ELSEVIER

BBA clinical

Pathogenicity in POLG syndromes: DNA polymerase gamma pathogenicity prediction server and database

CrossMark

Anssi Nurminen[a], Gregory A. Farnum[b], Laurie S. Kaguni[a,b,*]

[a] Institute of Biosciences and Medical Technology, University of Tampere, Tampere, Finland
[b] Department of Biochemistry and Molecular Biology and Center for Mitochondrial Science and Medicine, Michigan State University, East Lansing, MI, USA

ARTICLE INFO

ABSTRACT

DNA polymerase gamma (POLG) is the replicative polymerase responsible for maintaining mitochondrial DNA (mtDNA). Disorders related to its functionality are a major cause of mitochondrial disease. The clinical spectrum of POLG syndromes includes Alpers-Huttenlocher syndrome (AHS), childhood myocerebrohepatopathy spectrum (MCHS), myoclonic epilepsy myopathy sensory ataxia (MEMSA), the ataxia neuropathy spectrum (ANS) and progressive external ophthalmoplegia (PEO). We have collected all publicly available POLG-related patient data and analyzed it using our pathogenic clustering model to provide a new research and clinical tool in the form of an online server. The server evaluates the pathogenicity of both previously reported and novel mutations. There are currently 176 unique point mutations reported and found in mitochondrial patients in the gene encoding the catalytic subunit of POLG, *POLG*. The mutations are distributed nearly uniformly along the length of the primary amino acid sequence of the gene. Our analysis shows that most of the mutations are recessive, and that the reported dominant mutations cluster within the polymerase active site in the tertiary structure of the POLG enzyme. The POLG Pathogenicity Prediction Server (http://polg.bmb.msu.edu) is targeted at clinicians and scientists studying POLG disorders, and aims to provide the most current available information regarding the pathogenicity of *POLG* mutations.

1. Introduction

DNA polymerase gamma is the enzyme responsible for replicating and maintaining mitochondrial DNA (reviewed in [15,16]). The functional, holoenzyme form of POLG is a heterotrimer consisting of a catalytic subunit (POLGA) and a dimeric accessory subunit (POLGB). POLGA carries three catalytic activities: 5′ − 3′ DNA polymerase, 3′ − 5′ exonuclease and 5′-deoxyribose phosphate lyase. POLGB enhances DNA binding, catalysis and holoenzyme processivity. Mutations in *POLG* are associated with a wide range of human disorders that exhibit shared and progressive phenotypes (reviewed in [46]). The POLG disorders range from prenatally-fatal conditions and severe infantile onset disorders, such as Alpers disease (Alpers-Huttenlocher syndrome), to milder, late-onset conditions such as progressive external ophthalmoplegia, and are described collectively as POLG syndromes.

POLG mutations are relatively rare, with an estimated carrier frequency of 1/100 individuals in the Western world [11]. Most are recessive, and symptoms typically manifest only in compound heterozygous patients. Although the symptoms and severity of the conditions vary widely, all share a few common features: they largely affect tissues with high energy demand, such as in the nervous system, muscle and liver, and they are all progressive conditions that show direct correlation between age of onset and the severity of the condition [9].

To date, 176 unique *POLG* missense mutations in mitochondrial patients have been reported in the literature. Typically, POLG patients carry either homozygous or compound heterozygous *POLG* mutations, though some dominant pathogenic mutations exist. We have previously mapped the mutations onto the tertiary structure of POLGA in the apo-enzyme form (PDB ID: 3IKM), and shown that the pathogenic mutations cluster into five distinct regions that we have termed Alpers clusters 1–5 ([6,9], Fig. 1). Furthermore, we reported that the cluster mapping of *POLG* mutations shows a direct genotype-phenotype relationship [9]. Cluster 1 consists of the polymerase active site of POLGA and its environs. Cluster 2 includes residues of the upstream DNA binding

Fig. 1. Structural model of the POLG holoenzyme in complex with primer-template DNA. In the three-dimensional ternary complex of POLG, the pathogenic clusters in the catalytic core, POLGA, form five distinct functional regions [6]. Upper panel, the structure modeled after PDB structure 4ZTU [39] illustrates the pathogenic clusters of POLGA colored as green for Cluster 1, yellow for Cluster 2, red for Cluster 3, blue for Cluster 4 and cyan for Cluster 5. Lower panel, subclusters are distributed throughout the primary amino acid sequence of the POLGA polypeptide. Subclusters are assigned for each continuous block on the primary amino acid sequence. The accessory subunits, POLGB, are illustrated in light grey (proximal subunit) and dark grey (distal subunit). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

channel, mutations in which are likely to alter DNA binding affinity and the overall processivity of the holoenzyme. Though the functionality of Cluster 3 remains undetermined, its location in the structure suggests that it may play a role in partitioning the 3′-end of the DNA primer strand between the polymerase and exonuclease active sites. Mutations in Cluster 3 are likely to affect both DNA binding affinity and catalytic activity. A recent study has also suggested that this region may be involved in functional interaction with the mtDNA helicase [31]. Cluster 4 contains the fewest residues and is located at the interface of the catalytic subunit with the distal accessory subunit. This interaction has not been characterized extensively, but it is implicated to serve an important role in enhancing enzyme processivity [26]. Residues in Cluster 5 lie within the intrinsic processivity (IP) subdomain of the spacer domain of the catalytic subunit, and based on mutational analysis are suggested to be involved in protein-protein interactions [23]. The five Alpers clusters represent both structural and functional modules of POLG and are thus hot spots for pathogenic mutations. Conversely, mutations that lie outside of the clusters are substantially less likely to be pathogenic.

Our clustering model demonstrates that the age of onset of numerous symptoms correlates strongly with the severity of the syndrome and the range of symptoms that are manifest (see Fig. 5, [9]). For example, the most severe form of POLG syndrome, MCHS (typical age of onset from birth to three years), includes patients that present with symptoms such as hypotonia, developmental delay, gastrointestinal dysfunction and hepatopathy. Patients diagnosed with AHS (typical age of onset from a few months after birth to 16 years) share most of these symptoms, but are also more likely to experience epileptic seizures. MEMSA patients (typical age of onset 13 to 25 years of age) are in most cases diagnosed with myopathies, neuropathies and ataxia. Patients diagnosed with ANS (typical age of onset from 15 to middle-age) show a wide variety in their symptoms, including symptoms from both the MEMSA and PEO patient groups. Finally, patients

with PEO (typical age of onset from 30 years and above) represent the least severe end of the spectrum, and present with symptoms such as diplopia and ptosis.

2. Results

2.1. POLG pathogenicity prediction server

We have gathered all available patient case reports identifying *POLG* mutations from public sources and created an online database for providing easy access to these data. The database and an online server, known as the "POLG Pathogenicity Prediction Server," are available at http://polg.bmb.msu.edu. For each patient case report, we have collated the mutations identified, the reported age of onset of the symptoms manifested, the clinical description of the patients and source of the data. Patient gender and information on administration of valproic acid is also included where available. The current database contains 681 patient case reports, with 176 unique missense mutations and 215 unique combinations of pathogenic mutations (Table 1). The large majority of patient cases (546, 80.2%) identify either compound heterozygous or homozygous *POLG* mutations. The remaining 135 cases (19.8%) report heterozygous variants that are thus putatively-dominant mutations.

In combination with our pathogenic clustering model [6,9], the patient data has enabled us to make statistical predictions about the severity and age of onset of symptoms for patients with both known and novel mutations. The server features a mutation query interface where the user can enter the *POLG* mutations identified in a patient (Fig. 2). The server then displays the cluster mapping of the input mutations and shows any existing patient cases. The statistical information also includes the allelic combinations in which the mutations have been reported. If other members of the family are not available for analysis, this information may be helpful in determining whether the input

Table 1
Contents of the POLG Pathogenicity Prediction Server and distribution of mutations. The majority of patient cases report compound heterozygous mutations.

Total number of patient cases	660
Infantile onset cases	181
Childhood onset cases	103
Juvenile onset cases	85
Adult onset cases	284
Unknown age of onset cases	7
Unique missense mutations	176
Unique pathogenic mutation combinations	215
Compound heterozygous patients	323
Homozygous patients	152
Heterozygous patients	128
Reports referenced	182

mutations occur *in cis* or *in trans*. Finally, based on the existing cases with similar cluster-mapping mutations, the server displays an indicator of the most probable age of onset, which can be used as the basis for a diagnosis/prognosis for a patient. The server also displays the symptoms found in patients with similar cluster combinations of mutations. An alternate and powerful use of the data is the classification of mutations as likely non-pathogenic variants (SNPs).

The server provides access to the patient and mutation data in various ways (Fig. 3). For example, the patient cases can be viewed based on their source reference, mutations, cluster combinations, age of onset, mutated residue number or closest known mutation (either in primary sequence or three-dimensional structure). The server also features data compilation pages, displaying the frequencies of age of onset and reported symptoms for all existing cluster combinations. The symptoms are displayed in two ways: as reported in the original case reports, and grouped into broader categories based on affected tissue

types (as per [9]). Additionally, though they are not used directly in the statistical predictions of pathogenicity, the server includes primary sequence-based pathogenicity predictions from PON-P2 [28], and also the most current population exome sequencing data on *POLG* from the ExAC project (Table S1; [7]).

2.2. Adjusting the cluster boundaries

The recently published POLG holoenzyme ternary structure makes it possible to identify all of the POLGA residues that interact directly with the DNA ligand. Of the 56 residues shown to interact with DNA, only six map outside of our previously defined pathogenic clusters. Because these 56 residues have a clear role in the functionality of POLG to bind and position the DNA ligand, they have a high risk for producing deleterious effects if mutated. By extending slightly the borders of four of our subclusters: 2B, 496–517 (previously 497–517); 2D, 752–769 (previously 752–767); 1F, 1098–1138 (previously 1104–1138); and 3C, 795–807 (previously 804–807), these residues are accommodated within the defined high-risk locations for mutations in the POLG structure. Full cluster definitions as defined previously based on biochemical, structural and genetic studies [9] are included as supplementary Table S2.

A major difference between the crystal structures of the apo-holoenzyme (PDB ID: 3IKM; [19]) and holoenzyme ternary complex (PDB ID: 4ZTU; [39]) is the positioning of the region constituting pathogenic Cluster 3D (residues 1047–1096; Fig. 4). We have previously termed this region the partitioning loop. Cluster 3D carries several known pathogenic mutations including p.S1080I, p.I1079L, p.C1077G, p.P1073L, p.G1076V and p.R1096C. In the apoenzyme structure, the tip of this long looping region extends within ~5 Å of the polymerase active site, protruding inward from the surface of the enzyme. In contrast, in the holoenzyme ternary complex structure, the
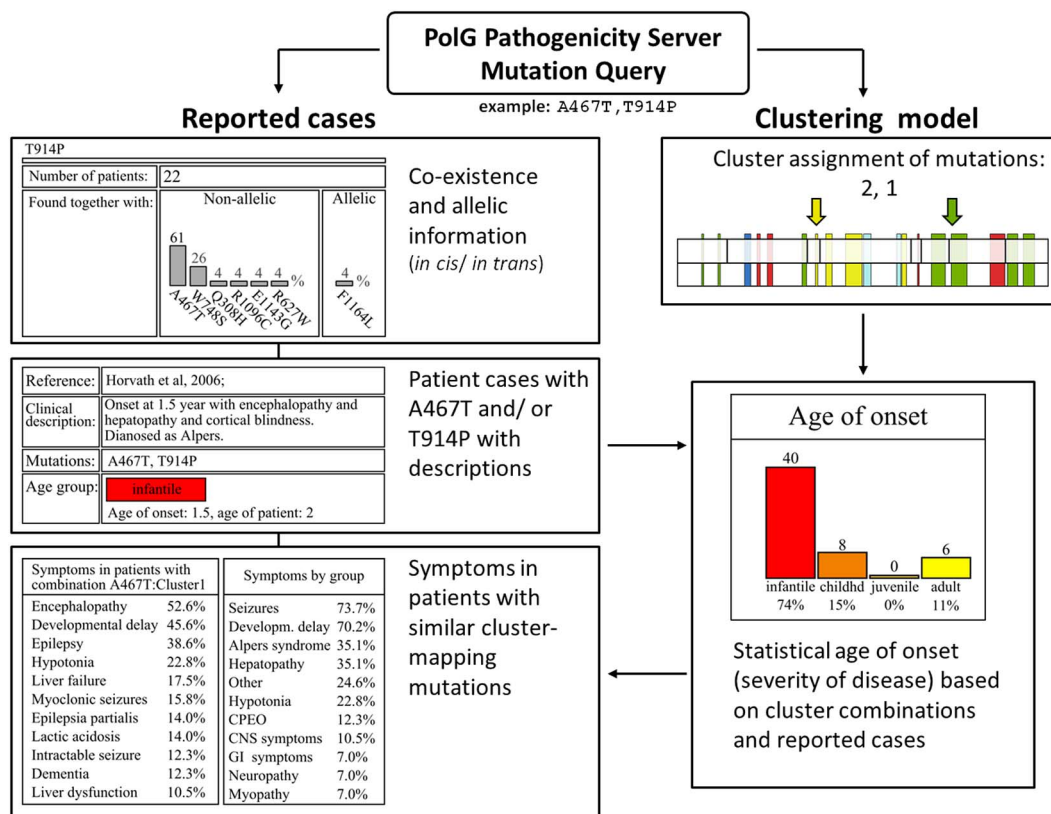
Fig. 2. The mutation query interface of the POLG Pathogenicity Prediction Server. The server provides statistical predictions of age of onset, and shows the typical symptoms for patients with mutations that map within pathogenic clusters 1–5. Information about allelic configuration of mutations can be helpful in cases for which pedigrees are not available. The server also provides direct access to clinical descriptions of patient cases that have similar mutations.
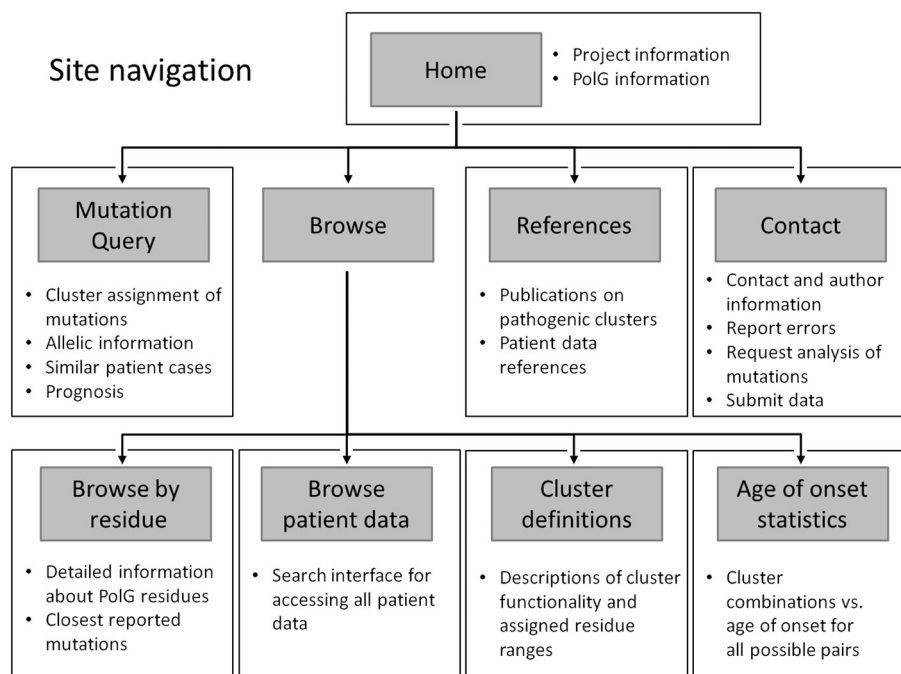
Fig. 3. Site navigation within the POLG Pathogenicity Prediction Server. The server provides access to collated data by various search interfaces, and provides statistical predictions of onset of symptoms based on the pathogenic clustering model.

same region of the loop closest to the polymerase active site in the apoenzyme structure is on the surface of the polymerase domain at a distance of > 20 Å from the polymerase active site. Though not discussed by Yin and collaborators [19,39], the crystal structure in the proximity of Cluster 3D is apparently highly disordered, and is missing two stretches of ~30 amino acids, residues 993–1024 and 317–340, suggesting that the structural evaluation of this region will remain a substantial challenge. In the apo-holoenzyme structure, most of the pathogenic mutations in Cluster 3D are located between the polymerase and exonuclease domains, suggesting that modulation of the exonuclease to polymerase ratio is a possible function of this region [6]. In addition, in the holoenzyme structure ternary complex, the 5′-end of the template strand of the DNA ligand exits the polymerase domain at a site in which interaction with Cluster 3D residues appears likely. A recent study by Qian et al. [31] reported that pathogenic mutation p.P1073L, located in cluster 3D, does not affect either the polymerase or exonuclease activities of POLG, but exhibits deleterious effects in an assay in which mitochondrial DNA helicase function is required, suggesting a possible functional interaction between them. Clearly, further investigation of the structure and function of this region is warranted.
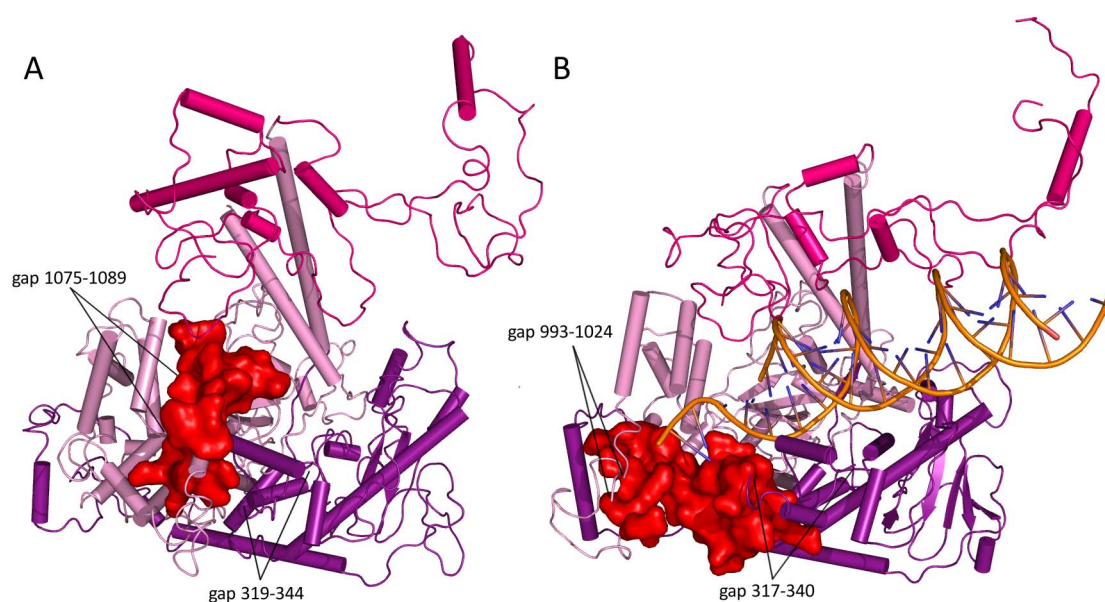
Fig. 4. Positioning of Cluster 3D in the POLG apo-holoenzyme as compared to the holoenzyme ternary complex. Cluster 3D (residues 1047–1096, red surface representation) is located in a substantially different position between the apo-holoenzyme (panel A, PDB ID: 3IKM) and holoenzyme ternary complex (panel B, PDB ID: 4ZTU). The repositioning of Cluster 3D appears unlikely to represent a conformational change. Rather, this region of the tertiary structure is likely to contain disordered regions, rendering difficult its structural evaluation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Common POLG mutations. The three most commonly reported POLG mutations, p.A467T, p.W748S and p.G848S, have been reported in all possible compound heterozygous combinations, as well as in homozygous form. Heterozygous case reports are rare for p.A467T and p.W748S, and there are none for p.G484S. Comparatively, p.G848S appears to occur in more severe cases, and p.W748S consistently shows a slightly milder phenotype in comparison to p.A467T.

Mutation	Clusters	Average age of onset (years)	Standard deviation	Number of reported cases
W748S + E1143G/wt	5 + SNP/–	59.0	4.0	2 (1 outlier[a])
A467T/wt	2/–	42.8	11.2	5 (1 outlier[b])
W748S/W748S	5/5	21.4	10.4	27 (1 outlier[c])
W748S + E1143G/W748S + E1143G	5 + SNP/5 + SNP	22.1	11.1	35
A467T/A467T	2/2	19.9	13.7	45
G848S/G848S	1/1	5.0	0.0	1
W748S/A467T	5/2	28.7	11.2	27
W748S + E1143G/A467T	5 + SNP/2	21.4	11.7	15
A467T/W748S + K561M	2/5 + 2	0.01	0.0	1
A467T/G848S	2/1	1.7	2.0	22
W748S/G848S	5/1	5.7	2.6	8
A467T/PNF	2/–	1.5	1.0	14 (1 outlier[d])
W748S/PNF	5/–	1.7	1.3	5
W748S + Q497H + E1143G/ W748S + Q497H + E1143G	5 + 2 + SNP/5 + 2 + SNP	19.0	4.0	2
W748S + Q497H + E1143G/A467T	5 + 2 + SNP/2	17.0	0.0	1

[a] Sarzi et al. [47], patient 36 (id: 404). [http://www.ncbi.nlm.nih.gov/pubmed/17452231].
[b] Galassi et al. [48], (id: 409). [http://www.ncbi.nlm.nih.gov/pubmed/18504126].
[c] Tzoulis et al. [49], (id: 552). [http://www.ncbi.nlm.nih.gov/pubmed/24841123].
[d] Martikainen et al. [50], (id: 666). [http://www.ncbi.nlm.nih.gov/pubmed/27111573].

2.3. Common mutations in the patient data

Three mutations represent the most commonly reported POLG mutations: p.A467T (c.1399G/A, 228 cases), p.W748S (c.2243G/C, 176 cases) and p.G848S (c.2542G/A, 64 cases). These map to three different clusters (p.G848S, Cluster 1; p.A467T, Cluster 2; p.W748S, Cluster 5). They have been reported in all possible compound heterozygous combinations as well as in homozygous form, making a comparative analysis of their observed phenotypes possible (Table 2). p.A467T and p.W748S have also been reported in heterozygous patients in eight and two cases, respectively. p.W748S is commonly found in cis with another mutation, p.E1143G. p.E1143G was assessed previously as non-pathogenic and potentially compensatory [3,30]. Our analysis corroborates this assessment as similar ages of onset are reported among patients homozygous for p.W748S (21 years, 30 cases) and p.W748S + p.E1143G (22 years, 35 cases). Furthermore, p.E1143G maps outside of the defined pathogenic clusters.

A comparison of the three mutations shows that p.G848S is associated with the most pathogenic phenotypes. When found in compound heterozygous form with p.A467T, the average age of onset is 1.7 years (± 2.0 years, infantile). By contrast, in compound heterozygous patients carrying p.G848S and p.W748S, the average age of onset is 5.7 years (± 2.6 years, childhood). A single patient carrying homozygous p.G848S has been reported with an age of five years at the time of examination [40]. Mutation p.G848S has been characterized biochemically to have a substantially reduced DNA-binding affinity and very low DNA polymerase activity [18]. A possible explanation for the lack of additional homozygous p.G848S patient case reports is that most of them result in prenatal mortality.

Patients carrying mutations p.A467T and p.W748S show similar ages of onset: the average age of onset for homozygous patients is 19.9 years (± 13.7) for p.A467T and 21.4 (± 10.4) years for p.W748S. These data corroborate the findings reported in a recent study on the extensive clinical heterogeneity of homozygous p.A467T patients [32]. Interestingly, compound heterozygous cases with genotype p.A467T/p.W748S manifest later, at ~25 years of age, suggesting that enzymes carrying deleterious mutations in different clusters could have compensatory properties. The heterozygous cases of p.A467T (five cases) and p.W748S (two cases) represent a small fraction of known carriers. The ExAC genome mapping project reports global population frequencies of 0.052% for p.A467T and 0.083% for p.W748S [7].

Moreover, in independent studies, p.A467T was found in the Belgian population at a particularly high frequency of 0.6% [43] and likewise, p.W748S was found in the Finnish population at a frequency of 0.8% [12]. By comparison, the frequency of p.W748S in the Finnish population is 0.57% in the ExAC dataset. These data alone argue that a dominant status for these mutations is unlikely in the absence of other genetic or environmental factors predisposing the patients to a mitochondrial disorder.

Biochemical studies show that in comparison to the wild type POLG, the p.A467T mutant enzyme exhibits moderate to substantially reduced DNA binding affinity resulting in similarly reduced polymerase processivity [25]. The p.W748S mutant enzyme was shown in one study to exhibit substantially reduced DNA binding affinity and polymerase processivity [3], whereas another reported enzymatic properties similar to the wild type enzyme [30]. A summary of the biochemical characterization of these mutations is included as Table S3.

2.4. Mutations yielding putatively non-functional POLG

Mutations that introduce frameshifts, premature stop codons, exon skipping or large deletions are likely to inactivate POLG function entirely and/or impact its folding, subunit interaction or stability. Structural perturbations would likely render the enzyme subject to cellular turnover. We have grouped together on the server mutations that would result in a putatively non-functional (PNF) POLG. Several of them have been characterized biochemically, corroborating such a non-functional status, though some may retain limited DNA binding capability or other partial functionality [17,34]. Patients with compound heterozygous PNF mutations that carry the three most commonly reported POLG mutations typically manifest with very severe, infantile onset conditions. For example, 14 patients with a compound heterozygous p.A467T/PNF genotype manifested symptoms at an average age of onset of 1.5 years. For four reported patients with a p.W748S/PNF genotype, the average age of onset is 1.7 years. No patients have been reported with p.G848S/PNF mutations. These data argue that the enzyme level produced from the single, partially-functional allele is a crucial factor in these patients, as compared with the more moderate phenotypes exhibited by the compound heterozygous and homozygous cases in which pA467T, p.W748S or p.G848S are present. Several exceptions have been reported. A single patient carrying A467T + E873X/E873X (A467T + PNF/PNF) survived until

10 years of age [27]. One possible explanation suggested by the authors of the report is that the premature stop codons are leaky, and some functional enzyme is still produced by the PNF allele. Similarly, Roos et al. [34] reported two brothers with a compound heterozygous genotype p.T914P/c.3104 + 3A > T (Cluster 1 mutation/splice site mutation). Molecular analysis showed that skipping of *POLG* exon 19 caused by c.3104 + 3A > T was incomplete, such that some wild type POLG was still produced, leading to an ameliorated phenotype with the patients manifesting symptoms at 20 and 50 years of age. Similar molecular analyses are warranted to determine which specific PNF mutations might still produce some wild type enzyme.

PNF mutations are all recessive, and the PNF carrying enzymes are likely incapable of competing for DNA binding with wild type POLG at the mtDNA replication fork. This hypothesis is corroborated by the presence of asymptomatic carriers in the pedigrees of patient reports. As such, the PNF enzymes also serve to establish an approximate baseline for assessing the severity and dominant versus recessive status of a mutation. Because PNF enzymes are putatively incapable of contributing to mtDNA replication, a mutation that is present in an individual in compound heterozygous form with a PNF mutation is likely to be a recessive mutation that produces an enzyme that is to some extent capable of mtDNA replication. This assessment supports a recessive status for two of the most common mutations, p.A467T and p.W748S, because they have been found in compound heterozygous form with PNF mutations. The patient data corroborates the biochemical evidence that p.G848S is only marginally more processive than a PNF enzyme [18], and thus a p.G848S/PNF genotype is not viable.

The patient data supports a model in which each deleterious *POLG* mutation affects the overall capacity of the enzyme to replicate and maintain mtDNA. Although recessive mutations may exert deleterious effects on enzyme functionality, symptoms are not manifest during the lifetime of an individual because a single wild type allele is typically sufficient. In both compound heterozygous and homozygous cases, each mutation may reduce the overall cellular capacity to maintain mtDNA, such that the most energy-demanding and mitochondria-dense tissues are affected. The patient data show that homozygous mutations are similar to compound heterozygous mutations, such that each allele with one or more deleterious mutations renders the condition more severe as aging progresses. Moreover, mutations affecting different critical functionalities (and mapping to different clusters) of POLG may show either compensatory or exacerbating effects.

The patient status in the first two-to-three years of life appears to be a critical determinant. The average survival time of patients that manifest with POLG syndromes under the age of three is on average under ten months with a maximum of five years after the time of onset of symptoms, as evidenced by 144 patient cases in which both the age of onset and age of death were reported (Fig. 5). The use of valproic acid to treat patients with epileptic seizures may account for shorter lifetimes reported at the more severe end of POLG syndromes, because treatment may lead to hepatotoxicity and acute liver failure [44].

Effects of environmental and cellular stress factors such as infections, unhealthy lifestyle, malnutrition, sleep deprivation and other conditions could not be controlled in the patient data evaluated in this study. In addition to the inconsistent reporting of patient data (age of onset/age at examination/age of death) to explain the high standard deviation seen in the data, it is likely that other currently unidentified genetic factors contribute to disease progression. In homozygous cases, the possible effects of consanguinity could also not be controlled.

2.5. Dominant POLG mutations

The POLG Pathogenicity Prediction Server contains 50 unique missense mutations that have been reported as heterozygous *POLG* mutations in a total of 131 individual patients. Patient case reports are available for examination through the mutation query and patient data access interfaces on the server, and a list of all pathogenic heterozygous
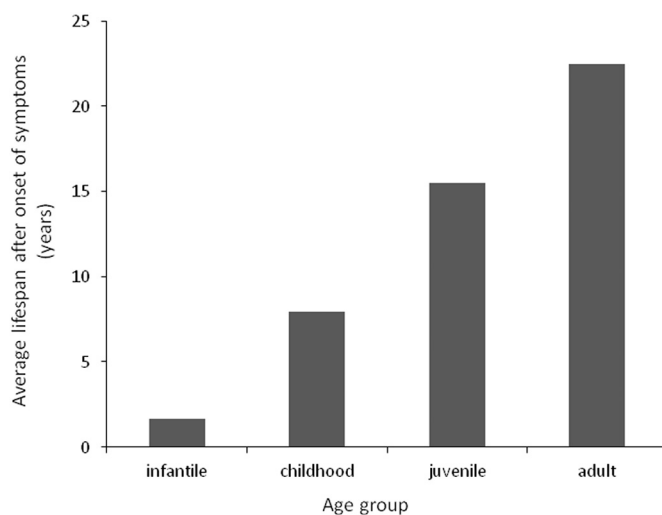

Fig. 5. Average lifespan after onset of symptoms. POLG syndromes are progressive conditions, and age of onset correlates directly with the severity of the disease. Patient data were analyzed in cases for which both the age of onset of symptoms and age of death were reported, and included 144 patient cases. The use of valproic acid to treat patients with epileptic seizures may account for shortened lifespans in the data in the more severe cases of POLG syndromes.

POLG mutations is also available with hyperlinks to the individual patient case reports. Based on the clinical data and in the absence of other explaining factors, these 50 mutations are tentatively assigned as pathogenic with a dominant inheritance mode. We have assessed the probability that they are the root cause of the symptoms described in the reported patient cases by taking into account the full patient dataset and their pathogenic cluster assignments.

We have subdivided the 50 heterozygous mutations into three groups as follows. Mutations in patient case reports that are not found at an elevated frequency in available population data and have been characterized as putatively-dominant in biochemical studies are categorized as most likely dominant. For 14 heterozygous mutations, only a single patient case has been reported, and additional data are required to confirm a putatively-dominant status. Mutations found in heterozygous patient cases without any other explanatory factors for the condition of the patients are categorized as having an unclear dominant pathogenicity. Mutations in which the case report indicated they are likely not the root cause for the condition of the patient have been categorized in the least likely dominant group. We have also taken into account biochemical data and reported asymptomatic carriers when assigning mutations into this last group. Our classification of the 50 heterozygous *POLG* mutations is discussed below, and is indicated on the list available on the POLG Pathogenicity Prediction Server.

The majority of the most-likely dominant pathogenic mutations surround the polymerase active site in the catalytic subunit, POLGA (Fig. 6). The residues in this region are involved in incoming nucleotide selectivity, base pairing with the template DNA strand and incorporation of the incoming deoxynucleotides into the primer DNA strand. These are the critical steps in DNA strand synthesis, and defects can impact both the efficiency and fidelity of mitochondrial DNA replication and lead to stalling of the DNA replication fork. Replication stalling can lead to mtDNA depletion, and result in base substitution, frameshift and deletion mutations and errors in nucleotide incorporation [37].

Notably, it was shown in a murine model that mitochondria can be relatively tolerant of point mutations, and a 500-fold increase in point mutations did not limit the lifespan of the animals [45]. This would argue that neither the reduced nucleotide selectivity nor defective exonucleolytic editing observed for some putatively-dominant mutations are likely the sole contributors in dominant pathogenicity. In that regard, exonucleolytic proofreading has been reported to increase the fidelity of POLG 20–100 fold depending on the assay used
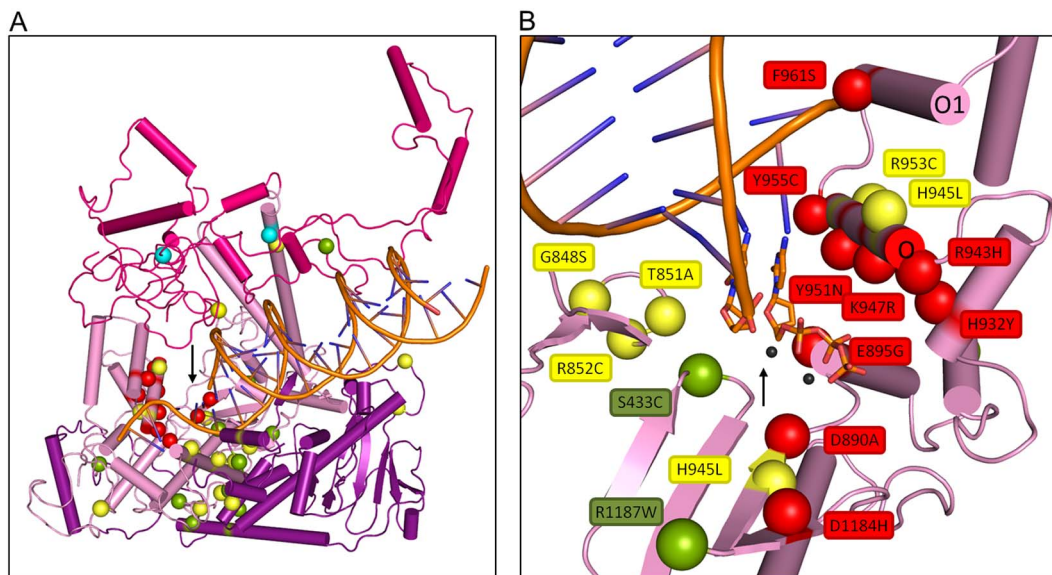
Fig. 6. Heterozygous POLG mutations. Panel A, model of the catalytic subunit of POLG illustrating mutations that have been reported to be heterozygous in the POLG Pathogenicity Prediction Server database. Mutations for which substantial corroborative evidence indicates dominant inheritance are shown in red spheres, and surround the polymerase active site (black arrow). Yellow spheres indicate mutations for which there is insufficient data to argue dominant status. Green spheres depict mutations least likely to be dominant. The color status is not indicative with respect to pathogenicity in compound heterozygous form. Two of the most common pathogenic *POLG* mutations, p.A467T and p.W748S, are shown in cyan spheres. Panel B, close-up of the active site of the polymerase domain of POLGA (black arrow). The putatively-dominant mutant residues (red spheres) interact with the incoming dNTP (orange), template DNA or the catalytic Mg-ions (black spheres). A hotspot for these is on the O-helix, in which every residue facing the polymerase active site has been reported in a heterozygous POLG patient. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

[14,21,22,29]. It also appears possible that proofreading activity is not the most critical function of the exonuclease domain; some mutations in the exonuclease domain result in reduced nucleotide polymerization rather than defects in proofreading [38], and it has been shown recently that exonucleolysis by POLGA is required for producing circular double-stranded DNA [26].

Though all mutations that are most likely to be dominant are located around the polymerase active site, not all mutations in close proximity to it are likely to be dominant. These include p.S433C and p.R1187W. There is only a single case report for p.S433C with an unspecified age of onset, and the mother of the patient was reported to be an asymptomatic carrier [13]. p.R1187W was reported in two cases in which the family histories do not support an assignment of dominant inheritance [33,35].

Three mutations, p.G848S, p.T851A and p.R852C, located in a highly conserved β-hairpin structure (residues 844–856) that lies between the polymerase and exonuclease domains, were reported as heterozygous in one in 66, one in four and one in 16 heterozygous patient cases, respectively. This region of the POLG structure is important for mispair recognition [39], and is likely to be involved in facilitating correct DNA binding near the polymerase active site, thus affecting both the polymerization rate and fidelity of the enzyme. Because there are many asymptomatic carriers for each of the three mutations, they are categorized as unlikely-dominant mutations. Moreover, they appear in the available population data with frequencies of 0.016% for p.G848S, 0.008% for p.T851A and 0.0066% for p.R852C [7]. Though unlikely dominant, they are nonetheless highly pathogenic when found in compound heterozygous form.

Mutations that lie near the polymerase active site and are categorized as having an unclear dominant status include p.D930N, p.H945L and p.R953C. Only a single patient case has been reported for p.H945L with matching symptoms and age of onset typical of dominant *POLG* mutations, but with no available family history [4]. p.R953C has been reported as a heterozygous mutation in only one patient case of four, without any family background [24] and is reported at a population frequency of 0.0016% in the ExAC database [7]. p.D930N has been described as exhibiting dominant-like pathogenicity in a yeast model,

without any corroborative patient data [1].

The most commonly reported *POLG* mutation presenting with a dominant inheritance pattern is p.Y955C, with 39 patient case reports. The onset of symptoms for these patients is in adulthood with an average age of 35 (± 14.5) years. Residue Y955 is located near the polymerase active site and maps within pathogenic Cluster 1. Its location in the POLG structure and the existing biochemical literature support a critical role in orienting and base pairing of the correct nucleotide with the template DNA strand [5]. Nucleotide polymerization by the variant POLG is reduced and misincorporation rate is elevated; kinetic analysis documents up to 1300-fold reduction in nucleotide incorporation fidelity [5]. Typical symptoms of patients carrying p.Y955C include PEO, muscle weakness and various neurological symptoms. Indeed, the most commonly reported symptoms for patients carrying putatively-dominant mutations are PEO/ptosis and muscle weakness, reported in 44% and 23% of the heterozygous patients, respectively (Fig. 7).

The biochemical characterizations reported to date for the puta-
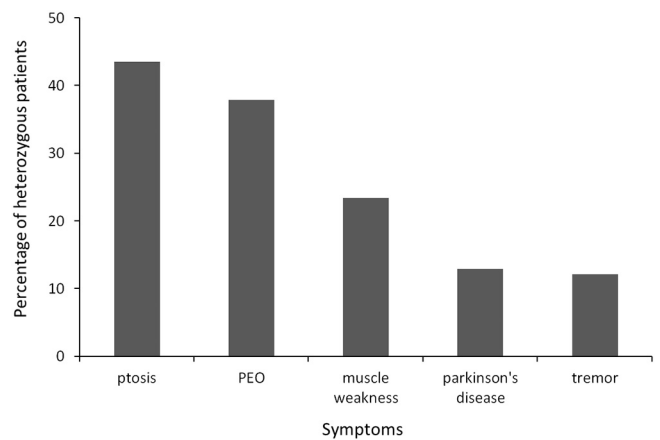
Fig. 7. The most commonly reported symptoms of heterozygous POLG patients. Symptoms in order of commonness in heterozygous patients include ptosis, progressive external ophthalmoplegia (PEO), muscle weakness, Parkinson's disease and tremors.

tively-dominant mutations [2,3,10,18, 25,30], combined with the patient case reports, highlight the fact that the most critical properties of dominant pathogenic mutations in POLG appear to be severely-reduced polymerase activity with sufficient DNA binding affinity to compete with wild type enzyme at the mtDNA replication fork (Table S3).

3. Discussion

We report the development of an online POLG Pathogenicity Prediction Server. The server represents a comprehensive, interactive database that complements and extends substantially the online Human DNA Polymerase gamma Mutation Database (http://tools.niehs.nih.gov/polg/). POLG data and patient cases are accessible in numerous ways as described under Results, and additional content is provided by including analysis by the PON-P2 algorithm [28] and ExAC population frequency data [7]. Most notably, the POLG Pathogenicity Prediction Server incorporates our previously described clustering model for pathogenicity prediction, which was derived by evaluation of biochemical, structural and genetic data on POLG and the related family A group DNA polymerases, bacterial DNA polymerase I and bacteriophage T7 DNA polymerase [6,9]. Because most POLG syndromes result from the incidence of compound heterozygous mutations, the clustering analysis of pairs of mutations based on the analysis of currently available mutations and mutational pairs gives insight into the likely pathogenicity of novel pairs, and of previously-identified alleles in combination with novel alleles. As such, we believe the prediction tool will be widely applicable as one parameter in the diagnosis of mitochondrial disorders.

Noteworthy challenges and/or limitations of the dataset include the varying methods and variety of tests employed in the examination of patients that have been reported and in particular, in publications prior to the generalized use of DNA sequencing to evaluate the spectrum of possible genes and genetic causes for mitochondrial disorders. Additionally, age of onset may not be reported as the first onset of symptoms; rather, the age of the patient at the time of examination has sometimes been reported. In view of these challenges, we highlight the importance of whole exome analysis (see for example [41]) and of reporting patient cases with full pedigrees and detailed descriptions of age of onset of symptoms. Though thorough clinical examination may not be possible for the immediate relatives of the patients, a description and overall assessment of their wellbeing would prove helpful. Finally, it is important to document the co-existence of potentially non-pathogenic polymorphisms such that any impact they might have in combination with a pathogenic mutation can be assessed appropriately.

Issues related to both intrinsic genetic and biochemical parameters must be evaluated further as new patient data on pathogenic mutations emerges. At present, the dataset of available mutations is characterized by the overrepresentation of the three common pathogenic mutations p.A467T, p.W748S and p.G848S that are present in composite in 71% of the 681 case reports. Whereas the relative frequency of new variants reported will likely remain similar, their absolute number will increase with streamlined reporting on an international scale. The POLG Pathogenicity Prediction Server features a contact form that can be used for pointing out errors in the existing data, notifying the database administrators of new publications and requesting confidential analysis of rare POLG mutations.

A potential shortcoming of the pathogenic clustering protocol is a direct reflection of the intrinsic structural and biochemical properties of POLG and indeed, of all enzymes. Whereas a new variant allele mapping to a pathogenic cluster is much more likely to be pathogenic than one that lies outside of a cluster, not all amino acid residues within a cluster or subcluster have a similar likelihood and/or level of pathogenicity. This inherent property is dependent strongly on both the biochemical function of the specific amino acid and the chemical nature of the amino acid change induced by the mutation. For example,

an amino acid that is crucial to POLG function such as the catalytic residues in the pol motifs A, B and C that are conserved among family A DNA polymerases and which lie in POLG subclusters 1D, 1E and 1F, respectively, may be either underrepresented or absent in the dataset because amino acid substitutions are not tolerated in live individuals. On the other hand, conservative amino acid changes within cluster-mapping residues may show no clinical phenotypes. In such situations, use of the residue browser function of the server will provide relevant information for further evaluation. Nonetheless, in the 'big picture' sense, it is clear that combinations of mutations mapping to the pol domain cluster 1 in combination with those mapping in either the DNA-binding channel cluster 2 or the putative protein: protein interaction cluster 5 constitute the bulk of the patient cases (96) manifesting infantile onset, and involve one or more of the most common mutations reported. Additionally, nine of 10 putatively-dominant mutations that we have classified as most-likely dominant (56 of 57 patient cases reported) map within the catalytic subclusters of the pol domain, 1D and 1E. This includes for example, the pathogenic mutation p.Y955C [10,20,42]. The single mutation mapping outside of cluster 1 is P765T in cluster 2; though its location in the POLG structure suggests a high likelihood of dominance, there is only a single case report and no existing biochemical data.

Biochemical characterization of mutations in conjunction with the patient data suggests that mutations that impair nucleotide polymerization by POLG without affecting DNA binding carry the highest risk for dominant inheritance. The low number of reported patient cases for some of the putatively-dominant mutations and lack of consistent testing and reporting in patient pedigrees obscures the evaluation of a likely dominant status. Based on the available genetic, biochemical and clinical data in our database we have re-evaluated the status of 49 putatively-dominant mutations and divided them into three classes. We suggest that the 10 mutations that fall within the most-likely dominant group be considered strongly by clinicians as a root cause of mitochondrial dysfunction.

It is well documented that the symptoms and tissue specificity of pathogenicity in patients suffering from POLG syndromes varies widely, even among patients carrying the same POLG mutations (see for example [32]). Nonetheless our analysis shows a clear consistency in age of onset that does not differ with gender. We also observed an apparently stronger correlation of symptoms within families. This observation highlights the likelihood that other genetic components are contributory and the importance of whole exome analysis in the evaluation of POLG syndromes, and of mitochondrial disorders in general to identify such components. Efforts such as Mitochondrial Disease Sequence Data Resource (MSeqDR) are underway to address the issue of nuclear genetic modifiers [8,36]. Future efforts should also address the possible contributions of mtDNA genetic background. At present, our understanding of the disease etiology of POLG syndromes together with the currently available diagnostic tools should be beneficial in family planning for couples undergoing genetic testing, even those with no familial history of mitochondrial disorders.

4. Materials and methods

Patient data deposited into the database and used in this study are anonymous and collected from publicly available journal articles. The source of each patient case is stated and available on the server. The data in the literature is reported in a non-standardized form and comprises a variety of different tests, examinations and details. The server categorizes each case based on the age of onset of the first symptoms reported into the following age groups: infantile, < 3 years; childhood, 3–13 years; juvenile, 13–20 years; or adult, ≥ 20 years. For reports in which only the age of the patient at the time of examination is reported, this age was used. If no age for the patient is reported, the age group is indicated as "unknown."

Symptoms were categorized based only on the diagnoses provided

in the report. For conditions such as Alpers syndrome, it might be the case that nearly all of the known symptoms for this condition have been reported in a patient case. However, if the diagnosis of Alpers has not been reported directly, the case has not been deposited as an Alpers entry. For other conditions such as ataxia, we have accepted terms such as "movement disorder," "ataxic gait," "gait disturbance" and "gait unsteadiness" as being synonymous.

The output of predictions of pathogenicity provided by the mutation query interface of the server are based on the pathogenic clustering model and statistics of existing patient cases with similar cluster-mapping mutations. The borders of the subclusters have been defined by available patient data, structural information and biochemical studies of mutations [6]. The statistical predictions reflect directly the contents of the database, and can be refined as new cases are added.

Transparency Document

The Transparency document associated with this article can be found, in online version.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbacli.2017.04.001.

References

[1] E. Baruffini, R. Horvath, C. Dallabona, B. Czermin, E. Lamantea, L. Bindoff, … T. Lodi, Predicting the contribution of novel POLG mutations to human disease through analysis in yeast model, Mitochondrion 11 (1) (2011) 182–190, http://dx.doi.org/10.1016/j.mito.2010.09.007.

[2] S.S.L. Chan, M.J. Longley, W.C. Copeland, The common A467T mutation in the human mitochondrial DNA polymerase (POLG) compromises catalytic efficiency and interaction with the accessory subunit, J. Biol. Chem. 280 (36) (2005) 31341–31346, http://dx.doi.org/10.1074/jbc.M506762200.

[3] S.S.L. Chan, M.J. Longley, W.C. Copeland, Modulation of the W748S mutation in DNA polymerase gamma by the E1143G polymorphism in mitochondrial disorders, Hum. Mol. Genet. 15 (23) (2006) 3473–3483, http://dx.doi.org/10.1093/hmg/ddl424.

[4] M. Delgado-alvarado, P. De, H. Jiménez-urbieta, B. Gago, A. Gabilondo, B. Bornstein, M.C. Rodríguez-oroz, Parkinsonism, cognitive deficit and behavioural disturbance caused by a novel mutation in the polymerase gamma gene, J. Neurol. Sci. 350 (1–2) (2015) 93–97, http://dx.doi.org/10.1016/j.jns.2015.02.011.

[5] P.A. Estep, K.A. Johnson, Effect of the Y955C mutation on mitochondrial DNA polymerase nucleotide incorporation efficiency and fidelity, Biochemistry 50 (29) (2011) 6376–6386, http://dx.doi.org/10.1021/bi200280r.

[6] L. Euro, G.A. Farnum, E. Palin, A. Suomalainen, L.S. Kaguni, Clustering of Alpers disease mutations and catalytic defects in biochemical variants reveal new features of molecular mechanism of the human mitochondrial replicase, Pol γ, Nucleic Acids Res. 39 (21) (2011) 9072–9084, http://dx.doi.org/10.1093/nar/gkr618.

[7] ExAC, Exome Aggregation Consortium (ExAC) browser, Retrieved from http://exac.broadinstitute.org, (2015).

[8] M.J. Falk, L. Shen, M. Gonzalez, J. Leipzig, M.T. Lott, A.P.M. Stassen, … L.J. Wong, Mitochondrial Disease Sequence Data Resource (MSeqDR): a global grass-roots consortium to facilitate deposition, curation, annotation, and integrated analysis of genomic data for the mitochondrial disease clinical and research communities, Mol. Genet. Metab. 114 (3) (2015) 388–396, http://dx.doi.org/10.1016/j.ymgme.2014.11.016.

[9] G.A. Farnum, A. Nurminen, L.S. Kaguni, Mapping 136 pathogenic mutations into functional modules in human DNA polymerase γ establishes predictive genotype–phenotype correlations for the complete spectrum of POLG syndromes, Biochim. Biophys. Acta, Bioenerg. 1837 (7) (2014) 1113–1121, http://dx.doi.org/10.1016/j.bbabio.2014.01.021.

[10] M.A. Graziewicz, M.J. Longley, R.J. Bienstock, M. Zeviani, W.C. Copeland, Structure-function defects of human mitochondrial DNA polymerase in autosomal dominant progressive external ophthalmoplegia, Nat. Struct. Mol. Biol. 11 (8) (2004) 770–776, http://dx.doi.org/10.1038/nsmb805.

[11] A.H. Hakonen, G. Davidzon, R. Salemi, L.A. Bindoff, G. Van Goethem, S. Dimauro, … A. Suomalainen, Abundance of the POLG disease mutations in Europe, Australia,

New Zealand, and the United States explained by single ancient European founders, Eur. J. Hum. Genet. 15 (7) (2007) 779–783, http://dx.doi.org/10.1038/sj.ejhg.5201831.

[12] A.H. Hakonen, S. Heiskanen, V. Juvonen, I. Lappalainen, P.T. Luoma, M. Rantamaki, … A. Suomalainen, Mitochondrial DNA polymerase W748S mutation: a common cause of autosomal recessive ataxia with ancient European origin, Am. J. Hum. Genet. 77 (3) (2005) 430–441, http://dx.doi.org/10.1086/444548.

[13] G. Hudson, M. Deschauer, R.W. Taylor, M.G. Hanna, D. Fialho, A.M. Schaefer, L.P. He, E. Blakely, D.M. Turnbull, P.F. Chinnery, POLG1, C10ORF2, and ANT1 mutations are uncommon in sporadic progressive external ophthalmoplegia with multiple mitochondrial DNA deletions, Neurology 66 (2006) 1439–1441, http://dx.doi.org/10.1212/01.wnl.0000210486.32196.24 https://www.ncbi.nlm.nih.gov/pubmed/16682683.

[14] A.A. Johnson, K.A. Johnson, Fidelity of nucleotide incorporation by human mitochondrial DNA polymerase, J. Biol. Chem. 44613 (512) (2001) 38090–38096, http://dx.doi.org/10.1074/jbc.M106045200.

[15] L.S. Kaguni, DNA polymerase γ, the mitochondrial replicase, Annu. Rev. Biochem. 73 (1) (2004) 293–320, http://dx.doi.org/10.1146/annurev.biochem.72.121801.161455.

[16] L.S. Kaguni, M.T. Oliveira, Structure, function and evolution of the animal mitochondrial replicative DNA helicase, Crit. Rev. Biochem. Mol. Biol. 9238 (2016) 1–12, http://dx.doi.org/10.3109/10409238.2015.1117056 https://www.ncbi.nlm.nih.gov/pubmed/26615986.

[17] M. Kaliszewska, J. Kruszewski, B. Kierdaszuk, A. Kostera-Pruszczyk, M. Nojszewska, A. Łusakowska, … K. Tońska, Yeast model analysis of novel polymerase gamma variants found in patients with autosomal recessive mitochondrial disease, Hum. Genet. 134 (9) (2015) 951–966, http://dx.doi.org/10.1007/s00439-015-1578-x.

[18] R. Kasiviswanathan, M.J. Longley, S.S.L. Chan, W.C. Copeland, Disease mutations in the human mitochondrial DNA polymerase thumb subdomain impart severe defects in mitochondrial DNA replication, J. Biol. Chem. 284 (29) (2009) 19501–19510, http://dx.doi.org/10.1074/jbc.M109.011940.

[19] Y.S. Lee, W.D. Kennedy, Y.W. Yin, Structural insight into processive human mitochondrial DNA synthesis and disease-related polymerase mutations, Cell 139 (2) (2009) 312–324, http://dx.doi.org/10.1016/j.cell.2009.07.050.

[20] W. Lewis, B.J. Day, J.J. Kohler, S.H. Hosseini, S.S.L. Chan, E.C. Green, … W.C. Copeland, Decreased mtDNA, oxidative stress, cardiomyopathy, and death from transgenic cardiac targeted human mutant polymerase gamma, Lab. Invest. 87 (4) (2007) 326–335, http://dx.doi.org/10.1038/labinvest.3700523.

[21] M.J. Longley, D.W. Mosbaugh, Properties of the 3′ to 5′ exonuclease associated with porcine liver DNA polymerase γ: Substrate specificity, product analysis, inhibition, and kinetics of terminal excision, J. Biol. Chem. 36 (1991) 24702–24711. https://www.ncbi.nlm.nih.gov/pubmed/1662214.

[22] M.J. Longley, D. Nguyen, T.a. Kunkel, W.C. Copeland, The fidelity of human DNA polymerase γ with and without exonucleolytic proofreading and the p55 accessory subunit, J. Biol. Chem. 276 (42) (2001) 38555–38562, http://dx.doi.org/10.1074/jbc.M105230200.

[23] N. Luo, L.S. Kaguni, Mutations in the spacer region of Drosophila mitochondrial DNA polymerase affect DNA binding, processivity, and the balance between Pol and Exo function, 280 (2005), pp. 2491–2497, http://dx.doi.org/10.1074/jbc.M411447200 https://www.ncbi.nlm.nih.gov/pubmed/1662214.

[24] P. Luoma, A. Melberg, J.O. Rinne, J.A. Kaukonen, N.N. Nupponen, R.M. Chalmers, … A. Suomalainen, Parkinsonism, premature menopause, and mitochondrial DNA polymerase γ mutations: clinical and molecular genetic study, Lancet 364 (9437) (2004) 875–882, http://dx.doi.org/10.1016/S0140-6736(04)16983-3.

[25] P.T. Luoma, N. Luo, W.N. Löscher, C.L. Farr, R. Horvath, J. Wanschitz, … A. Suomalainen, Functional defects due to spacer-region mutations of human mitochondrial DNA polymerase in a family with an ataxia-myopathy syndrome, Hum. Mol. Genet. 14 (14) (2005) 1907–1920, http://dx.doi.org/10.1093/hmg/ddi196.

[26] B. Macao, J.P. Uhler, T. Siibak, X. Zhu, Y. Shi, W. Sheng, … M. Falkenberg, The exonuclease activity of DNA polymerase γ is required for ligation during mitochondrial DNA replication, Nat. Commun. 6 (2015) 7303, http://dx.doi.org/10.1038/ncomms8303.

[27] R.K. Naviaux, K.V. Nguyen, POLG mutations associated with Alpers' syndrome and mitochondrial DNA depletion, Ann. Neurol. 55 (5) (2004) 706–712.

[28] A. Niroula, S. Urolagin, M. Vihinen, PON-P2: prediction method for fast and reliable identification of harmful variants, PLoS One 10 (2) (2015) 1–17, http://dx.doi.org/10.1371/journal.pone.0117380.

[29] M.W. Olson, L.S. Kaguni, 3′ - 5′ exonuclease in Drosophila mitochondrial DNA polymerase, J. Biol. Chem. 32 (1992) 223136–223142. https://www.ncbi.nlm.nih.gov/pubmed/1429661.

[30] E.J.H. Palin, A. Lesonen, C.L. Farr, L. Euro, A. Suomalainen, L.S. Kaguni, Functional analysis of H. sapiens DNA polymerase γ spacer mutation W748S with and without common variant E1143G, Biochim. Biophys. Acta 6 (2010) 545–551, http://dx.doi.org/10.1016/j.bbadis.2010.02.003 https://www.ncbi.nlm.nih.gov/pubmed/20153822.

[31] Y. Qian, J.L. Ziehr, K.A. Johnson, Alpers disease mutations in human DNA polymerase gamma cause catalytic defects in mitochondrial DNA replication by distinct mechanisms, Front. Genet. 6 (2015) 1–11, http://dx.doi.org/10.3389/fgene.2015.00135 https://www.ncbi.nlm.nih.gov/pubmed/25914719.

[32] S. Rajakulendran, R.D.S. Pitceathly, J.W. Taanman, H. Costello, M.G. Sweeney, C.E. Woodward, … S. Rahman, A clinical, neuropathological and genetic study of homozygous A467T POLG-related mitochondrial disease, PLoS One 11 (1) (2016) 1–16, http://dx.doi.org/10.1371/journal.pone.0145500.

[33] J. Reichenbach, R. Schubert, R. Horvàth, J. Petersen, N. Fütterer, E. Malle, … S. Zielen, Fatal neonatal-onset mitochondrial respiratory chain disease with T cell

immunodeficiency, Pediatr. Res. 60 (3) (2006) 321–326, http://dx.doi.org/10.1203/01.pdr.0000233252.60457.cf.

[34] S. Roos, B. Macao, J.M. Fusté, C. Lindberg, E. Jemt, E. Holme, ... M. Falkenberg, Subnormal levels of POLγA cause inefficient initiation of light-strand DNA synthesis and lead to mitochondrial DNA deletions and autosomal dominant progressive external ophthalmoplegia, Hum. Mol. Genet. 22 (12) (2013) 2411–2422, http://dx.doi.org/10.1093/hmg/ddt094.

[35] C. Rouzier, A. Chaussenot, V. Serre, K. Fragaki, S. Bannwarth, S. Ait-El-Mkadem, ... V. Paquis-Flucklinger, Quantitative multiplex PCR of short fluorescent fragments for the detection of large intragenic POLG rearrangements in a large French cohort, Eur. J. Hum. Genet. 22 (4) (2014) 542–550, http://dx.doi.org/10.1038/ejhg.2013.171.

[36] L. Shen, M.A. Diroma, M. Gonzalez, D. Navarro-Gomez, J. Leipzig, M.T. Lott, ... Xiaowu Gai, MSeqDR: a centralized knowledge repository and bioinformatics web resource to facilitate genomic investigations in mitochondrial disease, Hum. Mutat. 37 (6) (2016) 540–548, http://dx.doi.org/10.1002/humu.22974.

[37] J.D. Stumpf, W.C. Copeland, Mitochondrial DNA replication and disease: Insights from DNA polymerase γ mutations, Cell. Mol. Life Sci. 68 (2011) 219–233, http://dx.doi.org/10.1007/s00018-010-0530-4 https://www.ncbi.nlm.nih.gov/pubmed/20927567.

[38] K. Szczepanowska, F. Foury, A cluster of pathogenic mutations in the 3′–5′ exonuclease domain of DNA polymerase gamma defines a novel module coupling DNA synthesis and degradation, Hum. Mol. Genet. 19 (18) (2010) 3516–3529, http://dx.doi.org/10.1093/hmg/ddq267.

[39] M.R. Szymanski, V.B. Kuznetsov, C. Shumate, Q. Meng, Y.-S. Lee, G. Patel, ... Y.W. Yin, Structural basis for processivity and antiviral drug toxicity in human mitochondrial DNA replicase, EMBO J. 34 (14) (2015) 1959–1970, http://dx.doi.org/10.15252/embj.201591520.

[40] S. Tang, J. Wang, N.-C. Lee, M. Milone, M.C. Halberg, E.S. Schmitt, ... L.-J.C. Wong, Mitochondrial DNA polymerase gamma mutations: an ever expanding molecular and clinical spectrum, J. Med. Genet. 48 (10) (2011) 669–681, http://dx.doi.org/10.1136/jmedgenet-2011-100222.

[41] R.W. Taylor, A. Pyle, H. Griffin, E.L. Blakely, J. Duff, L. He, ... P.F. Chinnery, Use of whole-exome sequencing to determine the genetic basis of multiple mitochondrial respiratory chain complex deficiencies, JAMA 312 (2016) 68–77, http://dx.doi.org/10.1001/jama.2014.7184 https://www.ncbi.nlm.nih.gov/pubmed/25058219.

[42] G. Van Goethem, B. Dermaut, A. Löfgren, J.-J. Martin, C. Van Broeckhoven, Mutation of POLG is associated with progressive external ophthalmoplegia characterized by mtDNA deletions, Nat. Genet. 28 (2001) 211–212.

[43] G. Van Goethem, J.J. Martin, B. Dermaut, A. Löfgren, A. Wiball, D. Ververken, ... C. Van Broeckhoven, Recessive POLG mutations presenting with sensory and ataxic neuropathy in compound heterozygote patients with progressive external ophthalmoplegia, Neuromuscul. Disord. 13 (2003) 133–142. http://dx.doi.org/10.1016/S0, https://www.ncbi.nlm.nih.gov/pubmed/12565911.

[44] J.D. Stewart, R. Horvath, E. Baruffini, I. Ferrero, S. Bulst, P.B. Watkins, R.J. Fontana, C.P. Day, P.F. Chinnery, Polymerase γ gene POLG determines the risk of sodium valproate-induced liver toxicity, Hepatology 52 (2010) 1791–1796, http://dx.doi.org/10.1002/hep.23891 https://www.ncbi.nlm.nih.gov/pubmed/21038416.

[45] M. Vermulst, J.H. Bielas, G.C. Kujoth, W.C. Ladiges, P.S. Rabinovitch, T.A. Prolla, L.A. Loeb, Mitochondrial point mutations do not limit the natural lifespan of mice, Nat. Genet. 39 (4) (2007) 540–543, http://dx.doi.org/10.1038/ng1988.

[46] M.J. Young, W.C. Copeland, Human mtochondrial DNA replication machinery and disease, Curr. Opin. Genet. Dev. 38 (2016) 52–62, http://dx.doi.org/10.1016/j.gde.2016.03.005 https://www.ncbi.nlm.nih.gov/pubmed/27065468.

[47] E. Sarzi, A. Bourdon, D. Chrétien, M. Zarhrate, J. Corcos, A. Slama, V. Cormier-Daire, P. de Lonlay, A. Munnich, A. Rötig, Mitochondrial DNA depletion is a prevalent cause of multiple respiratory chain deficiency in childhood, J. Pediatr. 150 (2007) 531–534, http://dx.doi.org/10.1016/j.jpeds.2007.01.044 http://www.ncbi.nlm.nih.gov/pubmed/17452231.

[48] G. Galassi, E. Lamantea, F. Invernizzi, F. Tavani, I. Pisano, I. Ferrero, L. Palmieri, M. Zeviani, Additive effects of POLG1 and ANT1 mutations in a complex encephalomyopathy, Neuromuscul. Disord. 18 (2008) 465–470, http://dx.doi.org/10.1016/j.nmd.2008.03.013 http://www.ncbi.nlm.nih.gov/pubmed/18504126.

[49] C. Tzoulis, G.T. Tran, J. Coxhead, B. Bertelsen, P.K. Lilleng, N. Balafkan, B. Payne, H. Miletic, P.F. Chinnery, L.A. Bindoff, Molecular pathogenesis of polymerase γ-related neurodegeneration, Ann. Neurol. 76 (2014) 66–81, http://dx.doi.org/10.1002/ana.24185 http://www.ncbi.nlm.nih.gov/pubmed/24841123.

[50] M. Martikainen, Y. Ng, G. Gorman, C. Alston, E. Blakely, A. Schaefer, P. Chinnery, D. Burn, R. Taylor, R. McFarland, D. Turnbull, Clinical, genetic, and radiological features of extrapyramidal movement disorders in mitochondrial disease, JAMA Neurol. 73 (2016) 668–674, http://dx.doi.org/10.1001/jamaneurol.2016.0355 http://www.ncbi.nlm.nih.gov/pubmed/27111573.