

Doenças mitocondriais na era da sequenciação de nova geração: estudo de 450 doentes

Mitochondrial diseases in the next generation sequencing era: study of 450 patients

Célia Nogueira^{1,2}, Cristina Pereira², Lisbeth Silva^{1,2}, Mateus Laranjeira¹, Altina Lopes², Raquel Neiva², Esmeralda Rodrigues³, Teresa Campos³, Esmeralda Martins⁴, Anabela Bandeira⁴, Margarida Coelho⁴, Marina Magalhães⁴, Joana Damásio⁴, Ana Gaspar⁵, Patrícia Janeiro⁵, Levy Gomes⁵, Ana Cristina Ferreira⁶, Sandra Jacinto⁶, José Pedro Vieira⁶, Luísa Diogo⁷, Helena Santos⁸, Carla Mendonça⁹, Laura Vilarinho^{1,2}

celia.nogueira@insa.min-saude.pt

(1) Unidade de Investigação e Desenvolvimento, Departamento de Genética Humana, Instituto Nacional de Saúde Doutor Ricardo Jorge, Porto, Portugal

(2) Unidade de Rastreio Neonatal Metabolismo e Genética, Departamento de Genética Humana, Instituto Nacional de Saúde Doutor Ricardo Jorge, Porto, Portugal

(3) Unidade Local de Saúde de São João, Porto, Portugal

(4) Unidade Local de Saúde de Santo António, Porto, Portugal

(5) Unidade Local de Saúde Santa Maria, Lisboa, Portugal

(6) Unidade Local de Saúde São José, Lisboa, Portugal

(7) Unidade Local de Saúde de Coimbra, Coimbra, Portugal

(8) Unidade Local de Saúde Gaia e Espinho, Vila Nova de Gaia, Portugal

(9) Unidade Local de Saúde do Algarve, Faro, Portugal

_Resumo

As doenças mitocondriais (DM) são doenças raras, clínica e geneticamente heterogêneas, de difícil diagnóstico, para as quais não existe uma terapia eficaz. O desenvolvimento da tecnologia de sequenciação de nova geração (NGS) revolucionou o diagnóstico molecular deste grupo de doenças, permitindo a identificação de novos genes associados a estas patologias. Nesta nova era genética, através da utilização da tecnologia de NGS, estudamos um grupo de 450 doentes suspeitos de DM, sem etiologia molecular. A nossa estratégia combinada de NGS, englobou a sequenciação de um painel de 213 genes nucleares associados a DM e do DNA mitocondrial completo.

Neste estudo, identificamos variantes causais em 134 (30%) doentes analisados, 88 dos quais apresentaram variantes no DNA nuclear e 46 no DNA mitocondrial, tratando-se na maioria de doentes pediátricos (66%). Neste grupo de doentes, identificamos 72 variantes patogénicas descritas na literatura e 20 novas variantes provavelmente patogénicas, assim como 62 variantes de significado indeterminado.

Como laboratório nacional de referência para o estudo e investigação das DM, demonstramos o contributo da tecnologia de NGS para esclarecer a etiologia molecular destes doentes, para expandir o espectro mutacional associado a estas patologias e oferecer um diagnóstico pré-natal e aconselhamento genético aos casais em risco.

_Abstract

Mitochondrial diseases (MD) are rare disorders, clinically and genetically heterogeneous, making their diagnosis a challenge and with hitherto no effective therapy options. The development of Next Generation Sequencing (NGS) technologies has revolutionized the understanding of many genes involved in the pathogenesis of MD. In this new genetic era, using a NGS approach, we studied a cohort of 450 Portuguese patients suspected of MD, and without a molecular etiology. Our combined NGS approach was first based on analysis of a custom-made targeted mitochondrial panel of 213 nuclear genes followed by the whole mitochondrial genome.

In this study, we identified disease related variants in 134 (30%) analyzed patients, 88 with nuclear DNA and 46 with mitochondrial DNA variants, being most of them pediatric patients (66%). The molecular analysis of this cohort revealed 72 already described pathogenic and 20 novel probably pathogenic variants, as well as 62 variants of unknown significance.

As a national laboratory for the study and research of MD, we demonstrated the power of NGS to achieve a molecular etiology of these patients, to expand the mutational spectrum associated to MD and to propose a prenatal diagnosis as well as an accurate genetic counseling for affected families.

_Introdução

As mitocôndrias são organelos celulares que desempenham um papel crucial na fosforilação oxidativa, produzindo mais de 90% da energia celular sob a forma de ATP. Estes organelos possuem o seu próprio material genético, o genoma mitocondrial (mtDNA), que é constituído por uma molécula circular de dupla hélice com 16.569 pares de bases. Este genoma é constituído por 37 genes que codificam 13 polipeptídeos, 22 tRNAs e 2 rRNAs. No entanto, as restantes 1500 proteínas mitocondriais são codificadas por genes nucleares (nDNA), sendo inicialmente traduzidas no citoplasma e posteriormente importadas para a mitocôndria (1-3).

As doenças mitocondriais (DM) são uma causa prevalente de doenças hereditárias do metabolismo, afetando aproximadamente 1:5.000 indivíduos (4). Estas doenças apresentam uma elevada heterogeneidade genética, uma vez que podem resultar de variantes patogénicas ao nível do nDNA e do mtDNA, podendo por este motivo estar associadas a diversos padrões de hereditariedade (5). As taxas de morbilidade e mortalidade entre os indivíduos afetados variam significativamente devido à ação concertada entre os dois genomas, sendo geralmente mais elevadas nos casos de rápida progressão e início pediátrico (6).

Os tecidos mais afetados são os que apresentam elevada dependência energética, como o cérebro, o músculo e o coração, pelo que clinicamente os doentes normalmente manifestam um envolvimento multissistémico, embora certos doentes tenham apenas o comprometimento de um único sistema, frequentemente o sistema neurológico (7,8).

Na era pré-molecular, o diagnóstico de uma DM dependia tradicionalmente de técnicas invasivas, como a biópsia muscular, para a determinação da atividade enzimática dos complexos da cadeia respiratória mitocondrial (CRM).

A era molecular da medicina mitocondrial iniciou-se em 1988 com a descoberta da primeira variante patogénica no mtDNA (9). Atualmente, a tecnologia de sequenciação de nova geração (NGS) tem a capacidade de gerar uma enorme quantidade

de dados num curto espaço de tempo e a um custo mais acessível, tornando este método ideal para a análise de grandes grupos de doentes (10,11). Assim, esta tecnologia não só permite identificar variantes patogénicas descritas na literatura, novas variantes provavelmente patogénicas, variantes de significado indeterminado (VUS), mas também variantes em genes ainda não associados a DM (10). Até à data foram descritas variantes causais associadas a DM em mais de 400 genes, quer mitocondriais, quer nucleares (12).

A aplicação desta tecnologia está a simplificar o diagnóstico das DM e a permitir oferecer um diagnóstico pré-natal e aconselhamento genético às famílias afetadas (13,14).

_Objetivo

Identificar a etiologia genética de 450 doentes com suspeita clínica e/ou bioquímica de doenças mitocondriais utilizando a tecnologia de sequenciação de nova geração para confirmação do diagnóstico.

_Doentes e métodos

Este trabalho complementou um estudo anterior de 146 doentes (15) e avaliou as características fenotípicas e o espectro genético de 450 doentes com DM em Portugal.

O DNA genómico foi extraído a partir de sangue periférico no *Bio Robot EZ1* (QIAGEN).

O NGS foi realizado no sequenciador *MiSeq* (Illumina). Para capturar os genes nucleares, foi utilizado um painel personalizado com 213 genes associados a DM, de acordo com a metodologia *SureSelect XT HS* (Agilent Technologies). Para o mtDNA, fez-se a sequenciação do mtDNA completo, utilizando a metodologia *Nextera XT* (Illumina).

Os ficheiros FASTQ gerados foram alinhados com o genoma de referência usando, para o nDNA, o *software Surecall* (Agilent Technologies) e a deteção e anotação das variantes foi realizada pelos *softwares SureCall* e *wANNOVAR* (<https://wannovar.wglab.org/>), respetivamente. Relativamente

ao mtDNA, os ficheiros FASTQ gerados foram alinhados com o genoma de referência usando o *software SeqMan NGen* (DNASStar) a deteção e anotação das variantes foi realizada com os *softwares SeqMan Pro* e *SeqMan NGen* (DNASStar), respetivamente.

As variantes foram filtradas tendo em consideração os seguintes critérios: i) o tipo de mutação, ii) previsões *in silico*, iii) a sua presença em bases de dados, iv) a frequência populacional e v) heteroplasmia maior que 5% (para variantes no mtDNA).

As variantes já classificadas como patogénicas ou com previsões *in silico* sugestivas de patogenicidade, foram confirmadas por sequenciação tradicional de Sanger, utilizado o protocolo *BigDye Terminator Cycle Sequencing Version 3.1* (Applied Biosystems) e os resultados da sequenciação foram analisados utilizando o sequenciador *ABI 3130XL* (Applied Biosystems). Quando disponíveis amostras de outros membros da família, foram realizados estudos de co-segregação.

_Resultados

Este estudo envolveu doentes provenientes de sete hospitais nacionais, localizados no norte, centro e sul de Portugal. A idade de início da doença foi muito variável, dos 4 dias aos 78 anos de vida. Os sintomas neurológicos foram os mais frequentes (42,8%), seguidos dos musculares (12,9%), oftalmológicos (11,9%), cardíacos (11,2%) e auditivos (4,6%).

O grupo pediátrico foi o mais representado neste estudo, com 58% dos doentes estudados (262/450), sendo os restantes 42% (188/450) constituídos por adultos. Um total de 29 (6%) doentes faleceram, sendo 23 crianças e 6 adultos, permitindo a investigação *post mortem* um diagnóstico molecular de 62% destes casos.

Este estudo permitiu identificar uma causa molecular em 30% dos doentes estudados (134/450), dos quais 66% apresentaram variantes no genoma nuclear e 34% no genoma mitocondrial.

Neste grupo de doentes identificamos 72 variantes patogénicas descritas na literatura, 20 novas variantes provavelmente patogénicas e 62 VUS, segundo os critérios do *American College of Medical Genetics and Genomics* (ACMG) (16), algumas já previamente descritas pela Unidade de Rastreio Neonatal, Metabolismo e Genética do Instituto Nacional de Saúde Doutor Ricardo Jorge (URN-INSA). (15,17).

Genoma Nuclear

Variantes patogénicas e provavelmente patogénicas

Dos 450 doentes, 51 (11%) dos casos apresentavam variantes patogénicas no nDNA, das quais 58 já descritas na literatura e 14 novas variantes provavelmente patogénicas (gráfico 1).

O número de variantes identificadas em cada gene variou entre 1% e 3%, no entanto, os genes *EARS2* e *VARS2* apresentaram o maior número de variantes, 9% e 7% respetivamente.

Variantes de significado indeterminado

Neste trabalho foram identificadas 54 VUS no nDNA em 37/450 (8%) doentes. Entre estas, 30 foram previamente descritas na literatura pela URN-INSA (15,17), sendo as restantes 24 novas variantes (gráfico 2).

Gráfico 1: Distribuição das variantes patogénicas / provavelmente patogénicas no nDNA por genes afetados e de acordo com seus grupos funcionais.

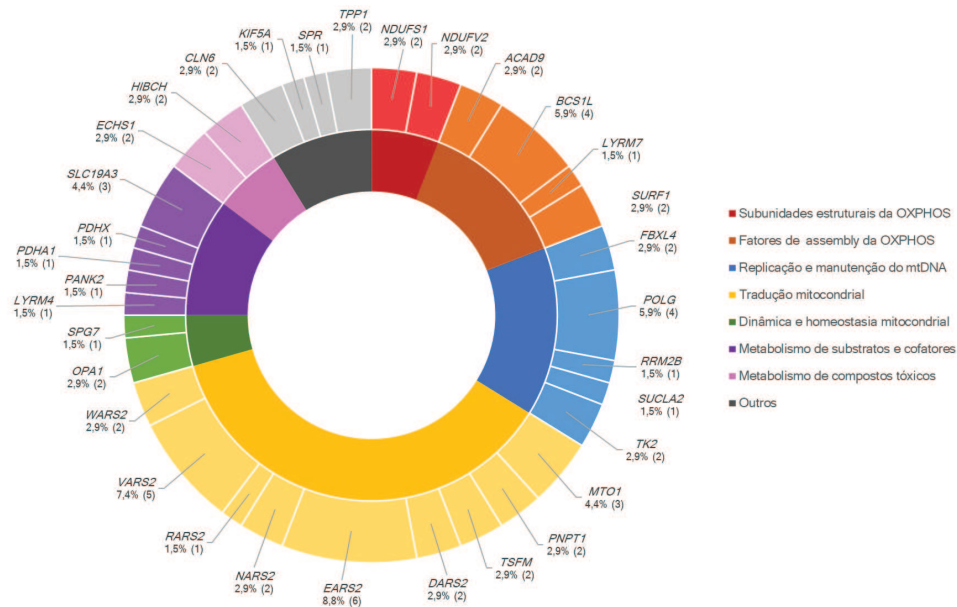
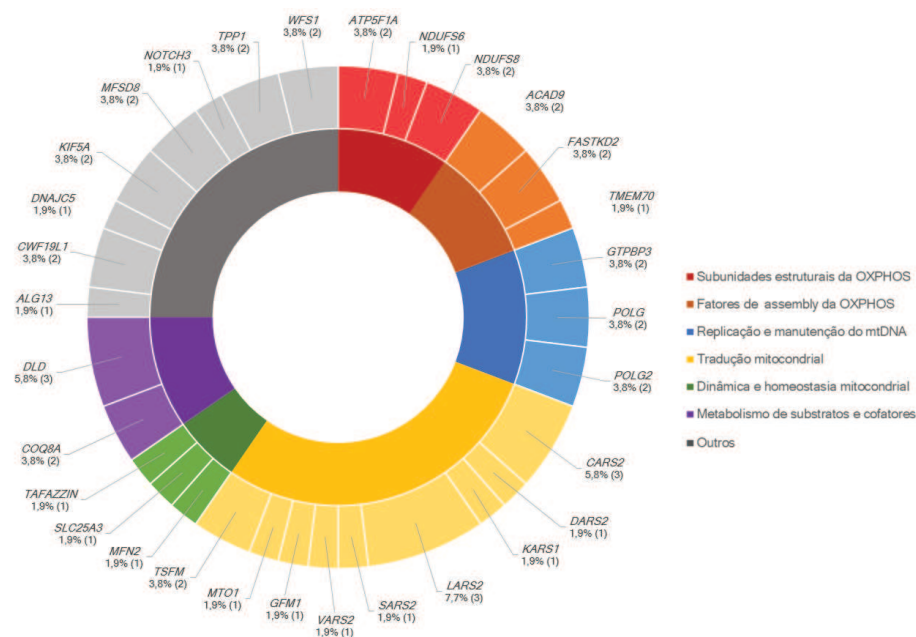


Gráfico 2: Distribuição das variantes de significado indeterminado no nDNA por genes afetados e de acordo com seus grupos funcionais.



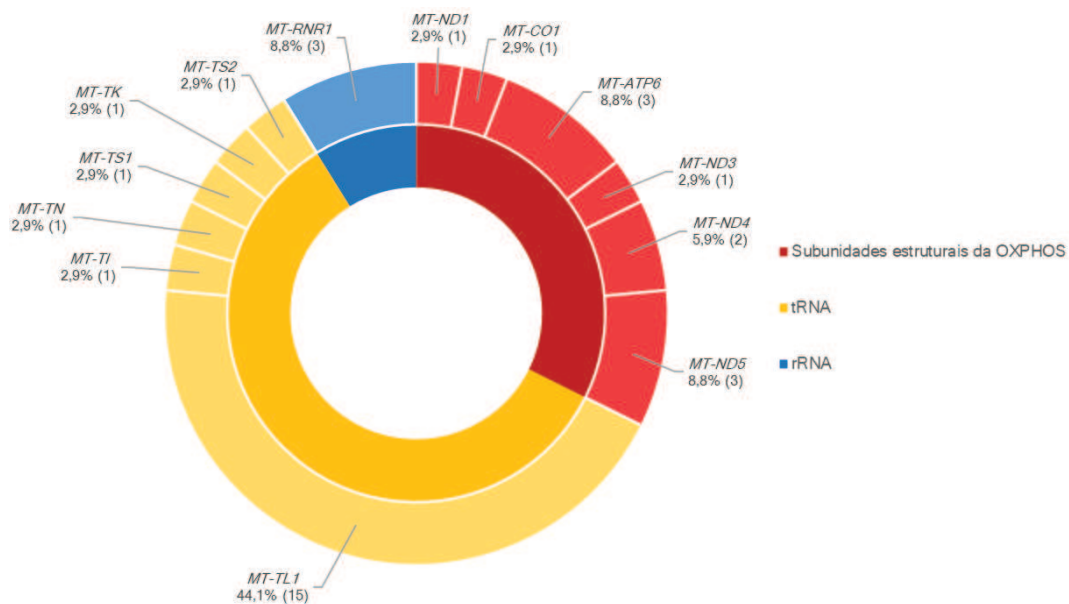
Genoma Mitocondrial

Variantes patogénicas e provavelmente patogénicas

A sequenciação do genoma mitocondrial revelou variantes patogénicas e provavelmente patogénicas bem como grandes deleções, em 38 dos 450 doentes (9%). Entre as variantes identificadas, 16 já tinham sido descritas na literatura, 11 das quais como patogénicas e cinco como provavelmente patogénicas, sendo duas novas variantes. Adicionalmente foram também identificadas quatro deleções simples de grandes dimensões e de diferentes tamanhos.

Os genes que codificam os tRNAs representam o maior grupo de genes com variantes identificadas, particularmente o *MT-TL1* devido à presença da mutação m.3243A>G detetada em 11 doentes. As variantes localizadas nos genes que codificam subunidades da OXPHOS estão equitativamente distribuídas entre eles (**gráfico 3**). A grande maioria das variantes foi detetada em heteroplasmia com percentagens que variaram entre os 5% e 95%.

Gráfico 3: Distribuição das variantes patogénicas / provavelmente patogénicas no mtDNA por genes afetados e de acordo com seus grupos funcionais.

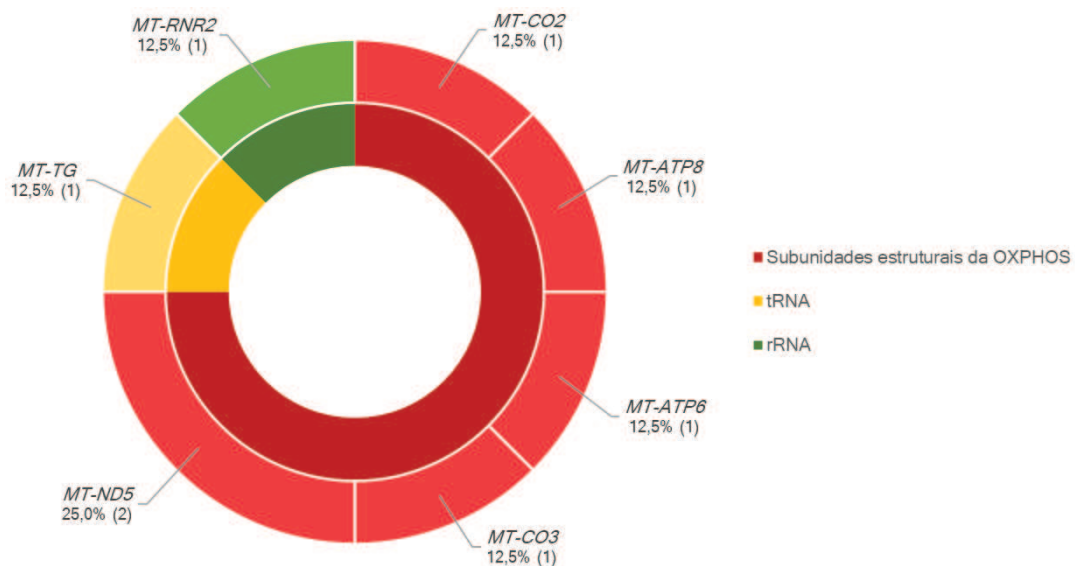


Variantes de significado indeterminado

A sequenciação completa do mtDNA revelou oito VUS em oito indivíduos distintos (8/450) distribuídas aleatoriamente por alguns genes estruturais da OXPHOS e pelos genes

MT-TG e *MT-RNR2* (gráfico 4). Todas as variantes foram identificadas em heteroplasmia exceto a variante localizada no gene *MT-CO3*.

Gráfico 4: Distribuição das variantes de significado indeterminado no mtDNA por genes afetados e de acordo com seus grupos funcionais.



_Discussão

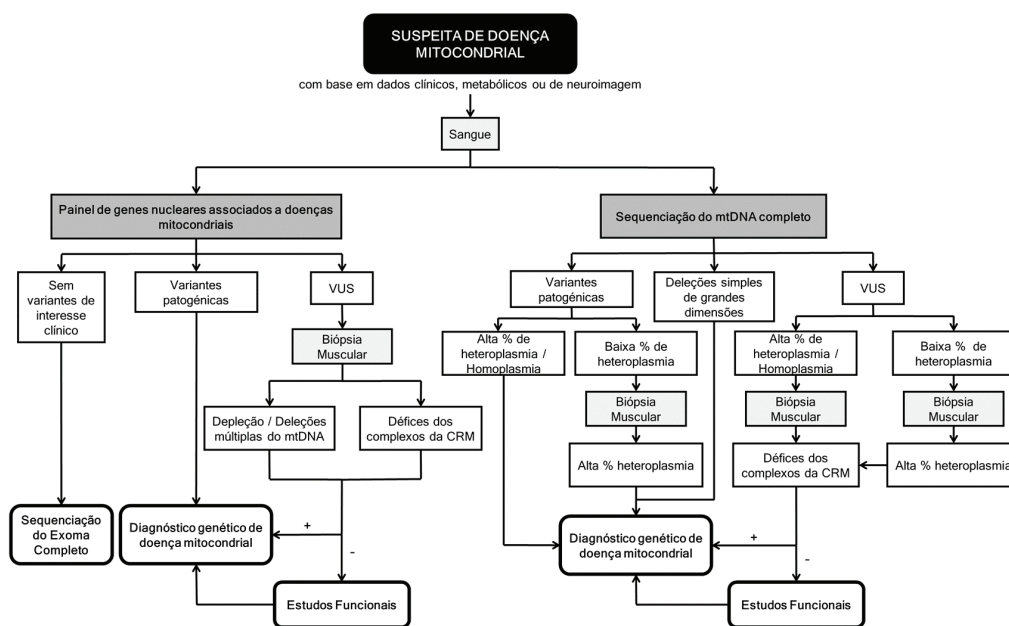
Até 2016, a sequenciação tradicional de Sanger foi utilizada para rastrear as variantes mais comuns do mtDNA bem como um grupo específico de genes nucleares associados a apresentações clínicas específicas e/ou défices enzimáticos dos complexos da CRM. Com os avanços tecnológicos, a abordagem molecular utilizada no laboratório da URN-INSA passou a incluir a tecnologia de NGS para a sequenciação de um painel de 213 genes nucleares associados a DM e do genoma mitocondrial completo. Inicialmente aplicamos esta estratégia a um grupo de 146 doentes com suspeita de DM, mas que não dispunham do diagnóstico molecular

confirmatório, como referido na nossa publicação anterior (15,17). Posteriormente, alargamos esta abordagem a mais 304 doentes suspeitos de DM, o que possibilitou um estudo mais abrangente de 450 doentes. Esta investigação permitiu-nos obter um diagnóstico molecular em 134/450 doentes (30%), 88 no nDNA e 46 no mtDNA, identificando um total de 154 variantes: 126 (82%) no genoma nuclear e 28 (18%) no genoma mitocondrial. Salientamos que, 72 eram variantes patogénicas descritas na literatura, 20 eram novas variantes e 62 eram VUS. As variantes no nDNA e mtDNA foram identificadas maioritariamente em crianças e em adultos, respetivamente, o que está de acordo com a literatura (18,19).

Neste estudo um dos fatores que influenciou a menor percentagem de doentes com variantes no mtDNA foi a idade, uma vez que 58% dos doentes estudados eram crianças, nas quais uma variante causal no nDNA é mais provável.

Com base nestes resultados propomos um algoritmo para as DM apresentado na **figura 1**.

Figura 1: Algoritmo de diagnóstico para suspeita de doença mitocondrial.



Atualmente, o sangue é a amostra de primeira linha para efetuar o painel de genes nucleares e a sequenciação do genoma mitocondrial completo. Apesar do tecido muscular ser rico em mitocôndrias, a análise da biópsia muscular, por se tratar de um procedimento invasivo, deve apenas ser considerada quando a apresentação clínica sugere a presença de uma grande deleção ou depleção do mtDNA. Pode ainda recorrer-se à biópsia muscular numa fase mais avançada da investigação para confirmar a patogenicidade de uma VUS, tornando-se assim relevante efetuar o estudo bioquímico da CRM. O tecido muscular é igualmente importante para determinar o nível de heteroplasmia quando a percentagem de mtDNA mutado detetada no sangue é

baixa, podendo não ser representativa do limiar de expressão da variante em estudo (20).

Salientamos ainda que deverão ser efetuados estudos de segregação familiar para confirmar as variantes patogénicas e/ou VUS, assim como, estudos funcionais em *zebrafish*, para determinar a patogenicidade das VUS identificadas.

O alargamento deste trabalho a um maior grupo de doentes permitiu aumentar a percentagem de diagnóstico definitivo de DM de 25% para 30%. No entanto, os 316 doentes (70%) que permaneceram sem caracterização molecular deverão continuar a ser investigados através da análise do exoma-clínico/total.

Conclusões

O presente estudo realça o espectro genético das doenças mitocondriais (DM) na era da sequenciação de nova geração (NGS), permitindo um avanço significativo na caracterização molecular destes doentes de difícil diagnóstico.

Os resultados do centro de referência para o estudo e investigação das DM do Instituto Nacional de Saúde Doutor Ricardo Jorge, permitiram esclarecer a etiologia molecular de 30% dos doentes estudados, alargar o espectro mutacional associado a estas patologias e oferecer em alguns casos um diagnóstico pré-natal e aconselhamento genético aos casais em risco.

Financiamento:

Fundação da Ciência e Tecnologia (PTDC/DTP-PIC/2220/2014, *Genetic Defects of Mitochondrial Diseases: a Next Generation Sequencing Approach*); Ao Programa Norte 2020 (NORTE-01-0246-FEDER-000014, DESVENDAR “DESCobrir, VENcer as Doenças rARas”).

Agradecimentos:

À UTI do DGH_INSA pela colaboração na confirmação das variantes identificadas, pelo método de Sanger. À Marisa Encarnação pelo apoio técnico no sequenciador MiSeq.

Referências bibliográficas:

- (1) Friedman JR, Nunnari J. Mitochondrial form and function. *Nature*. 2014 Jan 16;505(7483):335-43. <https://doi.org/10.1038/nature12985>
- (2) Mishra P, Chan DC. Mitochondrial dynamics and inheritance during cell division, development and disease. *Nat Rev Mol Cell Biol*. 2014 Oct;15(10):634-46. <https://doi.org/10.1038/nrm3877>
- (3) Kotrys AV, Szczesny RJ. Mitochondrial Gene Expression and Beyond—Novel Aspects of Cellular Physiology. *Cells*. 2019 Dec 19;9(1):17. <https://doi.org/10.3390/cells9010017>
- (4) Gorman GS, Schaefer AM, Ng Y, et al. Prevalence of nuclear and mitochondrial DNA mutations related to adult mitochondrial disease. *Ann Neurol*. 2015 May;77(5):753-9. <https://doi.org/10.1002/ana.24362>
- (5) Rahman S. Mitochondrial disease in children. *J Intern Med*. 2020 Jun;287(6):609-33. <https://doi.org/10.1111/joim.13054>
- (6) Gorman GS, Chinnery PF, DiMauro S, et al. Mitochondrial diseases. *Nat Rev Dis Primers*. 2016 Oct 20;2:16080. <https://doi.org/10.1038/nrdp.2016.80>
- (7) DiMauro S, Schon EA, Carelli V, et al. The clinical maze of mitochondrial neurology. *Nat Rev Neurol*. 2013 Aug;9(8):429-44. <https://doi.org/10.1038/nrneurol.2013.126>
- (8) Maresca A, Del Dotto V, Romagnoli M, et al.; ER-MITO Study Group. Expanding and validating the biomarkers for mitochondrial diseases. *J Mol Med (Berl)*. 2020 Oct;98(10):1467-78. <https://doi.org/10.1007/s00109-020-01967-y>
- (9) Wallace DC, Singh G, Lott MT, et al. Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science*. 1988 Dec 9;242(4884):1427-30. <https://doi.org/10.1126/science.3201231>
- (10) Stenton SL, Prokisch H. Genetics of mitochondrial diseases: Identifying mutations to help diagnosis. *EBioMedicine*. 2020 Jun;56:102784. <https://doi.org/10.1016/j.ebiom.2020.102784>
- (11) Riley LG, Cowley MJ, Gayevskiy V, et al. The diagnostic utility of genome sequencing in a pediatric cohort with suspected mitochondrial disease. *Genet Med*. 2020 Jul;22(7):1254-61. <https://doi.org/10.1038/s41436-020-0793-6>
- (12) Frazier AE, Thorburn DR, Compton AG. Mitochondrial energy generation disorders: genes, mechanisms, and clues to pathology. *J Biol Chem*. 2019 Apr 5;294(14):5386-95. <https://doi.org/10.1074/jbc.R117.809194>
- (13) French CE, Delon I, Dolling H, et al; NIHR BioResource—Rare Disease; Next Generation Children Project; Rowitch DH, Raymond FL. Whole genome sequencing reveals that genetic conditions are frequent in intensively ill children. *Intensive Care Med*. 2019 May;45(5):627-36. <https://doi.org/10.1007/s00134-019-05552-x>
- (14) Poulton J, Steffann J, Burgstaller J, et al.; workshop participants. 243rd ENMC international workshop: Developing guidelines for management of reproductive options for families with maternally inherited mtDNA disease, Amsterdam, the Netherlands, 22-24 March 2019. *Neuromuscul Disord*. 2019 Sep;29(9):725-33. <https://doi.org/10.1016/j.nmd.2019.08.004>
- (15) Nogueira C, Silva L, Pereira C, et al. Targeted next generation sequencing identifies novel pathogenic variants and provides molecular diagnoses in a cohort of pediatric and adult patients with unexplained mitochondrial dysfunction. *Mitochondrion*. 2019 Jul;47:309-317. <https://doi.org/10.1016/j.mito.2019.02.006>
- (16) Richards S, Aziz N, Bale S, et al.; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015 May;17(5):405-24. <https://doi.org/10.1038/gim.2015.30>
- (17) Nogueira C, Pereira C, Silva L, et al. Avanços no diagnóstico das doenças mitocondriais através da sequenciação de nova geração. *Boletim Epidemiológico Observações*. 2018;7(21):5-8. <https://repositorio.insa.pt/handle/10400.18/5546>
- (18) Pronicka E, Piekutowska-Abramczuk D, Ciara E, et al. New perspective in diagnostics of mitochondrial disorders: two years' experience with whole-exome sequencing at a national paediatric centre. *J Transl Med*. 2016 Jun 12;14(1):174. <https://doi.org/10.1186/s12967-016-0930-9>
- (19) Subathra M, Ramesh A, Selvakumari M, et al. Genetic Epidemiology of Mitochondrial Pathogenic Variants Causing Nonsyndromic Hearing Loss in a Large Cohort of South Indian Hearing Impaired Individuals. *Ann Hum Genet*. 2016 Sep;80(5):257-73. <https://doi.org/10.1111/ahg.12161>
- (20) Lawless C, Greaves L, Reeve AK, et al. The rise and rise of mitochondrial DNA mutations. *Open Biol*. 2020 May;10(5):200061. <https://doi.org/10.1098/rsob.200061>