

**VALIDAÇÃO DE ALTERAÇÕES SOMÁTICAS E *DE*
NOVO EM TUMORES DE WILMS ESPORÁDICOS
IDENTIFICADAS POR SEQUENCIAMENTO DE EXOMA**

MAYRA TOLEDO MENDES DE CASTRO

**Dissertação apresentada à Fundação Antônio
Prudente para obtenção do título de Mestre em
Ciências**

Área de Concentração: Oncologia

Orientadora: Dra. Dirce Maria Carraro

São Paulo

2013

FICHA CATALOGRÁFICA

Preparada pela Biblioteca da Fundação Antônio Prudente

Castro, Mayra Toledo Mendes

Validação de alterações somáticas e de novo em tumores de Wilms esporádicos identificadas por sequenciamento de exoma /

Mayra Toledo Mendes de Castro - São Paulo, 2013.

72p.

Dissertação (Mestrado)-Fundação Antônio Prudente.

Curso de Pós-Graduação em Ciências - Área de concentração:
Oncologia.

Orientadora: Dirce Maria Carraro

Descritores: 1. TUMOR DE WILMS. 2. EXOMA. 3. NEOPLASIAS RENAIIS. 4. CRIANÇA. 5. GENES DO TUMOR DE WILMS. 6. MUTAÇÃO.

"Fechei os olhos e pedi um favor ao vento:
Leve tudo o que for desnecessário.
Ando cansada de bagagens pesadas...
Daqui para frente apenas o que couber no bolso e no coração."

Cora Coralina

DEDICATÓRIA

À toda minha família,
de sangue e de coração,
que sempre estiveram comigo
durante este árduo trabalho, dedico.

AGRADECIMENTOS

À minha orientadora, Dra. Dirce Maria Carraro, pelo aprendizado, incentivo, oportunidade, paciência, apoio, dedicação e orientação durante o desenvolvimento deste trabalho.

À todos os integrantes do Laboratório de Genômica e Biologia Molecular, fundamentais para o desenvolvimento do trabalho: Camila, Andréa, Bruna, Elisa, Felipe, Cris, Márcia, Roberto, Carolina, Tatiana, Mabel, Dani, Vanina e Alex. Deixo ainda, um agradecimento especial à Giovana, pelo carinho, atenção, paciência, parceria e companheirismo durante o desenvolvimento deste projeto.

Às integrantes do Banco de Tumores, Eloisa, Louise, Bianca e Ana Paula.

Aos integrantes do setor de Biotecnologia e Bioinformática do A.C.Camargo Cancer Center, em especial ao Renato D. Puga, pela paciência, companhia, risadas, ensinamentos e colaboração para que eu chegasse aos resultados apresentados aqui.

Ao Dr. Pedro Alexandre Favoretto Galante e sua equipe do laboratório de Bioinformática do Hospital Sírio-Libanês pela assistência e colaboração na análise dos resultados obtidos.

Àqueles que estavam comigo, mesmo “de longe”, durante todos os dias, durante as aulas, almoços, passeios e viagens, com dias sempre regados à muita risada e boas conversas. Vocês sempre estiveram dispostos e pacientes a me ouvir e ajudar sempre que foi necessário: Bia Maia, Bia Nunes, Ju, Michelle, Érica, André, Júlio, Bruno, Mayara, Mari, Camila e Bruna. Sem vocês, com certeza, tudo seria mais difícil se não impossível.

Às minhas irmãs amadas que, independente de quanto tempo fiquemos sem nos encontrar, estão e estarão presentes sempre e pra sempre na minha vida! Um obrigada especial àquelas que estavam mais próximas e sempre me socorriam nos momentos mais aflitivos deste

percurso: Kolgat e K-xumba, vocês são extremamente essenciais para mim! Muito obrigada por tudo! A todas EssaKnas, meu muitíssimo obrigada pela nossa história! Uma vez EssaKna, sempre EssaKna!

Às minhas companheiras da vida, Thaís, Marcela e Natália, por todas as conversas, desabafos, risadas, passeios... por sempre se fazerem (e estarem!) presentes, independente da maneira. Por vocês serem extremamente essenciais, hoje e sempre.

Aos meus companheiros amarelos de todos os dias, todas as conversas, viagens, aprendizado, amor, companheirismo. Onde eu realmente aprendi a trabalhar em equipe e a valorizar nossos sonhos da maneira mais intensa possível. Pelos momentos inesquecíveis, pelas conversas sérias e palhaçadas, pelos conselhos, pelos pequenos gestos de enorme valor... por tudo que permanece comigo pra sempre. Fox, Fozzy, Marião, Pastel, Suli, Dod's, Pólo, Camis e Pilha. Vocês me ajudaram a recarregar as baterias e seguir em frente quando o cansaço e a vontade de desistir de tudo cresciam e tomavam conta de mim.

Ao Centro Internacional de Pesquisa e Ensino (CIPE) do A.C. Camargo Cancer Center por ter me oferecido toda estrutura para que o projeto fosse desenvolvido.

Aos responsáveis pela Pós-Graduação da Fundação Antônio Prudente-A.C. Camargo Cancer Center, Ana Maria Kurinari, Luciane Pitombeira, e Vanuza Barros.

À Suely pela formatação da tese e ao Jefferson pelo envio de papers.

À CAPES e à FAPESP, pelo suporte financeiro.

Por último, como a base de tudo, deixo um agradecimento especial aos meus pais, irmão, avós e padrinhos pelo contínuo apoio. Por fazerem com que eu tivesse forças para superar os obstáculos encontrados, por me fazer acreditar, lutar e conquistar, mesmo quando tudo parecia perdido, mesmo que para isso eu tivesse que atravessar caminhos tortuosos e cheios de espinhos. À Quilla, pelo colo, carinho e companhia nos momentos mais necessários! Vocês me fizeram acreditar, sempre! "Quem caminha sozinho, pode até chegar mais rápido. Mas aquele que vai acompanhado, com

certeza, chegará mais longe, e terá a indescritível alegria de compartilhar alegria... alegria esta, que a solidão nega a todos que a possuem” (Cecília Meireles). Sozinha, eu jamais conseguiria! Amo vocês!

RESUMO

Castro MTM. **Validação de alterações somáticas e *de novo* em tumores de Wilms esporádicos identificadas por sequenciamento de exoma.** São Paulo; 2013. [Dissertação de Mestrado-Fundação Antônio Prudente]

O tumor de Wilms (TW) é uma neoplasia renal infantil, de histologia trifásica, composta pelos componentes epitelial, estromal e blastematoso, que recapitula o rim fetal. Pode ser uni ou bilateral, desenvolvendo-se a partir de células do blastema metanéfrico que são incapacitadas de completar sua diferenciação durante a diferenciação do rim. A maioria dos TWs é esporádico, resultado de mutações somáticas, restritas ao tecido tumoral. Aproximadamente 30% dos TWs apresentam mutações nos genes *WT1*, *WTX* e *CTNNB1*. Porém, 70% dos casos de TW ainda continuam sem nenhuma alteração associada. Neste sentido, este projeto tem por objetivo a validação de alterações somáticas e *de novo* associadas ao desenvolvimento de TW, previamente identificadas por nosso grupo de pesquisa através do sequenciamento de exoma. O sequenciamento de exoma partiu de 4 amostras associadas a um caso de TW esporádico de histologia favorável: amostras tumoral e sanguínea do probando e amostras sanguíneas dos progenitores. Para este estudo, inicialmente, as variantes identificadas foram filtradas daquelas presentes no *dbSNP*. Para seleção das candidatas a variantes, foram estabelecidos os seguintes critérios: a base deveria ter sido representada pelo menos 10 *reads* e a base variante estar presente em pelo menos 20%, e não estar localizada em regiões de alta similaridade ou repetitivas do genoma. Para seleção dos candidatos a variantes *de novo* nenhum *read* da base variante poderia estar presente no sangue dos progenitores). Para as variantes que ocorreram dentro da região codificadora do gene, foram selecionadas as alterações *missenses* (alteração de aminoácido) classificadas como patogênicas em pelo menos um programa de predição (*SIFT* e/ou *Polyphen2*), variantes *nonsense* (sem

sentido - que codifica um códon de parada), deleções e inserções. Para as variantes que ocorrem fora da região codificadora foram selecionados as alterações em sítios de *splice*. Das 84 alterações *de novo* encontradas, selecionamos 11 variantes para a validação por sequenciamento capilar, duas que levavam a códon de parada prematuro na proteína, uma em sítio de *splice*, além de 8 variantes *missenses* (que levavam a alteração de amino ácido na proteína). Destas 11 variantes, 4 não foram confirmadas, 6 apresentaram a variante em pelo menos um dos progenitores (alteração herdada) e apenas 1 alteração foi confirmada como sendo uma variação *de novo*. A variante missense resultante da alteração (A>G) que leva a mudança de aminoácido (Asn>Asp), foi encontrada no gene *PARD6B* que codifica uma proteína citoplasmática envolvida na divisão celular e em processos de polarização celular, sendo importante para o desenvolvimento embrionário e remodelação tecidual. Com relação às alterações somáticas, 8 mutações *missense* foram selecionadas para validação. Após o sequenciamento capilar, 6 alterações não foram confirmadas na amostra tumoral. Para as duas alterações restantes, confirmadas em tecido tumoral, apenas uma foi confirmada como sendo uma alteração somática. A variação (G>A) com troca de aminoácido (Glu>Lys) foi encontrada no gene *DROSHA*, que apresenta um papel central na via de biogênese de microRNAs. Essa alteração ocorre exatamente no códon que abrange um dos sítios de ligação do cofator magnésio que é essencial para o funcionamento da proteína codificada por este gene. Esta alteração foi identificada em 9,3% das 97 amostras de TW avaliadas, mostrando tratar-se de uma mutação recorrente neste tumor.

SUMMARY

Castro MTM. **[Validation of somatic and *de novo* alterations in sporadic Wilms tumor identified by exome sequencing]**. São Paulo; 2013. [Dissertação de Mestrado-Fundação Antônio Prudente]

Wilms Tumor (WT) is a childhood kidney neoplasia that presents a triphasic histology, formed by epithelial, stromal and blastemal components that recapitulates the fetal kidney. It can be uni or bilateral, familial or sporadic, and is developed from metanephric blastemal cells that are unable to complete their differentiation during the kidney development. The majority of WTs is sporadic, resulting from somatic mutations that are restricted to the tumor tissue. About of 30% of cases presents mutations in *WT1*, *WTX* and *CTNNB1* genes. However, 70% of WTs remain without any associated alteration. In this sense, the aim of this project was the validation of somatic and *de novo* alterations possibly associated with WT onset that were previously identified by our group through exome sequencing. The exome sequencing was performed in 4 samples related to a favorable histology WT patient: patient's blood and tumor samples, in addition to mother's and father's blood samples. For this study, initially, the variants identified were filtered if present on *dbSNP*. For the selection of variant candidates, some criteria were established: the variant base should be represented by at least 10 reads and the variant should be presented in at least 20%, and not be located at highly similar or repetitive genome regions. For the selection of *de novo* variant any read of variant base could not be presented on proband's blood and absent at parents samples. For variations that occurred within the coding region of the gene, we selected the missense alterations (amino acid changes) classified as pathogenic in at least one of the prediction programs (SIFT and/or Polyphen2), nonsense alterations (that create a stop codon), deletions and insertions. For variations that occur outside of the coding region we selected changes in splice sites. From 84 *de novo* alterations

identified, we selected 11 variants for validation by capillary sequencing, two that lead to a stop codon in the protein, one at splice site and 8 missense that lead to amino acid change in the protein. From the 11 genes, 4 were not confirmed, 6 showed the variant at least in one of the parents (inherited alterations) and only 1 alteration was confirmed to be a *de novo* variation. The missense variant resulting from the alteration (A>G) that results in an amino acid change (Asn>Asp), was found in *PARD6B* gene, that encodes a cytoplasmatic protein involved in cellular division and polarization processes, been important for the embrionic development and tissue remodelation. Regarding somatic alterations, 8 were selected for validation. After capillary sequencing, 6 alterations were not confirmed in the tumor. From the remaining 2 alterations that were validated in the tumor sample, only one was confirmed as an acquired variation. The variation (G>A) with aminoacid change (Glu>Lys) was found yhe *DROSHA* gene, which has a central role in the microRNA processing pathway. This alteration occurs exactly at a codon that represents one of the binding sites for a magnesium cofactor, which is essential for the proper function of this protein. This mutation was found in 9.3% of 97 additional WT samples, suggesting it to be a recurrent somatic mutation in this tumor.

LISTA DE FIGURAS

Figura 1	Exemplo de análise do mapeamento dos dados de sequenciamento do exoma.....	30
Figura 2	Resultado da padronização de PCR.....	36
Figura 3	Exemplo da análise do sequenciamento capilar para o gene <i>MRPL37</i>	37
Figura 4	Exemplo da análise do sequenciamento capilar para o gene <i>ACTL7A</i>	38
Figura 5	Exemplo da análise do sequenciamento capilar para o gene <i>PKHD1L1</i>	39
Figura 6	Exemplo da análise do sequenciamento capilar para o gene <i>SLC7A7</i>	40
Figura 7	Exemplo da análise do sequenciamento capilar para o gene <i>DROSHA</i>	41
Figura 8	Análise do sequenciamento capilar para o gene <i>PARD6B</i>	48
Figura 9	Sequenciamento da alteração somática validada no gene <i>DROSHA</i> em amostras independentes.....	50
Figura 10	Modelo de biogênese do Mirna.....	60
Figura 11	Esquema do gene <i>DROSHA</i> , dividido entre seus principais domínios.....	61

LISTA DE TABELAS

Tabela 1	Características dos <i>primers</i> que flanqueiam todos os éxons do <i>WT1</i>	19
Tabela 2	Total de dados gerados.....	31
Tabela 3	Alterações <i>de novo</i> selecionadas para validação.....	34
Tabela 4	Alterações somáticas selecionadas para validação.....	34
Tabela 5	Sequência dos primers forward e reverse e tamanho dos amplicons das alterações de novo e somáticas selecionadas para validação.....	35
Tabela 6	Resultado da análise de sequenciamento capilar para validação das alterações <i>de novo</i> selecionadas.....	42
Tabela 7	Resultado da análise de sequenciamento capilar para validação das alterações somáticas selecionadas.....	43
Tabela 8	Sequências dos primers (forward e reverse) selecionados para validação de novas alterações de novo, seguido por seus amplicons.....	45
Tabela 9	Resultado da análise de sequenciamento capilar para validação de novas alterações <i>de novo</i> selecionadas.....	47
Tabela 10	Resultado da análise de sequenciamento do gene <i>DROSHA</i> para verificação da presença de outras alterações.....	51

LISTA DE SIGLAS E ABREVIATURAS

µg	Micrograma
µL	Microlitro
12T/21F	Denominação do probando
cDNA	Do inglês, DNA complementar ao RNA
CEP	Comitê de Ética em Pesquisa
dbSNP	Do inglês <i>Single Nucleotide Polymorphism Database</i>
DNA	Ácido Desoxirribonucleico
dNTP	Desoxirribonucleotídeo
dsRNA	RNA dupla fita
EDTA	Do inglês, <i>Ethylene Diamine TetrAcetic Acid</i>
ETOH	Etanol
EXO I	Exonuclease I
FISH	Do inglês, Fluorescence in situ hybridization
GBTTW	Grupo Cooperativo Brasileiro para Tratamento do Tumor de Wilms
ICE	Ifosfamida, Carboplatina e Etoposide
INCA	Instituto Nacional do Câncer
LOH	Perda de heterozigose
M	Molar
MgCl₂	Cloreto de Magnésio
MgSO₄	Sulfato de Magnésio
miRNAs	MicroRNAs
mM	Milimolar
mRNA	RNA mensageiro
NCBI	Do inglês, <i>National Center for Biotechnology Information</i>
ng	Nanograma
°C	Graus Celcius
pb	Pares de base
PCR	Do inglês, <i>Polimerase Chain Reaction</i>
pmol	Picomol

pre-miRNAs	Precursos de microRNAs
pri-miRNAs	MiRNAs primários
Refseq	<i>NCBI Reference Sequence Database</i>
RNA	Ácido Ribonucleico
rpm	Rotações por minuto
SAP	Do inglês, <i>Shrimp Alkaline Phosphatase</i>
SIOP	Do francês, <i>Société Internationale d'Oncologie Pédiatrique</i>
siRNA	RNA de interferência
SNV	Do inglês, <i>single nucleotide variation</i>
TCF/LEF	Família de fatores de transcrição
TCLE	Termo de consentimento livre esclarecido
TW	Tumor de Wilms
U	Unidade de enzima

ÍNDICE

1	INTRODUÇÃO	1
1.1	Tumor de Wilms: Aspectos epidemiológicos e clínicos	2
1.2	Aspectos morfológicos e moleculares de Tumores de Wilms.....	4
1.3	Alterações genéticas em Tumores de Wilms.....	5
1.3.1	Gene <i>WT1</i>	7
1.3.2	Gene <i>WTX</i>	8
1.3.3	Gene <i>CTNNB1</i>	8
1.3.4	Locus <i>WT2</i>	9
1.3.5	Gene <i>PLCG2</i>	10
1.3.6	Gene <i>APC</i>	11
1.3.7	Gene <i>HACE1</i>	12
1.4	Sequenciamento nova geração	13
2	JUSTIFICATIVA.....	14
3	OBJETIVOS	15
3.1	Objetivo Geral	15
3.2	Objetivos Específicos	15
4	METODOLOGIA	16
4.1	Comitê de ética.....	16
4.2	Seleção dos pacientes	16
4.3	Caracterização da amostra.....	17
4.4	Extração do DNA genômico	18
4.5	Etapas realizadas anteriormente a este projeto	18
4.5.1	Rastreamento de mutações em <i>WT1</i> por sequenciamento de Sanger .	18
4.5.2	Sequenciamento do exoma	20
4.5.3	Análise de Bioinformática	22
4.6	Seleção das alterações encontradas.....	24
4.6.1	Desenho dos <i>primers</i>	24
4.6.2	Reações de amplificação (PCRs).....	24
4.6.3	Tratamento com EXO/SAP e reações de sequenciamento capilar	25

4.6.4	Confirmação das alterações candidatas por sequenciamento capilar...	26
-------	---	----

5	RESULTADOS.....	28
5.1	Primeira seleção das alterações para validação	28
5.2	Novo sequenciamento e análise das bibliotecas de exoma	31
5.3	Segunda seleção das alterações para validação	32
5.4	Desenho dos <i>primers</i>	35
5.5	Validação por sequenciamento capilar.....	36
5.6	Análise e comparação das características das alterações selecionadas.....	43
5.7	Nova seleção de genes relacionados a alterações <i>de novo</i> e desenho de <i>primers</i>	45
5.8	Sequenciamento da alteração somática validada em amostras Independentes.....	49
6	DISCUSSÃO	52
7	CONCLUSÃO	63
8	REFERÊNCIAS BIBLIOGRÁFICAS	64

ANEXOS

Anexo 1 Termo de Consentimentos Livre Esclarecido (TCLE)
assinado pelos pacientes que cederam material biológico
utilizado neste projeto

Anexo 2 Resultados das alterações somáticas selecionadas para
validação, obtidos através das análises no Mozaik, CLC e
Sanger. Chr – Cromossomo

Anexo 3 Resultados das alterações de novo selecionadas
para validação, obtidos através das análises no Mozaik,
CLCBio e Sanger. Chr – Cromossomo

1 INTRODUÇÃO

A cada dia, a classificação dos tumores infantis ganha mais detalhamento, devido aos avanços nas técnicas laboratoriais baseadas na biologia molecular, imunologia e genética. Enquanto os tumores nos adultos estão, em geral, relacionados à exposição a vários fatores de risco como o tabagismo, estilo de vida, alimentação, ocupação e agentes carcinógenos específicos, a maior parte das causas dos tumores pediátricos ainda é desconhecida. Por isso, quando comparados aos tumores em adultos, os tumores pediátricos diferem destes em diversos aspectos, tais como características anatomopatológicas, local de acometimento e comportamento clínico (KOESTERS et al. 2003; REIS et al. 2007).

Dados de incidência mundial mostram que as neoplasias pediátricas representam entre 2% e 3% de todas as neoplasias, sendo que o percentual mediano dos tumores pediátricos observados nos Registros de Câncer de Base Populacional Brasileiros encontra-se próximo de 3%. No Brasil, as estimativas para o ano de 2012 serão válidas também para o ano de 2013 e apontam a ocorrência de aproximadamente 518.510 casos novos de câncer, incluindo os casos de pele não melanoma, reforçando a magnitude do problema do câncer no país. Sem os casos de câncer da pele não melanoma, estima-se um total de 385 mil casos novos. Como, para o Brasil, em 2012, à exceção dos tumores da pele não melanoma, estimam-se 384.340 casos novos de câncer, depreende-se, portanto, que ocorrerão

cerca de 11.530 casos novos de câncer em crianças e adolescentes até os 19 anos (Ministério da Saúde 2011).

Dentre os tumores sólidos pediátricos os mais comumente encontrados são os retinoblastomas, hepatoblastomas, neuroblastomas, sarcoma de Ewing, rabdomiossarcomas e os nefroblastomas ou tumores de Wilms (TW).

1.1 TUMOR DE WILMS: ASPECTOS EPIDEMIOLÓGICOS E CLÍNICOS

Os tumores renais representam de 5 a 10% dos tumores da infância (LITTLE 1999), sendo que aproximadamente 95% destes são TW, tendo uma maior incidência entre 2 e 3 anos de idade (BUCKLEY 2011), acometendo 1 em cada 10.000 crianças (RIVERA e HABER 2005; LITTLE 2005). A idade média do diagnóstico ocorre entre 43 e 48 meses, sendo entre 90 e 95% dos casos, apresentando-se de forma unilateral, e nos casos de bilateralidade, o diagnóstico é feito cerca e 1 ano antes dos tumores unilaterais (BRESLOW et al. 1988). Os casos familiares, aproximadamente 2% dos casos, geralmente estão associados com o aumento da frequência de tumores bilaterais e idade de acometimento menor ao diagnóstico (HUFF 2011).

Segundo resultados do primeiro estudo do Grupo Cooperativo Brasileiro para Tratamento do Tumor de Wilms (GCBTTW), o sexo feminino apresenta uma frequência um pouco maior de acometimento, enquanto as

crianças do sexo masculino tendem a desenvolver a doença mais precocemente (a idade média para os meninos é de 36 meses, enquanto que para as meninas é de 46 meses).

Uma combinação de cirurgia e quimioterapia é utilizada na maioria dos casos, salvo aqueles que encontram-se em estádios mais avançados, onde, além da quimioterapia, alia-se o tratamento radioterápico. Assim, mais de 85% dos pacientes diagnosticados com TW são curados e 15% apresentam recidiva da doença. Pacientes com recidiva apresentam uma taxa de cura de 50% com o uso de terapias modernas (GREEN et al. 1997).

Pesquisas clínicas e moleculares concentram-se, atualmente, na busca de marcadores que possam definir pacientes que se beneficiariam de uma terapia menos agressiva porém eficiente no tratamento de TW, resultando em uma excelente taxa de sobrevida (TAKAHASHI et al. 2002; WILLIAMS et al. 2004; ZIRN et al. 2006; HUANG et al. 2009; MASCHIETTO et al. 2011). Atualmente, o tratamento do TW está focado no desenvolvimento de terapias mais individuais, seja através de uma melhor adequação do tratamento aos grupos de risco, seja pela identificação de novos fatores de risco baseados em características tumorais próprias, levando em conta os diferentes tipos histológicos, possibilitando o desenvolvimento de novas ações terapêuticas, mais adequadas ao prognóstico individual, resultando na obtenção de um alto índice de cura (PICCOLI 2010).

1.2 ASPECTOS MORFOLÓGICOS E MOLECULARES DE TUMORES DE WILMS

O TW é uma neoplasia embrionária originada a partir de erros no processo de diferenciação celular em estágios precoces do desenvolvimento metanéfrico, na transição epitélio-mesênquima, durante a nefrogênese normal (LAI e TAY 1999; LI et al. 2002), envolvendo interações complexas entre as células epiteliais e mesenquimais.

Este tumor apresenta uma morfologia complexa, caracterizado por sua diversidade histológica e descrito como tumor embrionário trifásico, composto por células blastematosas, estromais (ou mesenquimais) e epiteliais, no rim fetal a nefrogênese (RIVERA e HABER 2005).

Essas evidências morfológicas dão suporte a premissa de que tais tumores originam-se através de uma célula precursora renal (MAITI et al. 2000; KUSAFUKA et al. 2002). Sugere-se, portanto, que o TW seja um tumor de células mesenquimais progenitoras, capazes de se diferenciar de forma aberrante, e que perderam a capacidade de responder aos sinais de controle de crescimento (HUFF 2011). Recentemente um trabalho do nosso grupo de pesquisa, demonstrou que é o componente blastematoso, e não o padrão de expressão do tumor como um todo, que recapitula os estágios iniciais da nefrogênese, refletindo com mais precisão a relação entre o bloqueio da diferenciação das células do blastema primitivo, que dá origem ao tumor e a nefrogênese (MASCHIETTO et al. 2008).

Apesar desta semelhança morfológica ter sido descrita há muito tempo (NICHOLSON 1931; WILLIS 1962), poucos trabalhos descreveram os genes que estão relacionados tanto com a nefrogênese quanto com a patogênese dos TW.

A regulação não apropriada de genes de vias de transdução de sinal é um fator importante para interrupção da diferenciação renal e, provavelmente, para o aparecimento do TW (SPARKS et al. 1998; MIYOSHI et al. 1998). Alguns estudos observaram uma alta concordância entre o padrão geral de expressão nos TW e de genes expressos nos primeiros estágios da diferenciação normal do rim, o que fortalece as evidências da conexão entre os dois processos: formação do tumor e nefrogênese (LI et al. 2002; LI et al. 2005; MASCHIETTO et al. 2008, 2011).

Devido a essa estreita relação, a caracterização molecular detalhada de ambos os mecanismos, tumor e nefrogênese, está levando cada vez mais a uma melhor compreensão dos eventos genéticos e moleculares envolvidos com o aparecimento do TW (LI et al. 2005, MASCHIETTO et al. 2008, 2011).

1.3 ALTERAÇÕES GENÉTICAS EM TUMORES DE WILMS

A elucidação dos mecanismos genéticos e biomoleculares ligados à tumorigênese de TW se faz extremamente necessária, uma vez que tais informações, atualmente, são escassas. O conhecimento mais detalhado relacionado às alterações moleculares dos TW pode resultar em uma melhor

compreensão da predisposição e do processo de formação tumoral, bem como sua progressão. Além disso, essa caracterização pode servir de base para a identificação de novos marcadores de prognóstico e alvos terapêuticos, favorecendo o desenvolvimento de tratamentos individualizados.

A maioria dos TW é esporádico, resultado de mutações somáticas restritas ao tecido tumoral, e apenas uma pequena porcentagem é consequência de mutações na linhagem germinativa, podendo ser herdada ou *de novo*. Menos de 10% dos TW estão associados a anomalias congênitas e síndromes, sendo 5 a 10% bilaterais ou multifocais e 1 a 2% associados à história familiar, relacionados à herança autossômica dominante com penetrância e expressividade variáveis (DOME e COPPES 2002).

Sabe-se hoje que diversos eventos genéticos estão envolvidos na gênese do TW esporádico, eventos estes que acabam por interferir na função de genes relacionados aos processos de proliferação e de diferenciação celular levando, conseqüentemente, ao desenvolvimento tumoral. Alguns genes envolvidos direta ou indiretamente nestes processos já foram identificados e parece haver uma participação preferencial de genes supressores de tumor com perda de função, à ativação de oncogenes durante a gênese do TW (GRUNDY et al. 2002).

1.3.1 Gene *WT1*

O primeiro gene identificado relacionado à tumorigênese do TW foi o *WT1* (NG_009272), gene que atua como supressor de tumor e exerce papel fundamental no desenvolvimento renal, intervindo na formação e maturação do néfron e na diferenciação e manutenção do glomérulo (MOORE et al. 1999; DAVIES et al. 2004). Este gene está localizado no cromossomo 11, na posição 11p13. Em casos diagnosticados de TW, esta região apresenta uma deleção de cerca de 345kb, onde se encontram os 10 éxons deste gene, alterações estas presentes em cerca de 20% dos casos diagnosticados com TW (HUFF 1998).

O *WT1* codifica um fator de transcrição que pode atuar como supressor tumoral ou como oncogene, dependendo do contexto em que a célula se encontra, o que sugere que ele desempenhe diferentes papéis no crescimento e desenvolvimento (RONG et al. 2006; SUGIYAMA 2010). A perda da função de *WT1* durante o desenvolvimento renal pode manter as células precursoras do néfron em um estado multipotente, favorecendo o surgimento do TW (RONG et al. 2006; SUGIYAMA 2010). A expressão de *WT1* é ampla nos componentes em desenvolvimento, e vai se tornando restrita, até que nos rins maduros, passa a ser expresso apenas em podócitos e glomérulos altamente diferenciados (HUFF 2011).

Mutações constitutivas em *WT1* já foram encontradas em famílias que apresentavam mais de um caso de TW (YUNIS e RAMSAY 1980), associadas ao aparecimento de más formações congênitas ou anomalias urogenitais no sexo masculino e/ou aniridia, doença rara que consiste na

falta congênita da íris ocular. Em famílias que não apresentavam anomalias congênitas, o *WT1* foi excluído como gene de susceptibilidade ao TW (DILLER et al. 1998).

1.3.2 Gene *WTX*

Outro gene que está alterado em TW é o *WTX* (NG_021345). Trata-se do primeiro gene supressor identificado no cromossomo X, localizado em Xq11.1 (RIVERA et al. 2007). Devido à sua localização em uma região sujeita à inativação por *imprinting*, homens e mulheres têm apenas uma cópia ativa do gene e, assim, um único ponto de mutação torna-se suficiente para inativar o gene (SCHEDL 2007). Tal característica faz com que o *WTX* se diferencie quanto ao processo de diferenciação, uma vez que na maioria dos genes supressores de tumor, este evento é bialélico (NUSSE 2007). Alterações em *WTX* do tipo “deleção” foram encontradas em aproximadamente 30% dos casos diagnosticados com TW (RIVERA et al. 2007).

Pesquisas recentes desenvolvidas por GUERTL e colaboradores (2010), propõe que a inativação do *WTX* seja um evento tardio na tumorigênese dos TWs. Para a maioria dos casos é observada a deleção completa do gene, e em alguns casos são encontradas mutações que levam a uma proteína truncada (RIVERA et al. 2007).

1.3.3 Gene *CTNNB1*

A proteína β -catenina é codificada pelo gene *CTNNB1* (NG_013302), e é um dos componentes da via de sinalização Wnt (POLAKIS 2000), exercendo papel de coativador dos membros da família TCF/LEF de fatores de transcrição durante o desenvolvimento embrionário, promovendo a ativação de genes envolvidos na proliferação, transformação celular e na adesão célula - célula (KORINEK et al. 1997).

Em TW, mutações no gene *CTNNB1* (3p22.1), estão localizadas no éxon 3, principalmente no códon 45 (*hot spot*). Foi demonstrado que mutações neste éxon ocorrem em 50% dos casos onde ocorre alteração de *WT1*, mas são raras se o *WT1* é selvagem (MAITI et al. 2000). No geral, deleções em β -catenina são encontradas em aproximadamente 15% dos casos de TW (KOESTERS et al. 1999). Além disso, em alguns TW que apresentam mutação em *WT1*, também são encontradas mutações nos éxons 7 e 8 do *CTNNB1*, mas ainda não se sabe quais efeitos essas alterações acarretam (RUTESHOUSER et al. 2008).

1.3.4 Locus *WT2*

Outra região cromossômica comumente alterada em TW esporádicos é a 11p15, denominada de *WT2*. Esta região foi identificada a partir da observação de que parte dos TW com *LOH* (perda de heterozigose) em 11p apresentavam envolvimento da região 11p15 e não 11p13 (região do gene *WT1*). Além disso, alterações constitutivas nesta região estão relacionadas com a síndrome de Beckwith-Wiedemann, que acarreta uma predisposição

ao TW. Foram identificados, nesta região, três genes com expressão parental preferencial, relacionados aos TW: os genes *IGF2*, *H19* e *CDKN1C* (DOME e COPPES 2002; HARUTA et al. 2008; HO et al. 2008).

A análise de famílias com vários indivíduos afetados pelo TW não relacionados a alterações em 11p13 e 11p15 levaram à identificação de outras duas regiões 17q e 19q, denominadas de FWT1 e FWT2 (*Familial Wilms Tumor 1 e 2*), respectivamente (DOME e COPPES 2002; GRUNDY et al. 2002; STRONG 2003).

1.3.5 Gene *PLCG2*

O gene *PLCG2* (NG_032019) está localizado em 16q24.1 e é responsável pela codificação de uma proteína que, quando analisada em um grupo de amostras de TW, apresentou expressão diminuída ou ausente (MASCHIETTO et al. 2008).

Diversos estudos relacionaram a *LOH* de 16q com o TW (MAW et al. 1992; SKOTNICKA-KLONOWICZ et al. 2000; MESSAHEL et al. 2009), evento observado em aproximadamente 20% dos casos e normalmente associada a um pior prognóstico. Em um estudo prévio realizado por nosso laboratório, verificou-se a diminuição ou ausência da expressão de *PLCG2* em TW, recapitulada nos estágios iniciais da nefrogênese indicando um possível envolvimento desse gene na formação desses tumores (MASCHIETTO et al. 2008).

Além disso, recentemente, um estudo desenvolvido por nosso grupo resultou na descoberta de uma variação *missense PLCG2*, que não foi

detectada em indivíduos controles ou banco de dados, podendo se tratar de alterações relacionadas ao desenvolvimento de TW (BARROS 2012).

1.3.6 Gene *APC*

O gene *APC* (NG_008481) está localizado em 5q21-q22, pertence à via de sinalização Wnt e, segundo estudos desenvolvidos por nosso grupo de pesquisa, possivelmente está relacionados à formação de TW, uma vez que a expressão deste gene encontra-se alterada em amostras tumorais (MASCHIETTO et al. 2008). Alterações no gene *APC* afetam a ativação da β -catenina, membros da via Wnt, e altera o processo de proliferação celular, processo este que tem papel chave no desenvolvimento de um amplo espectro de doenças malignas (MAITI et al. 2000).

O gene *APC* é um gene exaustivamente estudado e avaliado em diversos tipos de tumores. Mutações somáticas neste gene tem sido identificadas em casos de tumores colorretais, tumores de estômago, pancreas, tireóide, ovário e outros. A maioria das alterações resulta em uma proteína truncada de APC, quer por uma mutação *nonsense* (30%) ou por uma mutação *frameshift* (68%). Esta característica permite concluir que a maioria das alterações neste gene é verdadeiramente deletéria para a função APC (LAURENT-PUIG et al. 1998). Recentemente, um estudo desenvolvido por nosso grupo resultou na descoberta de duas alterações *missense* no gene *APC* em duas amostras de TW, que não foram identificadas em indivíduos controles ou banco de dados, podendo se tratar de alterações relacionadas ao desenvolvimento de TW (BARROS 2012).

1.3.7 Gene *HACE1*

O gene *HACE1*, localizado no braço longo do cromossomo 6, (NM_020771.3), foi identificado como envolvido em um rearranjo cromossômico em t(6;15)(q21;q21) de pacientes com TW. Este rearranjo está relacionado com o aumento de expressão do mRNA de *HACE1* quando comparado à expressão no tecido normal, mesmo quando a translocação não acontece diretamente no locus de *HACE1*. Através de análises utilizando o método de *FISH* (Fluorescence in situ hybridization) confirmou-se a localização do gene em humanos e que, em indivíduos com TW, esta região está realmente alterada (ZHANG et al. 2007).

Perdas genômicas envolvendo o cromossomo 6q21 vem sendo descritas e relacionadas aos mais variados tipos de tumor, como próstata, mama, e ovário, bem como leucemias e linfomas, sugerindo que esta região abrange um ou mais genes supressores tumorais (GNARRA et al. 1994; MARTIN et al. 2003).

A nível molecular, *HACE1* aparece como regulador da progressão do ciclo celular durante stress celular decorrente da interferência da degradação de ciclina D1. Em estudo recente, o grupo de ANGLESIO e colaboradores (2004) demonstrou que *HACE1* atua em conjunto com *p53* na prevenção quanto à formação tumoral e a mutação em ambos os genes aumenta drasticamente a incidência de câncer, bem como a variedade dos tumores quando comparados a camundongos com *p53* mutado em apenas um alelo.

1.4 SEQUENCIAMENTO NOVA GERAÇÃO

O desenvolvimento de novas técnicas de sequenciamento de DNA permitiu uma enorme evolução no campo da biologia molecular e da genética, sendo que atualmente é possível o sequenciamento completo do DNA genômico, com custo e tempo muito reduzidos. Em paralelo aos avanços das técnicas de sequenciamento, as estratégias de captura de regiões genômicas de interesse foram aprimoradas, para que se pudesse direcionar o sequenciamento às porções genômicas codificantes, o exoma, abrangendo grande parte das variações funcionais presentes em uma célula ou tecido (SARAH et al. 2010).

Assim, o sequenciamento do repertório gênico completo pode ser utilizado como uma ferramenta poderosa para que se possa definir alterações gênicas germinativas que predispõe a um determinado tumor, e as alterações somáticas presentes no tecido tumoral, dando pistas dos fatores genéticos associados à sua formação e desenvolvimento.

2 JUSTIFICATIVA

Um aspecto importante envolvido com o aparecimento do tumor é o da definição do painel de mutações associados ao TW, uma vez que apenas 30% dos casos de TW estão relacionados a alterações em genes já descritos (RUTESHOUSER et al. 2008), demonstrando que para a grande maioria dos casos de TW (70%) mutações gênicas ainda não foram devidamente associadas.

Assim, devido à necessidade de se identificar alterações moleculares responsáveis pelo surgimento e progressão tumoral, este projeto propõe a validação de alterações somáticas e *de novo* em pacientes com TW esporádico, incluindo DNA do tumor e do sangue do paciente, bem como a análise do DNA do sangue de seus respectivos progenitores.

3 OBJETIVOS

3.1 OBJETIVO GERAL

Este projeto tem por objetivo a identificação de alterações pontuais somáticas e *de novo* em amostras de tumores e de sangue de pacientes de Tumor de Wilms (TW) e seus progenitores.

3.2 OBJETIVOS ESPECÍFICOS

- 1) Confirmação da presença das alterações *de novo* e somáticas através de sequenciamento capilar identificadas previamente por sequenciamento do repertório gênico por sequenciamento do exoma.
- 2) Avaliação da frequência de variações somáticas identificadas em amostras de TW independentes.

4 METODOLOGIA

4.1 COMITÊ DE ÉTICA

O presente trabalho foi avaliado e aprovado pelo Comitê de Ética em Pesquisa (CEP) do A.C.Camargo Cancer Center-Fundação Antônio Prudente, sob registro 1396/10 B.

4.2 SELEÇÃO DOS PACIENTES

Foram selecionados quatro pacientes anteriormente participantes de projetos do mesmo grupo de pesquisa, uma vez que as amostras sanguíneas destes pacientes e das respectivas mães já haviam sido coletadas.

Em seguida, foi realizada uma avaliação clínica que culminou na exclusão de pacientes com tumores bilaterais ou com suspeita de síndromes associadas, incluindo hemihiperplasia, resultando em 3 amostras próprias para o estudo. Ao se avaliar previamente o material armazenado no Banco Tumores do Departamento de Patologia do A.C.Camargo Cancer Center referente aos três pacientes pré-selecionados, um deles era composto apenas por necrose. Dessa forma 2 pacientes foram selecionados para este estudo, porém, por conta de problemas encontrados na obtenção de novas amostras (sangue paterno), prosseguiu-se com apenas 1 paciente, do qual

há 4 amostras de DNA relacionadas (sangue e tumor do paciente, sangue da mãe e sangue do pai).

As amostras utilizadas foram obtidas através de um projeto anterior, desenvolvido no A.C.Camargo Cancer Center, coordenado pela Prof^a. Dra. Beatriz de Camargo. Este referido projeto está registrado no Programa de Integração de Gerenciamento de Amostras (AC Projects), do A.C.Camargo Cancer Center, sob registro DUB533 e aprovado no comitê de ética (CEP) sob registro 1304-09. Consta entre os documentos referentes a este projeto, o termo de consentimento livre esclarecido (TCLE), devidamente assinado pelos pacientes, permitindo que, caso houvesse excedente das amostras cedidas, as mesmas poderiam ser utilizadas em outros projetos da instituição (Anexo 1).

4.3 CARACTERIZAÇÃO DA AMOSTRA

As amostras utilizadas no presente estudo estão relacionadas ao paciente TW 12T/21F, atualmente com 7 anos. Tal paciente foi diagnosticado com TW aos 2 anos de idade, localizado no rim direito, em estadio I e de risco intermediário. O paciente foi então submetido ao protocolo *SIOP 2001* (quimioterapia pré-operatória) seguida por nefrectomia direita total.

Após decorridos seis meses da cirurgia, foram identificadas massas retroperitoneiais (1 pré-caval e outra para-vertebral), adjacentes ao leito renal, deslocando-se súpero-lateralmente à veia cava inferior e

anteriormente à cabeça pancreática, mas sem comprometimento linfonodal. Assim, o paciente foi submetido a novas sessões de quimioterapia (6 ciclos de ICE - Ifosfamida, Carboplatina e Etoposide) e radioterapia (14 sessões), resultando em resposta completa ao tratamento.

Foram coletadas, além das amostras correspondentes ao sangue e tumor do paciente, amostras sanguíneas de seus pais para que se realizasse a extração do DNA das mesmas.

4.4 EXTRAÇÃO DO DNA GENÔMICO

Para a extração do DNA genômico das amostras de sangue, foi utilizado o kit Puregene®-DNA purification gene (Qiagen - Duesseldorf, Alemanha) conforme recomendações do fabricante. O DNA dos tumores foi extraído pela metodologia de fenol/clorofórmio, conforme protocolo seguido pelo Banco de Macromoléculas do A.C.Camargo Cancer Center.

4.5 ETAPAS REALIZADAS ANTERIORMENTE A ESTE PROJETO

4.5.1 Rastreamento de Mutações em *WT1* por Sequenciamento de Sanger

Com o objetivo de buscar novas alterações nas amostras selecionadas para este projeto, excluindo aquelas que apresentassem mutações já bem descritas, foi realizado o sequenciamento completo do gene *WT1* pela Dra. Mariana Maschietto. A sequência genômica do *WT1*

(NG_009272.1) foi obtida do *National Center for Biotechnology Information* (NCBI – <http://www.ncbi.nlm.nih.gov>) para desenho dos oligonucleotídeos flanqueando os 10 éxons correspondentes às quatro variantes de *splicing* alternativo principais do gene em questão (Tabela 1), procedimento este, realizado por nosso grupo de pesquisa, anteriormente ao início deste projeto.

A amplificação de todos os fragmentos foi feita usando o kit *Platinum Taq DNA Polimerase* (Invitrogen – Califórnia, Estados Unidos) para uma reação com volume final de 25µl, seguindo o protocolo descrito pelo fabricante, realizando apenas pequenas variações na temperatura de anelamento e na concentração de magnésio, visando à obtenção de melhores resultados. Os produtos amplificados foram visualizados em gel de agarose 1% corado com *Sybr Safe* (Invitrogen – Califórnia, Estados Unidos).

Tabela 1 - Características dos *primers* que flanqueiam todos os éxons do *WT1*

<i>Primers</i>	Exons flanqueados	Tamanho esperado	<i>Primers</i>	Exons flanqueados	Tamanho esperado
WT1 1	1	467bp	WT1 7	7	292bp
WT1 2	1	608bp	WT1 8	8	309bp
WT1 3	2 e 3	767bp	WT1 9	9	354bp
WT1 4	4	292bp	WT1 10	10	538bp
WT1 5	5	253bp	WT1 11	10	527bp
WT1 6	6	286bp	WT1 12	10	622bp

As reações de sequenciamento foram realizadas conforme descrito no item 4.6.3. e as sequências geradas foram alinhadas na sequência referência e analisadas no programa *CLCBio Main Workbench*.

4.5.2 Sequenciamento do Exoma

A etapa de sequenciamento do exoma foi realizada pela Dra. Elisa Napolitano e Ferreira. O enriquecimento da região correspondente ao exoma foi utilizado o *SureSelect Human All Exon 50Mb Kit* (Agilent) em conjunto com o *SOLiD™ Fragment Library Barcoding Kit* (Applied Biosystems - Califórnia, Estados Unidos) para a construção das bibliotecas, seguindo as recomendações dos fabricantes.

Simplificadamente, o DNA genômico foi fragmentado por sonicação, ligado a adaptadores específicos, hibridizado com moléculas de RNA contendo biotina, correspondentes a região do exoma (50Mb) e capturados com partículas magnéticas de estreptavidina. Cada amostra foi marcada por um barcode único adicionado a adaptadores específicos, permitindo que se realizasse o sequenciamento em paralelo de amostras distintas.

Por fim, as bibliotecas resultantes das diferentes amostras foram misturadas em quantidades equimolares, amplificadas através de PCR em emulsão e sequenciadas utilizando o *SOLiD™ 4 System*. A tecnologia de sequenciamento do *SOLiD™* utiliza a amplificação em emulsão do DNA ligado a esferas. Após a amplificação, as esferas com fragmentos de DNA modificados na extremidade 3' são depositadas em uma *flowchip*.

O sequenciamento ocorre por múltiplas etapas de hibridização e ligação. Iniciadores são hibridados aos adaptadores no DNA molde e octâmeros fluorescentes hibridizam com a região adjacente do iniciador. Através da ação de uma ligase, ocorre a ligação destes fragmentos e então é feita a detecção da fluorescência presente no octâmero. A marcação fluorescente do octâmero depende do par de bases inicial, permitindo a interrogação das duas primeiras bases nas quais ele se hibrida. Após a hibridização e detecção, as três bases finais do octâmero são clivadas e ocorrem múltiplos ciclos de ligação, detecção e clivagem. Posteriormente, o produto de extensão é removido e um iniciador complementar à região n-1 do DNA molde inicia um novo ciclo de ligações. Cinco ciclos de iniciadores são completados para cada fragmento de DNA molde. A principal vantagem desta metodologia é que cada base é determinada em duas ligações independentes, resultando em uma leitura precisa (99,9%) que requer pouca cobertura para atingir um valor confiável de chamada de bases.

Foi realizado o sequenciamento do exoma completo do caso selecionado no estudo (paciente TW 12T/21F), incluindo DNA do sangue e do tumor do paciente e do sangue de seus progenitores partindo de 3µg de DNA.

Os resultados obtidos ao fim deste sequenciamento, não apresentaram cobertura suficiente para que a análise proposta fosse devidamente realizada. Sendo assim, foi realizado um segundo sequenciamento do exoma, utilizando o equipamento *SOLiD™ 5500XL* (Life

Technologies) para que se aumentasse a cobertura das amostras analisadas.

O protocolo utilizado para a realização deste segundo sequenciamento de exoma, segue as mesmas etapas daquele utilizado no sequenciamento utilizando o *SOLiD™ 4 System*, com diferenças significativas apenas no que diz respeito ao software utilizado para análise do sequenciamento.

4.5.3 Análise de Bioinformática

As análises bioinformáticas do sequenciamento do exoma foram realizadas em uma primeira fase através de colaboração com o grupo do Dr. John Weinstein do Departamento de Bioinformática e Biologia Computacional do M.D. Anderson Cancer Center (Universidade do Texas, Houston, EUA) e, em um segundo momento, com o grupo do Dr. Pedro Galante do Centro de Pesquisa do Hospital Sírio Libanez.

O Dr. Xiaoping Su do grupo do Dr. John Weinstein (M.D.Anderson Cancer Center) realizou a análise bioinformática utilizando a ferramenta *Mozaik Assembler* para mapeamento no genoma humano. Para a identificação das alterações pontuais e *indels*, utilizou-se a ferramenta *Gigabayes*.

A análise bioinformática realizada pelo Dr. Pedro Galante está descrita a seguir, de forma resumida: As leituras geradas foram alinhadas contra o genoma humano de referência (GRCh37/hg19) através do alinhador *Mapreads* do Bioscope (<http://www.lifescopelcloud.com/>). As sequências que

não apresentaram resultados positivos neste primeiro alinhamento foram alinhadas novamente, porém, utilizando-se da ferramenta *NovoalignCS* (<http://novocraft.com/>) que permite intervalos (*gaps*) nos alinhamentos, importantes por serem, em potencial, inserções ou deleções (indels).

As variações de nucleotídeo único (SNVs), assim como as inserções e deleções (indels) foram identificadas utilizando um conjunto de programas públicos em combinação com um conjunto de programas (*pipelines*) desenvolvidos pela equipe local (Laboratório de Bioinformática do Hospital Sírio-Libanês).

Partindo dos alinhamentos, as SNVs e indels foram identificadas utilizando o *SAMtools* (<http://samtools.sourceforge.net/>) e o *mpileup* (<http://samtools.sourceforge.net/mpileup.shtml>), que permitem a manipulação de alinhamentos em diferentes formatos, sendo capazes de converter, ordenar e mesclar alinhamentos, remover duplicatas de PCR, gerar informações por posições, entre outras funções.

Os resultados gerados foram processados por *pipelines* que identificaram o tipo das substituições (SNV ou Indel) bem como sua localização genômica (região codificadora, UTR, intrônica, em sítios de *splice* ou intergênica). Para as substituições em regiões codificadoras, o gene afetado foi devidamente identificado, baseado no conjunto de genes conhecidos do *RefSeq* (<http://www.ncbi.nlm.nih.gov/RefSeq/>, versão 55) e o tipo da substituição (sinônima, não sinônima *missense* e não sinônima *non-sense*).

4.6 SELEÇÃO DAS ALTERAÇÕES ENCONTRADAS

As etapas seguintes foram realizadas no escopo dessa dissertação com a participação fundamental da aluna de doutorado Giovana Torrezan.

4.6.1 Desenho Dos *Primers*

Baseando-se nas alterações encontradas, foram desenhados *primers* a partir da sequência do DNA genômico disponibilizado pelo *National Center for Biotechnology Information* (<http://www.ncbi.nlm.nih.gov>), objetivando a amplificação dos fragmentos de interesse, que variam de 147 a 492 pares de bases (pb).

Após o desenho dos *primers*, os mesmos foram avaliados pelo programa *Oligotech v1.0*, para exclusão daqueles que apresentassem, possivelmente, temperatura de anelamento maior do que 2°C de diferença entre ao pares, estruturas secundárias ou homodímeros com temperatura de até 10°C.

4.6.2 Reações De Amplificação (PCRs)

Inicialmente, as PCRs foram realizadas contendo 1 unidade da enzima *Platinum® Taq Polimerase High Fidelity* (Invitrogen – Califórnia, Estados Unidos) 1X tampão 2mM dNTP 2mM MgSO₄, 4pmol de *primer forward*, 4pmol de *primer reverse* e 12,5ng de DNA genômico. Porém, após alguns testes, verificou-se uma maior eficiência na utilização da enzima *KAPA2G Fast HotStart ReadyMix (2X)*, mix este composto pelo tampão da

referida enzima (1x), dNTP (0,2mM cada dNTP 1x), $MgCl_2$ (1,5mM a 1x) e a enzima *KAPA2G Fast HotStart DNA Polymerase* (0,5U). Para completar todos os itens necessários à reação, foram adicionados os *primers* e 1,5 μ l (18,75ng) de DNA. Para as amostras controle (positivo e negativo), foram utilizadas uma alíquota de DNA de indivíduo controle e uma alíquota de água Milli-Q respectivamente e todas as reações foram realizadas com um volume final de 20 μ l.

A reação foi incubada a 94°C por 2min, seguido de 40 ciclos de 94°C por 15s, 60°C por 18s e 72°C por 1s, seguido por 72°C por 1min, mantidas, ao fim, a 4°C. Pequenas modificações nestas condições foram realizadas para que se obtivesse otimização da reação para cada par de *primer*.

Em seguida, os produtos de PCR foram avaliados em gel de agarose 1% para checagem do tamanho das sequências obtidas.

4.6.3 Tratamento com EXO/SAP e Reações de Sequenciamento Capilar

Para as amplificações que obtiveram banda única quando visualizadas em gel de agarose, os produtos de PCR foram tratados com 1 unidade de SAP (*Shrimp Alkaline Phosphatase*) e 1 unidade de EXOI (Exonuclease I), juntamente com o tampão de eluição (1X), numa reação de volume total de 10 μ l. Os produtos foram então, tratados e incubados a 37°C por 30 minutos e 80°C por 15 minutos, mantidos, em seguida, a 4°C.

Após este processo, as amostras foram sequenciadas pelo método Sanger, com o *BigDye® Terminator v3.1 Cycle Sequencing Kit* (Life Technologies – Califórnia, Estados Unidos), 5pmol do oligonucleotídeo

específico para cada fragmento e água para que se obtivesse um volume final de 10µl da reação. Para a precipitação do DNA, foram adicionados a reação 1µl de acetato de sódio 3M pH 5.2 1µl de EDTA 125mM pH=8 e 30µL de ETOH 100% na temperatura ambiente, permanecendo por 15 minutos no escuro à temperatura ambiente. A reação foi centrifugada a 4.000rpm por 30 minutos à 4°C, seguido de um *spin* invertido. A reação foi então lavada com 30µl de ETOH 70% gelado, centrifugada novamente nas mesmas condições a 4.000rpm por 25 minutos à 4°C, seguida de *spin* invertido.

Ao fim desta etapa, foram adicionados 13µl de formamida e a reação foi desnaturada a 95°C por 4 minutos, sendo, em seguida, mantida no gelo até ser levada ao sequenciador ABI Prism 3130 XL (Applied Biosystems - Califórnia, Estados Unidos).

4.6.4 Confirmação das Alterações Candidatas por Sequenciamento Capilar

As alterações selecionadas, amplificadas através de PCR e sequenciadas por sequenciamento capilar, foram então avaliadas manualmente com o auxílio do programa *CLC Genomics 5*.

Para esta análise, foram pareadas a sequência referência do gene em questão, juntamente com as sequências relacionadas àquele gene e os produtos resultantes do processo de sequenciamento das amostras. Em seguida, através dos dados fornecidos pela análise bioinformática, a alteração foi localizada e analisada.

Para uma das alterações validadas, foram utilizadas, 69 amostras de DNA tumoral e 27 amostras de cDNA tumoral de probandos diagnosticados com TW, totalizando 96 amostras que estavam devidamente armazenadas no Banco de Macromoléculas do A.C.Camargo Cancer Center obtidas em colaboração com o Childrens Oncology Group (COG), referentes a outros casos de TWs de estadios III ou IV, apresentando apenas o componente blastematoso. Para estas amostras, realizou-se o sequenciamento dos quatro éxons do domínio proteico. Além disso, como controle negativo, foram utilizadas 272 amostras referentes a indivíduos saudáveis (sem câncer), para que se pudesse descartar a possibilidade de a mutação encontrada representar um polimorfismo.

5 RESULTADOS

5.1 PRIMEIRA SELEÇÃO DAS ALTERAÇÕES PARA VALIDAÇÃO

Os dados oriundos do sequenciamento do exoma das amostras (DNA do tumor e sangue do paciente, bem como do sangue de seus progenitores), foram analisados para busca de alterações somáticas (presentes apenas no tumor do paciente) e alterações *de novo* (presentes no sangue e tumor do paciente e ausente nos pais).

Inicialmente, utilizamos os resultados gerados pela análise bioinformática realizada em colaboração com o grupo do M.D. Anderson Cancer Center (Universidade do Texas, Houston, EUA). Os resultados obtidos apresentaram baixa cobertura, dificultando a chamada das variantes. Mesmo assim, algumas variantes foram selecionadas para dar prosseguimento às análises.

Assim, foram selecionadas para validação por sequenciamento capilar 18 alterações somáticas e 18 alterações *de novo*, sendo que nenhuma delas foi confirmada, sugerindo a ocorrência de problemas na análise de bioinformática na etapa de mapeamento contra o genoma humano utilizando a ferramenta *Mozaik Assembler*, ou na etapa de identificação de alterações pontuais e *indels* pela ferramenta *Gigabases*.

Optou-se, então, pela inspeção manual do mapeamento utilizando a ferramenta *CLC Genomics Workbench* (CLCbio), onde verificou-se que

apenas duas das 36 alterações reportadas pela análise de bioinformática apresentavam os resultados descritos. Destas, em 18 genes foram identificadas alterações do tipo “troca de uma base”, distintas daquelas previamente descritas, presentes a uma distância que varia de 1 a 3 nucleotídeos em relação à coordenada genômica reportada pela análise de bioinformática (Figura 1). Estas alterações foram avaliadas pelo sequenciamento capilar e 10 delas (71%) foram confirmadas (Anexos 2 e 3).

De forma geral, os resultados obtidos através do sequenciamento capilar e o mapeamento realizado com a ferramenta *CLC Genomics Workbench*, foram concordantes para 65% das coordenadas genômicas avaliadas (Anexos 2 e 3), sugerindo que a ausência de validação dos dados iniciais foi decorrente de artefatos gerados pelo alinhamento e pelos critérios utilizados na análise de bioinformática anterior, e não por artefatos gerados pelo sequenciamento na plataforma *SOLiD™4*.



Figura 1 - Exemplo de análise do mapeamento dos dados de sequenciamento do exoma. Gene *ASPM*. Entre as linhas verticais vermelhas, a posição da deleção esperada, descrita como resultado de análise no *Mozaik*. Em destaque (em roxo), uma troca de base (de T para C), presente na maioria das *reads* mapeadas, na posição vizinha àquela descrita pelo *Mozaik*. (chr1/consensus: sequência referência do gene. Sequências em verde – provenientes do sequenciamento da fita senso. Sequências em vermelho – provenientes do sequenciamento da fita antisenso.

Além disso, para se obter uma cobertura maior e consequentemente facilitar as análises, foi realizado novo sequenciamento, desta vez com o *SOLID™ 5500XL* (Tabela 2).

5.2 NOVO SEQUENCIAMENTO E ANÁLISE DAS BIBLIOTECAS DE EXOMA

Foi realizado um novo sequenciamento das quatro bibliotecas de exoma, na plataforma *SOLID™ 5500XL* (Life Technologies - Califórnia, Estados Unidos), seguindo as recomendações do fabricante, utilizando dois exomas por *lane* de sequenciamento. Foram depositadas, em média 200 milhões de *beads* em cada *lane*, utilizando a estratégia de sequenciamento *barcode paired-end* que realiza o sequenciamento de 35 bases a partir de uma das extremidades do fragmento, e de 75 bases a partir da outra extremidade.

Com os dados oriundos deste sequenciamento, pode-se verificar um aumento significativo quanto ao total de sequências resultantes para cada amostra, chegando a duplicar alguns resultados anteriormente obtidos na utilização da plataforma *SOLID™ 4*, como descrito na Tabela 2.

Tabela 2 - Total de dados gerados.

Amostra	Primeira leitura (<i>SOLiD™4</i>)	Segunda leitura (<i>SOLiD™ 5500XL</i>)	Aumento de cobertura (%)	Cobertura média da região alvo (%)	Cobertura média (X)
TW12	128.329.856	194.425.183	52%	90.58%	41X
21F	281.583.245	620.978.245	121%	93.68%	178X
21M	186.428.521	431.293.170	131%	93.42%	122X
21P	107.684.954	288.505.016	168%	93.18%	100X

Número total de sequências (leituras) gerados nas plataformas *SOLiD™4* e *SOLiD™ 5500XL* para cada amostra. (Amostra TW12 – amostra de DNA tumoral do probando. 21F – Amostra de DNA sanguíneo do probando. 21M – Amostra de DNA sanguíneo da mãe. 21P – Amostra de DNA sanguíneo do pai).

Os dados obtidos foram analisados em colaboração com o grupo do Dr. Pedro Galante do Hospital Sírio Libanês.

Todas as variações encontradas pela análise bioinformática foram carregadas em um banco de dados relacional (MySQL-<http://www.mysql.com/>) para que se pudesse realizar comparações entre as alterações identificadas em cada amostra. Para identificação das alterações somáticas, foram selecionadas apenas aquelas variações de nucleotídeo único exclusivamente reportadas na amostra tumoral. Para identificação das alterações *de novo*, foram selecionadas aqueles SNVs, reportadas tanto na amostra tumoral quanto na sanguínea do paciente e ausentes nas amostras dos progenitores.

5.3 SEGUNDA SELEÇÃO DAS ALTERAÇÕES PARA VALIDAÇÃO

As alterações SNVs *de novo* e somáticas identificadas que estavam presentes no banco de dados de SNP (*dbSNP*) foram descartadas, sendo selecionadas somente as alterações desconhecidas.

Em relação às alterações *de novo* desconhecidas, três foram classificadas como alterações de perda de função (duas *nonsense* e uma *splice site*), automaticamente selecionadas para validação (Tabela 3). Foram identificadas ainda, 80 alterações *missense* novas e, dentre elas 20 apresentaram 10 ou mais *reads* distintas no resultado obtido após o sequenciamento no *SOLID™ 5500XL*, reportando a variante descrita. Estas

20 alterações foram avaliadas em dois programas de predição dos efeitos decorrentes de alterações de troca de aminoácido (SIFTBlink - http://sift.jcvi.org/www/SIFT_BLink_submit.html e Polyphen - <http://genetics.bwh.harvard.edu/pph2/>), o que possibilitou a seleção daquelas descritas como deletérias em pelo menos um deles.

Como resultado desta análise, foram selecionadas 7 alterações não sinônimas de interesse, onde 2 foram excluídas por estarem localizadas em regiões de alta similaridade com outras porções do genoma, impossibilitando o desenho de *primers* específicos. Em resumo, após todos os critérios, foram selecionadas para validação 8 alterações *de novo* (5 *missense* 2 *nonsense* e uma *splice site*). Para estas alterações selecionadas, *primers* foram desenhados e analisados no programa Oligotech v1.0 (Tabela 3).

Para a seleção das alterações somáticas foram identificadas apenas 12 alterações de interesse candidatas, todas *missense* (não foi identificada nenhuma alteração *nonsense* ou *splice site*). Foram excluídas 4 alterações por estarem localizadas em regiões repetitivas. Por fim, obtivemos 8 alterações somáticas candidatas à validação, para as quais foram desenhados *primers* (Tabela 4) .

Tabela 3 - Alterações *de novo* selecionadas para validação

Gene	Chr	Posição (Hg19)	Base referência	Cobertura base referência	Base alterada	Cobertura base alterada	Percentual Base Alterada	Classificação	SIFT	Polyphen
<i>ACTL7A</i>	9	111625218	A	16	G	11	41%	<i>Missense</i>	Afeta a ação protéica	Benigna
<i>AGRN</i>	1	957638	G	40	T	11	22%	<i>Missense</i>	Afeta a ação protéica	Provavelmente danosa
<i>NAGPA</i>	16	5075484	C	10	A	13	57%	<i>Missense</i>	Afeta a ação protéica	Provavelmente danosa
<i>RPL3L</i>	16	1995846	G	11	A	12	52%	<i>Missense</i>	Afeta a ação protéica	Provavelmente danosa
<i>SAA2</i>	11	18269491	G	12	A	12	50%	<i>Missense</i>	Alteração tolerada	Possivelmente danosa
<i>APBA3</i>	19	3759854	G	6	A	9	60%	<i>Nonsense</i>	-	-
<i>IRX1</i>	5	3599338	G	5	C	6	55%	<i>Nonsense</i>	-	-
<i>MRPL37</i>	1	54681921	T	32	A	5	14%	<i>Splice site</i>	-	-

Chr – Cromossomo

Tabela 4 - Alterações somáticas selecionadas para validação

Gene	Chr	Posição (Hg19)	Base referência	Cobertura base referência	Base alterada	Cobertura base alterada	Percentual Base Alterada	Classificação
<i>DROSHA</i>	5	31421465	C	32	T	5	14%	<i>Missense</i>
<i>PKHD1L1</i>	8	110463332	A	28	C	6	18%	<i>Missense</i>
<i>SLC7A7</i>	14	23244714	G	6	A	5	45%	<i>Missense</i>
<i>TLR7</i>	X	12906217	T	26	C	5	16%	<i>Missense</i>
<i>TRIOBP</i>	22	38121115	G	20	A	8	29%	<i>Missense</i>
<i>TTK</i>	6	80750407	T	5	C	5	50%	<i>Missense</i>
<i>WDR27</i>	6	170049309	C	11	T	5	31%	<i>Missense</i>
<i>ZNF432</i>	19	52537713	C	24	T	8	25%	<i>Missense</i>

Chr – Cromossomo

5.4 DESENHO DOS PRIMERS

Os genes selecionados tiveram *primers* sintetizados, conforme visualizado na Tabela 5, seguido do tamanho de seus respectivos *amplicons*.

Tabela 5 - Sequência dos *primers forward* e *reverse* e tamanho dos amplicons das alterações *de novo* e somáticas selecionadas para validação.

Gene	Sequência <i>Primer Forward</i>	Sequência <i>Primer Reverse</i>	<i>Amplicon</i> (pb)
<i>de novo</i>			
<i>ACTL7A</i>	GACATGGCATCATCGTGGACTGG	GACCACGTAGGACACGCCATGG	492
<i>AGRN</i>	ATGACACCTACCGCCATGTC	GGTCTCCAAAGCCGCTGATCAC	147
<i>APBA3</i>	GGAGATTGCTGATGCCCAC	GCTCTTCAGGTGTGACTAG	278
<i>IRX1</i>	CCTTGAAGCGGCTGCTTC	GATGGCCAGCATGATCTTC	348
<i>MRPL37</i>	CACTCCACAGCTTCTGATTTG	CACTCCACAGCTTCTGATTTG	261
<i>NAGPA</i>	CACTGCAGCAAACCTGTCCTTG	ACGCCAGTGGCCTTGAAATTTCC	309
<i>RPL3L</i>	GGAAGTGAACAACGACTTCG	CAGGAAGGTCTTGCTCACCTAG	282
<i>SAA2</i>	GGCTCTCGTCGGAATATAAGTCC	ACACTCAGGACCAAGGAGCA	190
<i>Somáticas</i>			
<i>DROSHA</i>	ATCGAGGGGCCTTAGGGAAT	ACCCCTCTAATCTTCACTCTGC	325
<i>PKHD1L1</i>	AGGTTTTGGAAATGAATCAGGGC	ACTGTCCGGGCATAGGAGAA	321
<i>SLC7A7</i>	GAAGGGATGGCAGAGAGGGA	TTAGCATAGTGCCTTGCCCG	349
<i>TLR7</i>	CCTGGCCACAGATGTGACTT	TCTTCCAGTTTGGCCACCAG	333
<i>TRIOBP</i>	CAGGACAACCCAGAACTTGT	GCCCAGGGGATGTCTTTGTT	287
<i>TTK</i>	AGTTAACCAAATGGCCAAGGGA	CTGTACTGTATGCAGAAGAAGAA	301
<i>WDR27</i>	TGAGAGACACAGGAGGGGAG	AGTGTGGAAGGTCTAGGT	318
<i>ZNF432</i>	AGGCTTCACCACAAAACACT	GATGAGCATGCTTTTCACAGT	280

5.5 VALIDAÇÃO POR SEQUENCIAMENTO CAPILAR

As PCRs foram realizadas para cada gene utilizando as 4 amostras do estudo: DNA de sangue e tumor do probando e DNA de sangue dos progenitores. Os produtos de PCR foram avaliados em gel de agarose 1% (Figura 2) para verificação do tamanho dos fragmentos amplificados.

As sequências obtidas na análise de cada gene foram alinhadas contra a sequência referência do gene obtida no banco de dados *Genbank* com auxílio do *software CLC Genomics 5*.

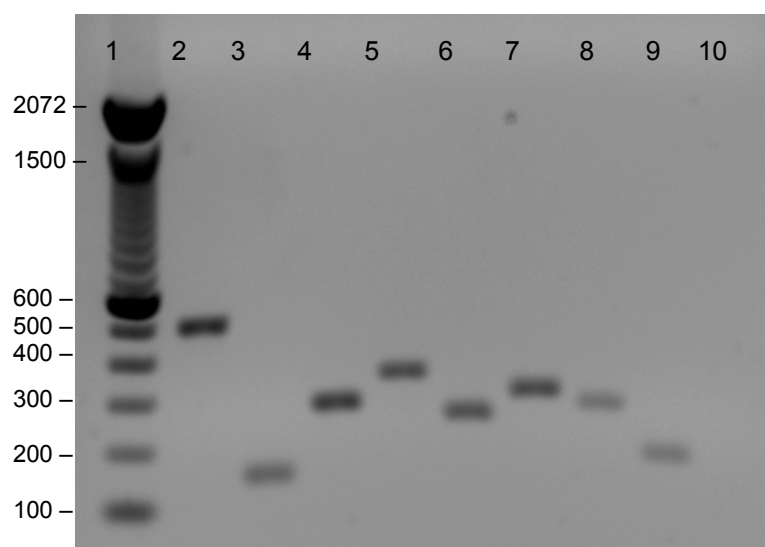


Figura 2 - Resultado da padronização de PCR. Gel de agarose, demonstrando o resultado da padronização da PCR, utilizando DNA controle para os 8 genes selecionados com alterações *de novo* (1- Ladder 100Kb 2 - ACTL7A, 3 - AGRN, 4 - APBA3, 5 - IRX1, 6 - MRPL37, 7 - NGPA, 8 - RPL3L, 9 - SAA2 10 – Controle Negativo).

Com relação às alterações previamente descritas como *de novo*, dos 8 genes sequenciados, 4 (50%) não apresentaram alteração descrita na análise bioinformática (Figura 3). Outros 4 genes (50%) apresentaram alteração tanto nas amostras tumoral e sanguínea, quanto em pelo menos

uma amostra dos progenitores (Figura 4) e, portanto, tratavam-se de alterações herdadas (Tabela 6).

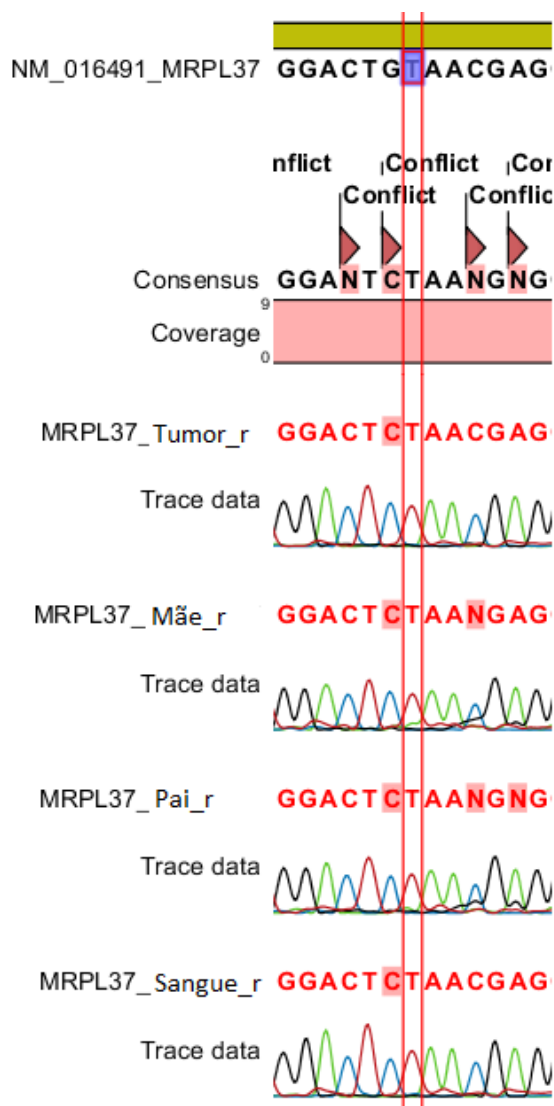


Figura 3 - Exemplo da análise do sequenciamento capilar para o gene *MRPL37* (alteração *de novo* – análise das sequências referentes ao sangue e tumor do probando, pai e mãe) Em destaque, a posição da alteração avaliada. NM_016491_MRPL37: sequência referência do gene. MRPL37_Sangue_r – DNA do sangue do probando, sequenciado com *primer reverse*; MRPL37_Mãe_r - DNA do sangue da mãe, sequenciado com *primer reverse* MRPL37_Pai_r - DNA do sangue do pai, sequenciado com *primer reverse*. MRPL37_Tumor_r – DNA do tumor do probando, sequenciado com *primer reverse*. Neste caso, pode-se observar que não há alteração de base nesta região, sendo que apenas a base referência (T) foi encontrada nas 4 amostras analisadas.

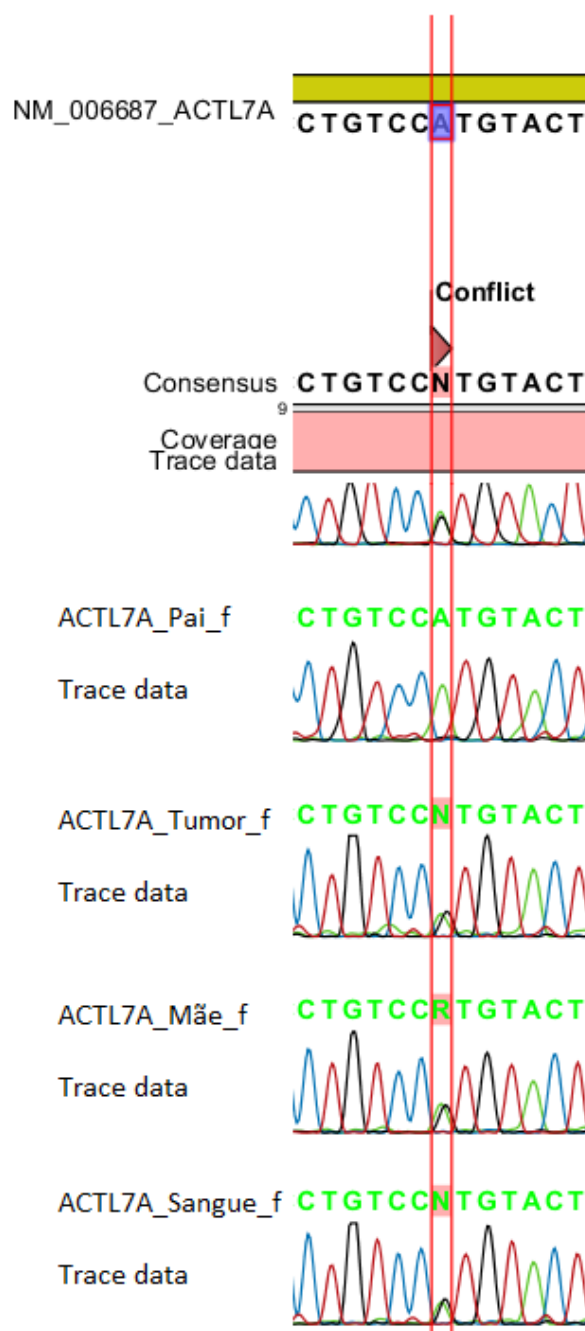


Figura 4 - Exemplo da análise do sequenciamento capilar para o gene *ACTL7A* (alteração *de novo* – análise das sequências referentes ao sangue e tumor do probando, pai e mãe) Em destaque, a posição da alteração avaliada. NM_006687_ACTL7A: sequência referência do gene. ACTL7A_Pai_f – DNA do sangue do pai, sequenciado com *primer forward*. ACTL7A_Tumor_f - DNA do tumor do probando, sequenciado com *primer forward*. ACTL7A_Mãe_f - DNA do sangue da mãe, sequenciado com *primer forward*. ACTL7A_Sangue_f - DNA do sangue do probando, sequenciado com *primer forward*. Neste caso, pode-se observar que a alteração está presente tanto na amostra sanguínea e tumoral do probando quanto na amostra materna.

Para as alterações previamente descritas como somáticas, verificou-se com o sequenciamento capilar que 6 genes (75%) não apresentaram a alteração esperada (Figura 5) e apenas 2 (25%) apresentaram a alteração descrita. Destes, em 1 caso a alteração estava presente não só na amostra tumoral, como na amostra sanguínea (Figura 6) enquanto que a outra alteração foi encontrada somente em tecido tumoral (Figura 7 e Tabela 7).

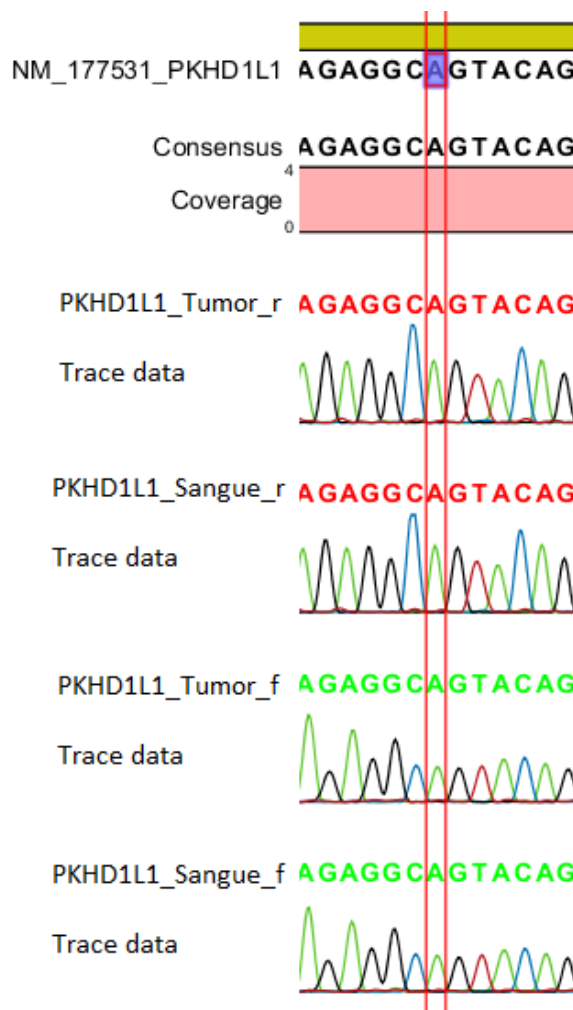


Figura 5 - Exemplo da análise do sequenciamento capilar para o gene *PKHD1L1* (alteração somática – análise das sequências referentes ao sangue e tumor do probando). Em destaque, a posição da alteração avaliada. NM_177531_PKHD1L1: sequência referência do gene. PKHD1L1_Tumor_r e PKHD1L1_Tumor_f - DNA do tumor do probando sequenciado com *primer reverse* e *forward* respectivamente; PKHD1L1_Sangue_r e PKHD1L1_Sangue_f - DNA do sangue do probando sequenciado com *primer reverse* e *forward* respectivamente. Neste caso, pode-se observar que não há alteração de base nesta região, sendo que somente a base referência (A) foi encontrada nas 2 amostras analisadas.

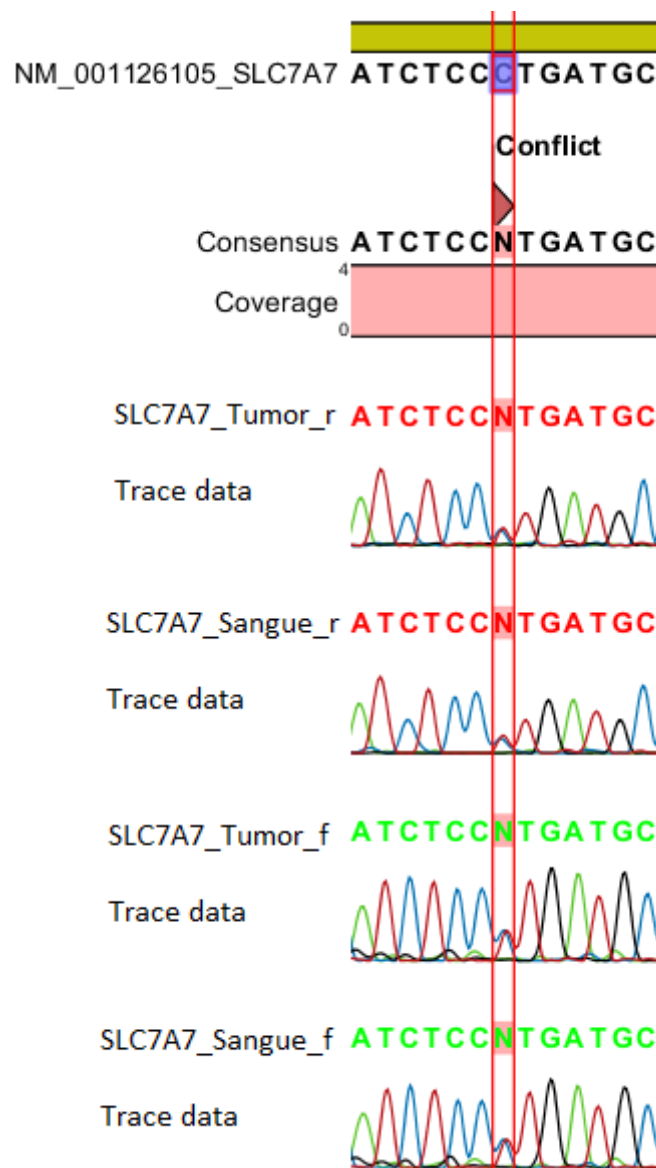


Figura 6 - Exemplo da análise do sequenciamento capilar para o gene *SLC7A7* (alteração somática – análise das sequências referentes ao sangue e tumor do probando). Em destaque, a posição da alteração avaliada. NM_001126105_SLC7A7: sequência referência do gene. SAA2_Tumor_r e SAA2_Tumor_f - DNA do tumor do probando sequenciado com *primer reverse* e *forward* respectivamente; SAA2_Sangue_r e SAA2_Sangue_f - DNA do sangue do probando sequenciado com *primer reverse* e *forward* respectivamente. Neste caso, pode-se observar que há uma troca de base na região de interesse em todas as amostras, demonstrado pela presença da base referência (C) e da base variante (G) nas 2 amostras analisadas, incluindo a amostra de sangue do probando, caracterizando uma alteração não somática.

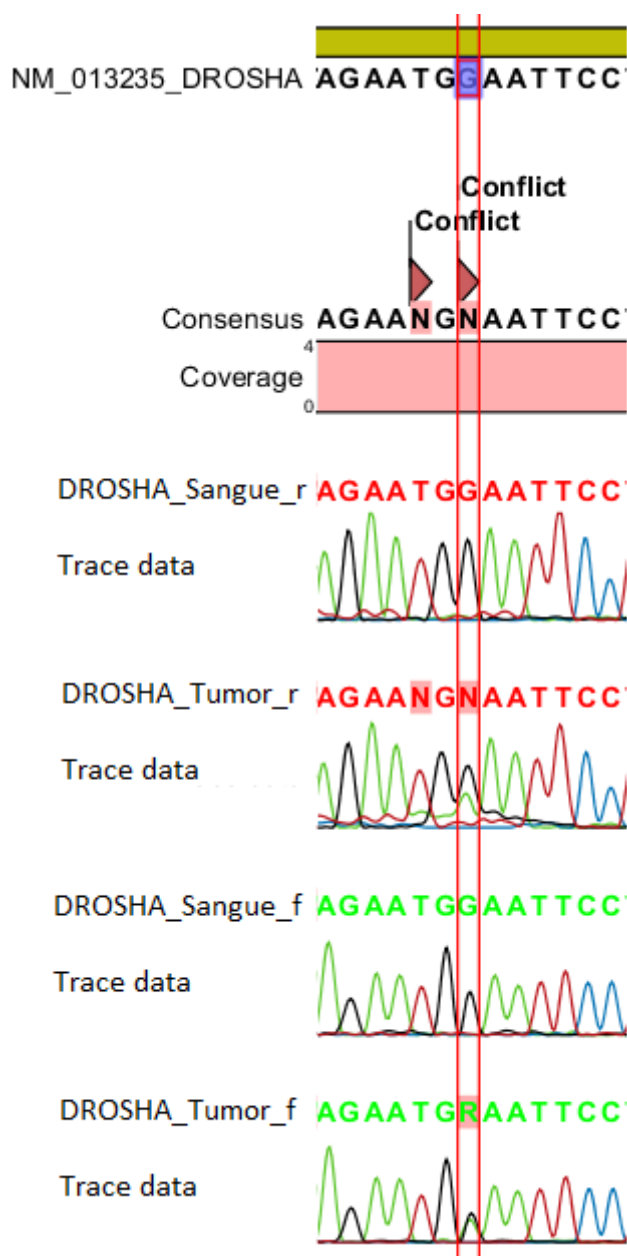


Figura 7 - Exemplo da análise do sequenciamento capilar para o gene *DROSHA* (alteração somática). Em destaque, a posição da alteração avaliada. NM_013235_DROSHA: sequência referência do gene. DROSHA_Tumor_r e DROSHA_Tumor_f - DNA do tumor do probando sequenciado com *primer reverse* e *forward* respectivamente; DROSHA_Sangue_r e DROSHA_Sangue_f - DNA do sangue do probando sequenciado com *primer reverse* e *forward* respectivamente. Neste caso, pode-se observar que há alteração de base (de G para A) na posição em questão, apenas nas amostras referentes ao DNA oriundo do tumor do probando. Nota-se, ainda, que a mesma alteração está ausente no sangue do probando.

Assim sendo, os resultados obtidos através da validação por sequenciamento capilar dos genes de interesse, na busca da alteração descrita em análise bioinformática, podem ser analisados nas Tabelas 6 e 7.

Tabela 6 - Resultado da análise de sequenciamento capilar para validação das alterações *de novo* selecionadas.

	Resultado sequenciamento Exoma					Resultado do sequenciamento capilar (Sanger)				
	Gene	Base Ref	Base Alt	Direção reads alteradas	Percentual Base Alterada	Filho (Sangue)	Filho (Tumor)	Pai	Mãe	Conclusão
de novo	ACTL7A	A	G	+/-	41%	A/G	A/G	A/A	A/G	Alteração Herdada
	AGRN	G	T	-	22%	G/G	G/G	G/G	G/G	Alteração não encontrada
	APBA3	C	A	+/-	60%	C/A	C/A	C/C	C/A	Alteração Herdada
	IRX1	G	C	+/-	55%	G/C	G/C	G/C	G/G	Alteração Herdada
	MRPL37	T	A	-	14%	T/T	T/T	T/T	T/T	Alteração não encontrada
	NAGPA	G	T	+/-	57%	G/T	G/T	G/G	G/T	Alteração Herdada
	RPL3L	G	A	+/-	52%	G/A	G/A	G/G	G/A	Alteração Herdada
	SAA2	C	T	+/-	50%	C/T	C/T	C/C	C/T	Alteração Herdada

A tabela apresenta os genes que foram sequenciados com as respectivas alterações selecionadas, o número de *reads* obtido no sequenciamento do exoma (% da base alterada em relação ao total de bases sequenciadas nesta posição), a direção das *reads* alteradas (fita senso (+), fita antisenso (-) ou em ambas (+/-)) e o resultado do sequenciamento capilar para cada amostra (Ref – Referência; Alt – Alterada)

Tabela 7 - Resultado da análise de sequenciamento capilar para validação das alterações somáticas selecionadas.

		Resultado sequenciamento Exoma				Resultado do sequenciamento capilar (Sanger)		
Gene		Base Ref	Base Alt	Direção <i>reads</i> alteradas	Percentual Base Alterada	Filho (Sangue)	Filho (Tumor)	Conclusão
Somáticas	<i>DROSHA</i>	G	T	-	14%	G/G	G/T	Alteração Validada
	<i>PKHD1L1</i>	A	C	+	18%	A/A	A/A	Alteração não encontrada
	<i>SLC7A7</i>	C	A	-	45%	C/A	C/A	Alteração não somática
	<i>TLR7</i>	T	C	+	16%	T/T	T/T	Alteração não encontrada
	<i>TRIOBP</i>	G	A	+	29%	G/G	G/G	Alteração não encontrada
	<i>TTK</i>	T	C	-	50%	T/T	T/T	Alteração não encontrada
	<i>WDR27</i>	G	T	-	31%	G/G	G/G	Alteração não encontrada
	<i>ZNF432</i>	G	T	+/-	25%	G/G	G/G	Alteração não encontrada

A tabela apresenta os genes que foram sequenciados com as respectivas alterações selecionadas, o número de *reads* obtido no sequenciamento do exoma (% da base alterada em relação à base referência) a direção das *reads* alteradas (fita senso (+), fita antisenso (-) ou em ambas (+/-)) e o resultado do sequenciamento capilar para cada amostra (Ref – Referência; Alt – Alterada)

5.6 ANÁLISE E COMPARAÇÃO DAS CARACTERÍSTICAS DAS ALTERAÇÕES SELECIONADAS

Devido à baixa frequência de validação obtida para as alterações selecionadas, que foi nula (0/8) para as alterações *de novo* e 1,25% para as alterações somáticas (1/8), fizemos uma avaliação das características de cada alteração reportada pela análise de bioinformática e realizamos uma inspeção manual dos dados gerado pelo sequenciamento de exoma.

Os dados mapeados das 4 amostras sequenciadas foram analisados manualmente com o auxílio da ferramenta *CLCBio Genomics*, possibilitando a identificação da presença ou não da alteração descrita mesmo que em frequência muito baixa, além de permitir uma avaliação visual da qualidade

das *reads* de cada região, como a presença de diversas alterações em uma mesma *read* e frequência de *reads* redundantes.

A partir dos dados de bioinformática, as alterações selecionadas foram avaliadas quanto à frequência da base alterada e quanto ao sentido das *reads* nas quais a alteração é reportada (fita senso (+), antisenso (-) ou em ambas (+/)) (Tabelas 6 e 7, quarta e quinta colunas). Esta avaliação permitiu verificarmos que, para as alterações *de novo*, todas as alterações que não foram identificadas em nenhuma amostra, foram reportadas em apenas uma das fitas e apresentaram frequências baixas (14 e 22%).

Por outro lado, as alterações que foram identificadas como herdadas, apresentavam *reads* contendo a alteração em ambos os sentidos e uma frequência acima de 40% (41 a 60%). Quando analisadas manualmente no *CLCBio Genomics*, foi observado que estas alterações estavam presentes em baixa frequência em pelo menos um dos progenitores e por este motivo não atingiram os critérios necessários na chamada de SNVs da análise de bioinformática.

Em relação às alterações somáticas, não foi possível observar um padrão entre as alterações não encontradas, encontradas no sangue e no tumor, ou validadas como somáticas.

5.7 NOVA SELEÇÃO DE GENES RELACIONADOS A ALTERAÇÕES DE NOVO E DESENHO DE *PRIMERS*

Após a observação do padrão de resultados descritos acima para as alterações *de novo*, foi realizada uma nova seleção de candidatos levando em consideração as informações de sentido das *reads* e a avaliação manual das sequências no software *CLCBio Genomics*. Para as 20 variantes *missense* selecionadas, realizou-se uma contagem das *reads* que cobriam a região onde se localizava a alteração. Aquelas variantes que apresentaram mais de 10 *reads* com a alteração descrita na análise bioinformática apenas na amostra do sangue do probando e ausente nos progenitores, foram selecionadas. Assim, chegou-se a três novos genes referentes a alterações *de novo*: *EPB41L1*, *ERBB2* e *PARD6B*. Deu-se prosseguimento através do desenho e confirmação da viabilidade dos *primers* (Tabela 8) para que se pudesse realizar eficientemente as PCRs, seguidas pelo sequenciamento dos produtos obtidos nas mesmas.

Tabela 8 - Sequências dos *primers* (*forward* e *reverse*) selecionados para validação de novas alterações *de novo*, seguido por seus amplicons.

Gene	Sequência <i>Primer Forward</i>	Sequência <i>Primer Reverse</i>	Amplicon (pb)
<i>PARD6B</i>	CTACCCACAGCAGATTGAAC	CTGCTTGCTATGGCTGCTA	219
<i>EPB41L1</i>	GTTACAGAGCATTCTCCCTC	GAGATTAAAAGAGCTGGCCAAG	239
<i>ERBB2</i>	CCCTGCCCTCTGAGACTGATGGCTACG	GCCGAACATCTGGCTGGTT	179

Todos os procedimentos realizados para estas últimas alterações selecionadas foram exatamente iguais aos realizados anteriormente. Os resultados obtidos ao fim do sequenciamento de Sanger permitiram concluir que, das três alterações selecionadas 2 (75%) não apresentavam a alteração previamente descrita na análise bioinformática e, portanto, apenas 1 (25%) alteração foi confirmada e pode ser validada (*PARD6B*) (Tabela 9).

A alteração (g.chr20: 49366843A>G, p.N313D) foi confirmada como uma alteração *de novo*, uma vez que foi encontrada apenas nas amostras referentes ao probando (sangue e tumor) e ausentes nas amostras dos progenitores (Figura 8).

Tabela 9 - Resultado da análise de sequenciamento capilar para validação de novas alterações *de novo* selecionadas.

Resultado sequenciamento Exoma					Resultado do sequenciamento capilar (Sanger)						
Gene		Base Ref	Base Alt	Direção <i>reads</i> alteradas	Percentual Base Alterada	Filho (Sangue)	Pai	Mãe	Conclusão	SIFT	Polyphen
de novo	<i>PARD6B</i>	A	G	+/-	56%	A/G	A/A	A/A	Alteração validada	Tolerada	Benigna
	<i>EPB41L1</i>	T	A	+	32%	T/T	T/T	T/T	Alteração não encontrada	Tolerada	Benigna
	<i>ERBB2</i>	C	T	+/-	56%	C/T	T/T	C/T	Alteração Herdada	Tolerada	Benigna

A tabela apresenta os genes que foram sequenciados com as respectivas alterações selecionadas, o número de *reads* obtido no sequenciamento do exoma (% da base alterada em relação ao total de bases sequenciadas nesta posição), a direção das *reads* alteradas (fita senso (+), fita antisense (-) ou em ambas (+/-)) e o resultado do sequenciamento capilar para cada amostra (Ref – Referência; Alt – Alterada)

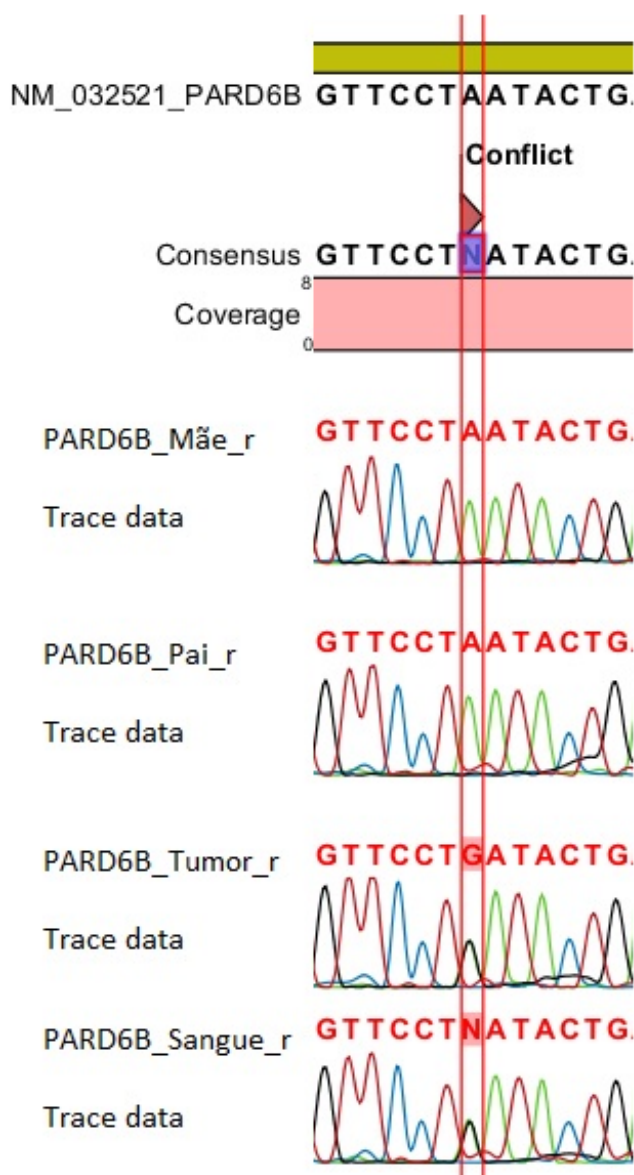


Figura 8 - Análise do sequenciamento capilar para o gene *PARD6B* (alteração *de novo* – análise das sequências referentes ao sangue e tumor do probando, pai e mãe) Em destaque, a posição da alteração avaliada. NM_032521_PARD6B: sequência referência do gene. PARD6B_Sangue_r – DNA do sangue do probando, sequenciado com *primer reverse*; PARD6B_Mãe_r - DNA do sangue da mãe, sequenciado com *primer reverse*. PARD6B_Pai_r - DNA do sangue do pai, sequenciado com *primer reverse*. PARD6B_Tumor_r – DNA do tumor do probando, sequenciado com *primer reverse*. Neste caso, pode-se observar que a alteração está presente apenas nas amostras referentes ao probando (A>G), confirmando sua classificação como alteração *de novo* e permitindo que a mesma fosse validada neste estudo.

5.8 SEQUENCIAMENTO DA ALTERAÇÃO SOMÁTICA VALIDADA EM AMOSTRAS INDEPENDENTES

Foi realizado o *screening* da alteração somática c.3439G>A, p.E1147K encontrada no gene *DROSHA* em outras amostras de TW. Das 96 amostras sequenciadas, 8 apresentaram a mesma mutação encontrada no probando selecionado para este estudo significando que esta alteração está presente em 9,3% de nossa casuística com TW (9/97) (Figura 9).

Da mesma forma, das 272 amostras de DNA oriundos de indivíduos saudáveis, nenhuma apresentou a mutação no gene *DROSHA*.

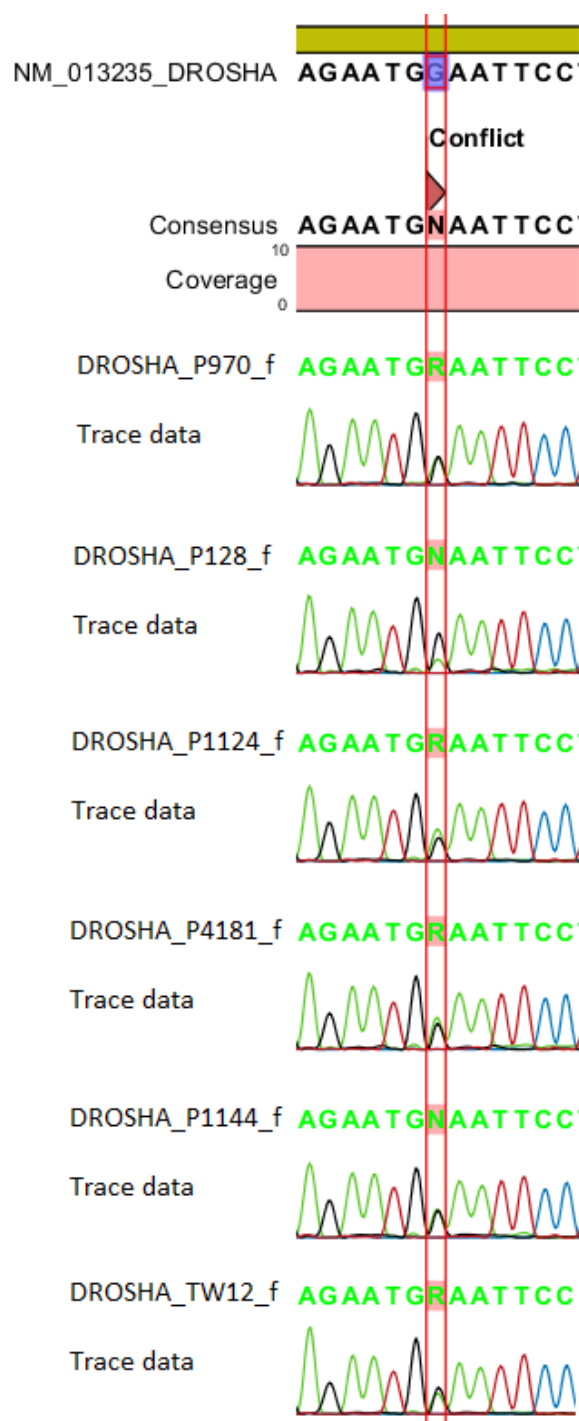


Figura 9 - Sequenciamento da alteração somática validada no gene *DROSHA* em amostras independentes. Em destaque, a posição da alteração avaliada. Exemplo de sequências referentes ao *primer forward* para 5 tumores (P970, P128, P1124, P4181, P1144) que apresentaram a mesma alteração somática identificada no tumor do probando do estudo (TW12). NM_013235_DROSHA: sequência referência do gene.

Visando a realização de análises futuras, foi realizado, também, o sequenciamento completo do gene *DROSHA* em amostras referentes a 3 amostras de TW, resultando na identificação de alguns polimorfismos (Tabela 10). Analisando as sequências obtidas, nenhuma outra mutação somática desconhecida foi encontrada no referido gene.

Tabela 10 - Resultado da análise de sequenciamento do gene *DROSHA* para verificação da presença de outras alterações.

Éxon	Posição da alteração (gene)	Posição da alteração (cDNA)	Alteração	Classificação <i>dbSNP</i>
7	962	321	C>T	descrita (rs55656741) – presente em 36.3% da população*
31	3665	1102	A>G	descrita (rs2287584) - presente em 44% da população*
34	4124	1255	T>C	descrita (rs2241337) - presente em 30.5% da população*

*População amostrada pelo projeto *HapMap*, que utilizou amostras de DNA oriundos de populações de descendência Africana, Européia e Asiática (<http://hapmap.ncbi.nlm.nih.gov/hapmappopulations.html.en>)

A tabela apresenta os resultados obtidos no sequenciamento, identificando as alterações encontradas.

6 DISCUSSÃO

O TW é uma neoplasia renal infantil que se desenvolve a partir de células do blastema metanéfrico incapacitadas de completar sua diferenciação durante a diferenciação do rim. Aproximadamente 30% dos TWs apresentam mutações nos genes *WT1*, *WT2*, *CTNNB1*, *APC* e *PLCG2*. Porém, 70% dos casos de TW ainda continuam sem nenhuma alteração associada.

Assim, este projeto teve por objetivo a validação de alterações somáticas e *de novo* associadas ao desenvolvimento de TW, alterações estas previamente identificados por nosso grupo de pesquisa através do sequenciamento de exoma. De início, partiu-se do sequenciamento de exoma das 4 amostras associadas a um caso de TW esporádico (amostras tumoral e sanguínea do probando e amostras sanguíneas dos progenitores), seguido pelo enriquecimento das regiões genômicas do DNA e sequenciamento destas regiões.

Inicialmente, para as análises de bioinformática, foi estabelecida uma colaboração com o Departamento de Bioinformática e Biologia Computacional do M.D. Anderson Cancer Center (Universidade do Texas, Houston, EUA). Devido a problemas na etapa de chamada de variantes dessa análise, nenhuma das 18 alterações somáticas e 18 alterações *de novo* selecionadas para validação por sequenciamento capilar puderam ser

confirmadas. Além disso, observamos que a cobertura do sequenciamento estava relativamente baixa.

Optou-se, então, por um segundo sequenciamento, no equipamento *SOLID™ 5500XL*, objetivando um aumento na cobertura das regiões alvo. Os dados obtidos nas duas corridas foram então analisados em colaboração com o grupo de bioinformática do Hospital Sírio Libanês.

Foram selecionadas novas alterações para validação por sequenciamento Sanger, ausentes no banco de dados *dbSNP*, dando-se preferência para aquelas classificadas como alterações de perda de função (*nonsense* e *splice site*) e de troca de aminoácido, por terem maiores chances de interferirem na função proteica e estarem relacionadas à formação tumoral. Para as alterações *missense*, foram selecionadas aquelas com maiores coberturas, e que foram classificadas como deletérias em programas de predição *in silico*.

Por fim, foram selecionadas 8 alterações somáticas e 11 alterações *de novo*, sendo que uma alteração de cada tipo pode ser validada. Por conta do número elevado de alterações *de novo* encontradas, foram utilizados parâmetros seletivos, para restringir e especificar a escolha das alterações candidatas à validação. Para as alterações somáticas, foram obtidos apenas 8 candidatos, dado este que concorda com os dados da literatura, onde verifica-se que tumores pediátricos no geral apresentam uma baixa frequência de mutações somáticas (VOGELSTEIN et al. 2013).

Segundo VOGELSTEIN et al. (2013), o número de mutações em certos tumores originários de tecidos renováveis, está diretamente

relacionado com a idade. Quando se realiza uma avaliação linear a respeito da regressão tumoral, verifica-se que mais da metade das mutações somáticas identificadas, ocorre durante a fase pré-neoplásica, ou seja, durante o crescimento de células normais utilizadas para reposição em tecidos epiteliais, gastrointestinais, genitourinário e outros tecidos. Todas essas mutações pré-neoplásicas são denominadas *passengers*, e se sugere que não apresentam nenhum efeito no processo neoplásico.

Esta linha de raciocínio também ajuda a explicar porque tumores pediátricos tem, no geral, muito menos mutações do que os tumores adultos. Tumores pediátricos geralmente ocorrem em tecidos não regenerativos, e aqueles que ocorrem em tecidos que se regeneram, como as leucemias, originam-se de células precursoras que não se renovam tão frequentemente quanto nos adultos. Estudos referentes ao sequenciamento do genoma em pacientes com leucemia suportam a idéia de que as mutações ocorrem de forma randômica em células precursoras normais, antes que estas células adquiram uma mutação inicial (WELCH et al. 2012).

Para a maioria dos genes sequenciados, a alteração descrita na análise bioinformática não foi confirmada pelo sequenciamento capilar. A baixa cobertura pode ser um dos motivos pelo qual a maioria das alterações não foi validada, uma vez que a cobertura referente às alterações somáticas e *de novo* validadas neste estudo foi 31% e 56% respectivamente, coberturas estas superiores à média correspondentes a todas alterações selecionadas (28% para as somáticas e 44% para as *de novo*). Para as alterações *de novo* foi verificado que as duas alterações não encontradas no

sangue do probando apresentavam as menores frequências entre as variantes. Ainda, pela avaliação manual dos mapeamentos, observamos que as variantes confirmadas como herdadas estavam presentes em um dos progenitores em baixa frequência.

Estes dados mostram a importância da cobertura das regiões alvo através do sequenciamento de nova geração, cobertura esta que, segundo a literatura, deve variar de 20 a 60x quando analisada em genomas grandes como o de mamíferos (DESAI e JERE 2012). Neste trabalho a amostra que apresentou a menor cobertura foi a tumoral (41X), o que poderia, em parte, contribuir o baixo número de SNVs reportadas nesta amostra. No entanto, é importante mencionar que é esperado que o número de alterações genéticas neste tipo de tumor seja bastante menor do que é encontrado em tumores esporádicos de indivíduos adultos.

Além disso, pode-se verificar que a presença da mutação nas *reads* correspondentes aos dois sentidos (*senso* e *antisense*) da molécula de DNA é uma informação importante e que deve ser considerada, uma vez que 26% das alterações fornecidas através da análise bioinformática puderam ser confirmadas após o sequenciamento capilar e verificação manual.

Por fim, até o momento, foram validadas neste trabalho uma alteração somática e uma alteração *de novo*. Visto que estudos prévios baseados em famílias descreveram a ocorrência de aproximadamente 143 alterações *de novo* ausentes no *dbSNP* em cada indivíduo (ROBINSON 2010), e que neste trabalho identificamos mais de 80 alterações *de novo* candidatas, o processo de seleção e validação continuará a ser realizado, buscando

validar um número maior de alterações, incluindo alterações do tipo inserções e deleções, que não foram avaliadas neste primeiro momento.

A alteração *de novo* validada neste estudo foi uma alteração *missense* encontrada no gene *PARD6B*, localizado no cromossomo 20. A alteração é do tipo “troca de bases” com troca de aminoácido (g.chr20: 49366843A>G, p.N313D). Atualmente, poucos estudos descrevam a relação entre este gene e processos tumorais, sendo que a maioria deles relaciona problemas na expressão deste gene com o desenvolvimento de câncer de mama (NOLAN et al. 2008). Conclui-se, portanto que faz-se extremamente necessário uma investigação mais profunda e específica, para que se elucide o papel do gene *PARD6B* na tumorigênese do TW.

O gene *PARD6B* codifica uma proteína citoplasmática que está envolvida na divisão celular e em processos de polarização celular, atuando como membro de um complexo multi-protéico (ALARCON 2010). Os processos de divisão celular assimétrica e polarização celular são fundamentais para o desenvolvimento embrionário e remodelação tecidual (ALARCON 2010). Além disso, a manutenção da estrutura dos tecidos também é dependente da polarização e da orientação da migração celular, e distúrbios nestes eventos são características de malignidade (CUNLIFFE et al. 2012).

PARD6B é necessária para a morfogênese do blastocisto, em especial no desenvolvimento de tecidos tropoectodérmicos, atuando ainda sob a polaridade apical-basal celular, formação de tecido conjuntivo, vedação de permeabilidade paracelular e superexpressão do fator de

transcrição Cdx2. Resultados obtidos por CUNLIFFE e colaboradores (2012), demonstraram que a proteína PARD6B é necessária para a diferenciação de células-tronco embrionárias em células de tecidos epiteliais e também para a maturação dos tecidos ecto e endodérmicos primitivos.

Em relação às alterações somáticas avaliadas no presente estudo, somente a alteração presente no gene *DROSHA* (g.chr5:31421465C>T) foi validada. Interessantemente esta mesma mutação foi encontrada em outros 8 TWs de nossa casuística, estando, portanto, presente em aproximadamente 10% dos tumores de Wilms avaliados neste estudo. Nossos resultados demonstram que as mutações *DROSHA* é um evento importante e frequente no TW. Além disso, a presença da mesma mutação em outras amostras de TW, somada à ausência de mutações de truncamento da proteína, indicam que estas mutações em *DROSHA* representam mais provavelmente um evento oncogênico. Esta conclusão é apoiada por uma revisão recentemente feita por Vogelstein e colaboradores (2013), onde foi sugerida a “regra do 20/20” para classificar os genes supressores de tumor e oncogenes. De acordo com proposta dos autores, para um gene ser classificado como um oncogene, mais de 20% de mutações descritas no mesmo devem estar em posições recorrentes e serem variantes *missense* - critério este preenchido pelas mutações no gene *DROSHA* descritas no nosso estudo.

A proteína DROSHA é uma RNase do tipo III envolvida nos estágios iniciais do processo de formação dos microRNAs (miRNAs) (LEE et al. 2003). MiRNAs constituem uma grande classe de RNA de fitas simples,

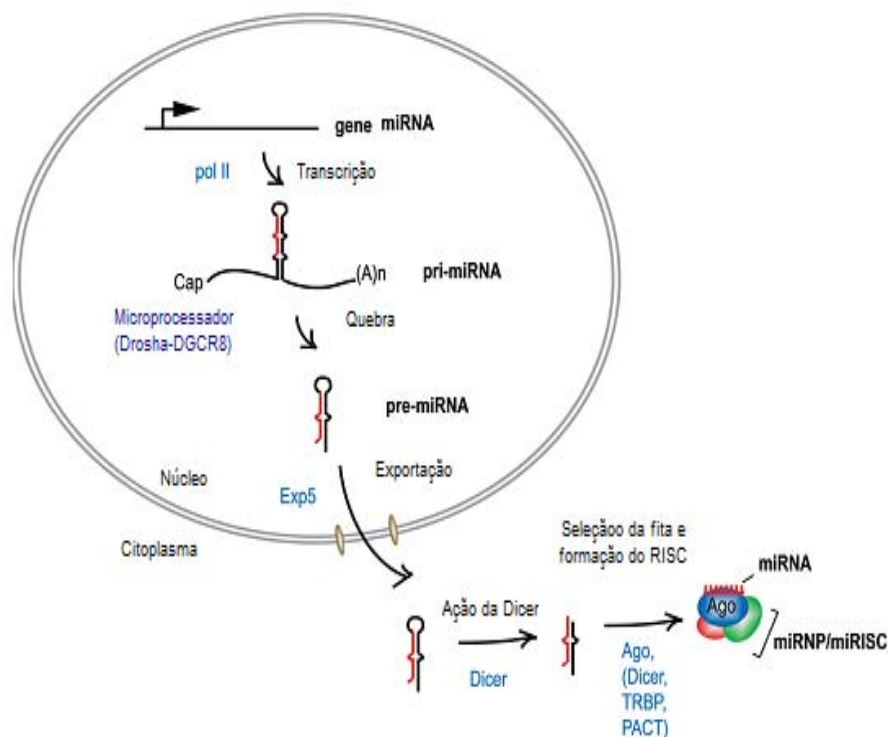
compostos por 19 a 25 nucleotídeos, filogeneticamente conservados. Estudos recentes demonstraram que miRNAs estão envolvidos na tumorigênese, e que centenas de miRNAs estão desregulados em células tumorais. Além disso, estudos funcionais confirmaram seu envolvimento em todos os processos moleculares e biológicos que levam à tumorigênese (NEGRINI et al. 2009).

Nos últimos 10 anos, tem-se visto um grande número de estudos baseando-se na relação de miRNAs e câncer. Apesar da grande variedade tanto em termos da doença quanto das abordagens experimentais adotadas, muitos princípios essenciais relacionados à atividade dos miRNAs na formação tumoral foram definidos (MENDELL e OLSON 2012). Já se sabe que alguns SNPs e mutações distribuídas em regiões de miRNAs primários, precursores de miRNAs e em miRNAs maduros, podem estar envolvidos na taxa de sobrevivência de câncer de pulmão (HU et al. 2008a), identificam variantes genéticas comuns em pré-miRNAs associadas com o crescimento do risco de câncer de mama em mulheres chinesas (HU et al. 2008b), além de que um SNP comum em pré-miR-146a, diminui a expressão de miRNAs maduros, predispondo indivíduos ao carcinoma de tireóide (JAZDZEWSKI et al. 2008).

O processo de biogênese dos miRNAs apresenta diversas etapas: primeiramente, os genes de miRNA são transcritos em longas moléculas de RNA denominadas miRNA primários (pri-miRNAs); no núcleo estas moléculas são processadas pela DROSHA, que cliva o pri-miRNA em estruturas em grampo de aproximadamente 70 nucleotídeos, denominados

precursores de miRNA (pré-miRNAs); estes, por sua vez, são processados e exportados ao citoplasma, onde são convertidos a miRNAs maduros através da clivagem da proteína DICER (Figura 10) (LEE et al. 2002). Assim, as duas proteínas da família RNase III, DROSHA e DICER, participam do processo de formação dos miRNAs e tem papéis chaves na regulação de genes mediados por miRNAs, em processos como desenvolvimento e diferenciação celular.

Nesse sentido, recentemente foram descritas alterações germinativas e somáticas no gene *DICER* associadas a tumores de Wilms. Trata-se de uma RNase que catalisa a clivagem de substratos do RNA dupla fita (dsRNA) para produzir microRNAs e RNAs de interferência (siRNAs) (CHAKRAVARTHY et al. 2010). Estudos desenvolvidos recentemente por WU et al. (2013) demonstram que alguns pacientes com TW apresentam mutações nos dois alelos de *DICER*, sugerindo que essa mutação possa estar relacionada com a patogênese desse tumor.



Fonte: Modificada LEE et al. (2006)

Figura 10 - Modelo de biogênese do miRNA, onde os genes são transcritos por uma RNA polimerase II para que ocorra a geração de um miRNA primário (pri-RNA). O início do processo dá-se através da ação do complexo Drosha-DGCR8 (complexo microprocessador). Drosha, assim como o DGCR8, está localizada no núcleo. O produto desta reação é um pre-miRNA de aproximadamente 70 nucleotídeos, com uma alça de aproximadamente 3 nucleotídeos. Esta estrutura funciona como uma assinatura que é reconhecida pelo fator de exportação nuclear, Exportin 5 (Exp5). Uma vez exportado, a RNase III citoplasmática Dicer participa da segunda etapa de maturação na produção do miRNA duplex. Em seguida, o duplex é separado e, normalmente, uma das fitas é maturada em miRNA e a outra fita é degradada.

A mutação somática encontrada neste trabalho, está localizada no éxon 32 do gene *DROSHA*, no domínio RNaseIII b (RIIIdb, como ilustrado na Figura 11) e é do tipo “troca de bases” (G>A) com troca de aminoácido (Glu>Lys). Tal troca ocorre exatamente no códon que abrange um dos três sítios de ligação do cofator magnésio (códon 1147), essencial para o funcionamento da enzima.

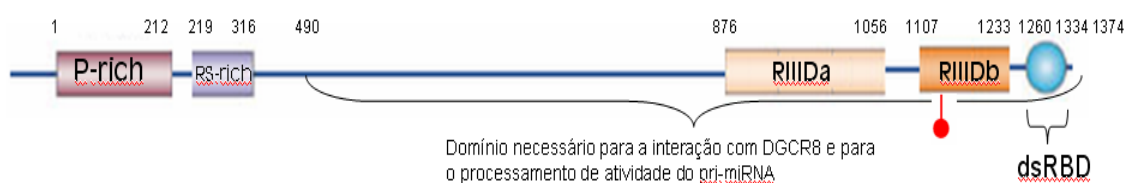


Figura 11 - Esquema do gene *DROSHA*, dividido entre seus principais domínios. P-rich - região rica em prolina. RS-rich – região rica em arginina e serina. RIIIDa e RIIIDb – domínios RNase. DsRBD – Domínio de ligação ao RNA. De 1 a 1374 representam as posições do aminoácido. Em destaque (vermelho) o local onde a alteração validada está localizada (aminoácido 1147 do domínio RIIIDB).

O domínio de RNase III (RIIID) é responsável pela reação endonucleolítica de enzimas de RNase III, como Dicer e Drosha. Os RIIIDs (denominados como RIIIDa e RIIIDb na Figura 11) são sítios bem conservados, encontrados em proteínas do tipo III-RNase de origem bacteriana, *Archea* e eucarióticas (HAN et al. 2004).

Muito ainda tem de ser esclarecido quanto à participação de miRNAs e de genes como *DROSHA* na participação na tumorigênese, tanto do TW quanto em vários outros tipos de câncer. Estudos subsequentes são necessários para avaliar o papel desta variação recorrente na tumorigênese de Wilms, e já estão em andamento no nosso grupo, fazendo parte de outro

projeto. Os resultados apresentados aqui mostram a necessidade de estudos futuros envolvendo o gene *DROSHA* para se elucidar o papel da biogênese dos miRNAs na tumorigênese de Wilms.

7 CONCLUSÃO

- ✓ Foram confirmadas as variações *de novo* e somática nos genes *PARD6B* e *DROSHA* respectivamente.
- ✓ A mutação no gene *DROSHA*, (g.chr5:31421465C>T) foi encontrada em 9,3% dos pacientes de TW avaliados, sugerindo desempenhar um papel fundamental e desencadeador da tumorigênese de Wilms.

8 REFERÊNCIAS BIBLIOGRÁFICAS

Alarcon VB. Cell polarity regulator PARD6B is essential for trophectoderm formation in the preimplantation mouse embryo. **Biol Reprod** 2010; 83:347-58.

Anglesio MS, Evdokimova V, Melnyk N, et al. Differential expression of a novel ankyrin containing E3 ubiquitin-protein ligase, Hace1, in sporadic Wilms' tumor versus normal kidney. **Hum Mol Genet** 2004; 13:2061-74.

Barros BDF. **Definição do perfil de alterações das sequências dos genes APC, CTNNB1, WT1, WTX e PLCG2 em tumores**. São Paulo; 2012. [Tese de Doutorado-Fundação Antônio Prudente].

Buckley KS. Pediatric genitourinary tumors. **Curr Opin Oncol** 2011; 23:297-302.

Breslow N, Beckwith JB, Ciol M, Sharples K. Age distribution of Wilms' tumor: report from the National Wilms' Tumor Study. **Cancer Res** 1988; 48:1653-7

Chakravarthy S, Sternberg SH, Kellenberger CA, et al. Substratespecific kinetics of *Dicer*-catalyzed RNA processing. **J Mol Biol** 2010; 404:392-402.

Cunliffe HE, Jiang Y, Meltzer PS. PAR6B is required for tight junction formation and activated PKC ζ localization in breast cancer. **Am J Cancer Res** 2012; 2:478-91.

Davies JA, Ladomery M, Hohenstein P, et al. Development of an siRNA-based method for repressing specific genes in renal organ culture and its use to show that the WT1 tumour suppressor is required for nephron differentiation. **Hum Mol Genet** 2004; 13:235-46.

Desai AN, Jere A. Next-generation sequencing: ready for the clinics ? **Clin Genet** 2012; 81:503-10.

De Camargo B, de Oliveira Santos M, Rebelo MS, et al. Cancer incidence among children and adolescents in Brazil: first report of 14 population-based cancer registries. **Int J Cancer** 2010; 126:715-20.

Diller L, Ghahremani M, Morgan J, et al. Constitutional WT1 mutations in Wilms' tumor patients. **J Clin Oncol** 1998; 16:11:3634-40

Dome JS, Coppes MJ. Recent advances in Wilms tumor genetics. **Curr Opin Pediatr** 2002; 14:5-11.

Guertl B, Leuschner I, Guelly C, Ebner B, Kronberger C, Hoefler G. Is predisposition for nephroblastoma linked to polymorphisms of the WTX gene? **Pathol Oncol Res** 2010; 16:189-91.

Gnarra JR, Tory K, Weng Y, et al. Mutations of the VHL tumor suppressor gene in renal carcinoma. **Nat Genet** 1994; 7:85-90.

Green DM, Coppes MJ, Breslow NE, et al. Wilms tumor. In: Pizzo PA, Poplack DG, editors. **Principles and practice of pediatric oncology**. Philadelphia: Lippincott-Raven; 1997. p.733-59.

Grundy PE, Green DM, Coppes MJ, et al. Renal tumors. In: Pizzo PA, Poplack DG, editors. **Principles and practice of pediatric oncology**. Philadelphia: Lippincott-Raven; 2002. p.865-93.

Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN. The Drosha–DGCR8 complex in primary microRNA processing. **Genes Dev** 2004; 18:3016-27.

Haruta M, Arai Y, Sugawara W, et al. Duplication of paternal IGF2 or loss of maternal IGF2 imprinting occurs in half of Wilms tumors with various structural WT1 abnormalities. **Gene Chromosomes Cancer** 2008; 47:712-27.

Ho J, Ng KH, Rosen S, Dostal A, Gregory RI, Kreidberg JA. Podocyte-specific loss of functional micrornas leads to rapid glomerular and tubular injury. **J Am Soc Nephrol** 2008; 19:2069-75.

Hu Z, Chen J, Tian T, et al. Genetic variants of miRNA sequences and non-small cell lung cancer survival. **J Clin Invest** 2008a; 118:2600-8.

Hu Z, Liang J, Wang Z, et al. Common genetic variants in pre-microRNAs were associated with increased risk of breast cancer in Chinese women. **Hum Mutat** 2008b; 30:79-84.

Huang CC, Gadd S, Breslow N, et al. Predicting relapse in favorable histology Wilms tumor using gene expression analysis: a report from the Renal Tumor Committee of the Children's Oncology Group. **Clin Cancer Res** 2009; 15:1770-8.

Huff V. Wilms tumor genetics. **Am J Med Genet** 1998; 79:260-7.

Huff V. Wilms' tumours: about tumour suppressor genes, an oncogene and a chameleon gene. **Nat Rev Cancer** 2011; 11:111-21.

Jazdzewski K, Murray EL, Franssila K, Jarzab B, Schoenberg DR, de la Chapelle A. Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. **Proc Natl Acad Sci USA** 2008; 105:7269-74.

Koesters R, Ridder R, Kopp-Schneider A, et al. Mutational activation of the betacatenin proto-oncogene is a common event in the development of Wilms' tumors. **Cancer Res** 1999; 59:3880-2.

Koesters R, Niggli F, von Knebel Doeberitz M, Stallmach T. Nuclear accumulation of beta-catenin protein in Wilms' tumours dagger. **J Pathol** 2003; 199:68-76.

Korinek V, Barker N, Morin PJ, et al. Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/colon carcinoma. **Science** 1997; 275:1784-7.

Kusafuka T, Miao J, Kuroda S, Udatsu Y, Yoneda A. Codon 45 of the β -catenin gene, a specific mutational target site of Wilms' tumor. **Int J Mol Med** 2002; 10:395-9.

Lai PS, Tay JSH. Wilms' tumor. In: Kurzrock R, Talpaz M, editors. **Molecular biology in cancer medicine**. Boca Raton: CRC Press; 1999. p.531-47.

Laurent-Puig P, Bérout C, Soussi T. APC gene: database of germline and somatic mutations in human tumors and cell lines. **Nucleic Acids Res** 1998; 26:269-70.

Lee Y, Jeon K, Lee JT, Kim S, Kim VN. MicroRNA maturation: stepwise processing and subcellular localization. **EMBO J** 2002; 21:4663-70.

Lee Y, Ahn C, Han J, et al. The nuclear RNase III Drosha initiates microRNA processing. **Nature** 2003; 425:415-9.

Lee Y, Han J, Yeom KH, Jin H, Kim VN. Drosha in primary microRNA processing. **Cold Spring Harb Symp Quant Biol** 2006; 71:51-7.

Li CM, Guo M, Borczuk A, et al. Gene expression in Wilms' Tumor mimics the earliest committed stage in the metanephric mesenchymal epithelial transition. **Am J Pathol** 2002; 160:2181-90.

Li W, Kessler P, Williams BR. Transcript profiling of Wilms tumors reveals connections to kidney morphogenesis and expression patterns associated with anaplasia. **Oncogene** 2005; 24:457-68.

Little J. **Epidemiology of childhood cancer**. Lyon: IARC; 1999. Introduction; p.1-9. (IARC Scientific Publications, n. 149).

Little M. Wilms' tumor: starting off the kidney all over again? **Prog Mol Subcell Biol** 2005; 40:107-32.

Maiti R, Alam R, Amos CI, Huff V. Frequent association of beta-catenin and WT1 mutations in Wilms tumors. **Cancer Res** 2000; 60:6288-92.

Martin ES, Cesari R, Pentimalli F, et al. The BCSC-1 locus at chromosome 11q23–q24 is a candidate tumor suppressor gene. **Proc Natl Acad Sci USA** 2003; 100:11517-22.

Maschietto M, de Camargo B, Brentani H, et al. Molecular profiling of isolated histological components of Wilms tumor implicates a common role for the Wnt signaling pathway in kidney and tumor development. **Oncology** 2008; 75:81-91.

Maschietto M, Trapé AP, Piccoli FS, et al. Temporal blastemal cell gene expression analysis in the kidney reveals new Wnt and related signaling pathway genes to be essential for Wilms tumor onset. **Cell Death Dis** 2011; 2:e224.

Maw MA, Grundy PE, Millow LJ, et al. A third Wilms' tumor locus on chromosome 16q. **Cancer Res** 1992; 52:3094-8.

Mendell JT, Olson EN. MicroRNAs in stress signaling and human disease. **Cell** 2012; 148:1172-87.

Messahel B, Williams R, Ridolfi A, et al. Allele loss at 16q defines poorer prognosis Wilms tumour irrespective of treatment approach in the UKW1-3 clinical trials: a Children's Cancer and Leukaemia Group (CCLG) Study. **Eur J Cancer** 2009; 45:819-26.

Ministério da Saúde. Instituto Nacional de Câncer. **Estimativa/2012: incidência do câncer no Brasil**. Rio de Janeiro: INCA; 2011.

Miyoshi Y, Iwao K, Nawa G, Yoshikawa H, Ochi T, Nakamura Y. Frequent mutations in the beta-catenin gene in desmoid tumors from patients without familial adenomatous polyposis. **Oncol Res** 1998; 10:591-4.

Moore AW, McInnes L, Kreidberg J, Hastie ND, Schedl A. YAC complementation shows a requirement for WT1 in the development of epicardium, adrenal gland and throughout nephrogenesis. **Development** 1999; 126:1845-57.

Negrini M, Nicoloso MS, Calin GA. MicroRNAs and cancer-new paradigms in molecular oncology. **Curr Opin Cell Biol** 2009; 21:470-9.

Nicholson GW. An embryonic tumor of the kidney in afoetus. **J Pathol Bacteriol** 1931; 34:711-30.

Nolan ME, Aranda V, Lee S, et al. The polarity protein par6 induces cell proliferation and is overexpressed in breast cancer. **Cancer Res** 2008; 68:8201.

Nusse R. Cancer: converging on β -Catenin in Wilms Tumor. **Science** 2007; 316:988-9.

Piccoli FS. **Identificação de potenciais marcadores de resposta terapêutica para o tumor de Wilms através de microarray usando plataforma de cDNA contendo genes pertencentes à via do ácido retinóico.** São Paulo; 2010. [Tese de Doutorado-Fundação Antônio Prudente].

Reis RS, Santos MO, Thuler LCS. Incidência de tumores pediátricos no Brasil. **Rev Bras Cancerol** 2007; 53:5-15.

Rivera MN, Haber DA. Wilms' tumour: connecting tumorigenesis and organ development in the kidney. **Nat Rev Cancer** 2005; 5:699-712.

Rivera MN, Kim WJ, Wells J, et al. An X chromosome gene, WTX, is commonly inactivated in Wilms tumor. **Science** 2007; 315:642-5.

Robinson PN, Whole-exome sequencing for finding *de novo* mutations in sporadic mental retardation. **Genome Biol** 2010; 11:144.

Rong Y, Cheng L, Ning H, et al. Wilms' tumor 1 and signal transducers and activators of transcription 3 synergistically promote cell proliferation: a possible mechanism in sporadic Wilms' tumor. **Cancer Res** 2006; 66:8049-57.

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; 47:461-70.

Sarah BN, Buckingham KJ, Lee C. et al. Exome sequencing identifies the cause of a mendelian disorder. **Nat Genet** 2010; 42:30-5.

Schedl A. Renal abnormalities and their developmental origin. **Nat Rev Genet** 2007; 8:791-802.

Skotnicka-Klonowicz G, Rieske P, Bartkowiak J, Szymik-Kantorowicz S, Daszkiewicz P, Debiec-Rychter M. 16q heterozygosity loss in Wilms' tumour in children and its clinical importance. **Eur J Surg Oncol** 2000; 26:61-6.

Sparks AB, Morin PJ, Vogelsteins B, Kinzler KW. Mutation analysis of the APC/beta-catenin/Tcf pathway in colorectal cancer. **Cancer Res** 1998; 58:1130-4.

Strong LC. The two-hit model for Wilms' tumor: where are we 30 years later?. **Genes Chromosomes Cancer** 2003; 38:294-9.

Sugiyama H. WT1 (Wilms' tumor gene 1): biology and cancer immunotherapy. **Jpn J Clin Oncol** 2010; 40:377-87.

Takahashi M, Yang XJ, Lavery TT, et al. Gene expression profiling of favorable histology Wilms tumors and its correlation with clinical features. **Cancer Res** 2002; 62:6598-605.

Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz Jr LA, Kinzler KW. Cancer genome landscapes. **Science** 2013; 339:1546-58.

Welch SR, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. **Cell** 2012; 150:264-78.

Williams RD, Hing SN, Greer BT, et al. Prognostic classification of relapsing favorable histology Wilms tumor using cDNA microarray expression profiling and support vector machines. **Genes Chromosomes Cancer** 2004; 41:65-79.

Willis RA. **The borderline of embryology and pathology**. Reino Unido; Butterworths; 1962.

Wu MK, Sabbaghian N, Xu B, et al. Biallelic DICER1 mutations occur in Wilms tumours. **J Pathol** 2013; 230:154-64.

Yunis JJ, Ramsay NK. Familial occurrence of the aniridia-Wilms tumor syndrome with deletion 11p13-14.1. **J Pediatr** 1980; 96:1027-30.

Zhang L, Anglesio MS, O'Sullivan M, et al. The E3 ligase HACE1 is a critical chromosome 6q21 tumor suppressor involved in multiple cancers. **Nat Med** 2007; 13:1060-9.

Zirn B, Hartmann O, Samans B, et al. Expression profiling of Wilms tumors reveals new candidate genes for different clinical parameters. **Int J Cancer** 2006; 118:1954-62.

Anexo 1 - Termo de Consentimentos Livre Esclarecido (TCLE) assinado pelos pacientes que cederam material biológico utilizado neste projeto.

HOSPITAL AC CAMARGO – Fundação Antonio Prudente

R: Professor Antônio Prudente, 211 – telefone: 21895000.

TERMO DE CONSENTIMENTO LIVRE ESCLARECIDO

(Obrigatório para Pesquisa Clínica em Seres Humanos – Resolução N. 196/96 e resolução CNS 251/97 do Ministério da Saúde)

DADOS DE IDENTIFICAÇÃO DO PACIENTE E/OU RESPONSÁVEL LEGAL

NOME DO PACIENTE: _____

Data de nascimento: ____/____/____

Documento de identidade nº. _____

Endereço: _____

Número: _____ Complemento: _____

Cidade: _____ Estado: _____

I- TÍTULO

ESTUDO DE POLIMORFISMO DOS GENES DO METABOLISMO DOS FOLATOS NOS TUMORES-EMBIONÁRIOS (NEUROBLASTOMA, TUMOR DE WILMS, RETINOBLASTOMA E MEDULOBLASTOMA)

II. OBJETIVO DO ESTUDO:

O Objetivo deste estudo é procurar entender os caminhos da formação dos tumores malignos através dos defeitos das células que possam estar relacionados com o início dos tumores e com diversas perguntas sobre a gestação, hábitos, estilo de vida procurando relação desses fatores com o desenvolvimento do tumor.

fatos que ocorreram durante a gravidez, e o modo de vida da Sra e de seu marido em casa e no trabalho.

IV. BENEFÍCIOS:

A sua participação neste estudo é voluntária, tendo o direito a retirar-se do estudo a qualquer momento. Sua recusa ou desistência não irá prejudicar o acompanhamento. Desta forma, não está prevista nenhuma forma de compensação econômica pela sua participação.

Todas as informações recebidas serão consideradas confidenciais e serão utilizadas somente para finalidade deste estudo, respeitando-se a não divulgação de quaisquer dados relacionados à identificação do paciente. Apenas os membros da equipe de saúde e da Comissão de Ética terão acesso aos registros.

Qualquer dúvida sobre o estudo, você poderá entrar em contato com a Enfermeira Débora Ugayama Bassi no telefone 9915-0957. Se o pesquisador principal (Médica Dra. Beatriz no 50528662) não fornecer as informações/esclarecimentos suficientes, por favor, entre em contato com o Coordenador do Comitê de Ética do HOSPITAL AC CAMARGO-SP, pelo telefone 2189 5000, ramais 5020 ou 2069.

As amostras coletadas serão processadas e armazenadas pelo Banco de macromoléculas do Biobanco do Hospital A C Camargo, e o excedente não utilizado por este projeto de pesquisa poderá ser utilizado em outros projetos no futuro, após a aprovação dos mesmos pelo Comitê de Ética em Pesquisa desta Instituição.

III. PROCEDIMENTOS:

Se você concordar em participar do estudo, será coletada 05 ml de sangue, ou 1 amostra de seu/sua filho (a) e do seu marido/ esposa de saliva (correspondente a uma colher de sopa). Desta retiraremos o DNA do material coletado, exclusivamente destinado a análise que nos permitam identificar os polimorfismos (isto é, alterações) de alguns genes, citados anteriormente. O Sr/ Sra também irá responder a um questionário, que pede informações sobre locais onde viveu com seu filho (a), aos

Declaro que fui esclarecido sobre os procedimentos, riscos e benefícios deste estudo; que tenho liberdade em retirar meu consentimento a qualquer momento, sem que isto traga prejuízo ao meu tratamento/acompanhamento; que não haverá remuneração financeira para este estudo; que minha identidade será preservada, mantendo-se todas as informações em caráter confidencial. Concordo em participar deste estudo.

São Paulo, ____ de ____ de ____

Assinatura do paciente ou responsável/representante local

Anexo 2 - Resultados das alterações somáticas selecionadas para validação, obtidos através das análises no Mozaik, CLC e Sanger. Chr – Cromossomo.

Gene	Chr	Mozaik Assembler				CLC Genomics Workbench						Sequenciamento capilar (Sanger)		
		Posição (Hg19)	Alteração	% alterado /referência	Classif	Posição (Hg19)	Alteração	Proteína	Reads alteradas	Cobertura	Freq.	Posição (Hg19)	Alteração	Comparação CLC/Sanger
ASPM	1	197072421	del T	13%	frameshift	197072420	T>C	Não ocorre troca de aminoácido (Gln-Gln)	55	107	51,4	197072420	T>C	Concordantes
C17orf87	17	5126690	del A	17%	frameshift	5126689	G>A	Não ocorre troca de aminoácido (Ile-Ile)	89	89	100	5126689	WT	Discordantes
EIF2B4	2	27591291	del A	13%	frameshift	27591291	WT	-	-	-	-	27591291	WT	Concordantes
FAM154A	9	18950896	del AT	13%	frameshift	18950895	T>C	p.[Lys27Glu];[=]	32	57	56,1	18950895	T>C	Concordantes
FAM48B1	X	24381388	del A	14%	frameshift	24381387	G>C	Não ocorre troca de aminoácido (Thr-Thr)	14	27	51,9			
FASTKD1	2	170403033	del G	13%	frameshift	170403030	A>G	p.Met467Val	66	66	100	170403030	A>G	Concordantes
HIVEP3	1	42046190	C>A	52%	non-sense	42046190	C>A	p.(Gly1427*)	10	99	10,1	42046190	WT	Discordantes
LAMA3	18	21456314	del T	40%	frameshift	21456314	WT	-	-	-	-	21456314	WT	Concordantes
PCDHAC1	5	140389457	del A	18%	frameshift	140389457	WT	-	-	-	-	140389457	WT	Concordantes
PKD1L1	7	47920346	del T	20%	frameshift	47920345	G>A	Não ocorre troca de aminoácido (Tyr-Tyr)	21	28	75	47920345	G>A	Concordantes
PTPLA	10	17636311	del G	19%	frameshift	17636309	G>A	p.[=];[His227Tyr]	25	39	64,1	17636309	G>A	Concordantes
RIF1	2	152292055	del A	19%	frameshift	152292055	WT	-	-	-	-	152292055	WT	Concordantes
SLC26A5	7	103019700	A>T	15%	non-sense	103019700	A>T	p.(Leu524*)	2	74	2,7	103019700	WT	Discordantes
UAP1	1	162549315	del CA	20%	frameshift	162549315	WT	-	-	-	-	162549315	WT	Concordantes
WDR72	15	54015022	G>T	19%	non-sense	54015022	WT	-	-	-	-	54015022	WT	Concordantes
ZNF208	19	22157303	del T	83%	frameshift	22157302	G>A	Não ocorre troca de aminoácido (Tyr-Tyr)	7	8	87,5	22157302	G>A	Concordantes
ZNF583	19	56935528	del T	13%	frameshift	56935527	A>G	Não ocorre troca de aminoácido (Ala-Ala)	57	57	100			
ZNF761	19	53957912	del T	27%	frameshift	53957912	WT	-	-	-	-	53957912	WT	Concordantes

Anexo 3 - Resultados das alterações *de novo* selecionadas para validação, obtidos através das análises no Mozaik, CLCBio e Sanger.
Chr – Cromossomo.

		Mozaik Assembler					CLC Genomics Workbench						Sequenciamento capilar (Sanger)		
Gene	Chr	Posição (Hg19)	Alteração	% alterado/referência		Classif	Posição (Hg19)	Alteração	Proteína	Reads alteradas	Cobertura	Freq	Posição (Hg19)	Alteração	Comparação CLC/Sanger
				Fs	Ft										
ELAVL2	9	23692800	del A	12%	3%	frameshift	23692800	WT	-	-	-	-			
ETAA1	2	67632126	del C	8%	15%	frameshift	67632125	C>T	p.Pro771Ser	30	30	100	67632125	C>T	Concordantes
EXPH5	11	108384208	del T	2%	16%	frameshift	108384207	C>T	p.[=];[Ser676Asn]	21	36	58,3	108384207	C>T	Concordantes
GGNBP2	17	34942496	del T	13%	13%	frameshift	34942496	WT	-	-	-	-	34942496	WT	Concordantes
GUCY2F	X	108708517	del G	1%	25%	frameshift	108708516	C>T	p.[Arg296Gln];[=]	12	20	60			
KCTD3	1	215793835	del TC	9%	3%	frameshift	215793834	G>A	Não ocorre troca de aminoácido (Ala-Ala)	44	44	100	215793834	G>A	Concordantes
KIF2B	17	51902253	del T	1%	15%	frameshift	51902252	C>T	Não ocorre troca de amoniácido (Leu-Leu)	37	37	100			
MAGI2	7	77708364	del T	2%	14%	frameshift	77708364	WT	-	-	-	-			
MGA	15	42054337	del A	1%	11%	frameshift	42054337	WT	-	-	-	-			
MUC16	19	9072743	del G	19%	13%	frameshift	9072742	T>C	p.[=];[Ile4902Val]	29	45	64,4	9072742	WT	Discordantes
OR51I2	11	5475434	del A	2%	15%	frameshift	5475434	WT	-	-	-	-			
PAFAH1B2	11	117031934	del T	9%	7%	frameshift	117031934	WT	-	-	-	-			
PSKH2	8	87076652	del G	1%	17%	frameshift	87076651	T>C	p.Gln132Arg	27	35	77,1	87076651	T>C	Concordantes
TAS2R19	12	11175083	del C	6%	13%	STOP CODON	11175083	WT	-	-	-	-			
TAS2R8	12	10958659	del G	10%	17%	STOP CODON	10958658	T>C	p.Met308Val	14	15	93,3	10958658	T>C	Concordantes
ZNF227	19	44739304	del C	9%	18%	frameshift	44739303	T>C	Não ocorre troca de aminoácido (Asn-Asn)	32	32	100	44739303	WT	Discordantes
ZNF619	3	40529355	del G	14%	5%	frameshift	40529355	WT	-	-	-	-	40529355	WT	Concordantes
ZNF845	19	53856637	del A	14%	4%	frameshift	53856637	WT	-	-	-	-			