

## Diagnóstico molecular de cancros hereditários por sequenciação de nova geração: cancro da mama e cancro colorretal

*Molecular diagnosis of hereditary cancers using next-generation sequencing: breast cancer and colorectal cancer*

Patrícia Theisen<sup>1</sup>, Catarina Silva<sup>2</sup>, Iris Pereira Caetano<sup>1</sup>, Pedro Rodrigues<sup>1</sup>, Glória Isidro<sup>1</sup>, Luís Vieira<sup>2</sup>, João Gonçalves<sup>1</sup>

joao.goncalves@insa.min-saude.pt

(1) Unidade de Genética Molecular; (2) Unidade de Tecnologia e Inovação, Departamento de Genética Humana, INSA.

### \_Resumo

O cancro da mama e o cancro colorretal constituem duas das principais causas de morte a nível mundial. Entre 5 a 10% destes casos estão associados a variantes germinais/hereditárias em genes de suscetibilidade para cancro. O objetivo deste trabalho consistiu em validar a utilização da sequenciação de nova geração (NGS) para identificar variantes previamente detetadas pelo método de Sanger em diversos genes de suscetibilidade para cancro da mama e colorretal. Foram sequenciadas por NGS 64 amostras de DNA de utentes com suspeita clínica de predisposição hereditária para cancro da mama ou colorretal, utilizando o painel de sequenciação *TruSight Cancer* e a plataforma *MiSeq* (Illumina). Estas amostras tinham sido previamente sequenciadas pelo método de Sanger para os genes *BRCA1*, *BRCA2*, *TP53*, *APC*, *MUTYH*, *MLH1*, *MSH2* e *STK11*. A análise bioinformática dos resultados foi realizada com os softwares *MiSeq Reporter*, *VariantStudio*, *Isaac Enrichment* (Illumina) e *Integrative Genomics Viewer* (Broad Institute). A NGS demonstrou elevada sensibilidade e especificidade analíticas para a deteção de variantes de sequência em 8 genes de suscetibilidade para cancro colorretal e da mama, uma vez que permitiu identificar a totalidade das 412 variantes (93 únicas, incluindo 27 variantes patogénicas) previamente detetadas pelo método de Sanger. A utilização de painéis de sequenciação de genes de predisposição para cancro por NGS vem possibilitar um diagnóstico molecular mais abrangente, rápido e custo-eficiente, relativamente às metodologias convencionais.

### \_Abstract

*Breast and colorectal cancers are two major causes of cancer-related deaths worldwide. 5-10% of these cases are associated with germline/hereditary variants in cancer susceptibility genes. The aim of this work was to validate a next-generation sequencing (NGS) cancer susceptibility gene panel for the identification of variants previously detected by Sanger sequencing in several breast and colorectal cancer susceptibility genes. DNA samples from 64 patients with a suspected inherited predisposition to breast or colorectal cancers were sequenced on a MiSeq using the TruSight Cancer Sequencing Panel (Illumina). These samples had been previously sequenced by the Sanger method for the BRCA1, BRCA2, TP53, APC, MUTYH, MLH1, MSH2 and STK11 genes. Bioinformatic analysis of NGS data included the MiSeq Reporter, VariantStudio, Isaac Enrichment (Illumina) and Integrative Genomics Viewer (Broad Institute) tools. High analytical sensitivity and specificity was obtained with NGS for the detection of sequence variants in 8 highly penetrant breast and colorectal cancer susceptibility genes, since it successfully identified all 412 sequence variants (93 unique variants, including 27 disease causing variants) previously detected by Sanger sequencing. Clinically useful NGS gene panels for cancer susceptibility will thus provide a more comprehensive and cost-effective molecular diagnosis approach, with a shorter turnaround time when compared to standard methodologies.*

### \_Introdução

O cancro da mama (CM) e o cancro colorretal (CCR) são responsáveis por uma elevada mortalidade associada ao cancro. Cerca de 5 a 10% destes casos são hereditários e estão associados a alterações germinais de elevada penetrância em genes de suscetibilidade para cancro. A identificação da causa genética subjacente aos cancros hereditários permite não só identificar os indivíduos com risco aumentado de desenvolver cancro como também oferecer uma medicina personalizada, mais eficaz na redução da incidência de cancro assim como da sua morbilidade e mortalidade.

Reconhecem-se, atualmente, cerca de 2 dezenas de genes de suscetibilidade para CM (1). Alterações germinais nos genes *BRCA1* e *BRCA2* têm elevada penetrância e são responsáveis por cerca de metade dos casos de CM hereditários (CMH). São também conhecidas alterações raras de elevada penetrância noutros genes, entre os quais se destacam *TP53*, *STK11*, *CDH1*, *CHEK2*, *PALB2* e *PTEN*. Recentemente foram também identificadas variantes germinais com penetrância moderada nos genes *BARD1*, *PALB2*, *RAD50*, *RAD51C* e *RAD51D*, entre outros.

Relativamente ao cancro colorretal hereditário (CCRH), estão identificadas diversas síndromes resultantes de alterações em vários genes, entre os quais se destacam *MLH1*, *MSH2*, *MSH6*, *PMS2* e *EPCAM* (síndrome de Lynch), *APC* (polipose adenomatosa familiar), *MUTYH* (polipose associada ao gene *MUTYH*) e *STK11* (síndrome de Peutz-Jeghers) (2).

O diagnóstico molecular de formas hereditárias de CM e CCR pelos métodos convencionais envolve, assim, a análise sequencial de múltiplos genes, tornando-se morosa e dispendiosa. A tecnologia de sequenciação de nova geração (NGS) veio possibilitar a pesquisa simultânea de alterações em múltiplos genes (painéis de genes), permitindo um diagnóstico molecu-

lar mais rápido, eficaz e com custo inferior ao da sequenciação pelo método de Sanger (3).

### \_Objetivo

Validação da NGS para identificação de variantes previamente detetadas pelo método de Sanger em genes de suscetibilidade para CM e CCR.

### \_Material e métodos

Foram analisados 64 indivíduos com suspeita clínica de CMH ou CCRH, aos quais tinha sido previamente solicitado consentimento informado. O DNA genómico foi extraído a partir dos leucócitos do sangue periférico utilizando o *kit Wizard DNA Extraction* (Promega) ou a plataforma robotizada *MagNA Pure LC* (Roche).

A análise molecular convencional incluiu a amplificação por PCR de todos os exões codificantes e sequências intrónicas flanqueantes dos genes *APC*, *MUTYH*, *MLH1*, *MSH2*, *STK11*, *BRCA1*, *BRCA2* e *TP53*, seguida de sequenciação bidirecional pelo método de Sanger utilizando o *kit BigDye Terminator Cycle Sequencing v1.1* e a plataforma de sequenciação ABI 3130xL (Applied Biosystems).

Para a análise por NGS, as bibliotecas de sequências-alvo foram preparadas a partir de DNA genómico pelo método de captura por hibridação, num protocolo que integrou o *TruSight Cancer Sequencing Panel* (que possibilita a análise de 94 genes que conferem predisposição para cancro) com o *kit TruSight Rapid Capture* (Illumina), e a sequenciação numa plataforma MiSeq com leituras *paired-end* de 150 bp. A análise bioinformática incluiu os softwares *MiSeq Reporter*, *VariantStudio*, *Isaac Enrichment* (Illumina) e *Integrative Genomics Viewer* (Broad Institute).

As variantes detetadas foram classificadas em função do seu significado clínico com recurso às bases de dados BIC, HGMD, LOVD, UMD e ClinVar.

### \_Resultados

Foram identificadas por NGS um total de 413 variantes de sequência nos genes analisados, correspondendo a 94 variantes únicas pontuais (substituição de um nucleótido ou pequenas inserções/deleções), das quais 27 são patogénicas (tabela 1).

Todas estas tinham sido previamente identificadas por sequenciação de Sanger, com a exceção de uma alteração no gene *STK11* (c.375-49G>A). Esta variante intrónica foi considerada um resultado falso positivo, resultando possivelmente de uma baixa cobertura por NGS (17 em 23 leituras na posição c.375-49 apresentaram a alteração G>A).

Adicionalmente, não foi possível detetar por NGS dois grandes rearranjos genómicos, patogénicos previamente identificados em duas amostras pela técnica de MLPA - *Multiplex Ligation-dependent Probe Amplification* (deleção do exão 5 do gene *MSH2*) e por PCR específica para a mutação fundadora portuguesa no gene *BRCA2* (c.156\_157insAlu).

Das 93 variantes únicas pontuais detetadas por NGS e confirmadas pelo método de Sanger, 80 corresponderam a alterações de um nucleótido e 13 a pequenas inserções/deleções (tabela 2).

Tabela 1: Classificação das 93 variantes únicas pontuais identificadas em genes de suscetibilidade para CM e CCR relativamente ao seu significado clínico.

Variantes	BRCA1	BRCA2	TP53	APC	MUTYH	MLH1	MSH2	STK11
SNP	12	16	2	4	2	3	5	1
BV	0	0	0	3	0	2	2	0
VUS	4	7	0	1	1	0	0	1
PV	2	4	1	5	4	6	5	0
Total*	137	160	5	43	22	21	21	3

SNP (*Single Nucleotide Polymorphism*): polimorfismo de nucleótido único; BV (*Benign Variant*): variante benigna; VUS (*Variant of Unknown/Uncertain Significance*): variante de significado clínico desconhecido/incerto; PV (*Pathogenic Variant*): variante patogénica.  
\* Somatório do número de vezes que as diferentes variantes únicas foram detetadas.

Tabela 2: Classificação das 93 variantes únicas pontuais identificadas em genes de suscetibilidade para CM e CCR relativamente ao tipo de alteração.

Variantes	BRCA1	BRCA2	TP53	APC	MUTYH	MLH1	MSH2	STK11
Pontuais (1 nt)	16	26	2	11	6	8	9	2
Indels	2	1	1	2	1	3	3	0
Total	18	27	3	13	7	11	12	2

nt - nucleótido; Indels - inserções/deleções.

O número de amostras analisadas por NGS implicou a realização de três ensaios independentes, cujos valores de cobertura vertical média por amostra variaram entre 35X e 362X com valores de qualidade (Q30) sempre superiores a 94%.

### \_Discussão

A análise por NGS demonstrou elevada sensibilidade, especificidade e repetibilidade analíticas para a deteção de variantes de sequência em 8 genes de suscetibilidade para CM e CCR, uma vez que permitiu identificar a totalidade das 412 variantes previamente detetadas pelo método de Sanger.

A limitação atual da deteção por NGS de grandes rearranjos genómicos é bem conhecida, sendo ultrapassada através da utilização de metodologias e técnicas complementares adequadas tais como, PCR específica para grandes rearranjos conhecidos ou MLPA para identificação de grandes deleções/inserções.

O aumento do número de genes analisados por NGS deverá resultar num incremento do número de variantes identificadas, sendo provavelmente desconhecido o significado clínico para muitas delas. Neste contexto, a interpretação dos resultados deverá ser cuidadosamente ponderada e, sempre que necessário, complementada com a realização de estudos *in silico* e com investigação adicional, que integre diversos estudos funcionais.

A aplicação da NGS no diagnóstico molecular permite uma resposta mais rápida e eficiente quando comparada com as metodologias convencionais. Contudo, todas as alterações identificadas por NGS e classificadas como variantes de significado clínico desconhecido, patogénicas e presumivelmente patogénicas, deverão ser confirmadas por sequenciação de Sanger. Os grandes rearranjos genómicos deverão continuar a ser pesquisados na rotina por metodologias complementares.

### \_Conclusões

Os ensaios de validação realizados comprovam a eficiência da NGS na deteção de alterações pontuais em 8 genes de elevada penetrância para CM e CCR. Neste contexto, a disponibi-

lização da sequenciação por NGS de painéis de genes de predisposição para estes dois tipos de cancro irá possibilitar um diagnóstico molecular mais abrangente, rápido e com custos reduzidos relativamente à sequenciação de Sanger. Complementarmente, tanto os utentes afetados como os seus familiares, com o devido aconselhamento genético e sob vigilância clínica especializada, beneficiarão destas novas tecnologias que têm um elevado impacto na saúde pública.

Dado que o painel de NGS em causa possibilita a análise simultânea até 94 genes associados a predisposição para cancro, evidencia-se o grande potencial de aplicação desta metodologia a outros tipos de cancros hereditários.

#### Referências bibliográficas:

- (1) Apostolou P, Fostira F. Hereditary breast cancer: the era of new susceptibility genes. *Biomed Res Int.* 2013;2013:747318. [www.ncbi.nlm.nih.gov/pmc/articles/PMC3618918/](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3618918/)
- (2) Lynch HT, Lynch JF, Lynch PM, et al. Hereditary colorectal cancer syndromes: molecular genetics, genetic counseling, diagnosis and management. *Fam Cancer.* 2008;7(1):27-39. Epub 2007 Nov 13.
- (3) LaDuca H, Stuenkel AJ, Dolinsky JS, et al. Utilization of multigene panels in hereditary cancer predisposition testing: analysis of more than 2,000 patients. *Genet Med.* 2014;16(11):830-7. [www.ncbi.nlm.nih.gov/pmc/articles/PMC4225457/](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4225457/)