Review

Genetics of mitochondrial diseases: Identifying mutations to help diagnosis

Sarah L. Stenton\textsuperscript{a,b}, Holger Prokisch\textsuperscript{a,b,*}

\textsuperscript{a} Institut für Humangenetik, Klinikum rechts der Isar, Technische Universität München, Trogerstraße 32, 81675 München, Germany
\textsuperscript{b} Institute of Neurogenomics, Helmholtz Zentrum München, Ingolstaedter Landstraße 1, D-85764 Neuherberg, Germany

\textbf{A R T I C L E  I N F O}

Article History:
Received 30 March 2020
Revised 21 April 2020
Accepted 22 April 2020
Available online xxx

Keywords:
Mitochondrial disease
Mutation
Diagnostic
Genetic
Multi-omic

\textbf{A B S T R A C T}

Mitochondrial diseases are amongst the most genetically and phenotypically diverse groups of inherited diseases. The vast phenotypic overlap with other disease entities together with the absence of reliable biomarkers act as driving forces for the integration of unbiased methodologies early in the diagnostic algorithm, such as whole exome sequencing (WES) and whole genome sequencing (WGS). Such approaches are used in variant discovery and in combination with high-throughput functional assays such as transcriptomics in simultaneous variant discovery and validation. By capturing all genes, they not only increase the diagnostic rate in heterogeneous mitochondrial disease patients, but accelerate novel disease gene discovery, and are valuable in side-stepping the risk of overlooking unexpected or even treatable genetic disease diagnoses.

© 2020 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license. (http://creativecommons.org/licenses/by-nc-nd/4.0/)

\section{1. Introduction}

The mitochondria are the so-called “powerhouses of the cell”. They are the principal generators of cellular energy as ATP. Our cells house hundreds to thousands of independently proliferating mitochondria, each with 2–10 copies of mitochondrial DNA (mt-DNA) in a network of fusing and budding organelles \cite{1}. Mitochondrial diseases result from disruption to the oxidative phosphorylation (OXPHOS) activity or to integral mitochondrial functions. The responsible mutations harbour in both the circular 16,569 base pair mt-DNA and the nuclear DNA. The mt-DNA encodes 37 genes including 13 encoded respiratory chain subunits. In the nuclear DNA over 1,000 mitochondrially localising proteins are encoded, translated in the cytoplasm, and translocated to the mitochondria by an elaborate protein import machinery.

Mitochondrial diseases are by far the largest class of inborn errors of metabolism (IEM) \cite{2} with a collective incidence of 1.6 in 5,000 \cite{3}. Organ involvement in the vast majority is multi-systemic with a predilection for the high-energy demanding tissues. These tissues depend on maintaining efficient energetic status and in times of metabolic stress, patients’ symptoms characteristically decompensate and regress. Taking mitochondrial respiratory chain complex I deficiency (MIM:252,010) as an exemplar, this is reflected by a phenotypic spectrum spanning adult-onset isolated eye involvement in Leber’s hereditary optic neuropathy (LHON, OMIM:535,000) to devastating infantile onset Leigh syndrome (infantile subacute necrotising encephalopathy, OMIM:256,000) and remarkably, such a spectrum may be observed in patients with the same mutation \cite{4}. Conversely, each clinical presentation can result from a defect in one of multiple disease genes, such as Leigh syndrome where over 90 genes are implicated (The Leigh Map, https://mseqdr.org/leighmap.php) \cite{5}. Moreover, in mt-DNA encoded disease, mutational heteroplasmy with either a “threshold effect” for disease expression, a correlation between increasing heteroplasmy level and disease severity \cite{6}, and conversely, a high heteroplasmic or even homoplasmic burden resulting in a lesser or no disease burden \cite{7} have all been reported. Clinical rating scales, such as the Mitochondrial Disease Criteria (MDC) score, indicate the likelihood of mitochondrial disease. However, they are unable to pinpoint precise genetic diagnoses as phenotypes with high weighting in such scores have equally high locus heterogeneity \cite{8}. Though especially prominent in mitochondrial disease, this phenotypic complexity inflicts the wider genetic field where in up to 90\% of patients diagnosed by whole genome sequencing (WGS), the phenotype is found to be a poor predictor of molecular cause \cite{9}. Moreover, there is no single reliable biomarker for mitochondrial disease. Measurement of OXPHOS enzyme activities, along with immunohistochemical and histoenzymatic assays of invasive muscle biopsy, and measurement of lactate, amino acids, and organic acids in blood and urine, have conventionally been at the

\textsuperscript{*} Corresponding author at: Institut für Humangenetik, Klinikum rechts der Isar, Technische Universität München, Trogerstraße 32, 81675 München, Germany.

E-mail address: prokisch@helmholtz-muenchen.de (H. Prokisch).

https://doi.org/10.1016/j.ebiom.2020.102784
2352-3964/© 2020 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license. (http://creativecommons.org/licenses/by-nc-nd/4.0/)
forefront of the diagnostic algorithm [10]. However, given their low specificity and the plethora of genes resulting in a secondary OXPHOS deficiency, they can only bring limited clarity to the molecular cause. Collectively, these features make the diagnosis of a mitochondrial diseases challenging. Despite the myriad of hurdles, motivation to provide a molecular diagnosis is high. It is invaluable to the patient in tailoring management and anticipatory care, stratifying disease risk, and providing accurate recurrence risk estimates for families. Moreover, it provides an opportunity to individually target treatment, such as replenishing a critical cofactor or restricting a dietary element to ameliorate the disease [11,12].

2. Genetic underpinning of mitochondrial disease

The genetic driven era of mitochondrial disease discovery commenced over 30 years ago with the report of the first mt-DNA encoded disease genes in LHON and Kearns-Sayre syndrome (KSS, OMIM: 530,000), shortly followed by the discovery of the first nuclear encoded defect in Pdhb1 [13–15]. Defective OXPHOS is the hallmark, and a wide-array of proteins with a myriad of functional roles converge to influence OXPHOS functionality. In this review, we focus on 338 disease genes, each demonstrated to cause a defect in an enzyme involved in aerobic energy metabolism (OXPHOS, Krebs cycle, pyruvate oxidation, fatty acid oxidation, and ketone bodies) (Fig. 1).

The mitochondrial proteome is under the control of two genomes. Though the mt-DNA is replicated, transcribed, processed, and translated within the mitochondrial itself, almost all of the proteins required are encoded in the nuclear genome. Accordingly, 36 (11%) of the disease genes are mt-DNA encoded and 302 (89%) are encoded in the nuclear DNA. All modes of inheritance are reported and a remarkable 22 genes are inherited by a combination of AR and AD inheritance according to OMIM (https://omim.org) (Fig. 1). The presence of more than one mode of inheritance may be explained by allele heterogeneity with differing consequences of the pathogenic variants, such as loss-of-function, gain-of-function, and dominant negative or variable magnitude of the pathogenic variant effect. SSBP1 is a recently described example, where missense variants span recessive, dominant, as well as de novo dominant inheritance, indicating that simple dosage reduction is not the only molecular mechanism underlying the disease [16].

3. Disease gene discovery

A definitive diagnosis is arguably only possible when genetic definition is reached. Conventionally, by collection of patients with similarity in clinical and biochemical phenotype, candidate gene sequencing with prior linkage analysis or homoyzosity mapping and mt-DNA sequencing were utilised. These methods led to the discovery of more than 150 genes over 25 years. An acceleration of disease gene discovery 10 years ago, starting with Acad9, marked the transition into the next generation sequencing (NGS) era (Fig. 2) [17]. Pre-NCS, approximately five disease genes were described per year, monumentally increasing to over 15 thereafter. Though the “low-hanging fruit” may have been picked, we predict many more disease genes await discovery given the vast size of the mitochondrial proteome and increasing use of unbiased WES, WGS, and “multi-omic” methodologies.

4. Genetic analyses over time

4.1. Candidate gene and mt-DNA sequencing

Candidate gene sequencing may be the method of choice in clear clinical syndromes where the majority of cases are explained by a handful of mutations (such as the 95% of LHON cases explained by mutations in MT-ND1, -ND4, and -ND6). Beyond these clear circumstances, frequent reassessment is often required, resulting in a protracted diagnostic period for the majority and reaching a diagnosis in just 10% [18].

4.2. Next generation sequencing

NGS encompasses targeted mt-DNA sequencing, panel sequencing, WES, and WGS. In facilitating the identification of patients presenting atypically for their molecular diagnosis, NGS is expanding our understanding of genotype-phenotype relationships in mitochondrial disease, the extent of which are not yet fully appreciated. This has led to an entourage of research articles expanding the clinical phenotype of diseases, most recently in MT-ATP6 [7].

4.3. Targeted mt-DNA NGS

mt-DNA sequencing screens the entire mt-DNA sequence and allows heteroplasmay level to be measured. While the easily accessible DNA samples such as from blood and urine are convenient measurement starting points, mt-DNA heteroplasmy, replication, and copy number vary from tissue to tissue, presumably to meet local energy demands [19]. Thus, a negative result does not exclude an mtDNA mutation and sampling of tissues such as the urinary epithelial cells or skeletal muscle is often necessitated. Moreover, some mtDNA variants, specifically deletions and mtDNA depletion, are challenging to detect and are detectable only in high energy demanding tissues from which the clinical phenotype arises. Therefore, even in the era of NGS, tissue biopsies are necessary in patients when the result of genetic testing is inconclusive. Such examples are progressive external ophthalmoplegia (PEO) and Kearns-Sayre Syndrome where single large-scale mt-DNA deletions are preferentially found in post-mitotic skeletal muscle and are often undetectable in blood [20] and mitochondrial myopathies secondary to muscle specific mtDNA mutations. These are important exceptions where muscle biopsy is deemed the diagnostic “gold standard”. In adults where a diagnosis in the mt-DNA is more likely (75%), it is therefore common practice to sequence the mt-DNA from muscle prior to WES or WGS. While in the paediatric population mt-DNA variants are often detectable in the blood and urine but for definite exclusion further tissues should be analysed [21].

4.4. NGS panel

Panels are targeted towards genes known to be involved in mitochondrial disease and candidate genes predicted to be involved in critical mitochondrial functions. In the literature, panels range from 100 genes implicated in complex I deficiency to more expansive panels targeting the mt-DNA and exons of over 1,000 nuclear genes encoding mitochondrial proteins, the overwhelming majority of which have yet to be assigned a role in mitochondrial disease. These panels achieved a diagnostic rate of 7%–31% [22–27] (Table 1). Today, applying a panel of 300 genes (MitoSure300, Agilent) in a heterogeneous cohort of suspected mitochondrial disease patients reaches a diagnostic rate of about 30% (Personal communication). The principal advantages of this focused approach are the high coverage (of over 200x in targeted nuclear, and 12,000x in mt-DNA sites) and relative ease and manageability of interpretation. A major caveat to the panel is its limited shelf-life due to the continual discovery of novel disease genes. Moreover, panels rely on correct a priori clinical suspicion as they are unable to appreciate variation beyond the targeted genes, and in cases of phenotypic mimicry the true diagnosis may not be captured.

Of note, “virtual panels” may be applied to WES data to facilitate the analysis of variants in a stepwise manner, focussing firstly on the disease gene panel of interest with the possibility for successive genetic analysis of the panel genes.
Mitochondrial disease genes. Mitochondrial disease genes (338) divided into six subsets according to their functional roles: (1) OXPHOS subunits, assembly factors, and electron carriers (102/338 genes), (2) mitochondrial DNA maintenance, expression, and translation (102/338 genes), (3) mitochondrial dynamics, homeostasis, and quality control (43/338 genes), (4) metabolism of substrates (40/338 genes), (5) metabolism of cofactors (41/338 genes), and (6) metabolism of toxic compounds (10/338 genes). Numerous genes have dual roles across these categories (indicated by an asterisk). The mode of inheritance is autosomal recessive in 262 genes, maternal in 36, autosomal dominant in 8, X-linked dominant in 6, X-linked recessive in 4, and 22 genes inherited by a combination of AR and AD inheritance.
inclusion of further panels. Virtual panels do however come with the notable disadvantage of lower coverage and blinkering of clinicians to important information beyond genes to which they are restricted, thereby potentially overlooking clearer diagnoses at other loci.

4.5. Whole exome sequencing

WES interrogates the protein coding regions of the genome which harbour approximately 400 protein modifying rare variants [28]. Moreover, the mt-DNA can be analysed from WES data in a holistic approach negating the need for complete mt-DNA sequencing prior or in parallel to WES [29]. Here, the mt-DNA is captured in “off-target” reads with high recall (>95%), precision (>99%), and estimation of heteroplasmy comparable to the “gold standard” targeted mt-DNA NGS (average difference of <10%). This is especially useful in the paediatric population where the higher frequency of nuclear DNA mutations (75%–90%) deems WES the prudent first-step. It is important to note that WES is typically performed on blood where heteroplasmy can be low, therefore a negative result does not exclude the presence of mt-DNA variants in other tissues.

In total, nine studies have evaluated diagnostic WES in suspected mitochondrial disease, achieving diagnostic yields of 35%–70% [24,30–37] (Table 1), an improvement in mean diagnostic yield from 14% in panel approaches to 47%. The considerably wide variation in diagnostic yield is in the most part a product of the patient selection criteria. Larger cohorts encompass less clearly defined patients representative of everyday clinical practice, in comparison to the small deeply phenotyped, homogenous, and biochemically confirmed

| Genes analysed | Publication | Size of cohort | Biochemical confirmation | Age group | mt-DNA analysis | Diagnostic rate | Proportion of mitochondrial diagnoses |
|---------------|------------|----------------|--------------------------|-----------|----------------|----------------|----------------------------------------|
| Panel         |            |                |                          |           |                |                |                                        |
| <500          | Calvo et al. [22] | 60            | +/-                       | P         | Included¹     | 22 (13)        | 100% (13)                            |
|               | Dele et al. [23] | 148           |                          | P and A   | Included¹     | 95% (13)       | 46% (6)                               |
|               | Legati et al. [24] | 125           | +/-                       | P and A   | Included¹     | 15% (19)       | 100% (19)                            |
| >500          | Calvo et al. [25] | 42            |                          | P         | Included¹     | 31% (13)       | 100% (13)                            |
|               | Vasta et al. [26] | 26            | +/-                       | P         | Excluded prior| 23% (6)        | 67% (4)                               |
|               | Lieber et al. [27] | 84            | +/-                       | P and A   | Excluded prior| 7% (6)         | 67% (4)                               |
|               | Panel summary  | 485           |                          |           |                | 14% (70)       | 84% (59)                              |
| WES           | Haack et al. [30] | 10            |                          | P         | Included¹     | 70% (7)        | 100% (7)                              |
|               | Taylor et al. [31] | 53            |                          |          | Excluded prior| 54% (28)       | 100% (28)                            |
|               | Ohtake et al. [32] | 104           |                          | P         | Excluded prior| 43% (45)       | 60% (27)                              |
|               | Wortmann et al. [33] | 109         |                          | P         | Excluded prior| 39% (42)       | 50% (21)                              |
|               | Legati et al. [24] | 10            |                          | P and A   | Included¹     | 60% (6)        | 50% (3)                               |
|               | Kohda et al. [34] | 142           |                          | P         | Included¹     | 35% (49)       | 86% (42)                              |
|               | Pronicka et al. [35] | 113          |                          | P         | Included¹     | 59% (67)       | 70% (47)                              |
|               | Pusepp et al. [36] | 28            |                          | P         | Included¹     | 57% (16)       | 25% (4)                               |
|               | Theunissen et al. [38] | 63          |                          | P and A   | Included¹     | 62% (39)       | 82% (32)                              |
|               | WES Summary    | 632           |                          |           |                | 47% (299)      | 71% (211)                             |

* Some patient had already undergone mt-DNA screening and were either considered negative or inconclusive (containing variants of uncertain significance) A: adult, P: paediatric.
intronic regions causing variants leading to splice aberrations are located in deep neuronal in illuminating three novel disease genes (COX6A1, TIMMDC1, and COQ5) respectively. Each resulted in splice aberration, captured by qPCR or transcriptomics, and subsequently disrupted gene function [42,47,48]. To date, WGS has not been assessed in a systematic way in suspected mitochondrial disease cohorts, it is therefore only possible to infer the benefits to diagnostic rate base on the wider rare disease field where a modest improvement of 2% is seen in comparison to WES (40% vs. 38%, respectively) [38].

4.7. The benefits of WES and WGS

The unbiased approaches of WES and WGS opportune exposure of pathogenic variants in genes beyond those initially suspected by the diagnostic work-up, which is about 30% in published studies (Fig. 4) (Table 1). Moreover, capturing data on all genes regardless of their disease-causing status allows reanalysis of negative cases. In negative WES, from studies totally over 1,000 reanalysed cases, diagnostic gains of up to 10%–20% are demonstrated just 1–3 years after the initial analysis, mostly due to the discovery of novel disease genes [49–52].

4.8. Trio sequencing

WES and WGS are ideally undertaken in trios, by simultaneous analysis of the proband and parents, to strengthen the diagnostic rate with identification of de novo variants. A modest 1% increase in diagnostic rate between proband-only and trio sequencing (39% and 40%, respectively) and an increase in de novo detection of 2% (47% and 49%, respectively) are reported in the rare disease field to date [38]. Despite these returns, trio sequencing increases concerns of high economic burden and analysis speed. Of the limited studies available for evaluation to date, French and colleagues demonstrate the application of trio WGS in acutely unwell paediatric patients, providing a timely molecular diagnosis (within 2–3 weeks) which influenced clinical management in 65% [9]. Recent studies also provide the first end-to-end cost-effectiveness analyses by prospectively ascertaining patients undergoing parallel WES and conventional diagnostic care. This not only resulted in decreased cost, but in a marked increase in Quality Adjusted Life Years (QALYs), a cost-effectivity which strengthened further still when benefits to first-degree relatives and
the benefits to parental reproductive outcomes were taken into consideration [52,53].

5. Variant of uncertain significance: “Innocent until proven guilty”

The overarching issue of all sequencing approaches is the discovery of variants of uncertain significance (VUS). With genomic advancement, came the shift in the diagnostic bottleneck from variant discovery to interpretation, which must be rigorous to avoid misinterpretation. Bringing clarity to large variant-rich genes, frequently harbouring synonymous and missense VUS, is notoriously challenging especially when different modes of inheritance are reported. **POLG** is a prime example, where an expertly curated database of reported variation is necessitated (Human DNA Polymerase Gamma Mutation Database, https://tools.niehs.nih.gov/polg/). As an initial step to structure diagnostics, the American College of Medical Genetics (ACMG) provides guidelines incorporating 28 criteria. These criteria span population, computational and predictive, functional, segregation, and allelic data to position a variant in a 5-tier system from “Pathogenic” through to “Benign” interpretation. Functional validation has prominent weighting in such criteria.

Here, we review the most profitable choices of functional assay in mitochondrial disease, for both validation of VUS and for variant discovery in unsolved WES and WGS cases. We focus on recent advances in the application of unbiased transcriptomic, proteomic, and metabolomic methodologies. Beyond these, methylomics, lipidomics, and glycomics could shed further light on the pathogenesis of mitochondrial disease and improve the outlook for the discovery of precision medicines in this largely “untreatable” collection of diseases.

5.1. Transcriptomics

Transcriptome analysis facilitates genome-wide DNA variant interpretation, both coding and non-coding. Non-coding variants are often located deep within the intronic sequence or in regulatory
regions and are influential in mRNA expression and splicing. Discarding lowly expressed reads, around 10–15,000 transcribed genes are detectable [42–46]. For the diagnosis of Mendelian disease, three aberrant events may be analysed: (1) aberrant splicing, (2) aberrantly low expression, (3) imbalance in allele specific expression, and variants often present as a combination of these events. Allele imbalance may arise secondary to a variant that results in non-sense mediated decay (NMD) of one allele, arises in a regulatory region resulting in decreased expression of one allele, or can be the consequence of phenomenon such as X-chromosome inactivation and imprinting. By capturing these events, RNA-sequencing provides functional validation and is often the missing link in a diagnosis, as it provides evidence of a second variant to complement a rare heterozygous mutation already identified by NGS [42,46].

To date, five systematic transcriptomic discovery studies in suspected Mendelian diseases have been described. The field is still evolving and a standardised protocol with concrete diagnostic criteria is not yet established. Given the differences in experimental design and in disease cohort, diagnostic rates range from 8% to 35% following negative WES or WGS analysis, with additional steps made in identifying strong candidates in the majority (Table 2). In the study of mitochondrial diseases, a median of five aberrant splicing events, one aberrantly expressed gene, and six mono-allelically expressed rare variants were called in patient-derived fibroblasts, a manageable number for manual curation and downstream validation studies [42]. Across all studies, diagnoses were made primarily by detection of aberrant splice events, with allele imbalance and aberrantly low expression playing a complementary role (Fig. 5). Such high throughput validation of splice variants by transcriptomics is essential due to the low specificity of splicing algorithms for true splice disrupting events [54], as often the causative variant is detected by WES or WGS but lost amongst a vast list of variants predicted to affect splicing. Moreover, in combination with WGS, transcriptomics was instrumental in establishing a new mitochondrial disease gene, TIMMD1, where a homozygous deep intronic variant led to pseudoexon creation, a downstream premature stop codon, and subsequent NMD [42].

The current challenges of RNA-sequencing arise in the tissue selection and data analysis. Tissues are known to demonstrate unique expression patterns. Therefore, the tissue selected needs to be representative of the expression patterns in the disease relevant tissue to circumvent an underpowered analysis. In mitochondrial disease, we are in a privileged position as the culprit genes are housekeeping, allowing high numbers to be captured across a range of different tissues (Fig. 5). Fibroblasts are obtainable by minimally invasive skin biopsy or as a by-product of a diagnostic muscle biopsy and can be utilised in further downstream functional validation, where required. Moreover, selection of such a non-affected tissue may have advantages, as the regulatory consequences on other genes are minimal allowing the causal defect to be a more prominent outlier. As we move away from the phenotype driven approach to diagnosis, negating the need for muscle biopsy, the most readily available tissue is blood. However, in corroboration with other studies, across tissues obtainable by biopsy we find blood to be the least informative (Fig. 5) [45,46]. Regarding the analysis, selection of control cohorts is an important consideration. Publicly available databases are rich in RNA-sequencing data, however given platform-specific biases, Kremer and colleagues argue for the optimal use of an in-house database where matched controls are generated overtime [42]. As outlined by the five studies, which each elected different control cohorts, both choices see success (Table 2). Key to all approaches is normalisation of the data, optimally for both known and unknown cofounders. OUTRIDER, a statistical method for detecting aberrantly expressed genes, makes steps in overcoming these hurdles [55].

### 5.2. Proteomics

Proteomic methodology was pivotal in cataloguing the 1500 predicted mitochondrial proteins [56]. It allows quantification of all detected proteins in one assay, removing the need for individual Western blot analyses of predefined proteins of interest. In patient-derived fibroblast cell lines approximately 5,000–8,000 proteins are quantified, encompassing over 65% of mitochondrial disease proteins [42,57,58]. Missense variants are by far the most commonly occurring variant in mitochondrial disease, accounting for over 50% of all mutations reported in the Human Gene Mutation Database (HGMD; http://www.hgmd.cf.ac.uk, accessed March 2020). Missense variants may have no functional consequence or may exercise their pathogenicity by disrupting interactions or destabilising the protein structure. Destabilisation occurs in approximately 40% of missense variants [59] and can be detected by proteomics as aberrantly low expression [57].

The application of proteomics as a complementary diagnostic tool to WES and WGS has not been studied systemically in mitochondrial disease to date, or indeed any rare disease cohort. However, there are a handful of examples in the literature of the role of proteomics in validation of VUS. Kremer and colleagues evaluated the consequence of diminished RNA expression on the protein level, confirming
complete or near complete (98%) loss of protein in patients with TIMMDC1 VUS [42]. Moreover, by the unbiased detection of cellular protein changes, the proteomic analysis revealed global reduction of complex I subunits consistent with the function of TIMMDC1 as a complex I assembly factor. Similarly, Lake and colleagues utilised proteomics to detect a mitochondrial translation defect caused by destabilisation of the small mitoribosomal subunit to validate biallelic VUS in the small ribosomal subunit encoding gene MRPS34, where the collective abundance of all detected small mitoribosomal subunit proteins was significantly decreased [58]. Interestingly, both studies remark on the ability of the reduction in subunits of specific OXPHOS enzymes on proteomics to closely resemble the residual OXPHOS enzyme activity in the respective cell line by biochemical measurement [42,58].

A limitation of proteomics arises in the lack of detection of the functional consequence of the majority of missense variants which do not lead to degradation of the defective protein. Indeed, a study systematically phenotyping a library of over 2850 rare disease-causing missense variants across almost 1150 genes (the human mutation open reading frame collection; hmORF), found that approximately 60% of missense variants resulted in preserved folding or stability [59].

5.3. Metabolomics

Measurement of lactate, amino acids, and organic acids in blood and urine reveal the initial biomarkers of mitochondrial disease. These metabolites provide diagnostic clues but have low sensitivity and specificity, and cannot pinpoint a diagnosis on the gene-level with the exception of select diagnoses, such as ethylmalonic aciduria indicative of ETHE1 mutations (Ethylmalonic encephalopathy, OMIM 602,473). Newborn screening is another example of a metabolite driven analysis, where a clear-cut answer is provided. However, with the exception of MCADD (medium-chain Acyl-CoA dehydrogenase deficiency, OMIM 607,008) due to ACADM mutations, severe and often lethal neonatal-onset mitochondrial disease are unfortunately inadequately represented in these screening initiatives.

Metabolomics is a mass spectrometry-based approach for the systematic capture and quantification of up to thousands of small molecule metabolites within a given tissue [60]. The approach allows analysis of the up- and downstream effects of mitochondrial dysfunction on cellular processes and can illuminate disease-distinguishing subsets of metabolites, termed “metabolite fingerprints”. The generation and analysis of the data is complex, not least due to the differing properties of the metabolites in concentration, charge,
hydrophobicity, and solubility, but in the need for complex enrichment and pathway analyses of the data. Moreover, the metabolome is under the combined influence of the primary genetic defect, the genetic background, and environmental exposures. The diagnostic value of metabolomics in certain subsets of genetic diseases such as inborn errors of metabolism (IEM) has proven to be high [61]. However, though its use has been reported in individual mitochondrial disease cases for the validation of VUS [42, 62, 63], a large systematic metabolomic study is needed to establish the true value of metabolomics as a complementary diagnostic tool to WES and WGS.

6. Focused functional validation assays

Beyond the global “omic” methodologies, advances in gene discoveries by WES and WGS have far outpaced the development of high-throughput functional studies. Often the choice of a specific functional validation assay is necessary and fluctuates from gene to gene. Details on specific assays are beyond the scope of this review and are explored in detail elsewhere [64]. Here, we highlight assays in the experimental toolkit capable of capturing the consequence of substantial subgroups of mitochondrial disease genes (Fig. 4).

Blue-native electrophoresis (BN-PAGE) is a traditional way to study the integrity of OXPHOS complexes in subunit and assembly factor defects, and more recently in combination with quantitative proteomics in complexome analysis [65]. Its development has provided valuable insight into the formation of the OXPHOS complexes, most notably the step-wise assembly of complex I, followed by the formation of suprastructures known as the “supercomplexes” [66]. OXPHOS enzyme assays measure the amount of residual enzyme activity of the respective complexes in a given tissue or cell line. Generally speaking, an isolated OXPHOS enzyme deficiency can be utilised as proof of pathogenicity for VUS within OXPHOS subunit encoding genes. These biochemical assays are complex procedures and must be performed in expert centres with good quality control, experience, and a certain turnover of samples. Measurement of oxygen consumption rate (OCR) by microscale oxygraphy provides an impression of the overall bioenergetic state of the mitochondrial in vivo and assesses the capability of the mitochondrial to respond to stress [67]. Notably, mutations can also be modelled in the yeast system without the need for a patient-derived tissue or cell line. Of the 338 currently described disease genes, 217 have orthologs in Saccharomyces cerevisiae according to the Saccharomyces Genome Database (SGD; http://www.yeastgenome.org) [68]. Moreover, where a functional validation assay reveals a phenotype, rescue complementation with wild-type cDNA can cement the pathogenic designation of a variant [69].

7. Outstanding questions

The days of rapid-paced gene discovery with gathering of the so-called “low-hanging fruit” have paved the way to the future diagnostic challenges. The complexities of inheritance almost untouched in mitochondrial disease research include the contribution of digenic, oligogenic, and polygenic inheritance, the potential for genetic modifiers in variable penetrance and phenotype expressivity, tissue specificity, and gene-environment interactions, amongst others. We foresee these to be the next frontiers in mitochondrial disease genetics.

With the exception of a number of mt-DNA encoded diseases, where penetrance is reported to be as low as 10%–50%, nearly all molecularly confirmed mitochondrial diseases result from highly penetrant alleles [67]. Such incomplete penetrance clouds the water for determination of variant pathogenicity given the absence of segregation with affected individuals in the pedigree, and higher population allele frequency than expected, two criteria pointing towards a likely benign variant. It is also proposed that milder phenotypes in mitochondrial disease may be attributable to a collective interaction of “weaker” alleles, akin to a complex disease, or from the influence of modifier loci. To delineate these subtle interactions, we require an immense volume of harmonised genotypic and phenotypic data, in addition to large pedigrees, underlying the importance of collaboration and data sharing. GENOMIT, a global registry initiative, works towards this objective in mitochondrial disease (www.GENOMIT.eu). Given the prevalence of the individual mitochondrial diseases, such analyses may struggle to reach the statistical power needed. Approaches are therefore emerging that utilise data from large genome-wide association studies (GWAS) of complex traits, and demonstrate the involvement of common polygenic factors in the variable expressivity of rare disease [71–73].

8. Conclusion

As we move further into the genomic and “multi-omic” era, the opportunity to understand the cause of each and every mitochondrial disease and to provide a diagnosis to the majority becomes ever more tangible. Though caution should be exercised in variant interpretation and actionability, the opportunity for every patient to be explored in the research environment may overcome the bottleneck in clinical translation of these methodologies. Moreover, to realise their true diagnostic capability we must increase willingness to pool resources, such as (un)solved WES and WGS data, to increase the signal of pathogenic variant recurrence. Overall, given the growing complexity of mitochondrial disease genetics, we endorse a wide-angle unbiased approach to diagnostics.

9. Search strategy and selection criteria

Data for this review were identified by a systematic search of PubMed using the term (“mitochondrial disease”) AND (mutation OR diagnostic OR omic) or were manually sought after from relevant articles. Preference was given to articles published over the last three years and each was academic and peer-reviewed.

Author contributions

Sarah L. Stenton wrote the manuscript and created the figures and tables. Holger Prokisch supervised the writing and editing of the manuscript.

Declaration of Competing Interest

The authors declare no competing interest.

Acknowledgements

This study was supported by a German Federal Ministry of Education and Research (BMBF) grant to the German Network for Mitochondrial Disorders (mitoNET, 01GM19068) and by the German BMBF and Horizon2020 through the E-Rare project GENOMIT (01GM1603). The funders had no role in design, literature analysis and interpretation, writing of the report.

References

[1] Keogh MJ, Chinnery PF. Mitochondrial DNA mutations in neurodegeneration. Biochim Biophys Acta Bioenerg 2015;1847(11):1401–11.
[2] Ferreira CR, van Karnebeek CD, Vockley J, Blau N. A proposed nosology of inborn errors of metabolism. Genet Med 2019;21(1):102–6.
Wagner M, Berutti R, Lorenz Lieber DS, Calvo SE, Shanahan K, Slate NG, Liu S, Hershman SG, et al. Targeted sequencing of mitochondrial respiratory chain disorders: exome sequencing for disease gene identification. Biochimica et Biophysica Acta (BBA)-General Subjects 2014;1840(4):1355–58.

Turner BM, Koenen BK, Stenman JA, van den Heuvel L, Rodenburg RJ. Whole exome sequencing of suspected mitochondrial patients in clinical practice. Journal of inherited metabolic disease 2015;38(3):437–43.

Kohda M, Tokuzawa Y, Kishita Y, Nyozyuki H, Miyazaki M, Mizuno Y, et al. A comprehensive genomics analysis reveals the genetic landscape of mitochondrial respiratory chain complex deficiencies. PLoS genetics 2016;12(1):e1005679.

Pronicka E, Piektowska-Abramczuk C, Ciara T, Trubicka J, Rozicki D, Karczewska P, Kowalska A, et al. Exome sequencing in diagnostic approach of mitochondrial disorders: two years’ experience with whole-exome sequencing at a national paediatric centre. Journal of translational medicine 2016;14(1):174.

Puussep S, Reinsin K, Pajusalu S, Murumets U, Oiglane-Shlik E, Reim R, et al. Effectiveness of whole exome sequencing in unselected patients with a clinical suspicion of a mitochondrial disorder in Estonia. Molecular genetics and metabolism report 2018;8:105–9.

Thussenis TE, Nguyen M, Kamps R, Hendrickx AT, Sallevecht SC, Gottschalk RW, et al. Whole exome sequencing is the preferred strategy to identify the genetic defect in patients with a probable or possible mitochondrial cause. Frontiers in genetics 2018;9:400.

Clark MM, Hildreth A, Batalov S, Ding Y, Chowdhury S, Watkins K, et al. Diagnosis of genetic diseases in seriously ill children by rapid whole-genome sequencing and automated phenotyping and interpretation. Science translational medicine 2019;11(489):eaat6177.

Ouweland WH. Whole-genome sequencing of rare disease patients in a national pediatric care system. BioRxiv 2019:572347. doi:10.1101/572347.

Talouin D, Harris KD, Kessler MD, Carlson J, Spiezie JA, Torres R, et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. BioRxiv 2019:563866. doi:10.1101/563866.

Ma M, Yu Y, Chuang LS, Hsu NY, Shi LS, Hakenberg J, et al. Disease-associated variants in different categories of disease located in distinct regulatory elements. BMC genomics 2015;16(58):53.

Krener LS, Bader DM, Mertes C, Kopajtich R, Pichler G, Iuso A, et al. Diagnostic advantages of personalized medicine vs Mendelian disorders via RNA sequencing. Nature communications 2017;8(1):15824.

Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, et al. Genetic diagnosis and molecular basis of mitochondrial respiratory chain disorders: exome sequencing data. The American journal of human genetics 2018;103(6):907–17.

Kohda M, Tokuzawa Y, Kishita Y, Nyuzuki H, Moriyama Y, Mizuno Y, et al. A comparative experience with whole-exome sequencing at a national paediatric centre. Journal of translational medicine 2016;14(1):174.

Puussep S, Reinsin K, Pajusalu S, Murumets U, Oiglane-Shlik E, Reim R, et al. Effectiveness of whole exome sequencing in unselected patients with a clinical suspicion of a mitochondrial disorder in Estonia. Molecular genetics and metabolism report 2018;8:105–9.

Thussenis TE, Nguyen M, Kamps R, Hendrickx AT, Sallevecht SC, Gottschalk RW, et al. Whole exome sequencing is the preferred strategy to identify the genetic defect in patients with a probable or possible mitochondrial cause. Frontiers in genetics 2018;9:400.

Clark MM, Hildreth A, Batalov S, Ding Y, Chowdhury S, Watkins K, et al. Diagnosis of genetic diseases in seriously ill children by rapid whole-genome sequencing and automated phenotyping and interpretation. Science translational medicine 2019;11(489):eaat6177.

Ouweland WH. Whole-genome sequencing of rare disease patients in a national pediatric care system. BioRxiv 2019:572347. doi:10.1101/572347.

Talouin D, Harris KD, Kessler MD, Carlson J, Spiezie JA, Torres R, et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. BioRxiv 2019:563866. doi:10.1101/563866.

Ma M, Yu Y, Chuang LS, Hsu NY, Shi LS, Hakenberg J, et al. Disease-associated variants in different categories of disease located in distinct regulatory elements. BMC genomics 2015;16(58):53.

Krener LS, Bader DM, Mertes C, Kopajtich R, Pichler G, Iuso A, et al. Diagnostic advantages of personalized medicine vs Mendelian disorders via RNA sequencing. Nature communications 2017;8(1):15824.

Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, et al. Genetic diagnosis and molecular basis of mitochondrial respiratory chain disorders: exome sequencing data. The American journal of human genetics 2018;103(6):907–17.

Kohda M, Tokuzawa Y, Kishita Y, Nyuzuki H, Moriyama Y, Mizuno Y, et al. A comparative experience with whole-exome sequencing at a national paediatric centre. Journal of translational medicine 2016;14(1):174.

Puussep S, Reinsin K, Pajusalu S, Murumets U, Oiglane-Shlik E, Reim R, et al. Effectiveness of whole exome sequencing in unselected patients with a clinical suspicion of a mitochondrial disorder in Estonia. Molecular genetics and metabolism report 2018;8:105–9.

Thussenis TE, Nguyen M, Kamps R, Hendrickx AT, Sallevecht SC, Gottschalk RW, et al. Whole exome sequencing is the preferred strategy to identify the genetic defect in patients with a probable or possible mitochondrial cause. Frontiers in genetics 2018;9:400.
Lake NJ, Webb BD, Stroud DA, Richman TR, Ruzzeneente B, Compton AG, et al. Biallelic mutations in MRPS34 lead to instability of the small mitoribosomal subunit and Leigh syndrome. The American Journal of Human Genetics 2017;101(2):239–54.

Sahni N, Yi S, Taipale M, Bass JI, Coulombe-Huntington J, Yang F, et al. Widespread macromolecular interaction perturbations in human genetic disorders. Cell 2015;161(1):647–60.

Zhang A, Sun H, Yan G, Wang P, Wang X. Mass spectrometry-based metabolomics: applications to biomarker and metabolic pathway research. Biomedical Chromatography 2016;30(1):7–12.

Coene KL, Kluitjtnmans LA, van der Heeft E, Engelke UF, de Boer S, Hoegen B, et al. Next-generation metabolic screening: targeted and untargeted metabolomics for the diagnosis of inborn errors of metabolism in individual patients. Journal of inherited metabolic disease 2018;41(3):337–53.

Wanders RJ, Vaz FM, Ferdinandusse S, van Kuilenburg AB, Kemp S, van Karnebeek CD, et al. Translational Metabolism: A multidisciplinary approach towards precision diagnosis of inborn errors of metabolism in the omics era. Journal of inherited metabolic disease 2019;42(2):197–208.

Barca JM, Fogazza M, Rugolo M, Chupin S, Del Dotto V, Ghelli AM, et al. Metabolomics hallmarks OPA1 variants correlating with their in-vitro phenotype and predicting clinical severity. Human Molecular Genetics 2020. doi:10.1093/hmg/ddaa047.

Thompson K, Collier JJ, Glasgow RJ, Robertson FM, Pyle A, Blakely EL, et al. Recent advances in understanding the molecular genetic basis of mitochondrial disease. Journal of Inherited Metabolic Disease 2020;43(1):36–50.

Heide H, Bleier L, Steger M, Ackermann J, Driese S, Schwamb-B, et al. Complexome profiling identifies TMEM126B as a component of the mitochondrial complex I assembly complex. Cell metabolism 2012;16(4):538–49.

Guerrero-Castillo S, Baertling F, Kownatzki D, Wessels HJ, Arnold S, Brandt U, et al. The assembly pathway of mitochondrial respiratory chain complex I. Cell metabolism 2017;25(1):128–39.

Yepez VA, Kremer LS, Iuso A, Gusic M, Kopańtich R, Kosarikov E, et al. OCR-Stats: Robust estimation and statistical testing of mitochondrial respiration activities using Seahorse XF Analyzer. PloS one 2018;13(7):e0199938.

Skrzypek MS, Nash RS, Wong ED, MacPherson KA, Hellerstedt ST, Engel SR, et al. Saccharomyces genome database informs human biology. Nucleic acids research 2018;46(D1):D736–42.

Kremer LS, Proksch H. Identification of disease-causing mutations by functional complementation of patient-derived fibroblast cell lines. Mitochondria. New York, NY: Humana Press; 2017. p. 391–406.

Giordano C, Carelli V. Reply: Mitochondrial DNA copy number differentiates the Leber’s hereditary optic neuropathy affected individuals from the unaffected mutation carriers. Brain 2016;139(1) e2–.

Oetjens MT, Kelly MA, Sturm AC, Martin CL, Ledbetter DH. Quantifying the polygenic contribution to variable expressivity in eleven rare genetic disorders. Nature communications 2019;10(1):4897.

Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nature genetics 2018;50(9):1219–24.

Khera AV, Chaffin M, Zekavat SM, Collins RL, Roselli C, Natarajan P, et al. Whole-genome sequencing to characterize monogenic and polygenic contributions in patients hospitalized with early-onset myocardial infarction. Circulation 2019;139 (13):1593–602.