*Essays in Biochemistry* (2018) **62** 399–408
https://doi.org/10.1042/EBC20170110

PORTLAND PRESS

Review Article

# Advancing genomic approaches to the molecular diagnosis of mitochondrial disease

**Sarah Louise Stenton**[1,2] **and Holger Prokisch**[1,2]

[1]Institute of Human Genetics, Technische Universität München, München 48559, Germany; [2]Institute of Human Genetics, Helmholtz Zentrum München, München 85764, Germany

**Correspondence:** Holger Prokisch (prokisch@helmholtz-muenchen.de)

Mitochondrial diseases present a diagnostic challenge due to their clinical and genetic heterogeneity. Achieving comprehensive molecular diagnosis via a conventional candidate-gene approach is likely, therefore, to be labour- and cost-intensive given the expanding number of mitochondrial disease genes. The advent of whole exome sequencing (WES) and whole genome sequencing (WGS) hold the potential of higher diagnostic yields due to the universality and unbiased nature of the methods. However, these approaches are subject to the escalating challenge of variant interpretation. Thus, integration of functional 'multi-omics' data, such as transcriptomics, is emerging as a powerful complementary tool in the diagnosis of mitochondrial disease patients for whom extensive prior analysis of DNA sequencing has failed to return a genetic diagnosis.

## Introduction

The molecular diagnosis of mitochondrial disease is subject to the complication of bi-genomic control of the energy-generating oxidative phosphorylation system (OXPHOS), as pathogenic variants can be identified in both the mtDNA and nuclear DNA (nDNA) [1]. In classical mitochondrial disease, the primary biochemical defect is located in the OXPHOS enzymes. Over 100 proteins are directly implicated in the biosynthesis and structural composition of the OXPHOS and its comprising complexes, 13 of which are encoded by the mtDNA with the remainder encoded by the nDNA. Additionally, mitochondrial function is subject to genes involved in regulating the expression of mtDNA, and those involved in mtDNA replication [2-4]. In total, ∼1500 genes are thought to be involved in healthy mitochondrial function [5] and to date, disease-causing variants have been identified in 309 genes implicated in mitochondrial metabolism (Figure 1) with presumably more pathogenic variants awaiting identification [6]. These disease-causing variants can have maternal, autosomal recessive, autosomal dominant, or X-linked inheritance patterns [7].

Defining the genetic aetiology of mitochondrial disease allows identification of ∼10% of mitochondrial diseases amenable to specific treatment strategies [8]. It is also crucial in enabling accurate genetic counselling, prenatal diagnosis and personalized disease surveillance for genotype-specific complications [9]. In addition to aid the understanding of the complex molecular basis of mitochondrial disease, mirrored by the expanding phenotypic spectrum.

Though conventional candidate gene sequencing may be the approach of choice in clearly defined phenotypes [10], the heterogeneity typical of mitochondrial disease supports the use of an unbiased approach early in the diagnostic pipeline [6] (Figure 2). Whole exome sequencing (WES) and whole genome sequencing (WGS) have evolved rapidly with the advent of high-throughput next-generation sequencing (NGS). As a consequence, the primary challenge has now shifted from the capacity to discover variants to the ability to interpret function and clinical impact [11,12] for which comprehensive American College of Medical Genetics and Genomics (ACMG) guidance has been developed [13]. A further promising approach to increase diagnostic yield is to integrate functional 'multi-omics' data [14] such as transcriptomics into the analysis pipeline. Thus facilitating detection of variants which have evaded detection or prioritization in WES and WGS approaches [15,16].

Received: 12 February 2018
Revised: 05 April 2018
Accepted: 08 May 2018

Version of Record published:
27 June 2018

*Essays in Biochemistry* (2018) **62** 399–408
https://doi.org/10.1042/EBC20170110

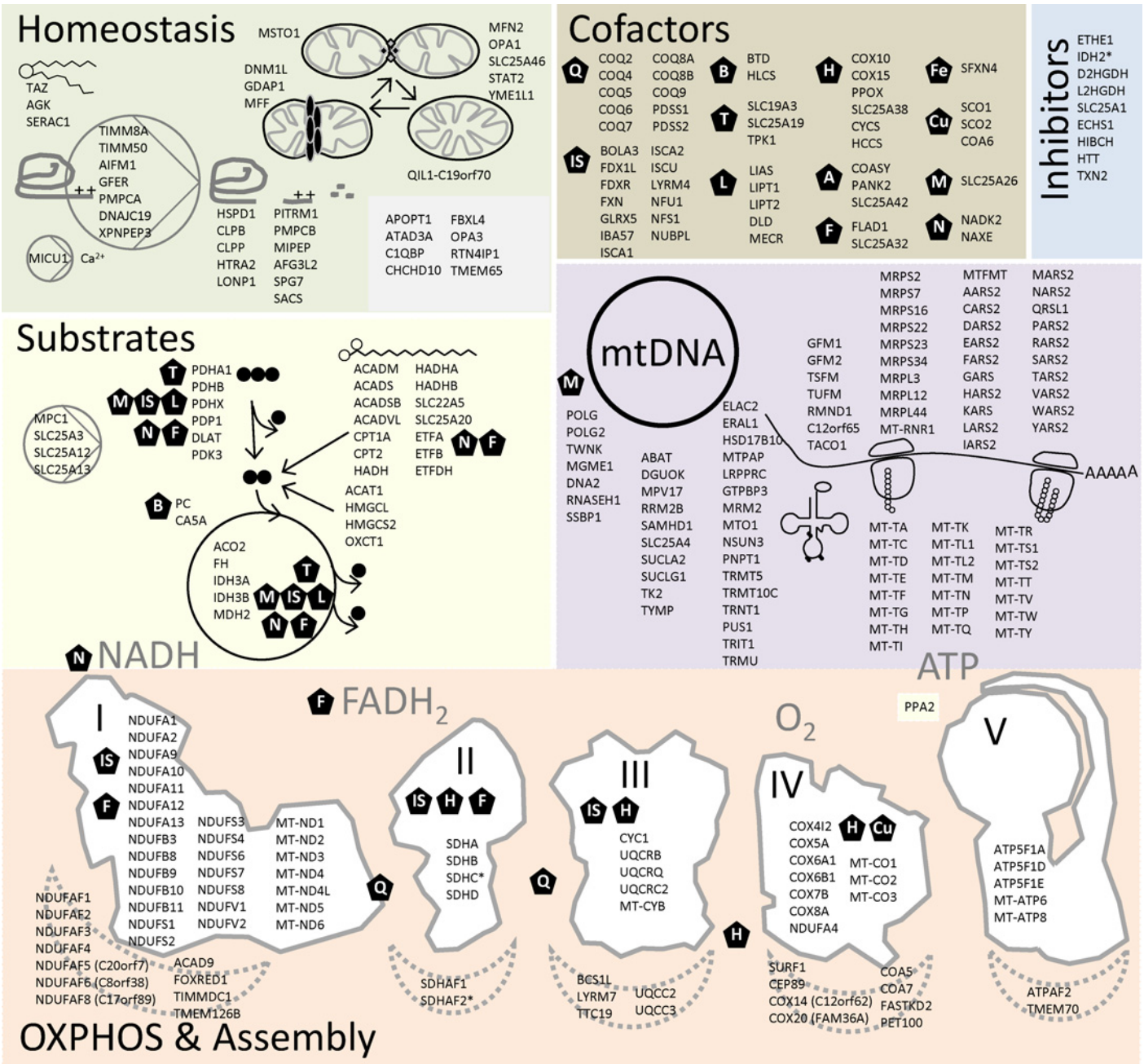


**Figure 1. Mitochondria displaying the involved structures and known disease genes (*n*=309) in different parts of mitochondrial energy metabolism**

A, coenzyme A; B, biotin; Cu, copper; F, riboflavin/FMN/FAD; Fe, iron; H, haeme; IS, iron–sulphur clusters; L, lipoic acid; M, S-adenosyl-methionine; N, NAD(P)H; Q, coenzyme Q10; T, thiamine pyrophosphate (figure modified from Wortmann et al. (2017) *Neuropediatrics*, with permission) [6].

## Candidate gene sequencing

Candidate gene sequencing has conventionally played an important role. In the 'biopsy first' diagnostic approach, clinical evaluation of the patient combined with the analysis of the respiratory chain complexes in muscle determined the subsequent Sanger sequencing of single candidate gene(s) [6]. It can therefore be utilized in confirmation of known variants and in identification of novel variants in known or putative disease genes, however, by definition cannot discover new disease-causing genes.

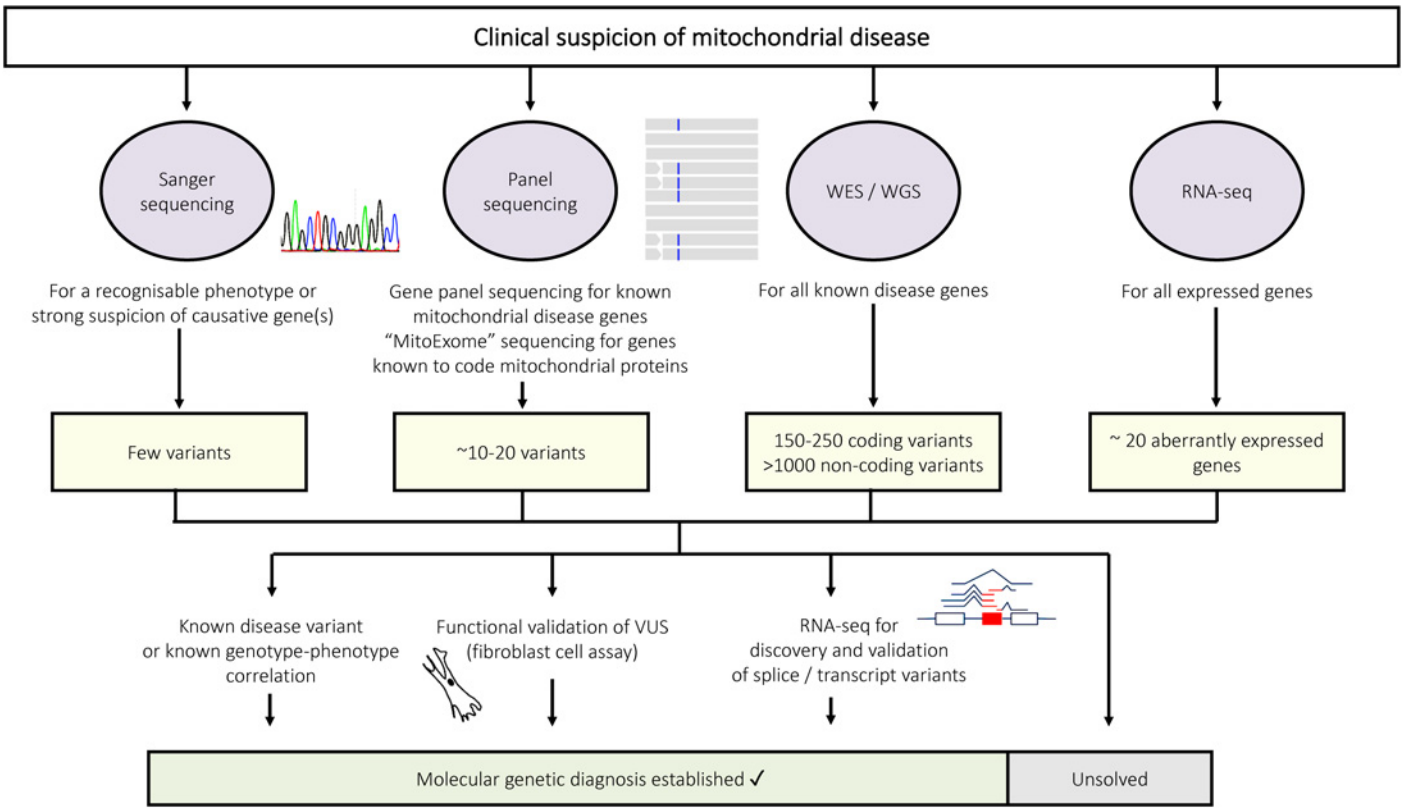


**Figure 2. Approach to establish the underlying genetic diagnosis in patient for whom mitochondrial disease is clinically suspected.**
Abbreviation: VUS, variant of unknown significance.

In this manner, large-scale mutation screening, exemplified by the high-resolution melting curve (HRMC) analysis and Sanger sequencing of 75 genes in a cohort of 152 patients with complex I deficiency, has been shown to achieve a relatively low rate of molecular genetic definition, 18% (28 of 152) of patients in the described study [17]. The reason for the low diagnostic yield in these approaches is three-fold. Firstly, gene-specific phenotypes, whereby a single gene is associated with a recognizable clinical presentation, are uncommon in mitochondrial disease making the selection of candidate genes challenging [17]. This is exemplified by mutations in *MT-TL1* which can cause, e.g. chronic progressive external opthalmoplegia (CPEO), mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episode (MELAS) syndrome and maternally inherited diabetes and deafness (MIDD), three strikingly different phenotypes [18]. Consequently, a strategy that sequentially sequences single genes is often laborious, expensive, time-consuming, and inefficient. Secondly, interpretation or artefacts of prior biochemical tests, necessitating invasive biopsy of affected tissue, can misdirect a candidate gene approach [19]. And thirdly, many patients are likely to carry mutations in genes so far not associated with mitochondriopathy. Therefore, the implementation of more exhaustive unbiased genome-wide strategies is necessary for complete molecular dissection [17].

Nevertheless, in the case of a recognizable phenotype and thus strong suspicion for a mutation in a single or limited number of genes, Sanger sequencing of these candidate genes, due to higher availability and increased speed may be preferred over genome-wide strategies. Exemplified in Leber's hereditary optic neuropathy (LHON) where over 95% of primary cases result from one of three mtDNA point mutations (m.11778G, m.14484T, m.3460G) [20], in MELAS syndrome where ~80% of all cases harbour a mutation in *MT-TL1* [21] and in Senger's syndrome, usually associated with *AGK* mutation [22].

## NGS

Technological advances have dramatically reduced the costs of NGS [23]. And, by adopting and leveraging NGS, clinical laboratories are performing an ever-expanding catalogue of genetic testing [13]. NGS for suspected mitochondriopathy can be divided into two fundamentally different approaches: targeted NGS gene panels, including sequencing of the 'MitoExome' and WES. At present, nine studies in the literature evaluate the diagnostic utility of

*Essays in Biochemistry* (2018) **62** 399–408
https://doi.org/10.1042/EBC20170110

**Table 1 Diagnostic yield of molecular genetic diagnostic studies for suspected mitochondriopathy**

| Patient cohort | Number of cases | Molecular genetic approach | Diagnostic yield |
|---|---|---|---|
| Paediatric mitochondriopathy patients with complex I deficiency (Haack et al. (2012)) [17] | 152 | HRMC analysis (75 genes) | 28 (18.4%) |
| Paediatric and adult patients with suspected mitochondrial disease based on clinical and/or biochemical findings (DaRe et al. (2013)) [29] | 148 | NGS panel (447 genes) | 13 (8.8%) |
| Paediatric and adult patients diagnosed with mitochondrial disease by integration of clinical, biochemical and morphological investigations (Ghezzi and Zeviani (2016)) [24] | 125 | NGS panel (132 genes) | 19 (15.2%) |
| Infantile mitochondrial disease patients with biochemical evidence of mitochondrial oxidative phosphorylation disease (Calvo et al. (2012)) [25] | 42 | MitoExome | 23 (54.8%) |
| Paediatric and adult patients with suspected mitochondrial disorder based on clinical, biochemical, and/or molecular findings (Lieber et al. (2013)) [26] | 102 | MitoExome | 22 (21.6%) |
| Paediatric and adult patients with biochemical evidence of multiple respiratory chain complex defects (Taylor et al. (2014)) [19] | 53 | WES | 28 (52.8%) |
| Paediatric mitochondrial respiratory chain disorder patients (Ohtake et al. (2014)) [27] | 104 | WES | 45 (43.3%) |
| Paediatric patients with clinical suspicion of a mitochondrial disorder (Wortmann et al. (2015)) [30] | 109 | WES | 42 (38.5%) |
| Childhood-onset mitochondrial respiratory chain complex deficiency patients (Kohda et al. (2016)) [28] | 142 | WES | 49 (34.5%) |
| Suspected mitochondrial disease patients from a paediatric reference centre (Pronicka et al. (2016)) [31] | 113 | WES | 67 (59.3%) |
| **Total** | **1090** | | **336 (30.8%)** |

NGS in mitochondrial disease, with patient cohorts selected based upon a particular biochemical signature of disease [19,24-28] or a diagnostic centre [29-31]. The diagnostic yields of these studies can be seen in Table 1 in addition to the previously discussed HRMC mutational screen analysis [17]. Of these publications, five utilize WES [19,27-29,31].

## NGS gene panels and MitoExome sequencing

The diagnostic success of gene panels is subject to the selection of candidate genes and pleiotropy, creating difficultly in design [32]. In one study of 148 patients with suspected mitochondrial disease, a targeted panel of 447 genes, encoding mitochondrial respiratory chain complexes, inducing secondary mitochondrial dysfunction or known to underlie disease mimicking mitochondrial disease, achieved identification of causative mutations in 13 patients (8.8%) [30]. In a second study of 125 patients, a targeted mitochondrial panel of 132 genes, associated with mitochondrial disease or known to be candidate genes in the same molecular pathways, achieved identification of causative mutations in 19 patients (15.2%) [24].

Sequencing the 'MitoExome' [25] is a more ambitious approach and involves exon capture and sequencing of the entire mitochondrial genome and all nuclear genes encoding mitochondrial proteins, based on the MitoCarta inventory [5,33]. Therefore capturing most genes implicated in mitochondrial biology. Fifteen genes in Figure 1 (*BTD*, *HLCS*, *HTT*, *MSTO1*, *RRM2B*, *SACS*, *SAMHD1*, *SLC19A3*, *SLC22A5*, *STAT2*, *TAZ*, *TPK1*, *TRMT5*, *TYMP*, *UQCC3*) are currently not captured in the 'MitoExome', emphasizing the limited shelf-life of gene panel approaches, even in such comprehensive approaches.

As with other approaches, diagnostic yield is subject to the stringency of the selection of the patient cohort. For example, an infantile biochemically proven mitochondrial disease cohort in contrast with a cohort of patients with late-onset and uncertain clinical diagnoses, whom have been shown to receive a molecular genetic diagnosis in 55 and 22% of cases respectively [25,26].

This result demonstrates that gene panel sequencing is a more effective alternative to the sequential testing of mtDNA and individual nDNA genes. However, alike to conventional approaches, in unsolved cases it is likely that the true causal variants reside in non-targeted genes or in non-targeted regions, such as the non-coding regulatory regions. If the pathogenic variant is missed, it is subsequently possible that a variant of uncertain significance (VUS) is incorrectly prioritized. Additionally, on a frequent basis, suspected mitochondriopathy patients are indeed not mitochondrial by genetic definition and would therefore inevitably evade detection by these panels [34].

## WES

WES has had a revolutionary impact on the molecular genetic diagnosis of mitochondrial disorders and Mendelian disorders as a whole. WES covers 2%, the exonic regions, of the genome [35] in which ∼85% of known monogenic disease-causing mutations [36] are located. Up to 97% of bases can be covered reliably at least 20 times, a high level of coverage for which studies have shown good concordance between WES and conventional Sanger sequencing [37]. The advent of WES has accelerated the capacity to identify variants explaining Mendelian disease in both known and novel disease genes, with a current success rate for detecting causal variants in 25–50% of cases [38-40].

WES reveals an estimated 30000 individual variants per patient with ∼400–500 protein-modifying rare variants [41,42], ∼100 genuine loss-of-function variants [43] and ∼60 novel protein-altering variants, currently not described in any variant database [44]. Procedures are therefore necessitated for variant prioritization. In suspected mitochondrial disease, the variant filtering pipeline comprises exclusion of variants with minor allele frequency (MAF) >0.01 and filtering for homozygous and compound heterozygous variants (assuming an autosomal recessive mode of inheritance) with prioritization of variants known to be associated with the respiratory chain defect observed in the patient, according to the Human Gene Mutation Database (HGMD) [45], and for genes with predicted mitochondrial localization according to the MitoP2 or MitoCarta database [5,17,33,46]. Familial 'trio' sequencing facilitates phasing of haplotypes, the detection of *de novo* [47] and autosomal dominant variants, as well as aiding in variant prioritization by reducing the number of candidates by ten-fold when compared with sequencing the proband alone [48]. By querying exome-wide, the WES approach also enables the identification of non-mitochondrial genes presenting clinically as mitochondriopathy [34].

The first application of exome sequencing in the setting of mitochondrial disease demonstrated how pathogenic mutations can be identified in a single patient [49]. Disease-causing variants in *ACAD9*, a member of the mitochondrial acyl-CoA dehydrogenase protein family, were identified in a compound heterozygous state in a patient with severe isolated respiratory chain complex I deficiency. The putative variants were confirmed by Sanger sequencing and segregation and validated by functional complementation in the affected patient fibroblast cell line.

This proof-of-principal study demonstrated the efficacy of WES, in combination with a functional cell assay to discover the molecular basis of complex I deficiency and offered a methodological paradigm to be exploited in the definition of the gene repertoire involved in mitochondrial disease as a whole.

In subsequent years, larger scale studies have described the success of WES in suspected and biochemically confirmed mitochondrial disease cohorts [19,27,28,30,31], with a diagnostic yield ranging from 35 to 59% [28,31], significantly higher than the 11% diagnostic rate achieved by conventional targeted Sanger sequencing [10]. The diagnostic yield is likely to be influenced by the stringency of selection of the patient cohort. For example, the study reporting the lowest yield of 39% [30] included patients with low to high clinical suspicion of mitochondrial disease according to the Nijmegen mitochondrial disease scoring system (Supplementary material (appendix)) [30,50], representative of the heterogeneous group of suspected mitochondrial patients investigated in diagnostic laboratories in daily practice.

The success of these studies has resulted in the accredited use of WES as a routine diagnostic tool [30] and has propelled WES into the first tier of clinical diagnostics. However, the diagnostic yield of WES is far from complete, with 41–65% [28,31] of patients not receiving a genetic diagnosis and for which there are a number of key explanations. Firstly, disease-causing variants may be detected by WES but elude prioritization or remain as VUS. And secondly, by definition WES can capture only the coding regions of the genome indicating that the disease-causing variants in unsolved cases may be located in non-coding regions to which WES is blind [30]. With new mitochondrial disease genes continually being discovered, mitochondriopathy patients are frequently diagnosed in the research setting, as clinical diagnostic pipelines are limited to known disease genes.

## WGS

WGS has the potential to reveal all genetic variants encompassing both the coding and non-coding regions of the genome. The number of variants is consequently large, with a typical genome differing from the reference human genome at 4.1–5.0 million sites [51]. WGS is therefore anticipated to increase diagnostic yield in mitochondrial disease due to identification of variants in the non-coding regions and large structural variants, including breakpoints. However, the vast number of variants, in combination with incomplete annotation, means variant prioritization remains a challenge [52,53].

## Variant prioritization

Sequencing technology has evolved rapidly and by virtue of the increased complexity of data, the shift in molecular diagnostic tools has been accompanied by new challenges in sequence interpretation. Multiple studies have

*Essays in Biochemistry* (2018) **62** 399–408
https://doi.org/10.1042/EBC20170110

raised concern with available variant annotation. A study of 460 literature-cited disease mutations found ∼27% of variants currently listed as 'pathogenic' in mutation databases to be common polymorphisms, sequencing errors or show no substantial evidence of pathogenicity [54]. Furthermore, when analysing the genomes of 179 participants for disease-causing variants in the 1000 Genome Project, researchers discovered that any given individual who carried ∼40–85 homozygous missense variants was predicted to be 'highly pathogenic' however, these individuals were healthy, thus suggesting variants of low penetrance or potentially incorrect annotation [55].

The ACMG workgroup formed with the goal of developing guidelines for interpretation of variants to aid in overcoming this challenge [13]. Pathogenicity should be determined by the entire body of evidence in aggregate in conjunction with uniform nomenclature informed by a standardized framework to ensure unambiguous designation of pathogenicity and to enable effective sharing and downstream use of genomic information. The objective framework for classifying these variants into a five-tier system ('pathogenic', 'likely pathogenic', 'uncertain significance', 'likely benign' and 'benign'), comprises population data, computational data, functional data (such as enzymatic assays and functional complementation in the muscle biopsy and fibroblast cell lines respectively) and segregation data.

# Overcoming the bottleneck of variant interpretation

## Repositories of patient data and data sharing

In the setting of rare disease, such as mitochondrial disease, power is in numbers and large datasets of mineable information are necessitated to be able to definitively label a variant as 'pathogenic' [14]. Examples of such databases are the Exome Aggregation Consortium (http://exac.broadinstitute.org) (ExAC) database, comprising the sequence data of 60000 individuals and the Genome Aggregation Database (gnomAD) comprising 123136 exomes and 15496 genomes from unrelated individuals [56]. The databases include healthy populations to allow identification of variants that may be rare however not pathogenic and ethnically diverse as rare variants are strongly geographically and ethnically constrained [57].

The entry of variants with corresponding phenotypic annotation from clinical diagnostic exome analysis into public archives, exemplified by ClinVar, to facilitate assertation of relationships between human variation and observed health status and interpretation of clinical relevance of variants, is of utmost importance [58]. The mitochondrial research community has recognized this need and has developed resources to address this objective (MSeqDR) [59]. Furthermore, establishment of mitochondrial networks, such as GENOMIT, a collaboration of international partners (http://genomit.eu) and the Mitochondrial Medicine Society (http://www.mitosoc.org), further enables data sharing and allows expertise to be shared synergistically.

## Integration of functional '-omics' data

Quantificative '-omics' analyses, such as proteomics, metabolomics and transcriptomics are predicted to play an increasing role in variant prioritization [14]. Interrogation of proteomic data has proved a powerful approach to elucidate the pathogenicity of novel variants due to detection of reduced or diminished protein level [15] and in detection of specific proteomic signatures [60] in mitochondrial disease. Metabolomics however, is limited by the small number of established gene–metabolite associations. Here we focus on RNA sequencing (transcriptomics), recently validated as a tool for the discovery of novel pathogenic variants in Mendelian disease.

# RNA sequencing

RNA-seq gives direct insight into the transcripts of a tissue for a specific snapshot in time [61]. In contrast with DNA, the actively transcribed RNA is highly dynamic, serving as a transient intermediary molecule between DNA and protein translation and revealing the link between the cellular phenotype and the underpinning genetics [62]. It has the capability to detect and quantify known pre-defined RNA species, in addition to rare and novel RNA transcript variants and isoforms [63]. Thus providing a direct insight into the transcriptional perturbations caused by putative VUS eluding detection or prioritization in WES and WGS (frequently synonymous or non-coding variants).

Three strategies for the systematic prioritization of candidate disease-causing variants have been described [15]. Firstly, genes with expression outside the physiological range can be identified as expression outliers (statistically significant deviations from population distribution) termed as aberrant expression. The cause of such aberrant expression includes rare variants in promoter and enhancer regions [64], and within both coding and intronic regions [15], in addition to post-transcriptional mechanisms. Secondly, RNA-seq can reveal mono-allelic expression (MAE), whereby one allele is silenced. Assuming an autosomal recessive mode of inheritance, MAE of heterozygous rare variants fits the recessive mode of inheritance assumption and allows their re-prioritization [15]. And thirdly, splicing of a gene can be affected by variants in positions susceptible to affect splicing. These variants result in an array of

aberrant splicing events (exon creation, skipping, extension and truncation to intronic inclusion) and are the most frequent aberrations detected in RNA-seq [15,16].

In parallel, two systematic studies assessed the diagnostic power of RNA-seq in cases for which WES did not yield a genetic diagnosis [15,16]. In a cohort of unsolved mitochondriopathy patients, RNA-seq of patient-derived fibroblasts established a diagnosis in 10% (5 of 48) of patients in addition to identification of candidate genes for the remainder [15]. In a similar-sized cohort of unsolved primary muscle disorder patients, RNA-seq of patient-derived muscle achieved a diagnostic yield of 35% (17 of 50) [16]. The analysis of disease-specific tissue(s) is a significant obstacle in RNA-seq due to the need for biopsy of tissues frequently unobtainable in a pre-mortem setting, such as the high energy demanding tissues affected in mitochondrial disease. While the study of primary muscle disorders focused on 190 neuromuscular disease associated candidate genes in the affected tissue only, the mitochondriopathy study demonstrated that RNA-seq can be performed genome-wide for all expressed genes in a given unaffected tissue. In fact, 2574 of the 3768 disease genes (68%) listed in OMIM were detected in fibroblast cell lines [65], and despite the presumably negligible physiological consequence of the variants on this cell type, they proved a valuable resource in this analysis. Evidence that tissue-specificity does not preclude RNA sequencing of unaffected tissues.

In the case of fibroblasts, a median of one aberrantly expressed gene, six mono-allelically expressed rare variants and five aberrant splicing events were detected. Identifying a practicable number of strong candidates, similar to the number of bi-allelic variants detected via WES, for subsequent manual curation to interpret pathogenicity, manual inspection and functional validation. *TIMMDC1*, a complex I assembly factor, was one such gene in which a deep intronic variant, activating a cryptic exon and introducing a premature stop codon, resulted in aberrant expression secondary to non-sense mediated decay [15] Given these results, RNA-seq is emerging as an invaluable complementary tool to genome-based molecular diagnostics.

## Concluding remarks

Molecular diagnostic approach in mitochondrial disease is rapidly evolving. Utilizing an unbiased analysis early in the diagnostic approach is desired and can frequently bypass the need for intensive, protracted and repeated investigative work-up. The most important challenge when utilizing high-yield NGS approaches moving forward, is variant prioritization and the reliable annotation of variants to distinguish true pathogenicity from the plethora of benign variants present between individuals. The emergence of standardized guidance for variant interpretation and integration of clinical, functional and 'multi-omics' data, hold the key for interpretation of these variants and in accelerating the diagnosis of mitochondrial diseases.

## Summary

- Mitochondrial diseases are a diagnostic challenge due to clinical and genetic heterogeneity even in the face of revolutionary advancements in molecular diagnostic technology.
- Disease-causing variants have been identified in over 300 genes (mtDNA and nDNA) and the number continues to grow.
- Molecular genetic diagnosis is paramount for identification of mitochondrial diseases amenable to treatment, disease surveillance, genetic counselling and prenatal diagnosis.
- Candidate gene approaches remain valuable for clearly defined gene-specific phenotypes but can otherwise be laborious, expensive and time-consuming.
- Genome-wide approaches (WES and WGS) have increased diagnostic yield due to their unbiased nature and can reveal 'non-mitochondrial' disease genes and non-coding variants.
- Increase in the complexity of sequencing data has escalated the challenge of variant interpretation.
- Integration of clinical, functional and 'multi-omics' data are powerful complementary tools to overcome the bottleneck of variant interpretation.

*Essays in Biochemistry* (2018) **62** 399–408
https://doi.org/10.1042/EBC20170110

- Repositories of patient data and data sharing increase potential to understand genotype–phenotype correlation in mitochondrial disease.

## Acknowledgements

We thank Dr Johannes Mayr for critically reading the manuscript and for the contribution of Figure 1 to this review.

## Funding

This work was supported by mitoNet German Network for Mitochondrial Diseases [grant number 01GM1113C (to H.P.)]. The GENOMIT European Network for Mitochondrial Disease [grant number 01GM1207 (to H.P. and S.L.S.)]. The SOUND EU Horizon2020 Collaborative Research Project [grant number 633974 (to H.P.)]. The PhD Program Medical Life Science and Technology scholarship, TUM School of Medicine (to S.L.S.).

## Competing interests

The authors declare that there are no competing interests associated with the manuscript.

## Abbreviations

ACMG, American College of Medical Genetics and Genomics; HRMC, high-resolution melting curve; MAE, mono-allelic expression; MELAS, mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episode; nDNA, nuclear DNA; NGS, next-generation sequencing; OXPHOS, oxidative phosphorylation system; VUS, variant of unknown significance; WES, whole exome sequencing; WGS, whole genome sequencing.

# References

1 Gorman, G.S., Chinnery, P.F., DiMauro, S., Hirano, M., Koga, Y., McFarland, R. et al. (2016) Mitochondrial diseases. *Nat. Rev. Dis. Primers* **2**, 16080, https://doi.org/10.1038/nrdp.2016.80
2 Mimaki, M., Wang, X., McKenzie, M., Thorburn, D.R. and Ryan, M.T. (2012) Understanding mitochondrial complex I assembly in health and disease. *Biochim. Biophys. Acta* **1817**, 851–862, https://doi.org/10.1016/j.bbabio.2011.08.010
3 Nouws, J., Nijtmans, L.G., Smeitink, J.A. and Vogel, R.O. (2011) Assembly factors as a new class of disease genes for mitochondrial complex I deficiency: cause, pathology and treatment options. *Brain* **135**, 12–22, https://doi.org/10.1093/brain/awr261
4 Smits, P., Smeitink, J. and van den Heuvel, L. (2010) Mitochondrial translation and beyond: processes implicated in combined oxidative phosphorylation deficiencies. *J. Biomed. Res.* **2010**, 737385, https://doi.org/10.1155/2010/737385
5 Pagliarini, D.J., Calvo, S.E., Chang, B., Sheth, S.A., Vafai, S.B., Ong, S.E. et al. (2008) A mitochondrial protein compendium elucidates complex I disease biology. *Cell* **134**, 112–123, https://doi.org/10.1016/j.cell.2008.06.016
6 Wortmann, S.B., Mayr, J.A., Nuoffer, J.M., Prokisch, H. and Sperl, W. (2017) A guideline for the diagnosis of pediatric mitochondrial disease: the value of muscle and skin biopsies in the genetics era. *Neuropediatrics* **48**, 309–314, https://doi.org/10.1055/s-0037-1603776
7 Koopman, W.J., Willems, P.H. and Smeitink, J.A. (2012) Monogenic mitochondrial disorders. *N. Engl. J. Med.* **366**, 1132–1141, https://doi.org/10.1056/NEJMra1012478
8 Distelmaier, F., Haack, T.B., Wortmann, S.B., Mayr, J.A. and Prokisch, H. (2016) Treatable mitochondrial diseases: cofactor metabolism and beyond. *Brain* **140**, e11, https://doi.org/10.1093/brain/aww303
9 Carroll, C.J., Brilhante, V. and Suomalainen, A. (2014) Next-generation sequencing for mitochondrial disorders. *Br. J. Pharmacol.* **171**, 1837–1853, https://doi.org/10.1111/bph.12469
10 Neveling, K., Feenstra, I., Gilissen, C., Hoefsloot, L.H., Kamsteeg, E.J., Mensenkamp, A.R. et al. (2013) A post-hoc comparison of the utility of Sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. *Hum. Mutat.* **34**, 1721–1726, https://doi.org/10.1002/humu.22450
11 MacArthur, D.G., Manolio, T.A., Dimmock, D.P., Rehm, H.L., Shendure, J., Abecasis, G.R. et al. (2014) Guidelines for investigating causality of sequence variants in human disease. *Nature* **508**, 469–476, https://doi.org/10.1038/nature13127
12 Goldstein, D.B., Allen, A., Keebler, J., Margulies, E.H., Petrou, S., Petrovski, S. et al. (2013) Sequencing studies in human genetics: design and interpretation. *Nat. Rev. Genet.* **14**, 460–470, https://doi.org/10.1038/nrg3455
13 Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J. et al. (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–423, https://doi.org/10.1038/gim.2015.30
14 Furness, L.M. (2017) Bridging the gap: the need for genomic and clinical-omics data integration and standardization in overcoming the bottleneck of variant interpretation. *Expert Rev. Precision Med. Drug Dev.* **2**, 79–89, https://doi.org/10.1080/23808993.2017.1322897
15 Kremer, L.S., Bader, D.M., Mertes, C., Kopajtich, R., Pichler, G., Iuso, A. et al. (2016) Genetic diagnosis of Mendelian disorders via RNA sequencing. *BioRxiv* 066738, https://doi.org/10.1101/066738
16 Cummings, B.B., Marshall, J.L., Tukiainen, T., Lek, M., Donkervoort, S., Foley, A.R. et al. (2017) Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci. Transl. Med.* **9**, eaal5209, https://doi.org/10.1126/scitranslmed.aal5209

17 Haack, T.B., Haberberger, B., Frisch, E.M., Wieland, T., Iuso, A., Gorza, M. et al. (2012) Molecular diagnosis in mitochondrial complex I deficiency using exome sequencing. *J. Med. Genet.* **49**, 277–283, https://doi.org/10.1136/jmedgenet-2012-100846
18 Nesbitt, V., Pitceathly, R.D., Turnbull, D.M., Taylor, R.W., Sweeney, M.G., Mudanohwo, E.E. et al. (2013) The UK MRC Mitochondrial Disease Patient Cohort Study: clinical phenotypes associated with the m. 3243A>G mutation—implications for diagnosis and management. *J. Neurol. Neurosurg. Psychiatry*, https://doi.org/10.1136/jnnp-2012-303528
19 Taylor, R.W., Pyle, A., Griffin, H., Blakely, E.L., Duff, J., He, L. et al. (2014) Use of whole-exome sequencing to determine the genetic basis of multiple mitochondrial respiratory chain complex deficiencies. *JAMA* **312**, 68–77, https://doi.org/10.1001/jama.2014.7184
20 Man, P.Y., Turnbull, D.M. and Chinnery, P.F. (2002) Leber hereditary optic neuropathy. *J. Med. Genet.* **39**, 162–169, https://doi.org/10.1136/jmg.39.3.162
21 Goto, Y.I., Nonaka, I. and Horai, S. (1990) A mutation in the tRNALeu (UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature* **348**, 651–653, https://doi.org/10.1038/348651a0
22 Mayr, J.A., Haack, T.B., Graf, E., Zimmermann, F.A., Wieland, T., Haberberger, B. et al. (2012) Lack of the mitochondrial protein acylglycerol kinase causes Sengers syndrome. *Am. J. Human Genetics* **90**, 314–320, https://doi.org/10.1016/j.ajhg.2011.12.005
23 Wetterstrand, K.A. (2013) DNA sequencing costs: data from the NHGRI Genome sequencing program (GSP). http://www.genome.gov/sequencingcosts.2013
24 Ghezzi, D. and Zeviani, M. (2016) New genes and pathomechanisms in mitochondrial disorders unraveled by NGS technologies. *Biochim. Biophys. Acta* **2016**, 1326–1335, https://doi.org/10.1016/j.bbabio.2016.02.022
25 Calvo, S.E., Compton, A.G., Hershman, S.G., Lim, S.C., Lieber, D.S., Tucker, E.J. et al. (2012) Molecular diagnosis of infantile mitochondrial disease with targeted next-generation sequencing. *Sci. Transl. Med.* **4**, 118ra10, https://doi.org/10.1126/scitranslmed.3003310
26 Lieber, D.S., Calvo, S.E., Shanahan, K., Slate, N.G., Liu, S., Hershman, S.G. et al. (2013) Targeted exome sequencing of suspected mitochondrial disorders. *Neurology* **80**, 1762–1770, https://doi.org/10.1212/WNL.0b013e3182918c40
27 Ohtake, A., Murayama, K., Mori, M., Harashima, H., Yamazaki, T., Tamaru, S. et al. (2014) Diagnosis and molecular basis of mitochondrial respiratory chain disorders: exome sequencing for disease gene identification. *Biochim. Biophys. Acta* **1840**, 1355–1359, https://doi.org/10.1016/j.bbagen.2014.01.025
28 Kohda, M., Tokuzawa, Y., Kishita, Y., Nyuzuki, H., Moriyama, Y., Mizuno, Y. et al. (2016) A comprehensive genomic analysis reveals the genetic landscape of mitochondrial respiratory chain complex deficiencies. *PLoS Genet.* **12**, e1005679, https://doi.org/10.1371/journal.pgen.1005679
29 DaRe, J.T., Vasta, V., Penn, J., Tran, N.T. and Hahn, S.H. (2013) Targeted exome sequencing for mitochondrial disorders reveals high genetic heterogeneity. *BMC Med. Genet.* **14**, 118, https://doi.org/10.1186/1471-2350-14-118
30 Wortmann, S.B., Koolen, D.A., Smeitink, J.A., van den Heuvel, L. and Rodenburg, R.J. (2015) Whole exome sequencing of suspected mitochondrial patients in clinical practice. *J. Inherit. Metab. Dis.* **38**, 437–443, https://doi.org/10.1007/s10545-015-9823-y
31 Pronicka, E., Piekutowska-Abramczuk, D., Ciara, E., Trubicka, J., Rokicki, D., Karkucińska-Więckowska, A. et al. (2016) New perspective in diagnostics of mitochondrial disorders: two years' experience with whole-exome sequencing at a national paediatric centre. *J. Transl. Med.* **14**, 174, https://doi.org/10.1186/s12967-016-0930-9
32 Scharfe, C., Lu, H.H., Neuenburg, J.K., Allen, E.A., Li, G.C., Klopstock, T. et al. (2009) Mapping gene associations in human mitochondria using clinical disease phenotypes. *PLoS Comput. Biol.* **5**, e1000374, https://doi.org/10.1371/journal.pcbi.1000374
33 Calvo, S.E., Clauser, K.R. and Mootha, V.K. (2015) MitoCarta2. 0: an updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Res.* **44**, D1251–D1257, https://doi.org/10.1093/nar/gkv1003
34 Panneman, D.M., Smeitink, J.A. and Rodenburg, R.J. (2018) Mining for mitochondrial mechanisms: linking known syndromes to mitochondrial function. *Clin. Genet.* **93**, 143–951, https://doi.org/10.1111/cge.13094
35 Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A. et al. (2011) Exome sequencing as a tool for Mendelian disease gene discovery. *Nat. Rev. Genet.* **12**, 745–755, https://doi.org/10.1038/nrg3031
36 Botstein, D. and Risch, N. (2003) Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat. Genet.* **33**, 228, https://doi.org/10.1038/ng1090
37 Sulonen, A.M., Ellonen, P., Almusa, H., Lepistö, M., Eldfors, S., Hannula, S. et al. (2011) Comparison of solution-based exome capture methods for next generation sequencing. *Genome Biol.* **12**, R94, https://doi.org/10.1186/gb-2011-12-9-r94
38 Yang, Y., Muzny, D.M., Xia, F., Niu, Z., Person, R., Ding, Y. et al. (2014) Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* **312**, 1870–1879, https://doi.org/10.1001/jama.2014.14601
39 Taylor, J.C., Martin, H.C., Lise, S., Broxholme, J., Cazier, J.B., Rimmer, A. et al. (2015) Factors influencing success of clinical genome sequencing across a broad spectrum of disorders. *Nat. Genet.* **47**, 717–726, https://doi.org/10.1038/ng.3304
40 Chong, J.X., Buckingham, K.J., Jhangiani, S.N., Boehm, C., Sobreira, N., Smith, J.D. et al. (2015) The genetic basis of Mendelian phenotypes: discoveries, challenges, and opportunities. *Am. J. Human Genet.* **97**, 199–215, https://doi.org/10.1016/j.ajhg.2015.06.009
41 Choi, M., Scholl, U.I., Ji, W., Liu, T., Tikhonova, I.R., Zumbo, P. et al. (2009) Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 19096–19101, https://doi.org/10.1073/pnas.0910672106
42 Ng, S.B., Turner, E.H., Robertson, P.D., Flygare, S.D., Bigham, A.W., Lee, C. et al. (2009) Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**, 272–276, https://doi.org/10.1038/nature08250
43 Robinson, P.N., Köhler, S., Oellrich, A., Wang, K., Mungall, C.J., Lewis, S.E. et al. (2014) Improved exome prioritization of disease genes through cross-species phenotype comparison. *Genome Res.* **24**, 340–348, https://doi.org/10.1101/gr.160325.113
44 Volk, A.E. and Kubisch, C. (2017) The rapid evolution of molecular genetic diagnostics in neuromuscular diseases. *Curr. Opin. Neurol.* **30**, 523–528, https://doi.org/10.1097/WCO.0000000000000478

45 Stenson, P.D., Mort, M., Ball, E.V., Evans, K., Hayden, M., Heywood, S. et al. (2017) The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum. Genet.* **136**, 665–677, https://doi.org/10.1007/s00439-017-1779-6

46 Elstner, M., Andreoli, C., Klopstock, T., Meitinger, T. and Prokisch, H. (2009) The mitochondrial proteome database: MitoP2. *Methods Enzymol.* **457**, 3–20, https://doi.org/10.1016/S0076-6879(09)05001-0

47 Ehmke, N., Graul-Neumann, L., Smorag, L., Koenig, R., Segebrecht, L., Magoulas, P. et al. (2017) *De novo* mutations in SLC25A24 cause a craniosynostosis syndrome with hypertrichosis, progeroid appearance, and mitochondrial dysfunction. *Am. J. Human Genet.* **101**, 833–843, https://doi.org/10.1016/j.ajhg.2017.09.016

48 McRae, J.F., Clayton, S., Fitzgerald, T.W., Kaplanis, J., Prigmore, E., Rajan, D. et al. (2016) Prevalence, phenotype and architecture of developmental disorders caused by *de novo* mutation. *BioRxiv* 049056, https://doi.org/10.1101/049056

49 Haack, T.B., Danhauser, K., Haberberger, B., Hoser, J., Strecker, V., Boehm, D. et al. (2010) Exome sequencing identifies ACAD9 mutations as a cause of complex I deficiency. *Nat. Genet.* **42**, 1131, https://doi.org/10.1038/ng.706

50 Morava, E., van den Heuvel, L.P., Hol, F., De Vries, M.C., Hogeveen, M., Rodenburg, R.J. et al. (2006) Mitochondrial disease criteria diagnostic applications in children. *Neurology* **67**, 1823–1826, https://doi.org/10.1212/01.wnl.0000244435.27645.54

51 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. *Nature* **526**, 68–74, https://doi.org/10.1038/nature15393

52 Dewey, F.E., Grove, M.E., Pan, C., Goldstein, B.A., Bernstein, J.A., Chaib, H. et al. (2014) Clinical interpretation and implications of whole-genome sequencing. *JAMA* **311**, 1035–1045, https://doi.org/10.1001/jama.2014.1717

53 Biesecker, L.G. and Green, R.C. (2014) Diagnostic clinical genome and exome sequencing. *N. Engl. J. Med.* **370**, 2418–2425, https://doi.org/10.1056/NEJMra1312543

54 Andreasen, C., Nielsen, J.B., Refsgaard, L., Holst, A.G., Christensen, A.H., Andreasen, L. et al. (2013) New population-based exome data are questioning the pathogenicity of previously cardiomyopathy-associated genetic variants. *Eur. J. Hum. Genet.* **21**, 918–928, https://doi.org/10.1038/ejhg.2012.283

55 Xue, Y., Chen, Y., Ayub, Q., Huang, N., Ball, E.V., Mort, M. et al. (2012) Deleterious-and disease-allele prevalence in healthy individuals: insights from current predictions, mutation databases, and population-scale resequencing. *Am. J. Human Genet.* **91**, 1022–1032, https://doi.org/10.1016/j.ajhg.2012.10.015

56 Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T. et al. (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291, https://doi.org/10.1038/nature19057

57 Gravel, S., Henn, B.M., Gutenkunst, R.N., Indap, A.R., Marth, G.T., Clark, A.G. et al. (2011) Demographic history and rare allele sharing among human populations. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11983–11988, https://doi.org/10.1073/pnas.1019276108

58 Landrum, M.J., Lee, J.M., Benson, M., Brown, G., Chao, C., Chitipiralla, S. et al. (2015) ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* **44**, D862–8, https://doi.org/10.1093/nar/gkv1222

59 Shen, L., Diroma, M.A., Gonzalez, M., Navarro-Gomez, D., Leipzig, J., Lott, M.T. et al. (2016) MSeqDR: a centralized knowledge repository and bioinformatics web resource to facilitate genomic investigations in mitochondrial disease. *Hum. Mutat.* **37**, 540–548, https://doi.org/10.1002/humu.22974

60 Lake, N.J., Webb, B.D., Stroud, D.A., Richman, T.R., Ruzzenente, B., Compton, A.G. et al. (2017) Biallelic mutations in MRPS34 lead to instability of the small mitoribosomal subunit and Leigh syndrome. *Am. J. Human Genet.* **101**, 239–254, https://doi.org/10.1016/j.ajhg.2018.03.015

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

62 Cieślik, M. and Chinnaiyan, A.M. (2018) Cancer transcriptome profiling at the juncture of clinical translation. *Nat. Rev. Genet.* **19**, 93, https://doi.org/10.1038/nrg.2017.96

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

64 Zhao, J., Akinsanmi, I., Arafat, D., Cradick, T.J., Lee, C.M., Banskota, S. et al. (2016) A burden of rare variants associated with extremes of gene expression in human peripheral blood. *Am. J. Human Genet.* **98**, 299–309, https://doi.org/10.1016/j.ajhg.2015.12.023

65 Kremer, L.S., Wortmann, S.B. and Prokisch, H. (2018) "Transcriptomics": molecular diagnosis of inborn errors of metabolism via RNA-sequencing. *J. Inherit. Metab. Dis.* 1–8, https://doi.org/10.1007/s10545-017-0133-4