

Estudo mutacional em tumor de Wilms por sequenciação de nova geração

Mutational study in Wilms' tumour using next-generation sequencing

Catarina Silva, Dina Carpinteiro, Luís Vieira

luis.vieira@insa.min-saude.pt

Unidade de Tecnologia e Inovação. Departamento de Genética Humana, INSA.

_Resumo

O Tumor de Wilms (TW) é o tumor renal mais comum da infância, com uma incidência de 1 em ~10000 crianças. Esta patologia é de etiologia genética complexa e diversificada. No entanto, cerca de um terço dos doentes apresenta mutações somáticas associadas aos genes *WT1*, *CTNNB1*, *TP53* e/ou *AMER1*. Assim, foi desenvolvido um painel de amplificações destes 4 genes para a identificação de mutações num grupo de doentes portugueses com TW, através de uma metodologia baseada na sequenciação de nova geração. As bibliotecas de DNA foram preparadas a partir de amostras de sangue periférico e tumor de 36 doentes com TW e sequenciadas no MiSeq. Foram identificadas alterações somáticas em 7 dos 36 (19,4%) doentes estudados. Conclui-se que a sequenciação de um painel de genes é um método rápido para a deteção de mutações somáticas quando desenhado com cuidado de forma a serem evitados problemas de perda de cobertura.

_Abstract

Wilms' tumor (WT) is the most common renal tumor of childhood affecting 1 in ~10,000 children. This tumor has a complex and diverse genetic etiology. About one third of the patients show somatic mutations in WT1, CTNNB1, TP53 and/or AMER1 genes. A new amplicon-based next-generation sequencing methodology was developed for the identification of mutations in those 4 genes in a cohort of Portuguese WT patients. DNA libraries were prepared from paired peripheral blood/tumor samples of 36 patients and sequenced on a MiSeq instrument. Somatic alterations were detected in 7 of 36 (19,4%) patients. It can be concluded that gene panel sequencing is a fast methodology for the detection of somatic mutations. However, care must be taken in the gene panel design in order to avoid low coverage targets.

_Introdução

O Tumor de Wilms (TW) é um tumor que tem origem nas células embrionárias envolvidas no desenvolvimento renal (restos nefrogénicos) e que permanecem após o nascimento (1). Sendo uma patologia rara, o TW é o tipo de cancro renal mais comum nas crianças afetando, nas populações ocidentais, 1 em ~10000 crianças, com diagnóstico por volta dos 3-4 anos de idade (2,3). A maioria dos casos tem apresentação unilateral e esporádica, havendo cerca de 5% de casos com manifestação bilateral (1,4). A maioria das mutações é de origem so-

mática estando sobretudo associada aos genes *WT1* (cr. 11), *CTNNB1* (cr. 3), *TP53* (cr. 17) e *AMER1* (cr. X), também conhecido como *WTX* (1,5). As mutações podem ocorrer singularmente ou em mais do que um gene no mesmo doente, de que é exemplo a associação das mutações do gene *CTNNB1* com as mutações em *WT1* (6). Neste trabalho, apresentam-se os resultados da aplicação da tecnologia de sequenciação de nova geração (NGS) a um painel de genes de TW. Este estudo foi aprovado pela Comissão de Ética para a Saúde do Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA).

_Objetivo

Este trabalho descreve a utilização da tecnologia de NGS aplicada à deteção de mutações somáticas em TW.

_Material e métodos

1) Desenho do painel de genes: Recorrendo ao *software Design Studio* (Illumina) foi criado um painel personalizado para os exões e zonas flanqueantes intrão/exão dos genes *WT1* (exões 2 a 11), *CTNNB1* (exão 4), *TP53* (exões 2 a 10) e *AMER1* (exões 1 e 2), num total de sequência cumulativa de 12306 bases. O painel foi composto por 83 amplicões de ~250 pb de tamanho, perfazendo um total de 20750 bases sequenciadas por amostra.

2) Preparação de bibliotecas de DNA: Neste estudo foram incluídas amostras de tumor e sangue periférico de 36 doentes Portugueses com TW, diagnosticados no Instituto Português de Oncologia de Francisco Gentil (Lisboa, n=30), Hospital Dona Estefânia (n=5) e Hospital de Santa Maria (n=1), entre 1988 e 2001. Após extração de DNA genómico, as bibliotecas foram preparadas usando o *kit TruSeq Custom Amplicon* (TSCA, Illumina), tendo-se introduzido um passo adicional para a avaliação qualitativa dos produtos indexados. As bibliotecas foram sequenciadas em duas corridas no MiSeq (Illumina) no modo de leitura *paired-end reads* de 250 pb.

3) Análise e tratamento de dados: A análise da qualidade global das corridas de sequenciação foi efetuada usando o *software Sequencing Analysis Viewer* (Illumina). As métricas de qualidade de cada amostra foram avaliadas recorrendo ao programa *FastQC*. A análise secundária das amostras foi realizada através do *MiSeq Reporter Software v.2.5.1*. As variantes somáticas foram filtradas tendo por base uma cobertura mínima superior a 30x, uma frequência do alelo mutado superior a 15% e a ausência na respetiva amostra de DNA de sangue periférico. As variantes foram visualizadas no *Integrative Genomics Viewer* (7) e a sua anotação foi efetuada usando bases de dados de variantes comuns. A previsão da patogenicidade foi efetuada utilizando o *MutationTaster2* (8) para as variantes constitucionais não anotadas.

4) Sequenciação de Sanger: As variantes somáticas identificadas por NGS foram comparadas com os resultados obtidos por sequenciação de Sanger para as respetivas amostras. Adicionalmente, no caso dos genes *WT1* (exões 7-10) e *CTNNB1* (exão 4), os resultados das amostras que não revelaram variantes somáticas por NGS foram comparados com os resultados previamente obtidos por sequenciação de Sanger.

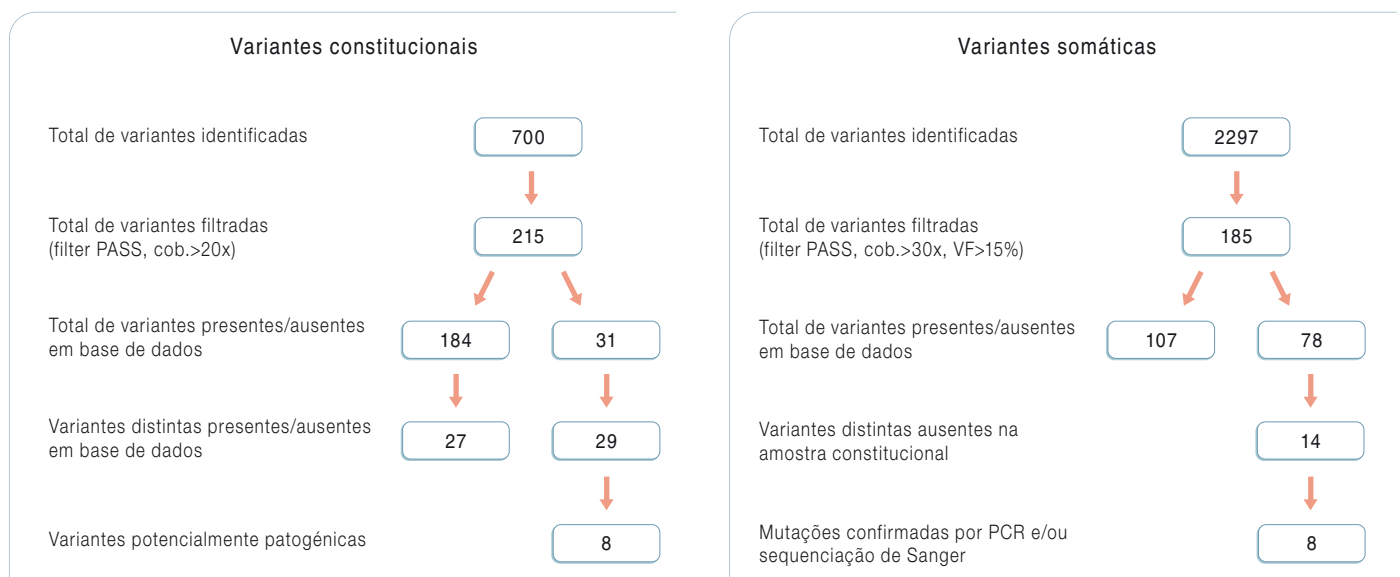
_Resultados

As duas corridas de sequenciação produziram, respetivamente, um total de 200 Mb (% \geq Q30 de 86,9%) e de 300 Mb (% \geq Q30 de 86,6%) de sequência.

A partir de um total de 700 variantes inicialmente identificadas nas amostras de sangue periférico, foram selecionadas 27 variantes distintas presentes em bases de dados (10 em *AMER1*, 9 em *TP53* e 8 em *WT1*), e 29 variantes distintas não descritas (18 em *AMER1*, 5 em *TP53*, 5 em *WT1* e 1 em *CTNNB1*), das quais 8 foram associadas a patogenicidade por darem origem a novos aminoácidos ou locais de *splicing* (figura 1). No entanto, o seu real efeito patogénico permanece desconhecido.

De um total de 2297 variantes identificadas nas amostras de tumor, foram filtradas 14 variantes distintas em 10 doentes. As variantes incluíram 9 alterações *missense*, 3 mutações *frameshift* e 1 deleção, distribuídas pelos genes *AMER1* (5 variantes), *CTNNB1* (4 variantes), *WT1* (3 variantes) e *TP53* (1 variante). Adicionalmente, foi detetada uma deleção completa do gene *AMER1* num doente do sexo masculino, refletida pela ausência de *reads* deste gene na amostra tumoral.

Figura 1: Representação esquemática dos passos principais do algoritmo utilizado para determinação das variantes constitucionais e somáticas.



Cob., cobertura; VF, variant frequency.

Tabela 1: ↓ Resumo das alterações somáticas encontradas em sete doentes com tumor de Wilms.

Doente	Genes em estudo		
	WT1 (cr. 11)	CTNNB1 (cr. 3)	AMER1 (cr. X)
WT6			Frameshift
WT16		Deleção_H36	
WT17		Missense_S45Y	
WT20		Missense_S45P	
WT24			Deleção génica
WT25	Frameshift		
WT36	Frameshift	Missense_S45F	

As 7 mutações confirmadas por sequenciação de Sanger e a deleção do gene *AMER1* encontram-se descritas na [tabela 1](#).

A análise do número médio de *reads* para cada um dos 83 amplicões sequenciados, no total de amostras, revelou uma elevada desproporção entre amplicões (intervalo: 0,01-874,97), sendo que 8 amplicões (9,6%), todos correspondentes ao gene *WT1*, apresentaram um número médio de *reads* inferior a 20. Em particular, a ausência total de *reads* do exão 7 do gene *WT1* em 2 doentes, não permitiu detetar a presença de mutações somáticas previamente identificadas pela metodologia de Sanger.

_Discussão

Neste estudo foi possível confirmar a existência de 8 alterações somáticas nos genes *WT1*, *CTNNB1* e *AMER1*, em 7 doentes com TW, utilizando a tecnologia de NGS. Com esta abordagem foi possível reduzir significativamente o tempo de pesquisa de mutações comparativamente à abordagem tradicional (sequenciação de Sanger), uma vez que todos os genes estudados foram sequenciados em paralelo.

No entanto, mais de 50% dos casos clínicos de TW não tinham até recentemente um diagnóstico molecular, o que tem levado à realização de estudos de sequenciação do exoma tumoral com o objetivo de identificar novos genes associados com esta pato-

tologia (9). Contudo, a análise dos dados obtidos ainda é um processo delicado e moroso, que requer uma devida *expertise* por forma a obterem-se os resultados desejados. Esta necessária experiência é ainda mais importante no caso das alterações somáticas, uma vez que a presença destas em níveis proporcionalmente baixos, relativamente aos respetivos alelos normais, pode conduzir à seleção de variantes não reais (falsos positivos).

No presente estudo, não foi possível evidenciar por sequenciação de Sanger 6 das 9 variantes *missense* selecionadas após aplicação do algoritmo de análise, cuja frequência do alelo mutante variou entre 18,5% e 24,5%. Estes resultados poderão dever-se à baixa cobertura (entre 45x e 54x) das respetivas posições nucleotídicas, que potenciaram a seleção de falsos positivos ou que sobrestimaram a frequência do alelo variante, levando a que não fosse detetado na sequenciação capilar (limite de deteção de ~15%).

Adicionalmente, a ausência/reduzido número médio de *reads* de alguns amplicões, observado no presente estudo, impediu a deteção de 2 mutações somáticas previamente identificadas no gene *WT1*. Assim, é importante assegurar que o desenho dos painéis de genes oferece uma cobertura homogénea dos vários amplicões e que, no caso de mutações somáticas, a profundidade de sequenciação seja mais elevada (>1000x), de forma a permitir uma distinção inequívoca entre variantes reais e com baixa representação (como nas situações de heterogeneidade tumoral) e variantes falsos positivos.

_Conclusões

A tecnologia de NGS associada ao uso de um painel de genes demonstrou ser uma ferramenta eficaz na deteção de variantes somáticas em doentes com TW, permitindo a identificação de vários tipos de mutações. A experiência agora adquirida permitirá alargar este painel a novos genes recentemente identificados em TW, possibilitando, num futuro próximo, estudar indivíduos com TW para os quais falta uma caracterização molecular.

Agradecimentos:

Os autores agradecem à Guida Boavida, aos médicos dos hospitais mencionados e às restantes pessoas envolvidas nos estudos de TW efetuados no INSA desde 1988, em particular à Sónia Matos e à Ana Jerónimo pela amplificação e sequenciação do exão 4 do gene *CTNNB1*, e dos exões 7 a 10 do gene *WT1*, respetivamente.

Referências bibliográficas:

- (1) Al-Hussain T, Ali A, Akhtar M. Wilms tumor: an update. *Adv Anat Pathol*. 2014;21(3):166-73. Review.
- (2) Huff V. Wilms' tumours: about tumour suppressor genes, an oncogene and a chameleon gene. *Nat Rev Cancer*. 2011;11(2):111-21.
www.ncbi.nlm.nih.gov/pmc/articles/PMC4332715/
- (3) Mahamdallie SS, Hanks S, Karlin KL, et al. Mutations in the transcriptional repressor REST predispose to Wilms tumor. *Nat Genet*. 2015;47(12):1471-4.
- (4) Hanks S, Perdeaux ER, Seal S, et al. Germline mutations in the PAF1 complex gene *CTR9* predispose to Wilms tumour. *Nat Commun*. 2014;5:4398.
www.ncbi.nlm.nih.gov/pmc/articles/PMC4143912/
- (5) Ruteshouser EC, Robinson SM, Huff V. Wilms tumor genetics: mutations in *WT1*, *WTX*, and *CTNNB1* account for only about one-third of tumors. *Genes Chromosomes Cancer*. 2008;47(6):461-70. www.ncbi.nlm.nih.gov/pmc/articles/PMC4332772/
- (6) Maiti S, Alam R, Amos CI, et al. Frequent association of beta-catenin and *WT1* mutations in Wilms tumors. *Cancer Res*. 2000;60(22):6288-92.
<http://cancerres.aacrjournals.org/content/60/22/6288.long>
- (7) Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*. 2013;14(2):178-92. Epub 2012 Apr 19.
www.ncbi.nlm.nih.gov/pmc/articles/PMC3603213/
- (8) Schwarz JM, Cooper DN, Schuelke M, et al. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods*. 2014;11(4):361-2.
- (9) Torrezan GT, Ferreira EN, Nakahata AM, et al. Recurrent somatic mutation in *DROSHA* induces microRNA profile changes in Wilms tumour. *Nat Commun*. 2014;5:4039.
www.ncbi.nlm.nih.gov/pmc/articles/PMC4062040/