

## Doenças genómicas no Departamento de Genética Humana do INSA: da resolução de mapa físico à medicina genómica personalizada

### Genomic diseases at the Department of Human Genetics of INSA: from physical map resolution to personalized genomic medicine

Dezso David<sup>1</sup>, Barbara Marques<sup>2</sup>, Joana Fino<sup>1</sup>, Andreas Gal<sup>3</sup>

dezso.david@insa.min-saude.pt

(1) Unidade de Investigação e Desenvolvimento. Departamento de Genética Humana, Instituto nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal

(2) Unidade de Citogenética. Departamento de Genética Humana, Instituto nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal

(3) Professor Emérito, Instituto de Genética Humana do Centro Médico Universitário de Hamburgo-Eppendorf, Hamburgo, Alemanha

#### \_Resumo

As doenças genómicas (DGs), resultantes de variantes estruturais (SVs) ou genómicas, apresentam diagnóstico clínico complexo. Podem ser simples ou complexas, equilibradas ou desequilibradas, associadas a doenças conhecidas ou a patologias ainda não descritas, constituindo recurso essencial para a anotação do genoma humano. A sua caracterização no DGH INSA acompanhou a evolução tecnológica, do mapa físico do genoma e arrays genómicos à sequenciação genómica (SG), culminando na medicina genómica de precisão (MGP).

A caracterização de SVs permitiu identificar genes afetados como: *ARHGEF6*, *TGFβ2*, *GLI2*, *TRPS1*, *PGK1*, *OAF*, *NECTIN1*, *FOXC1*, *ANKS1B*, *WDR26*, *YIPF5* e *CAMK2B*. Os mecanismos patogénicos mais comuns foram a disrupção génica e o efeito de posição.

A SG permitiu identificar, o espectro completo de variantes genómicas e genéticas, sendo a interpretação da patogenicidade o principal desafio. A implementação da MGP numa coorte de 28 casos simples e 13 complexos, integrando dados genómicos, expressão diferencial, análise de vias, correlação fenótipo-genótipo e fenotipagem reversa, permitiu o diagnóstico de mais de 80% dos casos, maioritariamente DGs, incluindo patologias ainda não descritas e mecanismos patogénicos adicionais. Esta abordagem poderá ser aplicada a todas as doenças raras e contribui para encurtar a odisseia do diagnóstico e para a melhoria dos cuidados de saúde dos doentes.

#### \_Abstract

Genomic disorders (GDs) arise from structural variants (SVs) or genomic variants, detectable or not by cytogenomic analysis, and pose significant clinical diagnostic challenges. They can be simple or complex, balanced or unbalanced, associated with OMIM diseases or previously undescribed pathologies, and constitute an essential resource for human genome annotation. Their characterization at the DGH has evolved with technological advances, from the physical genome map and genomic array painting to genome sequencing (GS), culminating in Precision Genomic Medicine (PGM).

Characterization of SVs identified the following genes as affected: *ARHGEF6*, *TGFβ2*, *GLI2*, *TRPS1*, *PGK1*, *OAF*, *FOXC1*, *NECTIN1*, *ANKS1B*, *WDR26*, *YIPF5*, *SPATC1L* and *CAMK2B*. Gene disruption and position effect were the most frequent pathogenic mechanisms.

The GS enabled nucleotide-resolution detection of the full spectrum of genomic and genetic variants, with interpretation of pathogenicity remaining the main challenge. Implementation of PGM in a cohort of 28 simple and 13 complex cases, integrating genomic data, differential expression, pathway analysis, phenotype-genotype correlation, and reverse phenotyping, allowed diagnosis in >80% of cases, mostly GDs, including previously undescribed conditions and additional pathogenic mechanisms. Short-insert GS captured all pathogenetically relevant breakpoints and variants. This approach could be extended to all rare diseases and contributes to improving patient care.

#### \_Introdução

As doenças genómicas (DGs) raras resultam de variantes cromossómicas estruturais (CSVs) ou de variações no número de cópias (CNVs), detetáveis ou não através de análise citogenómica. Estas variantes podem ser simples, com dois pontos de quebra, ou complexas, com múltiplos pontos de quebra, podendo apresentar-se de forma equilibrada ou desequilibrada. Frequentemente, estão associadas a doenças mendelianas e a síndromes clínicas descritas na base de dados OMIM (*Online Mendelian Inheritance in Man*), ou a condições ainda não descritas. O diagnóstico, sobretudo nos casos de CSV complexas ou de variantes potencialmente relacionadas com doenças até agora caracterizadas, continua a ser um desafio quando baseado apenas na avaliação clínica.

As variantes estruturais (SVs), nomeadamente as translocações recíprocas equilibradas, constituem recursos valiosos para a identificação de elementos genómicos patogénicos e dos fenótipos associados, contribuindo de forma significativa para a anotação do genoma humano mórbido, incluindo regiões codificantes e não codificantes. Durante décadas, a

cariotipagem convencional foi a abordagem de referência para deteção de CSVs e respetivos pontos de quebra, embora com resolução limitada (>5 Mb).

Antes da conclusão do Projeto Genoma Humano e do desenvolvimento das tecnologias de sequenciação de nova geração (NGS), a caracterização das CSVs baseava-se principalmente na hibridação fluorescente *in situ* (FISH), com sondas de dimensão variável. (1) Após o desenvolvimento dos *microarrays* cromossómicos, em amplicões genómicos de derivados e *microarray painting*. (2)

Na ausência de resolução nucleotídica dos pontos de quebra e de perfis genómicos abrangentes capazes de integrar CNVs, variantes de nucleotídeo único (SNVs) e pequenas inserções/deleções (indels), a interpretação clínica das SVs é limitada. A introdução de diferentes bibliotecas de sequenciação genómica (SG) - incluindo de insertos longos (li), pares *mate-pair* (mp) e insertos curtos (si) - possibilitou a deteção em larga escala de variantes estruturais com resolução nucleotídica, abrindo caminho ao desenvolvimento da medicina genómica personalizada (MGP). (3-5)

## \_Objetivo

Este trabalho descreve a evolução da investigação em doenças genómicas (DGs) no Departamento de Genética Humana do INSA, antes e após a disponibilização da sequência completa do genoma humano e da sequenciação genómica (SG). Esta linha de investigação teve como objetivo contribuir para a anotação do genoma mórbido, através da identificação de elementos genómicos patogénicos, culminando na abordagem atual de medicina genómica de precisão baseada em SG, que encurta a odisseia do diagnóstico e contribui para a melhoria dos cuidados de saúde de indivíduos com DGs e das respetivas famílias.

## \_Material e métodos

Para a implementação da MGP foi estudada uma coorte composta por 41 probandos e respetivos familiares com diagnóstico clínico de DGs, caracterizados predominantemente por alterações do neurodesenvolvimento e anomalias congénitas, com ou sem dismorfismos craniofaciais. Estas condições estavam associadas a variantes estruturais simples (n=28) ou complexas (n=13).

A análise genómica incluiu três tipos de bibliotecas de sequenciação - *li*, *mp* e *si* - complementadas com a sequenciação de mRNA e lncRNA obtidos a partir de sangue periférico. Os métodos utilizados seguiram protocolos padronizados, de acordo com as instruções dos fabricantes. A informação detalhada sobre as metodologias encontra-se resumida na [tabela 1](#) e descrita nas publicações referenciadas.

## \_Resultados e discussão

A investigação desenvolvida acompanhou os avanços tecnológicos, evoluindo da identificação de genes candidatos através de abordagem baseada no mapeamento físico para as estratégias assentes na SG e na medicina personalizada.

### Investigação de genes candidatos baseada no mapeamento físico

Na era do mapa físico genómico, os pontos de quebra nos cromossomas envolvidos em rearranjos associados a doença foram mapeados por FISH, recorrendo a cromossomas artificiais e a amplicões de PCR de longo alcance, específicos para cada marcador STS (*Sequence Tagged Sites*) ([figura 1A](#)). A identificação com resolução nucleotídica foi realizada por sequenciação de Sanger ([tabela 1](#)). Com esta abordagem foi possível identificar os seguintes genes:

– *ARHGEF6* (OMIM \*300267) – Identificado como interrompido no IVS 10 pelo ponto de quebra em Xq26, sendo responsável pela perturbação do desenvolvimento intelectual ligada ao X tipo 46 (XLID46; OMIM #300436). (6)

– *TGFβ2* (OMIM \*190220) – Identificado como provavelmente sujeito a um efeito de posição no ponto de quebra em 1q41, resultante da translocação familiar t(1;7)(q41;p21) associada à Anomalia de Peter familiar. Por seu turno, o gene interrompido pelo ponto de quebra em 7p21 foi o *HDAC9* (OMIM \*606543), atualmente associado à síndrome aurículo-condilar autossómica dominante (AD), tipo 4 (ARCND4, OMIM #620457). (1)

– *GLI2* (OMIM \*165230) – Num feto com ectrodactilia, foi identificada uma translocação t(2;11)(q14.2;q14.2)dn, com o ponto de quebra na região intergénica entre *INHBB* (OMIM \*147390) e *GLI2*. Este último está associado à síndrome AD Culler-Jones (CJS, OMIM #615849), caracterizada, entre outros por polidactilia pós-axial dos pés, e à HPE9 (OMIM #610829). O mecanismo proposto foi um efeito de posição sobre *GLI2*, com impacto na regulação espaço-temporal. (7)

Posteriormente, a separação dos cromossomas derivados por citometria de fluxo, seguida da amplificação dos mesmos e do mapeamento dos pontos de quebra através de pintura de *microarrays* cromossómicos, facilitou o mapeamento destes CSVs (figura 1B e tabela 1).

Esta abordagem permitiu identificar os seguintes genes afetados (tabela 1):

– *TRPS1* (OMIM \*604386) – Primeiro caso descrito de síndrome trico-rino-falângica, tipo I (TRPSI, OMIM #1903501) associada a t(8;13)(q23.3;q21.31)fam. Propôs-se como mecanismo patogénico a relocação de um elemento regulador em *cis* proveniente da região do cromossoma 13 para a região 5' do *TRPS1*, conduzindo à sua sobre-expressão. A gravidade e expansão fenotípica no doente índice permanecem por esclarecer. (2)

– *PGK1* (OMIM \*311800) – Foram necessárias análises moleculares e funcionais detalhadas para demonstrar que o fenótipo anteriormente atribuído à t(3;14)(q26.33;q12) e à disrupção do *NUBPL* (OMIM \*613621) se devia, na realidade, a uma nova SNV no gene *PGK1* (ENST00000373316.5: c.358G>A, p.E120K), causadora de deficiência de *PGK1* (OMIM #300653). Este caso evidencia as dificuldades do diagnóstico clínico de uma forma grave da deficiência e a complexidade em distinguir as manifestações diretamente associadas à deficiência de *PGK1* daquelas resultantes da evolução clínica da doença. (8)

– *OAF* (OMIM \*621070) e *NECTIN1* (OMIM \*600644) – Foi identificada uma translocação equilibrada t(11;18)(q23.3;q11.2)dn, num indivíduo com anomalias oculares e défice neurológico (disgenesia queratolenticular e *ectopia lentis*). O ponto de quebra em 11q23.3 resulta numa redução da expressão deste regulador transcricional *OAF*. Já a *NECTIN1* ligada a defeitos oculares congénitos e localizada 500 kb a montante deste ponto de quebra, apresenta níveis de expressão aumentada. (9)

– *FOXC1* (OMIM \*601090) – O cromossoma 6 em anel [r(6)], que resulta numa microdeleção de 1880 kb em 6p25.3, está associado a disgenesia do segmento anterior e outras anomalias. O fenótipo observado é principalmente atribuível a haploinsuficiência de *FOXC1*. (10)

Tabela 1: ⬇️ Variantes estruturais associadas a doenças genómicas descritas nos últimos 25 anos.

Case	Metodologias	Patologia	Gene(s) candidatos / efeito	Ref.
<b>Baseada em mapeamento físico</b>				
t(X;21)(q26;p11)dn	FISH, Sequenciação Sanger	XLID46 (OMIM #300436)	<i>ARHGEF6</i> – disrupção do gene	(6)
t(1;7)(q41;p21)fam	Citogenética, FISH	Anomalia de Peter	<i>TGFβ2</i> – efeito de posição	(1)
t(2;11)(q14.2;q14.2)dn	Citogenética, FISH	Ectrodactilia	<i>GLI2</i> – efeito de posição	(7)
t(8;13)(q23.3;q21.31)fam	Isolamento de der, Pintura de array	TRPS1 (OMIM #190350)	<i>TPRS1</i> – efeito de posição	(2)
t(3;14)(q26.33;q12)fam	Isolamento de der, Pintura de array, WES, Estudo funcional	Deficiência de PGK1 (OMIM #300653)	<i>PGK1</i> c.358G>A, p.E120K	(8)
t(11;18)(q23.3;q11.2)dn	FISH, Isolamento de der, Estudos de expressão em humanos e murganhos, TADs	Anomalias oculares e défice neurológico	<i>OAF</i> – disrupção; <i>NECTIN1</i> – efeito de posição	(9)
r(6)(p24q25)	CMA	ASGD3 (OMIM #601631)	<i>FOXC1</i> – deleção	(10)
<b>Baseada na SG e na MGP</b>				
t(12;17)(q23.1;q21.33)dn <sup>a</sup>	liGS, Sequenciação Sanger	Síndrome de haploinsuficiência do <i>ANKS1B</i>	<i>ANKS1B</i> – disrupção do gene	(11)
t(1;3)(q24.11;p25.3)dn <sup>a</sup>	liGS, Sequenciação Sanger	SKDEAS (OMIM #617616)	<i>WDR26</i> – disrupção do gene	
t(16;17)(q24;q21.3)dn <sup>a</sup>	liGS, Sequenciação Sanger	KBGS (OMIM #148050)	<i>ANKRD11</i> – disrupção do gene	
t(2;19)(p13;q13.1)dn <sup>a</sup>	liGS, Sequenciação Sanger	Assintomático	NA	
<b>MGP – SVs, CNVs, SNVs, indels, expansões de repetições e genoma mitocondrial</b>				
t(17;19)(p13;p13.3)	liGS, Sequenciação Sanger, Array CGH, WES	SIHIWES (OMIM #617159)	<i>CHD4</i> – c.4442delG, p.(G1481Vfs*21)	(12)
t(5;9)(q31.3;p13)	liGS, siGS; Sequenciação Sanger	Infertilidade	<i>YIPF5</i> – efeito de posição	(13)
t(4;21)(p14;q21.3)	liGS, SIGS, Sequenciação Sanger	Infertilidade	<i>SPATC1L</i> – efeito de posição	
t(7;22)(p13;q13.1)dn	mpGS, Sequenciação Sanger	Infertilidade; MRD54 (OMIM #617799) <sup>b</sup>	<i>CAMK2B</i> – disrupção do gene	
11 CSVs simples, 5 inversões simples, 1 cromossoma supranumerário	siGS, liGS, RNA-Seq, Análise fenotípica comparativa e reversa	Alterações do neurodesenvolvimento e Anomalias congénitas	trans-heterozigidade <sup>c</sup> variantes genéticas	(14) <sup>d</sup>

<sup>a</sup> Variantes identificadas no decurso de diagnóstico pré-natal convencional;

<sup>b</sup> A infertilidade é atribuível à translocação, enquanto a perturbação do desenvolvimento intelectual resulta da disrupção de *CAMK2B*;

<sup>c</sup> Corresponde a uma trans-heterozigotia fenotípica digénica, decorrente da disrupção de *ODAD2* em associação com uma SNV em *CFTR*;

<sup>d</sup> Manuscrito em processo de revisão

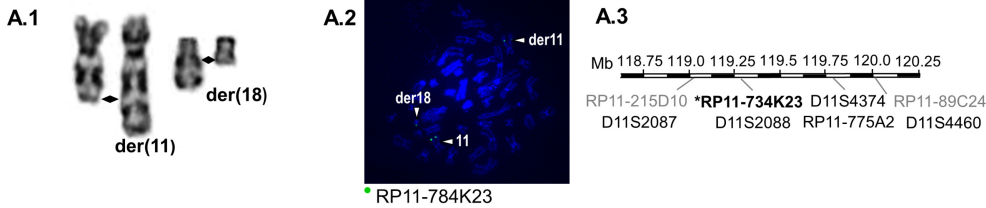
WES - Sequenciação do exoma completo; CMA - Análise de microarray cromossómico; dn - *de novo*; fam - familiar; FISH - Hibridação *in situ* Fluorescente;

Array CGH - Hibridação Genómica Comparativa por Array; der - cromossoma derivado; liGS - sequenciação genómica de insertos longos;

siGS - sequenciação genómica de insertos curtos; mpGS - sequenciação *made-pair*; CNV - Variação do número de cópias; RNA seq - sequenciação de RNA; NA - não aplicável.

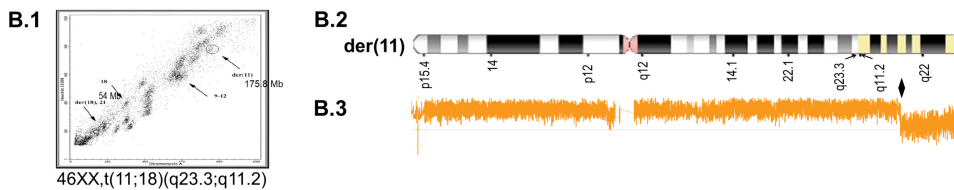
Figura 1: Identificação dos pontos de quebra por diferentes abordagens.

**A. Mapeamento por FISH com base no mapa físico genómico**



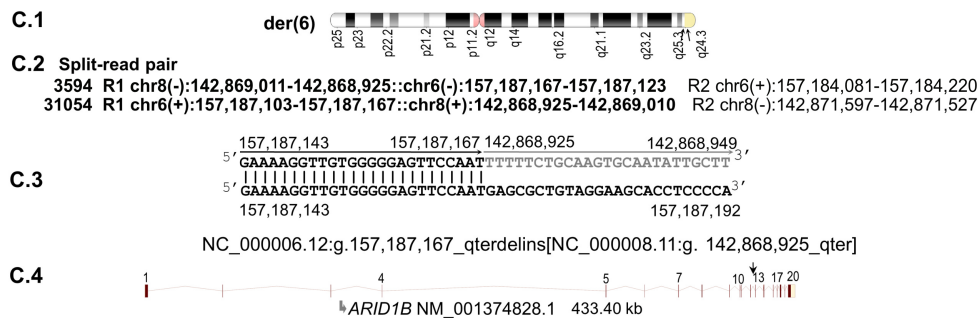
A) Mapeamento por FISH com base no mapa físico genómico.  
A.1) Cariótipo parcial da t(11;18)(q23.3;q11.2)dn.  
A.2) Mapeamento do ponto de quebra no cromossoma 11 por FISH utilizando o clone RP11-784K23.  
A.3) Esquema do mapa físico da região com marcadores STS e clones BAC utilizados em FISH.

**B. Mapeamento por isolamento dos derivados e microarrays**



B) Mapeamento por isolamento de derivados e microarrays.  
B.1) Cromossomas derivados isolados por citometria de fluxo. As setas indicam os grupos de cromossomas e os derivados, separados principalmente de acordo com os tamanhos dos mesmos.  
B.2) Ideograma do cromossoma der(11), com os pontos de quebra indicados por setas.  
B.3) Identificação dos pontos de quebra (losango) com amplicon genómico do der(11) de 175,8 Mb, utilizado como sonda para chromosome painting.

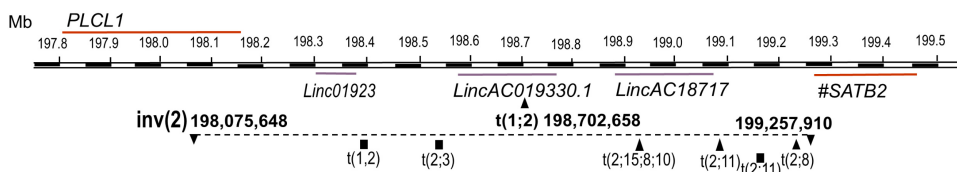
**C. Identificação do ponto de quebra por sequenciação genómica – Disrupção génica**



C) Sequenciação genómica.  
C.1) Ideograma do der(6).  
C.2) Pares de leitura que definem o ponto de quebra ao nível da sequência genómica, com a sequência de junção em negrito, enquanto a posição do ponto de quebra está indicada por dois-pontos duplos (::).  
C.3) Localização do ponto de quebra ao nível nucleotídico e designação de acordo com o sistema ISCN 2020.  
C.4) Disrupção do ARID1B no intrão 12.

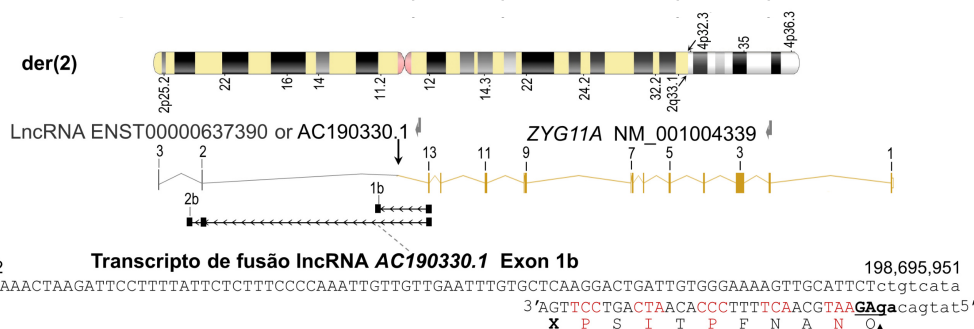
Figura 2: Exemplos de mecanismos patogénicos.

**A. Efeito de posição cromossómica sobre o SATB2**



A) Efeito de posição cromossómica em SATB2. Mapa físico da região com genes codificantes e não codificantes, pontos de quebra da t(1;2) e da inv(2) identificados neste estudo e os previamente publicados.

**B. Transcrito quimérico entre o ZYG11A (exão 13) e AC190330.1 (exão 1b)**



B) Transcritos quiméricos entre o exão 13 de ZYG11A e o exão 1b de lncRNA AC190330.1, presumivelmente não funcionais.

## Medicina genómica personalizada baseada em sequenciação genómica

Com as metodologias de SG que aplicamos, o número médio de CSVs e CNVs por caso, nas bibliotecas lIGS, mpGS e siGS, foi de 371389 e 1632, respetivamente, num total de 1851 CSVs e 17732 CNVs entre todos os indivíduos analisados; enquanto o número médio de SNVs e indels nas bibliotecas siGS e mpGS foi de 4162 milhões e 3808 milhões, respetivamente. Devido sobretudo à anotação ainda insuficiente do genoma humano mórbido, à sobreposição fenotípica de variantes genómicas e genéticas distintas que co-segregam no mesmo indivíduo e a potenciais efeitos cumulativos, a principal dificuldade atual não se encontra na identificação destas variantes, mas sim na interpretação da sua patogenicidade e relevância clínica.

A abordagem desenvolvida integrou dados genómicos e de expressão génica diferencial, análise de vias KEGG de interações moleculares, fenotipagem abrangente - incluindo a comparação sistemática dos fenótipos clínicos -, bem como dados genómicos convergentes e a correlação fenótipo-genótipo, culminando na confirmação através de fenotipagem reversa.

Dos 28 probandos analisados, 19 foram diagnosticados com doenças descritas na OMIM e cinco apresentaram patologias humanas ainda não descritas. Dois terços dos casos diagnosticados correspondem a DGs, enquanto os restantes foram atribuídos a variantes genéticas do tipo SNV e indels ou a outros mecanismos.

Os mecanismos patogénicos identificados (**figura 1C, 2A e 2B**) incluem: i) disrupção génica causada por pontos de quebra; ii) efeitos de posição, incluindo o causado por uma inversão críptica; iii) transcritos quiméricos ou de fusão presumivelmente não funcionais; iv) trans-heterozigotia digénica composta por diferentes tipos de variantes; e v) efeito de posição associado a inativação não aleatória, com expansão da inativação para a região autossómica. **(11,14)**

Entre os casos complexos, cinco foram reclassificados, tendo inicialmente sido considerados citogeneticamente simples; os restantes foram confirmados como complexos. Estes casos apresentaram, na sua maioria, desequilíbrios genómicos, envolvendo múltiplos tipos de variantes cromossómicas, diferentes pontos de quebra e genes afetados. Dada a sua complexidade e os potenciais efeitos cumulativos, a correlação fenótipo-genótipo revelou-se particularmente desafiante.

## \_Conclusões

Os objetivos deste trabalho consistiram em identificar genes candidatos associados a doenças e os respetivos mecanismos patogénicos, contribuindo simultaneamente para a anotação do genoma humano mórbido e para a otimização dos cuidados de saúde de indivíduos com DG e dos seus familiares.

Através da abordagem de MGP, implementámos a identificação do espectro completo de variantes estruturais e pequenas variantes genéticas, subjacentes aos fenótipos observados, elucidámos os seus mecanismos moleculares e avaliámos a utilidade diagnóstica de uma abordagem integrativa, incluindo RNA-Seq.

Os resultados demonstram que a abordagem de MGP proporciona uma taxa de diagnóstico elevada das DGs (acima de 80%), reduz a "odisseia diagnóstica" e contribui para a anotação funcional das regiões codificantes e não codificantes do genoma mórbido. Para além disso, a estratégia implementada é aplicável a outros grupos de patologias, nomeadamente doenças raras não diagnosticadas.

## Agradecimentos:

Agradecemos a todos os participantes destes estudos; aos membros do Grupo de Investigação em Doenças Genómicas do DGH do INSA pelo contributo na execução dos trabalhos; aos grupos responsáveis pela produção das bibliotecas e aquisição dos dados de sequenciação; aos nossos colaboradores nacionais e internacionais; à FCT e ao INSA, pelo financiamento da investigação.

Referências bibliográficas:

- (1) David D, Cardoso J, Marques B, et al. Molecular characterization of a familial translocation implicates disruption of HDAC9 and possible position effect on TGFBeta2 in the pathogenesis of Peters' anomaly. *Genomics*. 2003 May; 81(5):489-503. [http://doi.org/10.1016/s0888-7543\(03\)00046-6](http://doi.org/10.1016/s0888-7543(03)00046-6)
- (2) David D, Marques B, Ferreira C, et al. Co-segregation of trichorhinophalangeal syndrome with a t(8;13)(q23.3;q21.31) familial translocation that appears to increase TRPS1 gene expression. *Hum Genet*. 2013 Nov; 132(11):1287-99. <http://doi.org/10.1007/s00439-013-1333-0>
- (3) Talkowski ME, Ernst C, Heilbut A, et al. Next-generation sequencing strategies enable routine detection of balanced chromosome rearrangements for clinical diagnostics and genetic research. *Am J Hum Genet*. 2011 Apr 08; 88(4):469-81. <http://doi.org/10.1016/j.ajhg.2011.03.013>
- (4) Dong Z, Qian J, Law TSM, et al. Mate-pair genome sequencing reveals structural variants for idiopathic male infertility. *Hum Genet*. 2023 Mar;142(3):363-77. <http://doi.org/10.1007/s00439-022-02510-4>
- (5) Bentley DR, Balasubramanian S, Swerdlow HP, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 2008 Nov 6; 456(7218):53-59. <http://doi.org/10.1038/nature07517>
- (6) Kutsche K, Yntema H, Brandt A, et al. Mutations in ARHGEF6, encoding a guanine nucleotide exchange factor for Rho GTPases, in patients with X-linked mental retardation. *Nat Genet*. 2000 Oct; 26(2):247-50. <http://doi.org/10.1038/80002>
- (7) David D, Marques B, Ferreira C, et al. Characterization of two ectrodactyly-associated translocation breakpoints separated by 2.5 Mb on chromosome 2q14.1-q14.2. *Eur J Hum Genet*. 2009 Aug; 17(8):1024-33. <http://doi.org/10.1038/ejhg.2009.2>
- (8) David D, Almeida LS, Maggi M, et al. Clinical severity of PGK1 deficiency due to a novel p.E120K substitution is exacerbated by co-inheritance of a subclinical translocation t(3;14)(q26.33;q12), disrupting NUBPL Gene. *JIMD Rep*. 2015 Mar 27; 23:55-65. [http://doi.org/10.1007/8904\\_2015\\_427](http://doi.org/10.1007/8904_2015_427)
- (9) David D, Anand D, Araújo C, et al. Identification of OAF and PVRL1 as candidate genes for an ocular anomaly characterized by Peters anomaly type 2 and ectopia lentis. *Exp Eye Res*. 2018 Mar; 168:161-70. <http://doi.org/10.1016/j.exer.2017.12.012>
- (10) Corona-Rivera JR, Corona-Rivera A, Zepeda-Romero LC, et al. Ring chromosome 6 in a child with anterior segment dysgenesis and review of its overlap with other FOXC1 deletion phenotypes. *Congenit Anom (Kyoto)*. 2019 Sep; 59(5):174-78. <http://doi.org/10.1111/cga>
- (11) David D, Freixo JP, Fino J, et al. Comprehensive clinically oriented workflow for nucleotide level resolution and interpretation in prenatal diagnosis of de novo apparently balanced chromosomal translocations in their genomic landscape. *Hum Genet*. 2020 Apr; 139(4):531-43. <http://doi.org/10.1007/s00439-020-02121-x>
- (12) Da Silva JD, Oliva-Teles N, Tkachenko N, et al. Novel frameshift CHD4 variant leading to Sifrim-Hitz-Weiss syndrome in a proband with a subclinical familial t(17;19) and a large dup(2)(q14.3q21.1). *Biomedicines*. 2022 Dec 21; 11(1):12. <http://doi.org/10.3390/biomedicines11010012>
- (13) David D, Fino J, Oliveira R, et al. Balanced chromosomal rearrangements implicate YIPF5 and SPATC1L in non-obstructive oligoasthenozoospermia and oligozoospermia and of a derivative chromosome 22 in recurrent miscarriage. *Gene*. 2023 Dec 15; 887:147737. <http://doi.org/10.1016/j.gene.2023.147737>
- (14) David D, Fino J, Rodrigues M, et al. A personalized genomic medicine approach to rare genomic disorders associated with simple chromosomal structural variants. *HGG Adv*. In press 2025