991>

Petrovski, S., & Goldstein, D. B. (2016, 07 14). Unequal representation of genetic variation across ancestry groups creates healthcare inequality in the application of precision medicine. *Genome Biol*, 17 (1), 157. <https://doi.org/10.1186/s13059-016-1016-y>

- Plon, S. E., Eccles, D. M., Easton, D., Foulkes, W. D., Genuardi, M., Greenblatt, M. S., Hogervorst, F. B., Hoogerbrugge, N., Spurdle, A. B., Tavtigian, S. V., & Group, I. U. G. V. W. (2008, Nov). Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat*, 29 (11), 1282-1291. <https://doi.org/10.1002/humu.20880>
- Pohl, M., Bortfeldt, R. H., Grutzmann, K., & Schuster, S. (2013, Oct). Alternative splicing of mutually exclusive exons—a review. *Biosystems*, 114 (1), 31-38. <https://doi.org/10.1016/j.biosystems.2013.07.003>
- Psych, E. C., Akbarian, S., Liu, C., Knowles, J. A., Vaccarino, F. M., Farnham, P. J., Crawford, G. E., Jaffe, A. E., Pinto, D., Dracheva, S., Geschwind, D. H., Mill, J., Nairn, A. C., Abyzov, A., Pochareddy, S., Prabhakar, S., Weissman, S., Sullivan, P. F., State, M. W., Weng, Z., Peters, M. A., White, K. P., Gerstein, M. B., Amiri, A., Armoskus, C., Ashley-Koch, A. E., Bae, T., Beckel-Mitchener, A., Berman, B. P., Coetzee, G. A., Coppola, G., Francoeur, N., Fromer, M., Gao, R., Grennan, K., Herstein, J., Kavanagh, D. H., Ivanov, N. A., Jiang, Y., Kitchen, R. R., Kozlenkov, A., Kundakovic, M., Li, M., Li, Z., Liu, S., Mangravite, L. M., Mattei, E., Markenscoff-Papadimitriou, E., Navarro, F. C., North, N., Omberg, L., Panchision, D., Parikshak, N., Poschmann, J., Price, A. J., Purcaro, M., Reddy, T. E., Roussos, P., Schreiner, S., Scuderi, S., Sebra, R., Shibata, M., Shieh, A. W., Skarica, M., Sun, W., Swarup, V., Thomas, A., Tsuji, J., van Bakel, H., Wang, D., Wang, Y., Wang, K., Werling, D. M., Willsey, A. J., Witt, H., Won, H., Wong, C. C., Wray, G. A., Wu, E. Y., Xu, X., Yao, L., Senthil, G., Lehner, T., Sklar, P., & Sestan, N. (2015, Dec). The PsychENCODE project. *Nat Neurosci*, 18 (12), 1707-1712. <https://doi.org/10.1038/nn.4156>
- Qian, X., Wang, J., Wang, M., Igelman, A. D., Jones, K. D., Li, Y., Wang, K., Goetz, K. E., Birch, D. G., Yang, P., Pennesi, M. E., & Chen, R. (2021). Identification of Deep-Intronic Splice Mutations in a Large Cohort of Patients With Inherited Retinal Diseases. *Front Genet*, 12, 647400. <https://doi.org/10.3389/fgene.2021.647400>
- Rehm, H. L., Harrison, S. M., & Martin, C. L. (2017, Dec). ClinVar Is a Critical Resource to Advance Variant Interpretation. *Oncologist*, 22 (12), 1562. <https://doi.org/10.1634/theoncologist.2017-0246>
- Rentas, S., Rathi, K. S., Kaur, M., Raman, P., Krantz, I. D., Sarmady, M., & Tayoun, A. A. (2020, 05). Diagnosing Cornelia de Lange syndrome and related neurodevelopmental disorders using RNA sequencing. *Genet Med*, 22 (5), 927-936. <https://doi.org/10.1038/s41436-019-0741-5>
- Richards, C. S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H. L., & Group, A. A. I. o. S. V. W. (2018, Dec). Response to Biesecker and Harrison. *Genet Med*, 20 (12), 1689-1690. <https://doi.org/10.1038/gim.2018.43>
- Roberts, M. E., Susswein, L. R., Janice Cheng, W., Carter, N. J., Carter, A. C., Klein, R. T., Hruska, K. S., & Marshall, M. L. (2020, 08). Ancestry-specific hereditary cancer panel yields: Moving toward more personalized risk assessment. *J Genet Couns*, 29 (4), 598-606. <https://doi.org/10.1002/jgc4.1257>
- Rofes, P., Menendez, M., Gonzalez, S., Tornero, E., Gomez, C., Vargas-Parra, G., Montes, E., Salinas, M., Solanes, A., Brunet, J., Teule, A., Capella, G., Feliubadal, L., Del Valle, J., Pineda, M., & Lazaro, C. (2020, 12). Improving Genetic Testing in Hereditary Cancer by RNA Analysis: Tools to Prioritize Splicing Studies and Challenges in Applying American College of Medical Genetics and Genomics Guidelines. *J Mol Diagn*, 22 (12), 1453-1468. <https://doi.org/10.1016/j.jmoldx.2020.09.007>
- Ryan, M. C., Cleland, J., Kim, R., Wong, W. C., & Weinstein, J. N. (2012, Sep 15). SpliceSeq: a resource for analysis and visualization of RNA-Seq data on alternative splicing and its functional impacts. *Bioinformatics*, 28 (18), 2385-2387. <https://doi.org/10.1093/bioinformatics/bts452>
- Sanders, S. J., Schwartz, G. B., & Farh, K. K. (2020, 04 24). Clinical impact of splicing in neurodevelopmental disorders. *Genome Med*, 12 (1), 36. <https://doi.org/10.1186/s13073-020-00737-2>
- Scotti, M. M., & Swanson, M. S. (2016, Jan). RNA mis-splicing in disease. *Nat Rev Genet*, 17 (1), 19-32. <https://doi.org/10.1038/nrg.2015.3>

- Shamseldin, H. E., AlAbdi, L., Maddirevula, S., Alsaif, H. S., Alzahrani, F., Ewida, N., Hashem, M., Abdulwahab, F., Abuyousef, O., Kuwahara, H., Gao, X., Molecular Autopsy, C., & Alkuraya, F. S. (2021, Oct 13). Lethal variants in humans: lessons learned from a large molecular autopsy cohort. *Genome Med*, *13* (1), 161. <https://doi.org/10.1186/s13073-021-00973-0>
- Shao, L., Xing, F., Xu, C., Zhang, Q., Che, J., Wang, X., Song, J., Li, X., Xiao, J., Chen, L. L., & Ouyang, Y. (2019, 03 19). Patterns of genome-wide allele-specific expression in hybrid rice and the implications on the genetic basis of heterosis. *Proc Natl Acad Sci U S A*, *116* (12), 5653-5658. <https://doi.org/10.1073/pnas.1820513116>
- Shenker, S., Miura, P., Sanfilippo, P., & Lai, E. C. (2015, Jan). IsoSCM: improved and alternative 3' UTR annotation using multiple change-point inference. *RNA*, *21* (1), 14-27. <https://doi.org/10.1261/rna.046037.114>
- Shi, Y., Chen, Z., Gao, J., Wu, S., Gao, H., & Feng, G. (2018, Oct). Transcriptome-wide analysis of alternative mRNA splicing signature in the diagnosis and prognosis of stomach adenocarcinoma. *Oncol Rep*, *40* (4), 2014-2022. <https://doi.org/10.3892/or.2018.6623>
- Shvetsova, E., Sofronova, A., Monajemi, R., Gagalova, K., Draisma, H. H. M., White, S. J., Santen, G. W. E., Chuva de Sousa Lopes, S. M., Heijmans, B. T., van Meurs, J., Jansen, R., Franke, L., Kielbasa, S. M., den Dunnen, J. T., 't Hoen, P. A. C., consortium, B., & consortium, G. (2019, 03). Skewed X-inactivation is common in the general female population. *Eur J Hum Genet*, *27* (3), 455-465. <https://doi.org/10.1038/s41431-018-0291-3>
- Slavin, T. P., Van Tongeren, L. R., Behrendt, C. E., Solomon, I., Rybak, C., Nehoray, B., Kuzmich, L., Niell-Swiller, M., Blazer, K. R., Tao, S., Yang, K., Culver, J. O., Sand, S., Castillo, D., Herzog, J., Gray, S. W., & Weitzel, J. N. (2018, 10 01). Prospective Study of Cancer Genetic Variants: Variation in Rate of Reclassification by Ancestry. *J Natl Cancer Inst*, *110* (10), 1059-1066. <https://doi.org/10.1093/jnci/djy027>
- Stenton, S. L., & Prokisch, H. (2020, 06). The Clinical Application of RNA Sequencing in Genetic Diagnosis of Mendelian Disorders. *Clin Lab Med*, *40* (2), 121-133. <https://doi.org/10.1016/j.cll.2020.02.004>
- Szelinger, S., Malenica, I., Corneveaux, J. J., Siniard, A. L., Kurdoglu, A. A., Ramsey, K. M., Schrauwen, I., Trent, J. M., Narayanan, V., Huentelman, M. J., & Craig, D. W. (2014). Characterization of X chromosome inactivation using integrated analysis of whole-exome and mRNA sequencing. *PLoS One*, *9* (12), e113036. <https://doi.org/10.1371/journal.pone.0113036>
- Tahiliani, J., Leisk, J., Aradhya, K., & .., e. a. (2020). Utility of RNA Sequencing Analysis in the Context of Genetic Testing. *Curr Genet Med Rep*, *8* , 140-146.
- Talhouarne, G. J. S., & Gall, J. G. (2018, 08 21). Lariat intronic RNAs in the cytoplasm of vertebrate cells. *Proc Natl Acad Sci U S A*, *115* (34), E7970-E7977. <https://doi.org/10.1073/pnas.1808816115>
- Tan, A. C., Fan, J. B., Karikari, C., Bibikova, M., Garcia, E. W., Zhou, L., Barker, D., Serre, D., Feldmann, G., Hruban, R. H., Klein, A. P., Goggins, M., Couch, F. J., Hudson, T. J., Winslow, R. L., Maitra, A., & Chakravarti, A. (2008, Jan). Allele-specific expression in the germline of patients with familial pancreatic cancer: an unbiased approach to cancer gene discovery. *Cancer Biol Ther*, *7* (1), 135-144. <https://doi.org/10.4161/cbt.7.1.5199>
- Tang, S. J., Shen, H., An, O., Hong, H., Li, J., Song, Y., Han, J., Tay, D. J. T., Ng, V. H. E., Bellido Molias, F., Leong, K. W., Pitcheshwar, P., Yang, H., & Chen, L. (2020, 02 07). Cis- and trans-regulations of pre-mRNA splicing by RNA editing enzymes influence cancer development. *Nat Commun*, *11* (1), 799. <https://doi.org/10.1038/s41467-020-14621-5>
- Tanner, M. K., Tang, Z., & Thornton, C. A. (2021, 02 26). Targeted splice sequencing reveals RNA toxicity and therapeutic response in myotonic dystrophy. *Nucleic Acids Res*, *49* (4), 2240-2254. <https://doi.org/10.1093/nar/gkab022>

- Thompson, K., Collier, J. J., Glasgow, R. I. C., Robertson, F. M., Pyle, A., Blakely, E. L., Alston, C. L., Oláhová, M., McFarland, R., & Taylor, R. W. (2020, 01). Recent advances in understanding the molecular genetic basis of mitochondrial disease. *J Inherit Metab Dis*, *43* (1), 36-50. <https://doi.org/10.1002/jimd.12104>
- Truty, R., Ouyang, K., Rojahn, S., Garcia, S., Colavin, A., Hamlington, B., Freivogel, M., Nussbaum, R. L., Nykamp, K., & Aradhya, S. (2021, 04 01). Spectrum of splicing variants in disease genes and the ability of RNA analysis to reduce uncertainty in clinical interpretation. *Am J Hum Genet*, *108* (4), 696-708. <https://doi.org/10.1016/j.ajhg.2021.03.006>
- Ungerleider, N., & Flemington, E. (2019, May 8). SpliceV: analysis and publication quality printing of linear and circular RNA splicing, expression and regulation. *BMC Bioinformatics*, *20* (1), 231. <https://doi.org/10.1186/s12859-019-2865-7>
- Update., N. R. C. U. C. o. D. F. S. A. (1996). The Evaluation of Forensic DNA Evidence. . In. National Academies Press <https://www.ncbi.nlm.nih.gov/books/NBK232615/>
- Urbanski, L. M., Leclair, N., & Añczuków, O. (2018, 07). Alternative-splicing defects in cancer: Splicing regulators and their downstream targets, guiding the way to novel cancer therapeutics. *Wiley Interdiscip Rev RNA*, *9* (4), e1476. <https://doi.org/10.1002/wrna.1476>
- Urbina-Jara, L. K., Rojas-Martinez, A., Martinez-Ledesma, E., Aguilar, D., Villarreal-Garza, C., & Ortiz-Lopez, R. (2019, 10 10). Landscape of Germline Mutations in DNA Repair Genes for Breast Cancer in Latin America: Opportunities for PARP-Like Inhibitors and Immunotherapy. *Genes (Basel)*, *10* (10). <https://doi.org/10.3390/genes10100786>
- Vaske, O. M., Bjork, I., Salama, S. R., Beale, H., Tayi Shah, A., Sanders, L., Pfeil, J., Lam, D. L., Learned, K., Durbin, A., Kephart, E. T., Currie, R., Newton, Y., Swatloski, T., McColl, D., Vivian, J., Zhu, J., Lee, A. G., Leung, S. G., Spillinger, A., Liu, H. Y., Liang, W. S., Byron, S. A., Berens, M. E., Resnick, A. C., Lacayo, N., Spunt, S. L., Rangaswami, A., Huynh, V., Torno, L., Plant, A., Kirov, I., Zabokrtsky, K. B., Rassekh, S. R., Deyell, R. J., Laskin, J., Marra, M. A., Sender, L. S., Mueller, S., Sweet-Cordero, E. A., Goldstein, T. C., & Haussler, D. (2019, 10 02). Comparative Tumor RNA Sequencing Analysis for Difficult-to-Treat Pediatric and Young Adult Patients With Cancer. *JAMA Netw Open*, *2* (10), e1913968. <https://doi.org/10.1001/jamanetworkopen.2019.13968>
- Vaz-Drago, R., Custodio, N., & Carmo-Fonseca, M. (2017, Sep). Deep intronic mutations and human disease. *Hum Genet*, *136* (9), 1093-1111. <https://doi.org/10.1007/s00439-017-1809-4>
- Virlogeux, V., Graff, R. E., Hoffmann, T. J., & Witte, J. S. (2015, Jun). Replication and heritability of prostate cancer risk variants: impact of population-specific factors. *Cancer Epidemiol Biomarkers Prev*, *24* (6), 938-943. <https://doi.org/10.1158/1055-9965.EPI-14-1372>
- Wai, H. A., Lord, J., Lyon, M., Gunning, A., Kelly, H., Cibin, P., Seaby, E. G., Spiers-Fitzgerald, K., Lye, J., Ellard, S., Thomas, N. S., Bunyan, D. J., Douglas, A. G. L., Baralle, D., working, S. a. d., & group. (2020, 06). Blood RNA analysis can increase clinical diagnostic rate and resolve variants of uncertain significance. *Genet Med*, *22* (6), 1005-1014. <https://doi.org/10.1038/s41436-020-0766-9>
- Wang, K., Li, M., & Hakonarson, H. (2010, Sep). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*, *38* (16), e164. <https://doi.org/10.1093/nar/gkq603>
- Wang, Y., Mashock, M., Tong, Z., Mu, X., Chen, H., Zhou, X., Zhang, H., Zhao, G., Liu, B., & Li, X. (2020). Changing Technologies of RNA Sequencing and Their Applications in Clinical Oncology. *Front Oncol*, *10* , 447. <https://doi.org/10.3389/fonc.2020.00447>
- Woolfe, A., Mullikin, J. C., & Elnitski, L. (2010). Genomic features defining exonic variants that modulate splicing. *Genome Biol*, *11* (2), R20. <https://doi.org/10.1186/gb-2010-11-2-r20>
- Yang, H., & Wang, K. (2015, Oct). Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat Protoc*, *10* (10), 1556-1566. <https://doi.org/10.1038/nprot.2015.105>

Yuan, J., Kensler, K. H., Hu, Z., Zhang, Y., Zhang, T., Jiang, J., Xu, M., Pan, Y., Long, M., Montone, K. T., Tanyi, J. L., Fan, Y., Zhang, R., Hu, X., Rebbeck, T. R., & Zhang, L. (2020, 02). Integrative comparison of the genomic and transcriptomic landscape between prostate cancer patients of predominantly African or European genetic ancestry. *PLoS Genet*, 16 (2), e1008641. <https://doi.org/10.1371/journal.pgen.1008641>

Zhang, Y., Liu, X., MacLeod, J., & Liu, J. (2018, Dec 27). Discerning novel splice junctions derived from RNA-seq alignment: a deep learning approach. *BMC Genomics*, 19 (1), 971. <https://doi.org/10.1186/s12864-018-5350-1>

Zhang, Y., Qian, J., Gu, C., & Yang, Y. (2021, 02 24). Alternative splicing and cancer: a systematic review. *Signal Transduct Target Ther*, 6 (1), 78. <https://doi.org/10.1038/s41392-021-00486-7>



