We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,300

Open access books available

172,000

International authors and editors

190M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Recent Progress in Drug Repurposing Using Protein Variants and Amino Acids in Disease Phenotypes/Disorders

Michael P. Okoh and Lukman A. Alli

Abstract

Life is constituted of large group of macromolecule, functional and structural called “Protein,” made of amino acids (AA), and linked with peptide bonds with specific protein unique sequences. Variations in proteins are thought to have diverse effects with consequences on structure, stability, interactions, pH, enzymatic activity, abundance and other properties. Variants can be of genetic origin or it could occur *de novo* at the post-translational protein level. The sequence of amino acids defines protein structure and functions. Protein is involved in several critical functions like the physical cell-cell communication. Breakthrough in molecular science has shown that, to develop drugs for managing a disease-associated variations requires understanding of consequences of variants on the function of the affected protein and the impact on the pathways, in which protein is involved. Using biophysical/bioinformatics methods, immense amount of variation data generated is handled-connected to disease phenotypes. Obviously, there remain continuous needs for the combinations of genetic probing methods/bioinformatics, to predict single-nucleotide variations (SNV), for effective rational drug design that would embrace naturally occurring bioactive components of plant origin, towards the effective management of disease phenotype emanating from protein and amino acid variations. This, well thought out and synchronized concept, remains a way forward.

Keywords: protein variation, epigenetics, disease management, single nucleotide variation, protein variants, amino acids in disease phenotypes/disorders

1. Introduction

Variations in the genome and protein expression remain a driver of most diseases with their polygenic phenotypes. Diseases may manifest with time emanating from aberrant protein expression. These biochemical processes are complex in nature, as it involves molecular interactions at both the DNA-RNA/protein level. Within the context of protein-protein interaction (PPI), it has become essential to look at the long-term clinical goal, which could be to identify disease-specific patterns of PPIs, which could serve as a disease- or treatment-responsive biomarkers whose selective measurement may lead to improved diagnosis or prognosis for common human disorders [1].

Interests in the study of variations in DNA/protein have helped to espouse factors germane to influencing biological processes and genome stability. This is more so bearing; genomic integrity is particularly important as they provide the blueprint for the next generation [2]. During cell division, homeostasis is required, and for human health, the genome needs to be copied prudently such that a copy of each chromosome is passed on to the daughter cells. The polygenic nature of some disease phenotypes, controlled by a combination of several genes all playing together makes it essential to unravel biological processes in a piecemeal. For instance, most diseases with a number of persons with inherited predispositions including, heart disease, arteriosclerosis, and some cancers are thought to be polygenic [3].

Looking at these processes within the context of ribonucleic acid (RNA) on the other hand presents different facts. At some point in time, RNA was thought to be much less important than DNA since it did not carry any of the genetic characteristics of an organism. However, lately, it has become obvious that this might not be entirely so bearing, the life cycle of RNA viruses for instance is directed to transport, multiply, and deliver the viral RNA genome into other cells. Fortunately, not all of these viral genomes can encode all proteins in the cell that are required for these known processes to be accomplished. Thus, overcoming this limitation, viruses are known to hijack cellular RNA-binding proteins (RBPs) [4, 5].

Responding to such invasion, host cells do concertedly employ specialized RBPs as a detection mechanism for viral RNAs and their intermediates of replication through the recognition of the molecular signage such as the under-methylated, cap triphosphate ends, and double-stranded RNA (dsRNA) [6]. Beyond this, several other observations have been made [3, 4], highlighting the essential role that RBPs play in regulating the viral life cycle. For instance, it is thought that RBP sensing of viral RNA triggers the cellular antiviral state, which can suppress viral gene expression [6], leading to the inhibition of protein synthesis and the production of interferons [4, 5].

Recently, using multiple proteome-wide approaches [7] had identified RBPs involved in the SARS-CoV-2 life cycle whilst showing that the repertoire of cellular RBPs widely remodels in response to SARS-CoV-2 infection, via proteins involved in antiviral defenses, RNA metabolism, and other pathways.

In all of these processes, transcription factor (TF) mutations have been studied for decades, with RBPs being overlooked as drivers of disease and as therapeutically relevant targets. Now it is established that RBPs determine the fate of transcribed RNAs by regulating their splicing, polyadenylation, translation, subcellular localization, and turnover [8].

For drug repurposing, diseases that are driven by a known or combination of mutants at the protein level are of major attention for direct targeting. Moreover, changes in cellular growth rate and the identity that occur during diseases such as cancer, hemoglobinopathy, etc., are, known to be driven by specific gene expression signatures that are programmed by the activity of DNA-binding TFs and RBP [9]. From recent findings, it is now clear that RNA-binding proteins (RBPs) are critical regulators of post-transcriptional gene expression [9]. Within this context, Liu and Shi [10] earlier established the importance of RBP in Amyotrophic lateral sclerosis (ALS), disease progression. Establishing that the heterogeneous ribonucleoproteins (horn A2/B1) mutation in patients with ALS did not just disable the protein, but instead, the mutation conferred some new toxic properties that scrambled RNA processing, fast-tracking the death of motor neurons [10].

Some other known fact is that missense mutation is a mistake in the DNA and it could arise due to aberrant TF. Missense mutations for instance in tumor suppressors result in its loss of function (LOF) in a variety of manners including loss of stability of the protein or the disruption of a crucial ligand/DNA/protein binding site [11]. The Worldwide Protein Data Bank (wwPDB) have over 88,000 protein

structures, many of which play vital roles in critical metabolic pathways that may be regarded as potential therapeutic targets and specific databases containing structures of binary complexes [12]. Moreover, a recent breakthrough in molecular science has shown that the key to developing targeted therapy, for disease-associated variations is with the critical understanding of the consequences of that variant on the function of the affected protein, and the impact on the pathways in which that protein is involved [9]. Proteins are produced and recycled by some critical processes in their tissue sources and are degraded into necessary amino acids through very controlled bio-signaling and feedback systems. For instance, the salvage pathways are known as a major source of nucleotides for the synthesis of DNA, RNA, and enzyme co-factors.

The disproportion of protein demand, dietary supply, and productions do result in a variety of disease phenotypes due mainly to deficiency, occasioned sometimes by variation properties. A critically important enzyme of purine salvage in rapidly dividing cells for instance is adenosine deaminase (ADA), which catalyzes the deamination of adenosine to inosine. Deficiency in ADA results in the disorder called severe combined immunodeficiency (SCID). This is a genetic disease amongst many others that is characterized by the development of nonfunctional T and B cells due to genetic mutation resulting in heterogeneous clinical phenotype [13].

2. DNA-protein interactions

The cis-regulatory DNA elements' interactions with the transcription factors seem to be critical components of transcriptional regulatory networks [14]. The genome with the complete cDNA sequences contains large numbers of transcription factors with their binding DNA sequences. It is expected that a comprehensive analysis of DNA-transcription factor interactions will provide a deep understanding of the mechanisms of drug metabolism in critical processes such as cell proliferation, developmental processes in tissue morphogenesis, and disease manifestation [14]. The combined use of chromatin immunoprecipitation (ChIP) assay with DNA microarrays (ChIP-chip) [14, 15] are the most widely used high-throughput method for discovering non-coding region but important (cis-regulatory) DNA elements for a transcription factor [16]. Albeit, the development of high-throughput methods for discovering transcription factors for DNA regulatory elements remains in its infancy, even though the yeast one-hybrid method [17] and phage display [16] are attractive candidates, in this regard. However, these methods have some shortcomings including, they are not easily scalable because of the use of living cells. Further, the overexpression of transcription factors are thought to affect cellular metabolism, and as such, making transcription factors difficult to screen. Thus, to avoid these difficulties, focus totally on in-vitro mRNA display technology such as in-vitro virus (IVV) method [11, 16] for the discovery of DNA-protein interactions serve as a good alternative.

To map out the transcriptional regulatory networks at a wider genome level, a comprehensive analysis of DNA-protein interactions is important. Thus, the IVV method had been employed for in vitro selection of DNA-binding protein heterodimeric complexes [18]. Using improved selection conditions, enhanced with a TPA-responsive element (TRE) as a bait DNA, known interactors such as; c-fos and c-jun were simultaneously enriched about 100-fold from a model library (a 1:1:20000 mixture of c-fos, c-jun, and gst genes) after one round of selection [18]. Moreover, the AP-1 family genes, including c-jun, c-fos, junD, junB, atf2, and b-atf, were successfully selected from an IVV library constructed from a mouse

brain poly A+ RNA after six rounds of selection [19]. These results indicated that this method (IVV selection system) have the potential to identify a variety of DNA binding protein complexes in a single experiment. Since almost all transcription factors form hetero-oligomeric complexes towards binding with their target DNA, this method should be most useful to search for DNA-binding transcription factor complexes [11, 16], which will further illuminate the understanding of drug repurposing in disease state conditions.

3. Diseases arising from mutations

Numerous computational tools have been developed for the interpretation, analysis, and prioritization of variations and their effects [20]. Many DNA/protein variations and disease-causing mutation databases are now available for references. For instance, the locus specific variation database (LSVD) is present at Leiden Open Variation Database (LOVD) system for all human genes [21]. Although some of the databases seem to contain similar information, however, the LSDBs are listed at the Human Genome Variation Society (HGVS) Website (<http://www.hgvs.org/locus-specific-mutation-databases>), the LOVD site (http://grenada.lumc.nl/LSDB_list/lstdbs), the GEN2PHEN server (<http://www.gen2phen.org/data/lstdbs>) [22], and at the Web Analysis of the Variome (<http://bioinformatics.ua.pt/WAVE/>) [20, 23, 24].

Besides the above databases, there are many others that were most recently covered ([20, 25], and the references therein). Moreover, recent advances in genome-wide association studies, next-generation sequencing technologies coupled with genetic linkage analysis have enhanced output in the analysis of mutation-causing diseases. Many of these methods are useful for detecting single-nucleotide polymorphisms (SNPs), which are found to be common in aberrant gene functioning. However, it may also be noted, the majority of structural variations (SVs) that occur in the human genome are yet to be fully characterized by single short-read platforms [26]. Suffice, for many genetic diseases, association studies have relied most heavily upon short read, high throughput sequencing technologies [27, 28].

Some genetic variations with the consequence encoded proteins are known to manifest into disease phenotypes with the deleterious outcome to the patient. Within these are hemoglobinopathy including sickle cell disease (SCD), which are caused by a single germ-line mutation substituting (A to T) in the codon for amino acid 6. The change converts a glutamic acid codon (GAG) to a valine codon (GTG) [29, 30].

3.1 Single mutation as a lead cause of amyotrophic lateral sclerosis (ALS)

Most recently, due to the advances mentioned above, it led to the finding that a mutation in the C9orf72 gene (chromosome 9 open reading frame 72 genes) is the primary genetic cause of amyotrophic lateral sclerosis (ALS). These losses of function, induced by the mutation of the C9orf72 gene are thought to affect communication between motor neurons and muscles in people with ALS [31]. Further, this mutation is thought in part to be responsible for 40–50% of hereditary cases of ALS, and 5–10% of cases without family history. This mutation consists of an expansion of a sequence of hexanucleotide (GGGGCC) DNA bases, going from a few copies (less than 20 in a healthy person) to more than 1000 copies [30]. Until now, it still remains unclear how this GGGGCC base repeat expansions cause neurodegeneration in ALS. Although, mechanistically, the C9orf72 protein function in a complex with the WDR41 and SMCR proteins (guanine exchange factors (GEF)) for Rab8 and Rab39 [31].

In a more recent study, the gene C9orf72 role on the protein TDP-43 (transactive response DNA binding protein-43) was revealed. The TDP-43 protein plays an important role in ALS. It is thought that the C9orf72 gene may affect the protein TDP-43's location within the cell. "In approximately 97% of ALS patients, it is being observed that the TDP-43 protein is depleted from the nucleus, forming aggregates in the cytoplasm rather than being in the nucleus, as is the case in healthy people [26, 32, 33].

The average incidence rate of ALS worldwide is about one in 50,000 people per year and the average age of onset of the disease is about 60 years, with men at a slightly higher risk compared to women. FDA-approved treatments for ALS are only modestly effective and the disease still results in complete paralysis and death within the first 5 years after diagnosis [31, 32].

3.2 Troponin variation in cardiomyopathy

The calcium-mediated interaction between actin and myosin is controlled by cardiac regulatory proteins, cardiac troponin T (cTnT) and troponin I (cTnI). The cardiac forms of these regulatory proteins theoretically have the potential of being unique to the myocardium [34], as they are coded for by specific genes.

Cardiac troponins are detected in the serum by the use of monoclonal antibodies to epitopes of cTnI and cTnT. These antibodies are highly specific for cardiac troponin and have negligible cross-reactivity with skeletal muscle troponins. Indeed, cTnI has not been identified outside the myocardium [34]. Cardiac troponin T is expressed to a small extent in skeletal muscle; however, the current cTnT assay does not identify skeletal troponins [35].

The majority of cTnI and cTnT form part of the contractile apparatus within the myocardial cell with lower concentrations found in the cytoplasm [35]. Whenever there is myocardial ischemia resulting in myocardial necrosis, the cTn will be released from the cytosolic pool into the bloodstream within a few hours of the injury. This is typically followed by a more prolonged and sustained elevation of cTn due to degradation of the contractile apparatus, which may also be a reflection of the size of the infarct [35].

However, the release kinetics of cTn after the myocardial injury can differ between individuals and is also dependent on myocardial blood flow. It can also differ between cTnI and cTnT which are thought to have monophasic and biphasic concentration-time profiles respectively, and with the increase in cTnT tending to last for longer than that of cTnI [34].

After the onset of an acute coronary event, cardiac troponins may not be detected in the serum for up to 4 hours and should be repeated 12 hours after the first test, if the troponin concentration is not raised in an individual presenting with chest pain.

In the identification of cardiac muscle damage, the measurement of serum cTnI and cTnT are superior in terms of sensitivity and specificity to cardiac muscle enzyme measurements [36]. Elevated cardiac troponin concentrations are now an acceptable standard biochemical marker for the diagnosis of myocardial infarction [37].

In order to enhance the comparison of results for cTnT, from one laboratory to another, troponin T is measured using a single assay, and a cutoff value of 0.1 µg/liter is indicative of myocardial damage [38]. However, there are several cTnI assays with different sensitivities and cutoff values. According to the European Society of Cardiology and American College of Cardiology consensus criteria, serum cTnI values that indicate myocyte necrosis/myocardial damage range from 0.1 to 2 µg/liter [38].

In the management of patients with acute chest pain, the measurement of cardiac troponins as markers of myocardial damage has produced two important

beneficial effects on clinical practice [39]. The first beneficial effect is that more patients with chest pain who would not have been diagnosed as having myocardial damage with conventional muscle enzyme assays are being diagnosed with myocardial infarction, even in the absence of ST-segment elevation. The second beneficial effect is that mortality is reduced because many of these patients are at high risk of full-thickness myocardial infarction or even death within 6 month period [40, 41].

The *Universal Definition of Myocardial Infarction* requires at least one cTn concentration above the 99th percentile value of a normal reference population for the diagnosis of myocardial injury [38]. However, there have been some concerns regarding the use of a 99th percentile threshold value for hs-cTn because of its limitations [42]. Firstly, the 99th percentile varies with assay [43]. Secondly, the 99th percentile varies with reference population selection (age, gender, ethnicity, and definition of healthy status), reference population size, and the statistical method used to calculate it [44, 45]. Some studies have shown that elevations of hs-cTn can be seen in older adults, which may be independent of pathological conditions [46, 47]. Thirdly, detectable chronic elevations in cTn above the 99th percentile are commonly seen in conditions such as chronic renal or cardiac failure [48, 49]. In addition, the improved analytical sensitivity of these assays has resulted in the detection of elevated cTn in numerous cardiac and non-cardiac conditions that cause myocardial cell necrosis, such as myocarditis, arrhythmia, cardiac procedures, pulmonary embolism, and sepsis [34, 41]. Due to these challenges, international guidelines have sought to promote consistency by proposing recommendations for determining 99th percentiles [50, 51]. It would therefore seem that the 99th percentile should not be the only metric for diagnosing acute myocardial injury.

Cardiac troponins may also be elevated in many other conditions associated with secondary ischaemic injury [44], such as large pulmonary emboli, coronary spasm, cardiac arrhythmias [52], hypertrophic cardiomyopathy [52], idiopathic dilated cardiomyopathy [53, 54]. It can also be elevated in conditions that cause myocardial injuries, such as cardiac trauma, chemotherapy [55], myopericarditis [55, 56], septicemia [57].

Some studies also found that cTn was detectable in nearly all children, where concentrations increased with increasing age and left ventricular mass, thus supporting the notion that cTn release is not always pathological [58].

In addition, it has recently been demonstrated that cTn may exhibit diurnal variations [59, 60]. One study noted that cTnT concentrations exhibited a decreasing trend between morning and afternoon (0830 hours and 1430 hours) for healthy individuals and individuals requiring hemodialysis [59]. For cTnI concentrations, a decreasing trend during these hours was also noted in individuals requiring hemodialysis, however, the pattern was not apparent in healthy individuals [59]. Furthermore, another study in men with type 2 diabetes found that cTnT decreased during the day and then increased during the night, with peak concentrations in the morning at 0830 hours [58]. This was further confirmed in another study of healthy individuals, where cTnT exhibited diurnal variation but cTnI did not have such variation [60]. In other words, cTn can be described as organ-specific but not disease-specific.

4. RNA splicing in disease diagnosis

RNA splicing is a post-transcriptional process necessary to form a mature mRNA [61]. There are two main forms of splicing, that is, constitutive splicing and alternative splicing.

Constitutive splicing involves removal of introns from the pre-mRNA and joining the exons together to form a mature mRNA. Alternative splicing describes how exons can be included or excluded in different combinations to create a

diverse array of mRNA transcripts from a single pre-mRNA and therefore serves as a process to increase the diversity of the transcriptome. It was initially thought that about 5% of human genes were subjected to alternative splicing [62]. Now, after the implementation of next-generation sequencing technologies, it is now known that the vast majority, >95% of mRNAs, are subjected to alternative splicing [63]. However, the function of a large fraction of these splice isoforms is still unknown.

Splicing is more prevalent in multicellular than in unicellular eukaryotes because of the lower number of intron-containing genes in the latter [64]. As evolution progress, alternative splicing becomes more prevalent in vertebrates than in invertebrates. Skipping of a single exon in the RNA-binding protein (RBP) and polypyrimidine tract binding protein 1 (PTBP1) may be responsible for numerous alternative splicing changes between species, which suggest that one splicing event can augment the varieties observed in transcriptome between species [65].

The hypothesis that alternative splicing largely contributes to organism diversity is fueled by the observation that the total number of protein-coding genes does not differ much between species. And indeed, as we move up the phylogenetic tree, alternative splicing complexity increases, with the highest complexity in primates [66, 67].

4.1 Major and minor spliceosome

RNA splicing is performed by the spliceosome, a large and dynamic ribonucleo-protein complex composed of proteins and small nuclear RNAs (snRNAs), which assembles on the pre-mRNA (**Figures 1 and 2**).

Figure 1.
Two-step splicing reaction. Splicing occurs by a 2-step trans-esterification reaction to remove introns and join exons together. The first step, U1 small nuclear ribonucleoprotein (snRNP) assembles at the 5' splice site of an exon and U2 snRNP at the branch point sequence (BPS), just upstream of the 3' splice site of the adjacent/ downstream exon. This configuration is known as the pre-spliceosome. Hereafter, U1 and U2 are joined by the snRNPs U5 and U4–U6 complexes to form the pre catalytic spliceosome. Next, U4–U6 complexes unwind, releasing U4 and U1 from the pre-spliceosomal complex. This allows U6 to base pair with the 5' splice site and the BPS. The 5' splice site gets cleaved, which leads to a free 3' OH-group at the upstream exon, and a branched intronic region at the downstream exon called the intron lariat.

Figure 2. Major and minor splicing. (A) Major, and minor splicing. The major introns are spliced out, and minor introns are either retained (and the mRNA is most often subsequently degraded) or the minor intron is spliced out, and a mature mRNA is formed. (B) The 4 basic splicing signals are the 5' splice donor site, the 3' splice acceptor site, the branch point sequence (BPS), and the polypyrimidine tract (PT). Spliceosomal components recognize and bind to these sequences and mediate the splicing reaction. Intronic and exonic splicing enhancers and silencers determine the inclusion rate of exons. The BPS (major, YNYURAY; minor, UCCUUAACU) is located 20–50 bp upstream of the 3' splice site, and the PT (Y_{10–12}) is located in between the BPS and the 3' splice site (N=any nucleotide, Y=C or U, R=A, or G and S=C or G). (C) Minor splicing uses different 5' and 3' splice sites and BPS, and lacks the PT. ESE indicates exonic splicing enhancers; ESS, exonic splicing silencers; ISE, intronic splicing enhancers; and ISS, intronic splicing silencers. Note: The **Figures 1 and 2** are a modification from van den Hoogenhof et al. [68].

During the second step, U5 pairs with sequences in both the 5' and 3' splice sites, positioning the 2 ends together. The 3' OH-group of the upstream (5') exon fuses with the 3' intron-exon junction, thereby conjoining the 2 exons and excising the intron in the form of a lasso-shaped intron lariat. Finally, the spliceosome disassembles, and all components are recycled for future splicing reactions.

Recent evidence has shown that splicing does not occur after transcription, but happens during transcription; therefore, the vast majority of human introns are spliced out when transcription is still taking place [69].

4.2 RNA splicing in cardiomyopathy

Several mouse models suggest a role for splicing factors in postnatal heart development. One such example is the alternative splicing factor ASF/SF2 (or SFRS1), an SR protein that is ubiquitously expressed and acting as an alternative splicing regulator [70]. ASF/SF2 conditional knockout mice die 6–8 weeks after birth, due to hypercontractile cardiac phenotype caused by a defect in Ca^{2+} handling. When ASF/SF2 is deleted, it leads to mis-splicing of several genes, including cardiac troponin T (cTnT), LIM-domain binding 3 (LDB3), and Ca^{2+} /calmodulin-dependent protein kinase (CamkII δ). Atypical alternative splicing of CamkII δ , cTnT, and LDB3 can present 20 days after birth, even though ASF/SF2 was deleted at the early stages of cardiogenesis.

Mis-splicing of CamkII δ in ASF/SF2 knockout hearts can lead to perturbation of Ca^{2+} handling and severe excitation-contraction coupling defects, which in turn leads to dilated cardiomyopathy (DCM).

Embryonic lethality may occur in systemic deletion of SC35 in mice, even before the onset of cardiogenesis [71]. Attempt to bypass this problem by generating a heart-specific knockout of SC35 uncovered the role of SC35 in the heart, as cardiac hypertrophy and DCM developed in these mice at 5–6 weeks of age [71].

In conclusion, ablation of SC35 in the heart shows that proper expression of this splice factor during postnatal heart development is essential to maintain cardiac form and function.

Severe and lethal DCM has been reported to occur 2 weeks after birth in mice with deletion of hnRNP U in the mouse heart [72]. The importance of alternative splicing of Ca^{2+} -handling genes in early postnatal heart development can be observed in the role of heterogeneous nuclear ribonucleoprotein U (HnRNP U) in splicing of calcium/calmodulin-dependent protein kinase II δ (CamkII δ).

4.3 Role of alternative splicing in disease phenotype

Atypical alternative splicing has been documented to contribute to disease severity and susceptibility [73]; as observed in retinitis pigmentosa, Prader-Willi syndrome, and spinal muscular atrophy [74, 75]. Spinal muscular atrophy, for example, is caused by the loss of the survivor of the motor neuron-1 (SMN1) gene, which is required for proper assembly and transport of snRNP [74].

Kong et al. [76] used a genome-wide approach to study alternative splicing changes in the diseased heart. The splicing of 4 key sarcomeric genes, troponin T (TNNT)-2, TNNI3, MYH7, and FLNC, were significantly altered in human ischemic cardiomyopathy, DCM, and aortic stenosis.

5. Epigenetic DNA modifiers

Epigenetics and its attendant markers influence the proliferation of diseases and their phenotypes. Outside, DNA canonical structure, DNA folds into alternative structures including DNA hairpins, cruciforms, triplexes or G-quadruplexes (G4), and holiday junctions [77, 78]. Besides these DNA structural changes, epigenetics processes, using DNA methylation and histone modification as the driver, are another primary vehicle for changes in DNA. Changes due to epigenetics modification with time can alter our phenotypes profoundly. Known facts are that everything from what we eat, drink, and smoke to other factors within our immediate environment including, stress can interfere with the way our genes express themselves up and down the line with the finest totality [79]. The primary

vehicles for epigenetic changes are DNA methylation and histone modification; there are many known enzymes that act on histone modifications by either adding or removing the covalent modifications. Such changes influence the degree of interaction between DNA and histone, which have some profound effects on the ability of that DNA to be transcribed. Histone modifications are subject to rapid changes (in seconds/minutes), giving room for the cell to respond to external stimuli. Furthermore, many of the known enzymes responsible for modifying histone residues have numbers of non-histone substrates such as transcription factors [80, 81].

Some mechanisms for the function of histone modifications have been characterized including; the compression of chromatin, and the recruitment of non-histone proteins [82]. There are different types of modification and these determine the amino acid residue produced. The modifications of histone lead to either gene activation or repression, and the addition of acetyl groups, to the tail of histone H3, neutralizes the basic charge of the lysine, leading to the unfolding of the chromatin, allowing transcription to occur. Conversely, the removal of these acetyl groups results in chromatin compression, which prevents transcription [82]. These kinds of changes in chromatin structure help to prevent access by other proteins that can further modify the chromatin (e.g., remodeling ATPases).

Understanding the etiology of some of these diseases, from PPI, protein DNA/RNA interaction is important as it will herald in more robust drug treatments for patients

Figure 3. Factors influencing epigenetics (modified from [84]): complex interplay is required for a wholesome gene expression understanding; such complexes will further enhance the use of phytomedicine in disease management.

with specific disease phenotypes. Along this line, Okoh et al. [81] recently, using available data, espoused the need for the combination of herbal medicine to target some epigenetic markers by way of epigenetic engineering (site-specific DNA binding module fusions with DNA demethylating enzymes for epigenetic induction of for instance; fetal hemoglobin (HbF) for therapy of sickle cell disease (SCD)). This is in consonance with earlier postulation [83, 84], implying such technique may provide a better way to activate/or repress inherent gene expression, bearing transient modification of DNA and histones should remain stable over many cell divisions helping in delaying HbF switching [83, 84].

Moreover, Okoh et al. [81], suggested that the de-methylation of DNA at the CpGs site on both DNA strands may be possible using the combination of herbal medicine, foods rich in flavonoids could be vital in tweaking histone acetylation, which can modulate gene expression. The figure below postulates the complex interplay between, epigenetics and phyto-compound modifiers towards enabling gene transcription for proper protein translation (**Figure 3**).

6. Future perspective

CRISPR/Cas9-based therapy, are been used as a candidate to be administered systemically, via intravenous infusion, for precision editing of a gene in target tissue in humans [85]. Similarly using this technology, gene therapy was developed to treat the rare neurodegenerative condition, Dopamine Transporter Deficiency Syndrome (DTDS) using a personalized approach with a view to counter the exact genetic fault present in a patient's neurons [85]. Using a novel approach where, skin cells from patients, turned into pluripotent stem cells in the laboratory with the aim to get neuronal cells with the disease-causing mutation. A vector carrying adeno-associated virus gene therapy was created to target the neurological fault and its efficacy was tested in both neuronal human cell lines and a mouse model, with the corresponding loss of function mutations in *SLC6A3* [85]. This research ingenuity/approach has provided, promising results leading to some clinical trials that may put an end to this cruel disease.

DTDS is an area of unmet medical needs and the disease is also known as infantile parkinsonism-dystonia, due to it having neurodegenerative and movement symptoms similar to Parkinson's disease [85]. It is a very rare inherited condition known to affect around 50 children around the world. Although this might be due to under-diagnosis by clinicians bearing the symptoms are similar to other inherited movement disorders e.g., cerebral palsy [85].

Environmental factors are implicated in the formation of ROS affecting human health by directing epigenetics signature of the genome, such could also drive the addition of methyl group ($-CH_3$) to some nucleotides neighboring guanosine (CpG islands) of the genome. These are areas where drug repurposing becomes essential as they could target methylation processes which are amongst, inherent biochemical/epigenetics machinery of cells, containing necessary pathways that allow environmental agents to induce mutations. Bearing these epigenetic signatures play a significant role in genomic balance, they play a leading role in several diseases hence are the essential target for drug repurposing.

Many diseases present some inherent opportunities via epigenetics markers that required intelligent manipulation of phyto-compounds to access new therapy that is efficient and easily accessible. Phytochemicals are known to play vital roles in preventing oxidative stress with concomitant damages [2, 85]. At the cellular and molecular level, they inactivate Reactive Oxygen Species (ROS). And under specific low concentration, inhibit or delay oxidative processes by interrupting the radical

chain reaction of lipid peroxidation [2, 86]. Bioactive components with anti-oxidative capacity naturally present in food are of great interest due to their beneficial effects on human health as they offer protection against oxidative deterioration.

DNA processes such as replication, transcription, recombination, and repair are, known to be facilitated by several factors covered in this chapter and others such as supercoiling that help facilitate both the packaging of DNA and many fundamental genetic processes that enabled the enzymatic manipulation of DNA. Aberrant RBP-RNA interactions are now known to promote disease progression, as much as mutations in TFs. RBP's role in disease was initially understudied because of their systematic evaluation was limited by, lack of sensitive and efficient assays for phenotypic interrogation of individual RBPs.

There is profound evidence that suggests, consumption of food rich in phytochemicals may progressively reduce the risk of different diseases by modulating immune-inflammatory markers [87]. Using the combination of disparate molecular/biophysical tools we recently [88], compared the binding affinity of artesunate and azadirachtin to gephyrin E this is towards enabling insights into natural bioactive compounds useful for rational drug design, essential in the race to manage myriad of disease phenotypes. The results from our research and others are necessary as they, may provide, the impetus for more studies into bioactive components of plant origin towards the effective management of different disease phenotypes.

6.1 Next-generation sequence in disease diagnosis

Next-generation sequencing (NGS), is a massively parallel and a high-throughput DNA sequencing technology that enables the fast generation of data on thousands to millions of base pairs of DNA from an individual patient by sequencing large numbers of genes in a single reaction [89]. NGS can sequence millions of DNA fragments in a massively parallel fashion, instead of sequencing a single DNA fragment one at a time, as observed in traditional capillary electrophoresis sequencing. The general workflow of NGS includes four main steps:

- I. library preparation,
- II. cluster generation,
- III. sequencing, and
- IV. data analysis.

Sequence reads are produced from fragment libraries, a pool of adaptor-ligated and enriched DNA fragments. One advantage is that a small quantity of DNA, from a patient, is needed to produce a library.

In step 1, patient DNA is randomly fragmented by different methods and then prepared for sequencing by ligating specific adaptor oligonucleotides to both ends of each DNA fragment. Adapter-ligated fragments are further enriched with specific oligonucleotides designed for the target genes included in the NGS panel and are then amplified by polymerase chain reaction (PCR). The prepared library is loaded into a flow cell for cluster generation and subsequent sequencing.

During sequencing, short read lengths (35–250 bp, depending on the platform) sequences that are produced are then aligned to a reference genome with bioinformatics software [89].

During data analysis, variant calling can be achieved by various standard and in-house analysis pipelines. All detected variants are checked against standard databases (e.g., dbSNP137, 1000 Genomes Project, Exome Variant Server, ExAC Browser, OMIM catalog, ClinVar, Human Gene Mutation Database) to enable interpretation of the pathogenicity of a given variant.

Next-generation sequencing panels are now commonly used in clinical diagnosis to identify genetic causes of various monogenic disease groups, such as epilepsy [90], intellectual disability [91, 92], neurodevelopmental disorders [93], neuro-metabolic disorders [94], amongst others.

The use of NGS in clinical laboratories is increasing, with application in the diagnosis of immune disorders, infectious diseases, human hereditary disorders, in non-invasive prenatal diagnosis, and recently, in the therapeutic decision making for somatic cancers [95, 96].

Today two different NGS technologies are mainly used in clinical laboratories: Ion Torrent and Illumina systems [97].

The Ion Torrent exploited the emulsion PCR using native dNTP chemistry that releases hydrogen ions during base incorporation by DNA polymerase and a modified silicon chip detecting the pH modification [98], while Illumina technology is based on the existing *Solexa sequencing by synthesis* chemistry with the use of very small flow-cells, reduced imaging time and fast sequencing process [97].

6.2 Usefulness of NGS

NGS approaches will remain useful because:

1. It is highly accurate and cost-effective.
2. It has a wide application for use in clinically heterogeneous inherited disorders, resulting in an increase in the number of reported disease-causing genes.

NGS is appealing when there is a genetic contribution in heterogeneous and complex diseases, such as in cardiomyopathies, in cardiac arrhythmias, in connective tissue disorders, in mental retardation or autism, and where a large number of genes are involved in a large phenotypic syndrome [99, 100]. In these cases, NGS approaches allow us to test a large number of genes simultaneously in a cost-effective manner [101].

Two options of NGS are currently available [101]:

1. Targeted gene panels sequencing or
2. Whole-exome sequencing (WES).

Targeted sequencing is applicable for genetic disorders, such as non-syndromic deafness [98], common diseases, such as hypertension and diabetes [102], or in traditional cytogenetic and Mendelian disorder diagnosis [103]. The main limitation of targeted sequencing is the rigidity of testing only a selected number of genes. Since the genetic field is rapidly evolving, new genes may be associated with a clinical phenotype, and as such redesigning and revalidation of the panel is needed [101].

The WES application could be applicable for the identification of genes responsible for the dominant Freeman-Sheldon syndrome, the recessive Miller syndrome, and the dominant Schinzel-Giedion syndrome [104]. The shortcoming of WES is

that about 10% of targeted bases sequenced in WES do not get the 20 read depth [105], required for clinical confidence and interpretation, and approximately only 85% of genes associated with human diseases into the principle database (OMIM) receive the adequate coverage [106].

6.3 Challenges of NGS in disease diagnosis

In the NGS process, one limiting step is the complexity of genetic variation interpretation in whole-exome, due to the presence of thousands of rare single nucleotide variations without pathogenic effect. Moreover, in the majority of human diseases, the pathological phenotype may be caused by a pathogenic rare mutation with a strong effect or it may be caused by a co-presence of multiple genetic variations [107].

Another important challenge of the use of the NGS approach in clinical diagnosis is the management of the amount of data generated [108]. Indeed generation, analysis, and also storage of NGS data require sophisticated bioinformatics infrastructure [109], which could be capital intensive.

A skilled bio(chem)-informatics staff is needed to manage and analyze NGS data, therefore increasing the impact of computing infrastructure and manpower on costs of NGS applications in clinical diagnostics [110, 111].

7. Conclusions

The interest in studying protein interactions, their variations, and their constituent's effects on pathological conditions has grown within the last few decades. These interests are behind, for instance, the increasing examination of the application of mass spectrophotometer (MS)-based experimental analyses of model systems to explore heterogeneous PPI networks and protein complexes, which will promote drug repurposing within the context of human diseases.

The use of other robust technology such as NGS in the study of biological processes, which would have otherwise remained elusive, is the driver for the actualization of personalized medicine for which drug repurposing form a central aspect. Molecular interactions at both the DNA-RNA/protein level are sequined with the SNP, epigenetics, mutations causing variants, and other factors emanating from our immediate environment e.g. stress. Technological advancement discussed in this chapter is all part of the process for science and scientists to understand the biological phenomenon that governs life at the molecular level.

Many disease variants resulting from SNP, single point mutation, and RBP role in disease manifestation are discussed with a view to heighten our thinking towards drug repurposing, the uptake of phytomedicine, and bioactive compounds in the management of various disease phenotypes.

IntechOpen

IntechOpen

Author details

Michael P. Okoh* and Lukman A. Alli
Faculty of Basic Medical Sciences, Department of Medical Biochemistry, College of
Health Sciences, University of Abuja, Abuja, Nigeria

*Address all correspondence to: michael.okoh@uniabuja.edu.ng

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

References

- [1] Vihinen M. Functional effects of protein variants. *Biochimie*. 2021; **180**:104-120. DOI: 10.1016/j.biochi.2020.10.009. Epub 2020 Oct 23
- [2] Okoh MP. Antioxidants an epigenetics regulator for the prevention of diseases and aging process. *Journal of Genetic Engineering and Biotechnology Research*. 2019;**1**(1):1-5
- [3] Trinder M, Francis GA, Brunham LR. Association of monogenic vs polygenic hypercholesterolemia with risk of atherosclerotic cardiovascular disease. *JAMA Cardiology*. 2020;**5**(4):390-399. DOI: 10.1001/jamacardio.2019.5954
- [4] Embarc-Buh A, Francisco-Velilla R, Martinez-Salas E. RNA-binding proteins at the host-pathogen interface targeting viral regulatory elements. *Viruses*. 2021;**13**(6):952. DOI: 10.3390/v13060952
- [5] Kaur R, Lal SK. The multifarious roles of heterogeneous ribonucleoprotein A1 in viral infections. *Reviews in Medical Virology*. 2020;**30**:2. DOI: 10.1002/rmv.2097
- [6] Gebhardt A, Habjan M, Benda C, et al. mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3. *Nature Communications*. 2015;**6**:8192. DOI: 10.1038/ncomms9192
- [7] Kamel W, Noerenberg M, Cerikan B, Chen H, et al. Global analysis of protein-RNA interactions in SARS-CoV-2-infected cells reveals key regulators of infection. *Molecular Cell*. 2021;**81**(13): 2851-2867. DOI: 10.1016/j.molcel.2021.05.023
- [8] Hentze M, Castello A, Schwarzl T, et al. A brave new world of RNA-binding proteins. *Nature Reviews. Molecular Cell Biology*. 2018;**19**:327-341. DOI: 10.1038/nrm.2017.130
- [9] Einstein JM, Perelis M, Chaim IA, Meena JK, Nussbacher JK, Tankka AT, et al. Inhibition of YTHDF2 triggers proteotoxic cell death in MYC-driven breast cancer. *Molecular Cell*. 2021;**81**(15):3048-3064.e9. DOI: 10.1016/j.molcel.2021.06.014
- [10] Liu Y, Shi S. The roles of hnRNP A2/B1 in RNA biology and disease. *Wiley Interdisciplinary Review*. 2021;**12**(2): e-1612
- [11] Al-Numair NS, Martin AC. The SAAP pipeline and database: Tools to analyze the impact and predict the pathogenicity of mutations. *BMC Genomics*. 2013;**14**(Suppl. 3):S4
- [12] Yair T, Alejandra H, Victor A, Hector V, Claudia M. Protein-protein and protein-ligand docking. In: *Protein Engineering-Technology and Application*. 1st ed. 2013. p. 187
- [13] Gans MD, Bernstein L, Shliozberg J, Gavrilova T, Rubinstein A. Outcomes of 3 patients with adenosine deaminase deficiency on long term enzyme replacement therapy. *Annals of Allergy, Asthma & Immunology*. 2021;**126**(5): 593-595
- [14] Li Y, Chen C, Kaye AM, Wasserman WW. The identification of cis-regulatory elements: A review from a machine learning perspective. *Biosystems*. 2015;**138**:6-17
- [15] Badihian N, Daniali SS, Kelishadi R. Transcriptional and epigenetic changes of brain derived neurotrophic factor following prenatal stress: A systematic review of animal studies. *Neuroscience & Biobehavioral Reviews*. 2020;**117**:12, 18=INTECH Open Access
- [16] Kumar R, Paray HA, Shrivastava T, Sinha S, Luthra K. Phage display antibody libraries: A robust approach

for generation of recombinant human monoclonal antibodies. *International Journal of Biological Macromolecule*. 2019;**15**(135):907-918

[17] Sheehan J, Marasco WA. Phage and yeast display. *Microbiology Spectrum*. 2014;**3**(1):AID-0028

[18] Tabata N, Horisawa K, Yanagawa H. Applications of the in vitro virus (IVV) method for various protein functional analyses. In: *Protein Engineering-Technology and Application*. 2013

[19] Tateyama S, Yanagawa H. Application of mRNA display for in vitro selection of DNA-binding transcription factor complexes. In: Bina M, editor. *Gene Regulation. Methods in Molecular Biology (Methods and Protocols)*. Totowa, NJ: Humana Press; 2013. p. 977. DOI: 10.1007/978-1-62703-284-1_8

[20] Niroula A, Vihinen M. Variation interpretation predictors: Principles, types, performance, and choice. *Human Mutation*. 2016;**37**(6):579-597. DOI: 10.1002/humu.22987

[21] Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v2.0: The next generation in gene variant databases. *Human Mutation*. 2011;**32**:557-563

[22] Thorisson G, Muilu J, Brookes A. Genotype-phenotype databases: Challenges and solutions for the post-genomic era. *Nature Reviews. Genetics*. 2011;**10**:9-18. DOI: 10.1038/nrg2483

[23] Singh JA, Wells GA, Christensen R, Ghogomu T, et al. Adverse effects of biologics: A network meta-analysis and Cochrane overview. *The Cochrane Database of Systematic Reviews*. 2011;**2**:CD008794. DOI: 10.1002/14651858.CD008794.pub2

[24] Lopes P, Oliveira JL. An automated real-time integration and

interoperability framework for bioinformatics. *BMC Bioinformatics*. 2015;**16**:1. DOI: 10.1186/s12859-015-0761-3

[25] Vihinen M, den Dunnen JT, Dalglish R, Cotton RGH. Guidelines for establishing locus specific databases. *Human Mutation*. 2011;**33**(2):298-305. DOI: 10.1002/humu.21646

[26] Mejzini R, Flynn LL, Pitout IL, Fletcher S, Wilton SD, Akkari PA. ALS genetics, mechanisms, and therapeutics: Where are we now? *Frontiers in Neuroscience*. 2019;**13**:1310. DOI: 10.3389/fnins.2019.01310

[27] Fogh I, Ratti A, Gellera C, Lin K, Tiloca C, et al. A genome-wide association meta-analysis identifies a novel locus at 17q11.2 associated with sporadic amyotrophic lateral sclerosis. *Human Molecular Genetics*. 2014;**23**(8):2220-2231. DOI: 10.1093/hmg/ddt587

[28] Zou Z-Y, Zhou Z-R, Che C-H, Liu C-Y, He R-L, Huang H-P. Genetic epidemiology of amyotrophic lateral sclerosis: A systematic review and meta-analysis. *Journal of Neurology, Neurosurgery, and Psychiatry*. 2017;**88**:540-549

[29] Kaul DK, Fabry ME, Nagel RL. Microvascular sites and characteristics of sickle cell adhesion to vascular endothelium in shear flow conditions: Pathophysiological implications. *Proceedings of the National Academy of Sciences of the United States of America*. 1989;**86**:3356-3360

[30] Kaur H, Aurif F, Kittaneh M, Chio JPG, Malik BH. Cardiomyopathy in sickle cell disease. *Cureus*. 2020; **12**(8):e9619

[31] Butti Z, Pan YE, Giacomotto J, Patten SA. Reduced C9orf72 function leads to defective synaptic vesicle release and neuromuscular dysfunction in zebrafish. *Communications Biology*.

2021;**4**(1). DOI: 10.1038/s42003-021-02302-y

[32] Stanga S, Caretto A, Boido M, Vercelli A. Mitochondrial dysfunctions: A red thread across neurodegenerative diseases. *International Journal of Molecular Sciences*. 2020;**21**(10):3719. DOI: 10.3390/ijms21103719

[33] Rouillard ME, Sutter PA, Durham OR, Willis CM, Crocker SJ. Astrocyte-derived extracellular vesicles (ADEVs): Deciphering their influences in aging. *Aging and Disease*. 2021;**12**(6):1462-1475

[34] Thygesen K, Alpert JS, Jaffe AS, Chaitman BR, Bax JJ, Morrow DA, et al. Fourth universal definition of myocardial infarction. *Journal of the American College of Cardiology*. 2018;**72**:2231-2264

[35] Park KC, Gaze DC, Collinson PO, Marber MS. Cardiac troponins: From myocardial infarction to chronic disease. *Cardiovascular Research*. 2017; **113**:1708-1718

[36] Katus HA, Remppis A, Looser S, Hallermeier K, Scheffold T, Kübler W. Enzyme linked immuno assay of cardiac troponin T for the detection of acute myocardial infarction in patients. *Journal of Molecular and Cellular Cardiology*. 1989;**21**:1349-1353

[37] Alpert JS, Thygesen K, Antman E, Bassand JP. Myocardial infarction redefined—a consensus document of The Joint European Society of Cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction. *Journal of the American College of Cardiology*. 2000;**36**:959-969

[38] Thygesen K, Alpert JS, White HD Joint ESC/ACCF/AHA/WHF Task Force for the Redefinition of Myocardial Infarction. Universal definition of myocardial infarction. *European Heart Journal*. 2007;**28**:2525-2538

[39] Jishi F, Hudson PR, Williams CP, et al. Troponin I, laboratory issues, and clinical outcomes in a district general hospital: Crossover study with “traditional” markers of myocardial infarction in a total of 1990 patients. *Journal of Clinical Pathology*. 2004;**57**:1027-1032

[40] Ottani F, Galavani M, Nicolini FA. Elevated cardiac troponin levels predict the risk of adverse outcome in patients with acute coronary syndromes. *American Heart Journal*. 2000;**140**:917

[41] Heidenreich PA, Alloggiamento T, Melsop K, et al. The prognostic value of troponin in patients with non-ST elevation acute coronary syndromes: A meta-analysis. *Journal of the American College of Cardiology*. 2001;**38**: 478-485

[42] Apple FS, Sandoval Y, Jaffe AS, Ordonez-Llanos J. IFCC Task Force on Clinical Applications of Cardiac Bio-Markers, Cardiac troponin assays: Guide to understanding analytical characteristics and their impact on clinical care. *Clinical Chemistry*. 2017;**63**:73-81

[43] Apple FS, Ler R, Murakami MM. Determination of 19 cardiac troponin I and T assay 99th percentile values from a common presumably healthy population. *Clinical Chemistry*. 2012;**58**:1574-1581

[44] Collinson PO, Heung YM, Gaze D, Boa F, Senior R, Christenson R, et al. Influence of population selection on the 99th percentile reference value for cardiac troponin assays. *Clinical Chemistry*. 2012;**58**:219-225

[45] Gore MO, Seliger SL, Defilippi CR, Nambi V, Christenson RH, Hashim IA, et al. Age- and sex-dependent upper reference limits for the high-sensitivity cardiac troponin T assay. *Journal of the American College of Cardiology*. 2014;**63**:1441-1448

- [46] Eggers KM, Apple FS, Lind L, Lindahl B. The applied statistical approach highly influences the 99th percentile of cardiac troponin I. *Clinical Biochemistry*. 2016;**49**:1109-1112
- [47] Monneret D, Gellerstedt M, Bonnefont-Rousselot D. Determination of age- and sex-specific 99th percentiles for high-sensitive troponin T from patients: An analytical imprecision- and partitioning-based approach. *Clinical Chemistry and Laboratory Medicine*. 2018;**56**:685-696
- [48] Kociol RD, Pang PS, Gheorghiade M, Fonarow GC, O'Connor CM, Felker GM. Troponin elevation in heart failure prevalence, mechanisms, and clinical implications. *Journal of the American College of Cardiology*. 2010;**56**:1071-1078
- [49] Corte Z, García C, Venta R. Biological variation of cardiac troponin T in patients with end-stage renal disease and in healthy individuals. *Annals of Clinical Biochemistry*. 2015;**52**:53-60
- [50] Apple FS, Jaffe AS, Collinson P, Mockel M, Ordonez-Llanos J, Lindahl B, et al. IFCC educational materials on selected analytical and clinical applications of high sensitivity cardiac troponin assays. *Clinical Biochemistry*. 2015;**48**:201-203
- [51] Goldhaber SZ. Pulmonary embolism. *Lancet*. 2004;**363**:1295-1305
- [52] Sato Y, Yamada T, Taniguchi R, et al. Persistently elevated serum concentrations of cardiac troponin T in patients with idiopathic dilated cardiomyopathy are predictive of adverse outcomes. *Circulation*. 2000;**103**:369-374
- [53] Sato T, Taniguchi R, Nagai K. Measurement of cardiac troponin I in patients with hypertrophic cardiomyopathy. *Heart*. 2003;**89**:659-660
- [54] Lauer B, Niedrau C, Kuhl, et al. Cardiac troponin T in patients with clinically suspected myocarditis. *Journal of the American College of Cardiology*. 1997;**30**:1354-1359
- [55] Imazo M, Demichelis B, Cecci E, et al. Cardiac troponin I in acute pericarditis. *Journal of the American College of Cardiology*. 2003;**42**: 2144-2148
- [56] Ammann P, Fehr T, Minder EI, et al. Elevation of troponin I in sepsis and septic shock. *Intensive Care Medicine*. 2001;**27**:965-969
- [57] Cardinale D, Sandri MT, Martinoni A, et al. Left ventricular dysfunction predicted by high early troponin I release after high dose chemotherapy. *Journal of the American College of Cardiology*. 2000;**36**:517-522
- [58] Klinkenberg LJ, van Dijk JW, Tan FE, van Loon LJ, van Dieijen-Visser MP, Meex SJ. Circulating cardiac troponin T exhibits a diurnal rhythm. *Journal of the American College of Cardiology*. 2014;**63**: 1788-1795
- [59] Aakre KM, Røraas T, Petersen PH, Svarstad E, Sellevoll H, Skadberg Ø, et al. Weekly and 90-minute biological variations in cardiac troponin T and cardiac troponin I in hemodialysis patients and healthy controls. *Clinical Chemistry*. 2014;**60**:838-847
- [60] Klinkenberg LJ, Wildi K, van der Linden N, Kouw IW, Niens M, Twerenbold R, et al. Diurnal rhythm of cardiac troponin: Consequences for the diagnosis of acute myocardial infarction. *Clinical Chemistry*. 2016;**62**:1602-1611
- [61] Chow LT, Gelinas RE, Broker TR, Roberts RJ. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*. 1977;**12**:1-8

- [62] Sharp PA. Split genes and RNA splicing. *Cell*. 1994;**77**:805-815
- [63] Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, et al. Alternative isoform regulation in human tissue transcriptomes. *Nature*. 2008;**456**:470-476. DOI: 10.1038/nature07509
- [64] Kornblihtt AR, Schor IE, Alló M, Dujardin G, Petrillo E, Muñoz MJ. Alternative splicing: A pivotal step between eukaryotic transcription and translation. *Nature Reviews. Molecular Cell Biology*. 2013;**14**:153-165. DOI: 10.1038/nrm3525
- [65] Gueroussov S, Gonatopoulos-Pournatzis T, Irimia M, Raj B, Lin ZY, Gingras AC, et al. An alternative splicing event amplifies evolutionary differences between vertebrates. *Science*. 2015;**349**:868-873. DOI: 10.1126/science.aaa8381
- [66] Barbosa-Morais NL, Irimia M, Pan Q, et al. The evolutionary landscape of alternative splicing in vertebrate species. *Science*. 2012;**338**:1587-1593. DOI: 10.1126/science.1230612
- [67] Merkin J, Russell C, Chen P, Burge CB. Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. *Science*. 2012;**338**:1593-1599. DOI: 10.1126/science.1228186
- [68] van den Hoogenhof MMG, Pinto YM, Creemers EE. RNA splicing: Regulation and dysregulation in the heart. *Circulation Research*. 2016; **118**(3):454-468
- [69] Tilgner H, Knowles DG, Johnson R, Davis CA, Chakraborty S, Djebali S, et al. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Research*. 2012; **22**:1616-1625. DOI: 10.1101/gr.134445.111
- [70] Xu X, Yang D, Ding JH, et al. ASF/SF2-regulated CaMKII δ alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. *Cell*. 2005;**120**:59-72. DOI: 10.1016/j.cell.2004.11.036
- [71] Ding JH, Xu X, Yang D, Chu PH, Dalton ND, Ye Z, et al. Dilated cardiomyopathy caused by tissue-specific ablation of SC35 in the heart. *The EMBO Journal*. 2004;**23**:885-896. DOI: 10.1038/sj.emboj.7600054
- [72] Ye J, Beetz N, O'Keeffe S, Tapia JC, Macpherson L, Chen WV, et al. hnRNP U protein is required for normal pre-mRNA splicing and postnatal heart development and function. *Proceedings of the National Academy of Sciences of the United States of America*. 2015;**112**:E3020-E3029. DOI: 10.1073/pnas.1508461112
- [73] Wang GS, Cooper TA. Splicing in disease: Disruption of the splicing code and the decoding machinery. *Nature Reviews. Genetics*. 2007;**8**:749-761. DOI: 10.1038/nrg2164
- [74] Lefebvre S, Bürglen L, Reboullet S, Clermont O, Burlet P, Viollet L, et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*. 1995;**80**:155-165
- [75] Kishore S, Stamm S. The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science*. 2006;**311**:230-232. DOI: 10.1126/science.1118265
- [76] Kong SW, Hu YW, Ho JW, Ikeda S, Polster S, John R, et al. Heart failure-associated changes in RNA splicing of sarcomere genes. *Circulation. Cardiovascular Genetics*. 2010;**3**:138-146. DOI: 10.1161/CIRCGENETICS.109.904698

- [77] Bacolla A, Wells RD. Non-B DNA conformations, genomic rearrangements, and human disease. *Journal of Biological Chemistry*. 2004;**279**:47411-47414
- [78] Kosiol N, Juranek S, Brossart P, Heine A, Paeschke K. G-quadruplexes: A promising target for cancer therapy
- [79] Crewsa D, Gillettea R, Samuel VS, et al. Epigenetic transgenerational inheritance of altered stress responses. *PNAS*. 2012. Available from: <http://www.pnas.org/content/early/2012/05/15/1118514109.full.pdf+html?withds=yes>
- [80] Bassett SA, Barnett MPG. The role of dietary histone deacetylases (HDACs) inhibitors in health and disease. *Nutrients*. 2014;**6**:4273-4301. DOI: 10.3390/nu6104273
- [81] Okoh MP, Alli LA, Tolvanen ME, Nwegbu MM. Herbal drug use in sickle cell disease management; trends and perspectives in Sub-Saharan Africa. *Current Drug Discovery Technologies*. 2019. DOI: 10.2174/1570163815666181002101611
- [82] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. *Nucleic Acids Research*. 2000;**2000**(28):235-242
- [83] Gilbert ER, Liu D. Flavonoids influence epigenetic-modifying enzyme activity:structure - function relationships and the therapeutic potential for cancer. *Current Medicinal Chemistry*. 2010;**17**(17):1756-1768
- [84] Busch C, Burkard M, Leischner C, Lauer UM, Frank J, Venturelli S. Epigenetic activities of flavonoids in the prevention and treatment of cancer. *Clinical Epigenetics*. 2015;**7**:64. DOI: 10.1186/s13148-015-0095
- [85] Ng J, Barral S, Barrigon C, et al. Gene therapy restores dopamine transporter expression and ameliorates pathology in iPSC and mouse models of infantile parkinsonism. *Science Translational Medicine*. 2021;**13**:594. DOI: 10.1126/scitranslmed.aaw1564
- [86] Halliwell B, Whiteman M. Measuring reactive species and oxidative damage in vivo and in cell culture: How should you do it and what do the results mean? *British Journal of Pharmacology*. 2004;**142**:231-255
- [87] Suroowan S, Mahomoodally F. Common phyto-remedies used against cardiovascular diseases and their potential to induce adverse events in cardiovascular patients. *Clinical Phytoscience*. 2015;**1**:1. DOI: 10.1186/s40816-015-0002-3
- [88] Okoh MP, Singla RK, Madu C, Soremekun O, Adejoh J, Alli LA, et al. Phytomedicine in disease management: In-silico analysis of the binding affinity of artesunate and azadirachtin for malaria treatment. *Frontiers in Pharmacology*. 2021;**12**:751032. DOI: 10.3389/fphar.2021.751032
- [89] Mardis ER. The impact of next-generation sequencing technology on genetics. *Trends in Genetics*. 2008; **24**(3):133-141
- [90] de Koning TJ, Jongbloed JD, Sikkema-Raddatz B, Sinke RJ. Targeted next-generation sequencing panels for monogenetic disorders in clinical diagnostics: The opportunities and challenges. *Expert Review of Molecular Diagnostics*. 2015;**15**(1): 61-70
- [91] Ajmabadi H, Hu H, Garshasbi M. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature*. 2011;**478**(7367):57-63
- [92] Hu H, Wienker TF, Musante L. Integrated sequence analysis pipeline provides one-stop solution for identifying disease-causing mutations.

Human Mutation. 2014;**35**(12): 1427-1435

[93] Okamoto N, Miya F, Tsunoda T. Targeted next-generation sequencing in the diagnosis of neurodevelopmental disorders. *Clinical Genetics*. 2015; **88**(3):288-292

[94] Ciara E, Rokicki D, Halat P. Difficulties in recognition of pyruvate dehydrogenase complex deficiency on the basis of clinical and biochemical features. The role of next-generation sequencing. *Molecular Genetics and Metabolism Reports*. 2016;**18**(7):70-76

[95] Winchester B. Lysosomal diseases: Diagnostic update. *Journal of Inherited Metabolic Disease*. 2014;**37**(4):599-608

[96] Carrera P, Di Resta C, Volonteri C, Castiglioni E, Bonfiglio S, Lazarevic D, et al. Exome sequencing and pathway analysis for identification of genetic variability relevant for broncho pulmonary dysplasia (BPD) in preterm new-borns: A pilot study. *Clinica Chimica Acta*. 2015;**451**:39-45. DOI: 10.1016/j.cca.2015.01.001

[97] Quail M, Smith ME, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: Comparison of ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics*. BioMed Central. 2012;**13**:341. DOI: 10.1186/1471-2164-13-341

[98] Merriman B, Rothberg JM. Progress in ion torrent semiconductor chip based sequencing. *Electrophoresis*. 2012;**33**:3397-3417. DOI: 10.1002/elps.201200424

[99] Harripaul R, Noor A, Ayub M, Vincent JB. The use of next-generation sequencing for research and diagnostics for intellectual disability. *Cold Spring Harbor Perspectives in Medicine*. 2017;**7**:a026864. DOI: 10.1101/cshperspect.a026864

[100] Di Resta C, Pietrelli A, Sala S, Della Bella P, De Bellis G, Ferrari M, et al. High-throughput genetic characterization of a cohort of Brugada syndrome patients. *Human Molecular Genetics*. 2015;**24**:5828-5835. DOI: 10.1093/hmg/ddv302

[101] Williams ES, Hegde M. Implementing genomic medicine in pathology. *Advances in Anatomic Pathology*. 2013;**20**:238-244. DOI: 10.1097/PAP.0b013e3182977199

[102] Szopa M, Ludwig-Gałęzowska A, Radkowski P, Skupień J, Zapala B, Piatek T, et al. Genetic testing for monogenic diabetes using targeted next-generation sequencing in patients with maturity-onset diabetes of the young. *Polskie Archiwum Medycyny Wewnętrznej*. 2015;**125**:845-851. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26552609>

[103] Ma ESK, Wan TSK, Au CH, Ho DN, Ma SY, Ng MHL, et al. Next-generation sequencing and molecular cytogenetic characterization of ETV6-LYN fusion due to chromosomes 1, 8 and 12 rearrangement in acute myeloid leukemia. *Cancer Genetics*. 2017;**218-219**:15-19. DOI: 10.1016/j.cancergen.2017.09.001

[104] Boycott KM, Vanstone MR, Bulman DE, MacKenzie AE. Rare-disease genetics in the era of next-generation sequencing: Discovery to translation. *Nature Reviews. Genetics*. 2013;**14**:681-691. DOI: 10.1038/nrg3555

[105] Yao R, Zhang C, Yu T, Li N, Hu X, Wang X, et al. Evaluation of three read-depth based CNV detection tools using whole-exome sequencing data. *Molecular Cytogenetics*. BioMed Central. 2017;**10**:30. DOI: 10.1186/s13039-017-0333-5

[106] Singh RR, Luthra R, Routbort MJ, Patel KP, Medeiros LJ. Implementation of next generation sequencing in clinical

molecular diagnostic laboratories:
Advantages, challenges and potential.
Expert Review of Precision Medicine
and Drug Development. 2016;**1**:109-120.
DOI: 10.1080/23808993.2015.1120401

[107] Frebourg T. The challenge for the
next generation of medical geneticists.
Human Mutation. 2014;**35**:909-911.
DOI: 10.1002/humu.22592

[108] Stein LD. The case for cloud
computing in genome informatics.
Genome Biology. 2010;**11**:207.
DOI: 10.1186/gb-2010-11-5-207

[109] Pabinger S, Dander A, Fischer M,
Snajder R, Sperk M, Efremova M, et al.
A survey of tools for variant analysis of
next-generation genome sequencing
data. Briefings in Bioinformatics.
2014;**15**:256-278. DOI: 10.1093/
bib/bbs086

[110] de Koning TJ, Jongbloed JDH,
Sikkema-Raddatz B, Sinke RJ. Targeted
next-generation sequencing panels for
monogenetic disorders in clinical
diagnostics: The opportunities and
challenges. Expert Review of Molecular
Diagnostics. 2015;**15**:61-70.
DOI: 10.1586/14737159.2015.976555

[111] Rehm HL, Bale SJ, Bayrak-
Toydemir P, Berg JS, Brown KK,
Deignan JL, et al. ACMG clinical
laboratory standards for next-
generation sequencing. Genetics
in Medicine. 2013;**15**:733-747.
DOI: 10.1038/gim.2013.92