Marilia Moreira Montenegro

Estudo do transcriptoma na síndrome de Bloom

Tese apresentada à Faculdade de Medicina da
Universidade de São Paulo para obtenção do título
de Doutor em Ciências

Programa de Patologia

Orientadora: Dr.^a Leslie Domenici Kulikowski

Coorientadora: Prof.^a Dr.^a Chong Ae Kim

São Paulo

2016

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Montenegro, Marília Moreira

Estudo do transcriptoma na síndrome de Bloom / Marília Moreira
Montenegro. -- São Paulo, 2016.

Tese(doutorado)--Faculdade de Medicina da Universidade de São Paulo.
Programa de Patologia.

Orientadora: Leslie Domenici Kulikowski.

Coorientadora: Chong Ae Kim.

Descritores: 1.Síndrome de Bloom 2.Instabilidade cromossômica
3.Transcriptoma 4.Análise de sequência de RNA 5.Sequenciamento de
nucleotídeos em larga escala 6.Expressão gênica

USP/FM/DBD-486/16

Esta dissertação ou tese está de acordo com as seguintes normas, em vigor no momento desta publicação:

Referências: adaptado de *International Committee of Medical Journals Editors* (Vancouver).

Universidade de São Paulo. Faculdade de Medicina. Divisão de Biblioteca e Documentação. *Guia de apresentação de dissertações, teses e monografias*. Elaborado por Anneliese Carneiro da Cunha, Maria Julia de A. L. Freddi, Maria F. Crestana, Marinalva de Souza Aragão, Suely Campos Cardoso, Valéria Vilhena. 3a ed. São Paulo: Divisão de Biblioteca e Documentação; 2011.

Abreviaturas dos títulos dos periódicos de acordo com *List of Journals Indexed in Index Medicus*.

*Dedico este trabalho aos meus pais que
sempre lutaram comigo e me criaram de uma
forma mais que especial, ao meu marido que
me apoia incondicionalmente e ao nosso filho
maravilhoso que nos ensinou o maior
significado do amor... “Sempre juntos.
Simples assim!”*

AGRADECIMENTOS

Agradeço,

Ao meu marido, **Alexandre (Vivi)**, por me apoiar demais em tudo, por me mandar tantas energias positivas, por dizer todos os dias que não estou sozinha, por tantos momentos de alegria e risadas, por se sacrificar e sempre entender que às vezes preciso ficar estudando, por ser meu porto seguro, meu amigo, meu amor, minha vidinha. Obrigada pelas risadas e pelos momentos quietinhos também. E, principalmente, agradeço todos os dias por ter me dado o melhor de você, que é nosso pequeno *Guga*. Obrigada por investir tanto em mim e cuidar tão carinhosamente da nossa família. Te amo demais... simples assim!

Ao nosso filho, **Gustavo**, um bebê lindo, esperto, feliz e risonho. Muitas vezes precisei sacrificá-lo com minha ausência para finalizar este trabalho, mas mesmo assim sempre fui recebida com um abraço delícia, beijos em forma de mordidas e muita festa. Nem sei como pode caber tanto amor em meu coração, você foi o melhor que a vida podia me dar. Te amo com toda força!

Aos meus pais, **Áurea e Andriw**, por sempre me apoiarem e darem a maior força para que eu corresse atrás de todo tipo de sonho. Independentemente dos meus erros ou acertos sempre acreditaram na minha capacidade, até mesmo quando eu duvidava. Obrigada também por aguentarem meu humor... Sem contar o amor e carinho com que cuidam do Guga para que eu possa estudar! “Eu também amo vocês”!

Aos meus familiares, **Todos! Especialmente minha Madrinha Ilanil**, por tanto amor, ensinamentos, momentos de risadas e por pegar frequentemente no meu pé para eu

investir sempre mais nos meus estudos - “Tu tens que ralar Guria”! E ao meu **Tio Marcão** pelas conversas, conselhos, e “crases” (rs)... **Obrigada!!!**

Às minhas Avós, **Hylanou e Cyrene**, por rezarem todos os dias para que tudo na minha vida dê certo!

À família Cabral Montenegro, **Malu, Rogério** (*in memorian*), **Edu, Jô e Bê**, por torcerem pela minha carreira de pesquisadora e por ótimos momentos juntos. E especialmente à **Malu** por me ajudar a cuidar do Guga para que eu pudesse estudar e por fazer tantos bolos deliciosos.

À minha querida amiga e Orientadora, **Dra. Leslie Domenici Kulikowski**, por chegar de mansinho e se tornar uma enorme professora e amiga. Por se esforçar tanto para fazer nosso grupo crescer, por me fazer amar a pesquisa ainda mais, pelas grandes contribuições para este trabalho, pelas incansáveis horas de “brain storming”, por sempre me puxar para cima, por sempre ter uma palavra de carinho e “ouvidos” maravilhosos, e também pelas broncas e críticas construtivas que me ajudaram a crescer como pessoal e profissionalmente! Obrigada por nunca desistir de nós!!! Agradeço também o **Jo**, pela paciência e por sempre tornar nossos encontros e congressos muito mais animados, principalmente os internacionais, rs!

À amiga e co-orientadora, **Dra. Chong Ae Kim**, por ter me acolhido há 10 anos com tanta confiança e pela paciência de esperar meu amadurecimento acadêmico. Por ter me ensinado muitas lições de vida, por fazer da equipe de genética uma enorme família e por ter dado tantas “tarefinhas” que me ajudaram a crescer em todos os aspectos da minha vida. Obrigada por ter esse coração tão grande!

Aos queridos amigos Citogenômicos, **Fabricia, Thais, Cintia, Samar, Juliana e Larissa**, por me ajudarem sempre que preciso. E, especialmente à “velha guarda”: **Evelin**, pelo

“delicioso” pão na chapa de quase todas as manhãs, por tanto carinho, risadas e fofocas; **Alê**, por ser nosso nono querido e grande conciliador; **Gil**, pelo imenso companheirismo, risada, fofoca e “maravilhosos” almoços no bandejão; **Bob**, por sempre ter grandes ideias e ser tão alto astral; e **Amon**, pelas ótimas piadocas. Obrigada por se tornarem uma extensão da minha família, pelas experiências trocadas, por me socorrerem sempre que foi preciso, por embarcarem comigo em todas as “tarefinhas” que arrumo e naquelas arrumadas por vocês também. À **Yanca**, “minha” estagiária, que me ensinou muito mais do que pude ensinar a ela, por me proporcionar a alegria de vê-la crescer com inteligência e graça a cada dia! Também agradeço demais ao **Julian**, que, mesmo me conhecendo há pouco tempo, me ajudou demais nesta tese ao revisar meus resultados e me ajudar com os gráficos. Obrigada amigos, sem vocês não teria tido tanta graça!

Aos queridos amigos **médicos e residentes** da equipe de genética, por avaliarem os pacientes deste estudo e estarem sempre acessíveis para me ajudar com inúmeras discussões de caso. E, especialmente à **Raquel** e ao **Diogo** pela amizade, desabafos, trabalhos tão produtivos, risadas e correções de artigos. Obrigada por sempre estarem disponíveis para me ajudar com qualquer coisa.

À **Profa. Dra. Magda Carneiro Sampaio**, por sempre acreditar em nosso grupo, nos ajudar a crescer e nos disponibilizar um local para que nossos estudos pudessem ser realizados. Muito obrigada!

Ao **Prof. Carlos Alberto Moreira Filho**, por gentilmente ceder um lugar para nossos estudos.

Aos **colegas do LIM36**. Em especial as queridas amigas, **Paty, Ana e Layla**, pela amizade e companheirismo, almoços animados, por estarem sempre disponíveis para discutir resultados comigo e me ensinarem imunologia. À **Andréa**, por ter me ajudado e ensinado a fazer o sequenciamento Sanger e analisar os resultados. À **Dra. Estela**, por sempre mandar ótimas energias e enfatizar que sou boa no que faço, sem me deixar

esquecer disso e desanimar. À **Thamiris**, aluna da Dra. Estela que me ajudou com a validação do gene estudado neste trabalho e por ser uma fofa. À **Fernanda**, técnica do LIM36, pelos grandes momentos de risadas sem parar, por ser sempre extremamente prestativa, por dividir muitos cosméticos comigo (rs), mas principalmente por sua amizade de longa data. Obrigada Meninas, sem vocês não teria dado tudo tão certo!

Ao **Prof. Dr. Raymundo Soares de Azevedo Neto**, por ser meu mentor durante minha representação discente e pela confiança depositada nesse projeto. Muito obrigada.

Ao funcionário **Thiago** do Departamento de Patologia, por ser sempre muito atencioso mesmo quando eu não parava de mandar e-mails, ou fazer ligações. Agradeço também à funcionária **Welluma**, por colaborar com a minha entrada no Doutorado.

Ao Bioinformata **Daniel Silveste**, por ter se disponibilizado a processar todos os dados deste trabalho e proporcionado horas de “viagem” em cima dos dados.

Ao **Prof. Luiz Leman Coutinho**, por gentilmente ter disponibilizado seu laboratório para que pudéssemos aprender as técnicas de sequenciamento utilizadas neste trabalho e em muitos outros. E a toda sua equipe, especialmente à **Marcela Paduam**, **Sônia Andrade** e **Gustavo Gasparin**, por serem muito receptivos comigo e me ajudarem a realizar os sequenciamentos e as análises de dados de expressão.

À **Dra. Maria Isabel Melaragno e sua equipe**, pela colaboração no desenvolvimento deste projeto e parceria em todos os congressos. E, em especial, à **Anelise** por me auxiliar com as extrações no Qiacube. Obrigada, vocês foram muito importantes.

À **Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)** pelo apoio financeiro.

Agradeço principalmente a *Todos os pacientes e seus familiares*, que lutam incansavelmente contra o tempo e tentam viver da melhor forma possível.

E por fim, agradeço à *vida*, por ser tão generosa comigo. Como sou sortuda!!!

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SÍMBOLOS

RESUMO

ABSTRACT

1 INTRODUÇÃO	01
1.1 Aspectos clínicos	02
1.2 Aspectos moleculares	04
1.2.1 Integridade celular e instabilidade genômica	04
1.2.2 A SB e o gene <i>BLM</i>	08
1.2.3 Mutações envolvendo o gene <i>BLM</i>	14
1.3 Aspectos citogenéticos	15
1.3.1 Quebras cromossômicas e cromatídicas	15
1.3.2 Configuração quadrirradial	17
1.3.3 Troca entre cromátides irmãs	17
1.4 Transcriptoma aplicado ao estudo da síndrome de Bloom	20
2 OBJETIVOS	22
2.1 Objetivos específicos	23
3 METODOLOGIA	24
3.1 Casuística	25
3.1.2 Quadro Clínico	26
3.1.2.1 Paciente K1	26
3.1.2.2 Paciente M1	27
3.2 Métodos	27
3.2.1 Estudo citogenético	29
3.2.2 Pesquisa de mutações do gene <i>BLM</i> por sequenciamento Sanger	30
3.2.3 Análise do transcriptoma pela metodologia de RNA-seq	32
3.2.3.1 Abordagem de análise A1 – Expressão diferencial (programas Bowtie2 e EdgeR)	34
3.2.3.2 Abordagem de análise A2 – Análise de variantes genômicas (programas Rsubread e DESeq2)	35
3.2.4 Validação da expressão do gene <i>BLM</i> pela metodologia de RT-qPCR	36
4 RESULTADOS	38
4.1 Análise citogenética	39
4.2 Análise mutacional pela metodologia Sanger	42

4.3 Análise do transcriptoma pela metodologia de RNA-seq	47
4.3.1 Expressão diferencial (abordagem A1 de análise)	48
4.3.2 Análise de variantes (abordagem A2 de análise)	50
4.3.2.1 Paciente K1	50
4.3.2.2 Paciente M1	57
4.4 Validação do gene <i>BLM</i> pela metodologia de RT-qPCR	61
5 DISCUSSÃO	63
5.1 Análise citogenética	64
5.2 Análise mutacional pela metodologia de Sanger	65
5.3 Análise do transcriptoma	67
5.3.1 Análise de expressão diferencial entre SB x controles	67
5.3.2 Análise de variações genômicas pela metodologia de RNA-seq	70
5.3.3 Transcriptoma aplicado a SB	70
5.4 Validação do gene <i>BLM</i> por RT-qPCR	72
6 CONSIDERAÇÕES FINAIS	73
7 CONCLUSÕES	75
ANEXO A – Parecer aprovado da comissão de Ética para Análise de Projeto de Pesquisa (Cappesq)	78
ANEXO B – Termo de Consentimento Livre e Esclarecido	79
ANEXO C – Análises de predição de proteína para as mutações encontradas na Paciente K1	83
ANEXO D - Análises de predição de proteína para as mutações encontradas na Paciente M1	98
9 REFERÊNCIAS	106
APÊNDICE A – Artigo submetido para publicação	
APÊNDICE B – Artigo publicado durante o desenvolvimento deste doutorado	
APÊNDICE C – <i>Abstracts</i> de artigos publicados durante este doutorado, que não estão diretamente relacionados a temática do presente estudo

LISTA DE FIGURAS

Figura 1 -	Danos sofridos pelo DNA, mecanismos de reparo de DNA e as consequências sofridas pela célula. a , Agentes danificadores de DNA (topo); Exemplos de lesões que ocorrem no DNA induzidas por tais agentes (meio) e os mecanismos de reparo de DNA mais relevantes para a remoção e correção destas lesões (abaixo). b , Efeitos agudos do DNA danificado na progressão do ciclo celular levando a parada do ciclo nas fases G1, S, G2 e M (topo), os efeitos no metabolismo de DNA (meio) e as consequências a longo prazo (abaixo), incluindo mudanças permanentes na sequência de DNA e seus efeitos biológicos. (Adaptado de Hoeijmakers, 2001).....	05
Figura 2 -	Mecanismos de reparo de quebra na dupla fita de DNA: Reparo por recombinação homóloga (HR – do inglês <i>Homologous Recombination</i>) (à esquerda) e o reparo por junção de extremidades não homólogas (NHEJ – do inglês <i>Non-Homologous End Joining</i>) (à direita). (Adaptado de Hoeijmakers, 2001).....	07
Figura 3 -	Representação esquemática da família RecQ helicase em diferentes espécies, desde humanos até leveduras. Os domínios estão alinhados em um código de cores, como indicado na parte inferior da figura: (Verde) domínio exonuclease; (Vermelho escuro) domínio ácido; (Amarelo) domínio helicase, altamente conservado entre as espécies; (Preto) domínio HRDC; (Vermelho claro) sinal de localização nucleolar e (Rosa) domínio RQC. O tamanho em aminoácidos da proteína de cada espécie está indicado à direita. O nome das... proteínas e das espécies de origem está indicado à esquerda. (Adaptado de Singh et al., 2009)....	09
Figura 4 -	Estrutura gráfica da proteína BLM.....	11
Figura 5 -	Representação esquemática da via de reparo por recombinação homóloga em procariotos e eucariotos. Em eucariotos é possível observar o momento em que o gene <i>BLM</i> é considerado atuante, marcado por uma estrela em vermelho, incluindo o processamento das extremidades quebradas da dupla fita (estrela 1), a abertura do D-loop que foi formado para a invasão da fita simples na forquilha de replicação da cromátide irmã (estrela 2), processamento, dissolução de <i>Holliday Junction</i> (estrela 3) e supressão de produtos com crossing over na síntese dependente de anelamento de fita simples (estrela 4 e 5). (Adaptado de www.genome.jp/kegg).....	12
Figura 6 -	Alterações cromossômicas observadas como resultado de instabilidade em células de portadores de SB: (A) Quebra cromossômica; (B) Quebra cromatídica; (C) Gap cromossômico; (D) Gap cromatídico; (E) Cromossomo dicêntrico; (F) Fragmentos; (G) Rearranjo; (H) Pulverização cromossômica, (I) Configuração quadrirradial e (J) Troca entre cromátides irmãs. (Laboratório de Citogenômica do LIM03).....	15
Figura 7 -	Demonstração de troca entre cromátides irmãs. À esquerda (A) é possível observar um controle normal, com apenas uma troca. À direita (B) observa-se um controle positivo para a síndrome de Bloom, demonstrando frequência aumentada de trocas entre cromátides irmãs. (Killen, 2011).....	18
Figura 8 -	Esquema gráfico demonstrando a estrutura química do BrdU (acima, no centro da figura), e sua incorporação em uma das fitas da dupla fita de DNA por dois ciclos celulares. (https://bio-ggs.blogspot.com.br/2010/05/meet-chickens-sister-chromatid.html).....	19

Figura 9 -	Fluxograma da metodologia utilizada no presente estudo.....	28
Figura 10 -	Exemplo de alterações cromossômicas observadas sob análise citogenética convencional da paciente K1. (a) Gap cromatídico – chtg e (b) Cromossômo dicêntrico.....	40
Figura 11 -	Exemplo de uma metáfase sob coloração diferencial de cromátides irmãs da paciente K1. A figura demonstra a frequência aumentada de trocas entre cromátides irmãs.....	40
Figura 12 -	Exemplo de alterações cromossômicas observadas sob análise citogenética convencional da paciente M1. (a) Gap cromatídico – chtg e (b) Cromossômo dicêntrico.....	41
Figura 13 -	Exemplo de uma metáfase sob coloração diferencial de cromátides irmãs da paciente M1.....	41
Figura 14 -	Eletroferograma da análise do sequenciamento Sanger, demonstrando uma troca de uma Adenina (A) por Guanina (G) na posição 3102 do éxon 16 do gene <i>BLM</i> (c.3102G>A:p.Thr1034=), observada na paciente K1.....	42
Figura 15 -	Eletroferograma da análise do sequenciamento Sanger, demonstrando uma troca de uma Citosina (C) por Adenina (A) na posição 3531 do éxon 18 do gene <i>BLM</i> (c.3531C>A:p.Ala1177=), observada na paciente K1.....	43
Figura 16 -	Eletroferograma da análise do sequenciamento Sanger, demonstrando uma troca de uma Citosina (C) por Timina (T) na posição 3945 do éxon 21 do gene <i>BLM</i> c.3945C>T:p.Leu1315=), observada na paciente K1.....	43
Figura 17 -	Eletroferograma da análise do sequenciamento Sanger, demonstrando uma troca de uma Timina (T) por Guanina (G) na posição 3358 no íntron entre os éxons 17 e 18 do gene <i>BLM</i> (c.3358+32T>G), observada na paciente K1.....	44
Figura 18 -	Eletroferograma da análise do sequenciamento Sanger, demonstrando uma variação do tipo <i>missense</i> na posição 3164 do éxon 16 do gene <i>BLM</i> (c.3164G>C:p.Cys1055Ser), observada na paciente M1.....	45
Figura 19 -	Eletroferograma da análise do sequenciamento Sanger, demonstrando uma variação do tipo <i>missense</i> na posição 3625 do éxon 19 do gene <i>BLM</i> (c.3625T>A:p.Ser1209Thr), observada na paciente M1.....	45
Figura 20 -	<i>Maplot</i> demonstrando a distribuição de <i>foldchange</i> para os 11.939 genes adequados para análise de expressão diferencial. Cada ponto representa um gene. Genes qualificados como diferencialmente expressos estão em vermelho...	49
Figura 21 -	Mapa do tipo <i>circus plot</i> demonstrando uma visão geral das variações genômicas observadas pelo estudo do transcriptoma pela metodologia de RNA-seq pela abordagem de análise A2. Lendo-se de fora para dentro: Cada um dos cromossomos 1-22, X e Y estão representados pelo ideograma disposto circularmente, demarcado por uma escala de megabase na borda exterior da figura; logo em seguida, em vermelho, estão as Indels de alto impacto, em laranja as de médio impacto e em amarelo as de baixo impacto; em seguida encontram-se os tipos de variantes encontradas (círculo cinza), seguido pelos SNPs (círculo azul), impacto dos SNPs (círculo amarelo) e os tipos das variantes encontradas (círculo cinza mais interno). A legenda desta figura está descrita por completo na Figura 22	52
Figura 22 -	Legenda do <i>circus plot</i> da Figura 21	53
Figura 23 -	Mapa do tipo <i>circus plot</i> demonstrando uma visão geral das variações genômicas observadas, apenas no gene <i>BLM</i> , pelo estudo do transcriptoma pela metodologia de RNA-seq pela abordagem de análise A2. A legenda desta figura está descrita por completo na Figura 24	54

Figura 24 -	Legenda do <i>circus plot</i> da Figura 23	55
Figura 25 -	Mapa do tipo <i>circus plot</i> demonstrando uma visão geral das variações genômicas observadas pelo estudo do transcriptoma pela metodologia de RNA-seq pela abordagem de análise A2. Lendo-se de fora para dentro: Cada um dos cromossomos 1-22, X e Y estão representados pelo ideograma disposto circularmente, demarcado por uma escala de megabase na borda exterior da figura; logo em seguida, em vermelho, estão as Indels de alto impacto, em laranja as de médio impacto e em amarelo as de baixo impacto; em seguida encontram-se os tipos de variantes encontradas (circulo cinza), seguido pelos SNPs (circulo azul), impacto dos SNPs (circulo amarelo) e os tipos das variantes encontradas (circulo cinza mais interno). A legenda desta figura está descrita por completo na Figura 26	58
Figura 26 -	Legenda do <i>circus plot</i> da Figura 25	59
Figura 27 -	Gráfico demonstrando a análise da quantificação relativa da expressão do gene <i>BLM</i> observada nas pacientes K1 e M1. Os níveis relativos estão representados por barras.....	62
Figura 28 -	Figura demonstrando o esquema de chromotripsis: quando um evento de catástrofe cromossômica (como uma explosão) causa a fragmentação do DNA que é subsequentemente reparado levando a um rearranjo do cromossomo. Esse rearranjo pode levar a perda de algum segmento. (Tubio e Estivill, 2011).....	65

LISTA DE TABELAS

Tabela 1 -	Descrição das amostras que compõe a casuística do presente estudo.....	25
Tabela 2 -	Estudo citogenético para confirmação diagnóstica da síndrome de Bloom, baseado em 20 células analisadas para coloração convencional e bandamento G, e em 10 células analisadas para trocas entre cromátides irmãs.....	39
Tabela 3 -	Descrição das variantes observadas no gene <i>BLM</i> pela metodologia de sequenciamento Sanger para as pacientes K1 e K2.....	46
Tabela 4 -	Dados brutos gerados pela Abordagem de análise A1.....	47
Tabela 5 -	Dados brutos gerados pela Abordagem de análise A2.....	48
Tabela 6 -	Distribuição das variantes encontradas nos transcritos do gene <i>BLM</i>	56
Tabela 7 -	Distribuição das variantes encontradas nos transcritos do gene <i>BLM</i> da paciente M1	60

LISTA DE SIGLAS

A	Adenina
BG	Bandamento G
BrdU	5-bromo-2'-desoxiuridina
C	Citosina
cDNA	DNA complementar
CHG	Coloração Hoechst-Giemsa
CNVs	Variação no número de cópias (do inglês – <i>Copy Number Variations</i>)
CT	Número de ciclos de PCR
CV	Coloração convencional
DAVID	<i>Database for Annotation, Visualization and Integrated Discovery</i>
DE	Diferencialmente expressos
DNA	Ácido desoxirribonucleico
DSB	Quebra na dupla-fita de DNA (do inglês – <i>Double-Strands Break</i>)
E.coli	<i>Escherichia Coli</i>
ESALQ	Escola Superior de Agronomia Luiz de Queiroz
FMUSP	Faculdade de Medicina da Universidade de São Paulo
G	Guanina
GO	<i>Gene Ontology</i>
HD	Hipótese Diagnóstica
HJ	<i>Holliday Junctions</i>
HR	Recombinação Homóloga (do inglês – <i>Homologous Recombination</i>)
HRDC	Helicase e RNaseD C terminal
ICr-HCFMUSP	Instituto da Criança do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo
Indels	Inserções e Deleções
ISCN	<i>International System of Cytogenetic Nomenclature</i>
KCl	Cloreto de potássio
mRNA	RNA mensageiro
NHEJ	Junção de extremidades não homólogas (do inglês – <i>Non-Homologous End Joining</i>)
NLS	Sinal de localização Nuclear (do inglês – <i>Nucleolar Localization Signal</i>)
NMD	Decaimento mediado por mutações nonsense (do inglês – <i>Nonsense-mediated decay</i>)
pb	Cc c Pares de bases
PCR	Reação em cadeia de polimerase (do inglês – <i>Polimerase Chain Reaction</i>)
RIN	<i>RNA Integrity Number</i>
RNA	Ácido ribonucleico
rpm	Rotação por minuto

RQC	RecQ <i>Family C terminal</i>
RT-qPCR	Reação em cadeia de polimerase quantitativa em tempo real (do inglês – <i>Real-Time quantitative polymerase chain reaction</i>)
SB	Síndrome de Bloom
SNPs	Polimorfismos de um único nucleotídeo (do inglês – <i>Single Nucleotide Polymorphism</i>)
SNVs	Variação de um único nucleotídeo (do inglês – <i>Single Nucleotide Variants</i>)
ssDNA	Fita simples de DNA (do inglês – <i>Single-Strand DNA</i>)
T	Timina
TCI	Troca entre cromátides irmãs
Teste K-S	Teste de Kolmogorov - Smirnov
USP	Universidade de São Paulo
UV	Luz ultravioleta
VEP	<i>Ensembl's Variant Effect Predictor</i>

LISTA DE SÍMBOLOS

%	Porcentagem
°C	Graus Celsius
<	Menor que
>	Maior que
cm	Centímetros
g	Gramas
h	Hora
M	Molar
mg	Miligramas
mL	Mililitros
ng	Nanogramas
μl	Microlitros

RESUMO

Montenegro MM. *Estudo do transcriptoma na síndrome de Bloom* [Tese]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2016.

A síndrome de Bloom (SB) é uma síndrome de instabilidade cromossômica rara, transmitida por herança autossômica recessiva. Caracteriza-se por deficiência de crescimento pré e pós-natal, microcefalia, hipoplasia malar, eritema telangiectásico em face e comprometimento do sistema imunológico, entre outras manifestações clínicas. Os pacientes com SB apresentam predisposição aumentada para o desenvolvimento de neoplasias em idade precoce, sendo esta, a principal causa de óbito. Ao estudo citogenético observa-se aumento de quebras cromossômicas espontâneas e trocas entre cromátides irmãs (TCI), que são utilizadas como marcador diagnóstico para a SB. Além disso, a literatura mostra que a maioria dos pacientes também apresenta mutações no gene *BLM*, que estão relacionadas a defeitos no mecanismo de reparo do DNA. No entanto, os mecanismos fisiopatológicos não são completamente compreendidos. Nesse sentido estudamos o transcriptoma de duas pacientes portadoras da síndrome de Bloom e de três controles utilizando a metodologia RNA-seq (Illumina, Inc., San Diego, CA). A análise de expressão diferencial revelou 216 genes diferencialmente expressos relacionados a vias relacionadas à resposta imune como replicação negativa da regulação da replicação do genoma viral, regulação positiva da proliferação de células B, via de sinalização mediada por *interferon* gama, ativação de células B, resposta a vírus, resposta imune adaptativa e processo efetor imune, e nenhuma diferença da expressão em genes de reparo de DNA. Concomitantemente, observamos a hiperexpressão do gene *BLM* para ambas as pacientes, contribuindo para a desestabilização de genes envolvidos em vias imunológicas, fenômeno também observado em alguns tumores. Dessa forma, sugerimos que a combinação de defeitos de proliferação linfocitária e defeitos de sinalização celular somados a outros, como perda celular e expressão alterada do gene *BLM*, podem contribuir diretamente para as principais características observadas na síndrome de Bloom, como a deficiência de crescimento e o elevado risco de câncer. Futuramente, o estudo do transcriptoma, aplicado a outros portadores da SB e outras síndromes de instabilidade, possibilitará uma análise mais acurada das interações gênicas relevantes para a desestabilização do genoma.

Descritores: síndrome de Bloom; instabilidade cromossômica; transcriptoma; análise de sequência de RNA; sequenciamento de Nucleotídeos em larga escala; expressão gênica.

ABSTRACT

Montenegro MM. *Transcriptome study in Bloom's syndrome*. [Thesis]. São Paulo: "Faculdade de Medicina, Universidade de São Paulo"; 2016.

Bloom Syndrome (BS) is a rare chromosome instability syndrome, with recessive autosomal inheritance. The main clinical manifestations are pre and postnatal growth deficiency, microcephaly, malar hypoplasia, telangiectasic facial erythema and compromised immune system, among others. Patients with BS present increased risk to the development of neoplasias at an early age, which is the main cause of death. Cytogenetic test is used as a diagnostic marker for BS since the patient's cells present increase in spontaneous chromosomal breaks and sister chromatid exchange (SCE). In addition, the literature reveals that most patients also present mutations in the *BLM* gene, which are related to defects in the DNA repair mechanism; however, it is still not completely understood. In this sense, we studied the transcriptome of two patients with Bloom's syndrome and three controls using the RNA-seq methodology (Illumina, Inc., San Diego, CA). Differential expression analysis revealed 216 differentially expressed genes related to immunological pathways such as: negative regulation of the regulation of the viral genome replication, positive regulation of B cells proliferation, gama-interferon mediated signalization pathway, B cells activation, virus response, adaptive immune response and immune effector process, and absence of difference of DNA repair genes expression. At the same time, we observed the hyperexpression of the *BLM* gene for both patients contributing for the destabilization of genes involved in immunological pathways, a phenomenon also observed in some tumors. Thus, we suggest that the combination of lymphoid proliferation defects and cell signaling defects added to others such as cell loss and altered expression of the *BLM* gene may contribute directly to the main characteristics observed in Bloom's syndrome, such as growth failure and high risk of cancer. In the future, the study of the transcriptome applied to other BS carriers and other instability syndromes, will allow a more accurate analysis of the relevant gene interactions to the destabilization of the genome.

Descriptors: Bloom syndrome; chromosomal instability; transcriptome; sequence analysis; RNA; high-throughput nucleotide sequencing; gene expression.

1 INTRODUÇÃO

A síndrome de Bloom (SB – OMIM #210900) é uma síndrome de instabilidade cromossômica rara, transmitida pelo padrão de herança autossômico recessivo (German, 1969; German, 1993; Gretzula et al., 1987). Foi descrita pela primeira vez por David Bloom com o relato de um paciente de nove anos do sexo masculino e de origem judaica, apresentando baixa estatura e eritema telangiectásico em face semelhante a lúpus eritematoso (Bloom, 1954). Em seguida (1964), o geneticista James German observou uma instabilidade cromossômica em pacientes com a doença, passando a identificar novos pacientes e descrever seus fenótipos (German, 1969).

A SB caracteriza-se principalmente por deficiência de crescimento pré e pós-natal, elevada fotossensibilidade caracterizada por erupções na face, sobretudo em região malar (semelhante a asas de borboleta como observado no lúpus eritematoso) e comprometimento do sistema imunológico, como o aumento da frequência de infecções de repetição, particularmente do trato respiratório superior e do trato gastrointestinal (Sanz e German, 2006). Nos indivíduos afetados pela síndrome, ocorre uma predisposição aumentada para o desenvolvimento de neoplasias epiteliais, linfóide, do sistema nervoso, trato gastrointestinal, entre outras. Essas neoplasias constituem a principal causa de morte entre os afetados (German e Passarge, 1989).

A frequência da SB é desconhecida na população em geral, porém a ocorrência para a principal mutação conhecida (*blm*^{Ash}) é maior em judeus de origem Ashkenazi e seus descendentes (German et al., 1977).

Em 1966, foi criado nos Estados Unidos, um programa denominado de Registro da síndrome de Bloom, visando a caracterização clínica e monitoração de indivíduos afetados. Por meio desse registro, foram obtidas as histórias clínicas e genéticas de pessoas diagnosticadas como portadoras da SB, no período de 1960 a 2009, e o curso da doença foi monitorado nesses indivíduos, facilitando o acúmulo de dados clínicos e experimentais da síndrome. Esse registro, que ainda está em vigor, tem permitido a associação de um leque de características relacionadas à síndrome (German, 1989; Sanz e German, 2006; http://www.med.cornell.edu/bsr/data_from_registry/).

Segundo o programa, há a descrição de apenas 265 casos pertencentes a 222 famílias (140 do sexo masculino e 125 do sexo feminino). Destes, 55 (24,8%) são pacientes com ancestral Judeu de origem Ashkenazi e 16 casos são brasileiros sem ancestral Judeu. A SB também foi descrita em outros grupos étnicos (Gretzula et al., 1987; German, 1989; German, 1994; German, 1999; Rezende et al., 2007).

Assim como outras doenças de mecanismo autossômico recessivo, a SB é encontrada com maior frequência quando há consanguinidade entre os pais dos afetados.

1.1 Aspectos clínicos

Descrevemos abaixo as principais manifestações clínicas encontradas em pacientes com SB:

Anormalidades do crescimento - A baixa estatura proporcionada é a característica clínica mais marcante presente nos pacientes com SB, de início pré-natal estendendo-se ao pós-natal (German, 1969; Gretzula et al., 1987; Sanz e German, 2006; Resende et al., 2007). Ao nascimento, o peso dificilmente excede os 2300 g e o comprimento médio é de 44 cm. A estatura final dos pacientes raramente ultrapassa os 145 cm em meninos e os 130 cm em meninas (German, 1969). A investigação de hormônio de crescimento apresenta-se normal (Gretzula et al., 1987).

Face – Os portadores da síndrome apresentam fácies típico caracterizado por microcefalia, face estreita com hipoplasia malar, nariz proeminente e eritema malar (German, 1969; Rezende et al., 2007). Ao nascimento, não apresentam alterações na pele. No entanto, com a exposição à luz solar, surgem lesões cutâneas caracterizadas por eritema telangiectásico, produzido por dilatação dos vasos (Bloom, 1954; German, 1969; Gretzula et al., 1987; Sanz e German, 2006; Resende et al., 2007). Essas erupções normalmente acometem a face, na região malar, nariz e lábios, assemelhando-se a asas de borboleta (lembrando lúpus eritematoso) podendo ainda afetar outras partes do corpo fotoexpostas, como o dorso das mãos e antebraços. Os eritemas telangiectásicos tendem a desaparecer com o tempo e podem ser substituídos por atrofia e hipocromia (German, 1993).

Pele – As manchas *café-au-lait* estão presentes em cerca de metade dos pacientes afetados (German, 1969). Manchas hipocrômicas também são comuns (Resende et al., 2007).

Anormalidades ocasionais – Outros achados, menos frequentes, são: hipodontia; cardiopatia congênita; fístula traqueoesofágica; estreitamento do meato uretral; oligodactilia, polidactilia, sindactilia ou clinodactilia; prega palmar única; luxação congênita de quadril e pé torto congênito (German, 1969; Sanz e German, 2006).

Sistema imunológico – A função imunológica dos pacientes portadores da SB é anormal com redução das concentrações de uma ou mais imunoglobulinas plasmáticas como a IgA, IgG, IgE e IgM (Taniguchi et al., 1982; German, 1993; Cunniff et al., 2016). São observados graus variados de imunodeficiência celular e humoral de etiologia desconhecida (Sanz e German, 2006; Resende et al., 2007). Na infância, são comuns diarreias e infecções recorrentes do trato respiratório. Cerca de 20% dos pacientes desenvolvem algum tipo de infecção bacteriana no trato respiratório. Quando não tratadas adequadamente, essas infecções podem progredir para bronquiectasia crônica. A doença pulmonar crônica é a segunda maior causa de óbito em pacientes com SB (Sanz e German, 2006).

Diabetes Mellitus – De acordo com o registro internacional da SB, 15,8% dos portadores desenvolve diabetes *mellitus* tipo II precocemente, entre 4 e 45 anos (média de 26 anos). Geralmente não são insulino dependentes. Anormalidades envolvendo mecanismos de liberação de insulina e tolerância à glicose são ocasionalmente observados (Sanz e German, 2006).

Fertilidade – Os homens afetados pela SB apresentam testículos hipoplásicos e podem ser inférteis devido à azoospermia (German, 1993). No entanto, os adolescentes desenvolvem os caracteres secundários normalmente. Também são observados hipospádia e criptorquidia (German, 1969). As mulheres podem apresentar menarca tardia e menopausa precoce. Os ciclos menstruais são irregulares e a fertilidade parece ser diminuída (Gretzula et al., 1987; Bachrati e Hickson, 2003; Sanz e German, 2006; Resende et al., 2007; Cunniff et al., 2016). Contudo, a literatura relata três gravidezes de

sucesso em portadoras de SB, cujas crianças nasceram normais e o exame citogenético excluiu a doença sendo as crianças, portanto, obrigatoriamente heterozigotas (Mulcahy et al., 1981; German, 1993).

Desenvolvimento cognitivo – Em geral a inteligência é normal, porém alguns pacientes podem apresentar deficiência cognitiva leve com dificuldade de aprendizado, problemas de leitura, déficit de atenção e de memória, observados principalmente na infância (German, 1969; German, 1993; Sanz e German, 2006).

Câncer – A neoplasia é a principal causa de morte dos pacientes acometidos pela doença (German e Passarge, 1989; Liu e West, 2008). Aproximadamente 42% dos afetados desenvolvem neoplasias em idade precoce e cerca de 10% destes desenvolvem mais de um tipo de neoplasia – principalmente linfomas, carcinomas epiteliais, hematopoiéticos e, em alguns casos, tumor de Wilms e retinoblastomas (German, 1997; Ellis et al., 1998; Ma et al., 2001; Bhisitkul e Rizem, 2004; Sanz e German, 2006; Resende et al., 2007). A SB é a única síndrome de instabilidade cromossômica em que os afetados podem apresentar qualquer tipo de neoplasia (Hickson et al., 2001). A idade média do surgimento varia entre 2 e 49 anos (média de 24,7 anos), e a idade média de óbito é de 24 anos (Ma et al., 2001, Bachrati e Hickson, 2003). Os pacientes acometidos pela SB são mais suscetíveis a toxicidades dos agentes quimioterápicos, sendo a radioterapia uma estratégia preferível ao tratamento (Martin et al., 2010).

1.2 Aspectos Moleculares da SB

1.2.1 Integridade celular e instabilidade genômica

Devido ao fato do DNA conter as informações genéticas em cada célula viva, sua integridade e estabilidade são essenciais (Clancy, 2008). Contudo, a molécula de DNA é continuamente danificada devido a agressões causadas por processos endógenos e exógenos à célula, tais como danos decorrentes do metabolismo oxidativo e da replicação, radiação ionizante, agentes químicos, entre outros (Payam e Hickson, 2001;

Sing et al., 2009). Para conviver com esses danos contínuos, as células desenvolveram uma gama de respostas, que inclui diferentes mecanismos de reparo de DNA e atraso do ciclo celular, para que os danos sejam corrigidos (Payam e Hickson, 2001; Hoeijmakers, 2001; Hoeijmakers, 2007) (Figura 1). Quando não reparados, tais danos podem se acumular, levando a alterações na informação genética, modificando a estrutura celular, comprometendo sua integridade e estabilidade, causando doenças (Hoeijmakers, 2001).

Figura 1 – Danos sofridos pelo DNA, mecanismos de reparo de DNA e as consequências sofridas pela célula. **a**, Agentes danificadores de DNA (topo); Exemplos de lesões que ocorrem no DNA induzidas por tais agentes (meio) e os mecanismos de reparo de DNA mais relevantes para a remoção e correção destas lesões (abaixo). **b**, Efeitos agudos do DNA danificado na progressão do ciclo celular levando a parada do ciclo nas fases G1, S, G2 e M (topo), os efeitos no metabolismo de DNA (meio) e as consequências a longo prazo (abaixo), incluindo mudanças permanentes na sequência de DNA e seus efeitos biológicos (Adaptado de Hoeijmakers, 2001).

Dentre todos os danos que uma célula pode sofrer, a quebra na dupla fita de DNA (DSBs – do inglês *Double-Strands Breaks*), (Martin et al., 2010) é a mais prejudicial, e pode surgir em decorrência de radiação ionizante, raios X, metabolização de radicais livres, produtos químicos e replicação de uma fita simples gerada por uma quebra

(Hoeijmakers, 2001). Para minimizar o impacto nocivo desse tipo de alteração, as células evoluíram dois tipos de mecanismos reparo: a junção de extremidades não homólogas (NHEJ – do inglês *Non-Homologous End Joining*) e a recombinação homóloga (HR – do inglês *Homologous Recombination*) (Figura 2).

Em mamíferos, a cicatrização ocorre preferencialmente por NHEJ, sendo comum advirem inserções ou deleções no ponto de quebra, quando o reparo é efetuado por essa via. Ocorrendo mais de uma quebra em uma mesma célula, esse mecanismo poderia unir extremidades não contíguas, levando a rearranjos estruturais (Hoeijmakers, 2001; Payne e Hickson, 2008).

O reparo por HR foi inicialmente estudado em leveduras, que se utilizam preferencialmente dessa via. Entretanto as células dos mamíferos também são capazes de reparar seu DNA por meio da HR, a qual restaura a integridade celular praticamente livre de erros, pois envolve o uso de uma sequência de DNA homóloga como modelo para sua reconstrução. Em leveduras, a HR pode ser efetuada utilizando-se do cromossomo homólogo, passando a operar em todas as fases do ciclo celular. Em contraste, nas células dos mamíferos, a HR utiliza apenas cromátides irmãs como modelo, restringindo-se as fases de pós-replicação do ciclo celular, preferencialmente as fases S e G2 (Deckbar et al., 2011, Madlenov e Iliakis, 2011).

Mutações em um gene específico com atividade de helicase que atua no reparo de DNA por HR, chamado *BLM*, frequentemente levam à desestabilização do genoma e consequentemente a um aumento na frequência de mutações em outros *loci*, levando ao fenótipo de SB (Ellis et al., 1995; Payam e Hickson, 2001).

Figura 2 – Mecanismos de reparo de quebra na dupla fita de DNA: Reparo por recombinação homóloga (HR – do inglês *Homologous Recombination*) (à esquerda) e o reparo por junção de extremidades não homólogas (NHEJ – do inglês *Non-Homologous End Joining*) (à direita). (Adaptado de Hoeijmakers, 2001).

1.2.2 A síndrome de Bloom e o gene *BLM*

O gene responsável pela SB, *BLM* (*Bloom syndrome*; *RECQ Protein-like*; *RECQL3* - #604610), identificado em 1995 por Ellis e colaboradores (Ellis et al., 1995), é um supressor tumoral localizado na região cromossômica 15q26.1 (Shastri e Schmidt, 2015). É composto por 22 éxons, que em condições normais produzem 11 transcritos, sendo dois codificadores de proteínas – BLM-001 (ENST00000355112.7) e BLM-002 (ENST00000560509.5). A principal proteína é codificada pelo transcrito BLM-001, denominada BLM. A BLM é constituída por 1417 aminoácidos e pertence à subfamília das DNA helicases, especificamente ao grupo das helicases RecQL homólogas à RecQ helicase de *Escherichia coli* (Ellis et al., 1995; Hickson et al., 2001).

A família RecQ é definida por um domínio helicase de aproximadamente 450 aminoácidos, que está localizado centralmente na proteína e é altamente conservado em eucariotos e procariotos, com homologia de 40 – 50% de aminoácidos com a RecQ de *E. coli*. (Figura 3) (Seki et al., 2008; Singh, 2009). Na espécie humana, há cinco RecQ homólogas identificadas, sendo RECQL1, RECQL2 (WRN), RECQL3 (BLM), RECQL4 e RECQL5. Três delas (WRN, BLM e RECQL4) estão associadas às síndromes de Werner, Bloom e Rothmund-Thomson, respectivamente (Liu e West, 2008).

Figura 3 – Representação esquemática da família RecQ helicase em diferentes espécies, desde humanos até leveduras. Os domínios estão alinhados em um código de cores, como indicado na parte inferior da figura: (Verde) domínio exonuclease; (Vermelho escuro) domínio ácido; (Amarelo) domínio helicase, altamente conservado entre as espécies; (Preto) domínio HRDC; (Vermelho claro) sinal de localização nucleolar e (Rosa) domínio RQC. O tamanho em aminoácidos da proteína de cada espécie está indicado à direita. O nome das proteínas e das espécies de origem está indicado à esquerda. (Adaptado de Singh et al., 2009).

A estrutura da proteína BLM humana (Figura 4), compreende os domínios: N-terminal que contém regiões com blocos de resíduos ácidos; o domínio helicase que contém sete motivos (assinaturas) que determinam sua função, como o motivo sequência de ligação de ATP e o DexH-box e o domínio C terminal que inclui o RQC (do inglês **RecQ** Family **C** terminal) que consiste de um domínio de zinco (Zn^{2+}) responsável por mediações proteína-proteína e o domínio de ligação ao DNA HRDC (**H**elicase e **R**Nase **D** C terminal), além de uma região de sinal de localização nuclear (NLS – do inglês **N**ucleolar **L**ocalization **S**ignal) (Ellis et al., 1995; Hickson et al., 2001; Bachrati e Hickson 2003; Shastri e Schmidt, 2015).

Figura 4 – Estrutura gráfica da proteína BLM.

A proteína BLM helicase utiliza energia derivada da hidrólise de ATP para separar a dupla fita de DNA, deslocando-se no sentido 3' → 5' (Bachrati e Hickson, 2003), desempenhando um papel essencial no processo de recombinação, transcrição, replicação e reparo de DNA, garantindo a estabilidade genômica (Payam e Hickson, 2001; Owen et al., 2014). Em condições adequadas, é considerada uma proteína chave anti-recombinante, pois uma de suas principais funções é organizar a migração e dissolução de estruturas intermediárias de DNA geradas durante o reparo por HR, como as Junções *Holliday* (Wu e Hickson, 2003). Neste caso, o processo de reparo por HR (Figura 5) é iniciado pela ligação do complexo MRN (Mre11/Rad51/Nbs1) às extremidades de DNA livres resultantes de quebra na dupla fita de DNA, promovendo a ressecção das extremidades de DNA danificadas. Este processo é necessário para a formação de extremidades de fitas simples de DNA (ssDNA – do inglês *Single-Strand DNA*). Subsequente à ressecção, as ssDNAs são estendidas pelas nucleases Exo1 e Dna2, apoiadas pela BLM. Entretanto, a exata sincronia entre as nucleases e a helicase não é clara (Liu e West, 2009). A extremidade 3' resultante, que pode se estender em mais de centenas de bases, é estabilizada pela proteína de replicação A (RPA – do inglês *Replication Protein A*), a qual é sequencialmente substituída pela recombinase Rad51. Após encontrar seu ponto de homologia, este filamento nucleoproteico (ssDNA+Rad51) invade a forquilha de replicação da cromátide irmã. Utilizando a extremidade 3' livre como *primer*, a síntese de DNA é realizada pela DNA polimerase e a ligação ocorre pela DNA-ligase I. Deste processo formam-se junções *Holliday* (HJ – do inglês *Holliday Junctions*), que são finalmente resolvidas pelo complexo BLM/TopIIIα/RMI1/RMI2, formando produtos com *crossing overs* (Helleday, 2003; Shastri e Schmidt, 2015). Outras formas resultantes da HR também servem como substrato para a BLM, como os G-quadruplex e os D-loops.

HOMOLOGOUS RECOMBINATION

Prokaryotic type

Eukaryotic type

Figura 5 – Representação esquemática da via de reparo por recombinação homóloga em procariotos e eucariotos. Em eucariotos é possível observar o momento em que o gene *BLM* é considerado atuante, marcado por uma estrela em vermelho, incluindo o processamento das extremidades quebradas da dupla fita (estrela 1), a abertura do D-loop que foi formado para a invasão da fita simples na forquilha de replicação da cromátide irmã (estrela 2), processamento, dissolução de *Holliday Junction* (estrela 3) e supressão de produtos com crossing over na síntese dependente de anelamento de fita simples (estrela 4 e 5). Adaptado de www.genome.jp/kegg.

Dessa forma, a proteína BLM é essencial para o reparo por recombinação homóloga e consequentemente para a manutenção da integridade genômica. Contudo, a função da BLM pode sofrer alterações em decorrência de mutações encontradas no gene *BLM*. Como resultado, células com uma proteína BLM não funcional apresentam níveis elevados de recombinação mitótica, quebras cromossômicas e cromatídicas, configurações quadrirradiais e aumento da frequência entre cromátides irmãs (Shastri e Schmidt, 2015).

1.2.3 Mutações envolvendo o gene *BLM*

Aproximadamente 70 variantes já foram descritas para o gene *BLM*, a maioria resultando em uma proteína BLM truncada, devido a uma variante *nonsense* (sem sentido) ou a uma variante *frameshift* (que afeta o quadro de leitura), que podem conduzir a um códon de parada prematuro (Bachrati e Hickson, 2003; Shastri e Schmidt, 2015). Em geral, essas mutações ocorrem entre os éxons 7 a 18, os quais estão envolvidos na codificação do centro Helicase do gene *BLM*, responsável pela função helicase da proteína (Shastri e Schmidt, 2015).

Dentre todas, a variante mais comum é do tipo *frameshift*, uma deleção de 6pb e uma inserção de 7pb (c.2207_2212delinsTAGATTC), localizada no éxon 10, na posição 2281 do gene *BLM* (Ellis et al., 1998). Essa mutação é referida como *blm^{Ash}*, e é observada com maior frequência entre Judeus de origem Ashkenazi e seus descendentes, sendo considerada uma mutação de efeito fundador. Esta mutação está estimada em 1:107 heterozigotos na cidade de Nova York (Li et al., 1998), podendo também ser observada em portadores de SB sem ancestral Ashkenazi. Não há relatos dessa mutação em portadores de SB brasileiros, com ou sem ancestral Ashkenazi.

Em 2007, German identificou 64 tipos diferentes de mutações, além da *blm^{Ash}*, em 125 de 134 portadores de SB, sendo que 54 destas mutações resultavam em uma proteína prematura truncada e 10 eram do tipo *missense*. Além das mutações, foram detectados 17 polimorfismos que não foram considerados causadores da doença nestes

pacientes, uma vez que a mudança causada no DNA foi frequente, silenciosa e estava presente em mais de uma pessoa, entre as que já possuíam uma mutação causadora da doença.

1.3 Aspectos Citogenéticos

Mutações no gene *BLM* perturbam o mecanismo de reparo por HR, gerando instabilidade genômica, que se reflete em alterações cromossômicas (Figura 6), principalmente quebras e *gaps* cromossômicos e cromatídicos, rearranjos estruturais (cromossomos dicêntricos, fragmentos, pulverização, rearranjos) e trocas recíprocas entre segmentos cromossômicos homólogos (configurações quadrirradiais). Todavia, o que distingue as células dos portadores da SB de outras síndromes de instabilidade, é a característica única de aumento da frequência de trocas entre cromátides irmãs (TCI).

Figura 6 – Alterações cromossômicas observadas como resultado de instabilidade em células de portadores de SB: (A) Quebra cromossômica; (B) Quebra cromatídica; (C) Gap cromossômico; (D) Gap cromatídico; (E) Cromossomo dicêntrico; (F) Fragmentos; (G) Rearranjo; (H) Pulverização cromossômica, (I) Configuração quadrirradial e (J) Troca entre cromátides irmãs. Laboratório de Citogenômica do LIM03.

1.3.1 Quebras cromossômicas e cromatídicas

De acordo com a fase do ciclo celular em que ocorrem, as quebras podem ser cromossômicas ou cromatídicas. Se uma quebra ocorrer durante a fase G1, com ambos

os fragmentos replicados durante a fase S, pode-se observar a quebra cromossômica, ou seja, quebra em ambas as cromátides (Figura 6 A). Se a quebra ocorrer durante a fase G2 do ciclo celular, pode-se observar a quebra cromatídica (Figura 6 B), ou seja, quebra em apenas uma das cromátides (Kuhn e Therman, 1986).

Quando as quebras ocorrem durante a fase S ou G2 do ciclo celular também podem originar fragmentos (Figura 6 F), que são segmentos cromossômicos de origem desconhecida. Estes podem ser: (1) fragmentos acêntricos, que são segmentos cromossômicos sem centrômero e (2) fragmentos cêntricos, que são segmentos cromossômicos com centrômero. Ambos são considerados cromossomos marcadores, e podem aparecer de forma aleatória nas células dos indivíduos portadores da SB (Kuhn e Therman, 1986).

Também é observada um tipo raro de anormalidade, em que, as extremidades de dois segmentos cromossômicos centroméricos, que sofreram quebras, se fundem. É o chamado cromossomo dicêntrico (Figura 6 E). Esse mecanismo leva à perda dos fragmentos acêntricos. A grande maioria dos cromossomos dicêntricos observados nas células de indivíduos portadores da SB, são simétricos (isodicêntricos) e decorrem de dois cromossomos homólogos que sofreram quebras em pontos idênticos (Kuhn e Therman, 1986).

Ainda, decorrentes de quebras cromossômicas e cromatídicas, são observadas translocações, deleções, duplicações, inversões e rearranjos (Figura 6 G) (Kuhn e Therman, 1986).

As falhas ou *gaps* cromossômicos ou cromatídicos (Figura 6 C e D) são alterações cujos mecanismos ainda não foram bem elucidados, que ocorrem entre a fase G2/S do ciclo celular. Correspondem a lesões acromáticas com pequenas mudanças estruturais sendo estas lesões, em geral, menores do que a largura da cromátide (Kuhn e Therman, 1986; Savage, 2004).

1.3.2 Configuração quadrirradial

Configurações quadrirradiais (Figura 6 I) são anormalidades cromossômicas que ocorrem como consequência de lesões sofridas pelas cromátides de cromossomos homólogos. Podem ser simétricas ou assimétricas, dependendo da posição do centrômero. As configurações assimétricas não são equilibradas. As configurações simétricas (aparentemente equilibradas) resultam da troca entre cromátides que tomam lugar em sítios similares nos dois cromossomos homólogos durante a mitose. Os centrômeros estão em raios opostos entre os quatro braços cromossômicos da figura quadrirradial. Esse tipo de configuração é característica da SB (German et al., 1974).

1.3.3 Trocas entre cromátides irmãs

A SB se diferencia das outras síndromes de instabilidade cromossômica por apresentar um aumento de TCI (Figura 6 J) quando comparado a culturas celulares de indivíduos normais.

Durante a fase S do ciclo celular, o DNA é replicado e cada cromossomo se duplica em duas cromátides irmãs, que segregam unidas pelo centrômero durante o ciclo celular. As cromátides irmãs são visualizadas durante a metáfase mitótica, quando os cromossomos estão completamente condensados. As TCIs são trocas recíprocas que ocorrem naturalmente em eventos associados à replicação normal do DNA, em que as cromátides irmãs sofrem quebras com subsequente reunião e troca recíproca dos segmentos homólogos. Esse processo ocorre predominantemente como parte do mecanismo de HR e é considerado livre de erros, já que nenhuma informação é alterada, apenas ocorre a troca de posição.

As TCIs ocorrem em baixas taxas nos indivíduos normais, sendo estimada em 1 – 4 trocas por célula (Figura 7 A) (Wilson e Thompson, 2007; Payne et al., 2009). Já as células dos portadores da SB possuem esse mecanismo de RH imperfeito, elevando de

40 a 100 vezes a frequência de TCI (Figura 7 B) (German et al., 1974, Sanz e German, 2006).

Figura 7 – Demonstração de troca entre cromátides irmãs. À esquerda (A) é possível observar um controle normal, com apenas uma troca. À direita (B) observa-se um controle positivo para a síndrome de Bloom, demonstrando frequência aumentada de trocas entre cromátides irmãs (Killen, 2011).

Este evento é facilmente observado ao microscópio óptico em células em cultura tratadas com 5-bromo-2'-desoxiuridina (BrdU). O BrdU é um análogo sintético de timidina que se incorpora ao DNA durante a replicação, quando adicionado à cultura celular (Chaganti et al., 1974). Como a replicação do DNA é feita de forma semiconservativa, o BrdU se incorpora em uma das fitas da dupla fita de DNA da célula filha. Durante o segundo ciclo de crescimento celular, em meio de cultura contendo BrdU, as duas cromátides irmãs se diferenciam em uma cromátide com a fita de DNA original (sem incorporação de BrdU) e outra com incorporação de BrdU, em ambas as fitas de DNA (Figura 8). Para que essa incorporação do BrdU ao DNA celular seja reconhecida, utiliza-se a técnica Hoeschst-Giemsa (Perry e Wolff, 1974). Esta técnica consiste na ligação do fluorocromo Hoechst 33258 ao BrdU através da denaturação do DNA por exposição à luz UV e posterior coloração com corante Giemsa, tornando possível observar a cromátide diferenciada onde o BrdU foi incorporado.

Figura 8 – Esquema gráfico demonstrando a estrutura química do BrdU (acima, no centro da figura), e sua incorporação em uma das fitas da dupla fita de DNA por dois ciclos celulares (<https://bioggs.blogspot.com.br/2010/05/meet-chickens-sister-chromatid.html>).

A análise cromossômica, com a utilização do agente BrdU, pode ser realizada em qualquer tipo de linhagem celular. Porém, a cultura de linfócitos é a mais utilizada (Ray e German, 1984). Para o diagnóstico pré-natal, pode-se utilizar linhagens celulares provenientes de fibroblastos, obtidos através de culturas de líquido amniótico, vilosidades coriônicas e células linfocitárias de sangue do cordão umbilical (Sanz e German, 2006).

O diagnóstico da SB é baseado em achados clínicos. No entanto, o estudo citogenético para pesquisa TCI é fundamental para a confirmação diagnóstica, sendo considerado padrão ouro de diagnóstico, uma vez que essa é uma característica patognomônica da SB (Sanz e German, 2006).

1.4 Transcriptoma aplicado ao estudo da síndrome de Bloom

Diferentes técnicas moleculares têm sido aplicadas na busca da compreensão dos mecanismos envolvidos na SB, como a pesquisa de mutações e ensaios enzimáticos. No entanto, essas estratégias ainda não conseguiram resultados totalmente hábeis para a compreensão da maquinaria molecular associada à etiopatologia da SB.

Metodologias mais recentes de análise das variantes do DNA, incluindo CNVs e mutações, e de expressão gênica como os sequenciamentos de nova geração, podem ser empregadas para identificar alterações de vias biológicas que convergem na manifestação do quadro clínico. No momento, a técnica de sequenciamento completo mais específica, que pode nos fornecer diversas informações ao mesmo tempo e que apresenta boa relação custo-benefício em relação as demais técnicas de alto rendimento, é o estudo do transcriptoma de RNA mensageiro (mRNA) pela metodologia de RNA-seq.

O transcriptoma é definido como o complemento de moléculas de RNA geradas por uma célula ou população celular, em um determinado tecido e em um momento específico (McGettigan, 2013). Ao selecionar todos os RNAs mensageiros, ou seja, RNAs expressos do genoma, podemos ter um perfil indicativo de atividade gênica, sendo possível mensurar e comparar os níveis de expressão de cada gene codificador de proteína (Adams, 2008). Dessa forma, a análise do transcriptoma é uma boa abordagem para o entendimento de variáveis funcionais do genoma. O transcrito contém operons, sítios de início transcricional, RNAs *antisense* e RNAs mensageiros, sendo todos esses aspectos muito importantes do transcriptoma que podem revelar valiosos dados a respeito das funções e regulações gênicas em doenças (Ju et al., 2011; Filliatrault, 2011). Além disso, essa metodologia pode permitir a detecção de variáveis genômicas, como Indels (**I**nserções e **D**eleções) e SNPs (polimorfismos de um único nucleotídeo, do inglês *Single Nucleotide Polymorphism*) (Minoche et al., 2011).

Aplicado à SB, tal estudo pode permitir uma melhor compreensão sobre a associação da estrutura molecular e as manifestações clínicas, e ainda permitir a

construção de uma rede de interações gênicas relacionada à doença, possibilitando uma análise mais abrangente, precisa e profunda dos mecanismos envolvidos na SB (Hurley et al., 2011).

Nesse contexto, o presente estudo propõe a investigação do transcriptoma de pacientes com SB, utilizando a metodologia de RNA-seq, com o intuito de descrever o perfil de expressão, detectar variações genômicas e observar possíveis interações gênicas relevantes no aparecimento e desenvolvimento dessa doença.

2 OBJETIVOS

O presente estudo tem como objetivo geral:

- ✓ Realizar o sequenciamento do transcriptoma sanguíneo de RNAs mensageiros (mRNAs), por meio da técnica de RNA-Seq, em pacientes com a síndrome de Bloom.

2.1 Objetivos específicos

- ✓ Implantar a técnica de sequenciamento do transcriptoma sanguíneo de RNAs mensageiros (mRNAs), por meio da técnica de RNA-Seq, no laboratório de Citogenômica (HCFMUSP);
- ✓ Obter o perfil de expressão gênica diferencial para esses pacientes em relação ao grupo controle;
- ✓ Realizar o estudo de enriquecimento funcional dos genes diferencialmente expressos;
- ✓ Estudar as possíveis variantes que possam estar relacionadas com vias biológicas significativas para a síndrome de Bloom;

3 METODOLOGIA

3.1 Casuística

Este projeto foi aprovado pelo Comitê de Ética para Análise de projetos de Pesquisa da Diretoria Clínica do Hospital das Clínicas e da Faculdade de Medicina da Universidade de São Paulo (Cappesq – HCFMUSP nº 63954/12) (ANEXO A). Os pacientes ou seus responsáveis legais assinaram um termo de consentimento livre e esclarecido (ANEXO B).

Trata-se de um estudo retrospectivo e prospectivo de duas pacientes do sexo feminino com quadro clínico compatível com a síndrome de Bloom em acompanhamento no ambulatório de Genética do Instituto da Criança do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (ICr – HCFMUSP). Embora existam outros pacientes portadores de síndrome de Bloom, que já realizaram atendimento na Unidade de Genética, apenas estas duas pacientes aceitaram participar do estudo.

O critério de inclusão foi de pacientes clinicamente compatíveis com a síndrome de Bloom. Demais pacientes foram excluídos.

Foram também incluídas três amostras controles de doadores considerados saudáveis, sendo um do sexo feminino e dois do sexo masculino, totalizando cinco indivíduos estudados. Todas as amostras utilizadas estão descritas na Tabela 1.

Tabela 1 – Descrição das amostras que compõe a casuística do presente estudo

Amostras	HD	Sexo	Idade (Anos)
K1	SB	Feminino	28
M1	SB	Feminino	12
C1	Saudável	Masculino	29
C2	Saudável	Feminino	29
C3	Saudável	Masculino	31

HD – Hipótese diagnóstica; SB – Síndrome de Bloom

3.1.2 Quadro clínico

Os dados clínicos das pacientes foram obtidos por meio da revisão de prontuários do Hospital e da Unidade de Genética sob supervisão da Dra. Chong Ae Kim. Em todas as consultas realizadas pela Unidade de Genética, foi preenchido um protocolo que consistiu em anamnese, exame físico detalhado e exames complementares de diferentes especialidades incluindo análises imunológicas, cardiológicas, endocrinológicas, gastroenterológicas e oftalmológicas. O *follow-up* das pacientes será descrito em seguida.

3.1.2.1 Paciente K1

Paciente K1, sexo feminino, 28 anos de idade, brasileira, é a quinta filha de pais consanguíneos e sem ancestral Judeu de origem Ashkenazi. Nasceu de 35 semanas de uma gestação sem intercorrências. Ao nascimento apresentava baixo peso, de 1800 g (< 5 percentil) e um comprimento pequeno, de 40 cm (< 5 percentil). Apesar de apresentar desenvolvimento neuropsicomotor normal durante o primeiro ano de vida, foi observado um discreto atraso no decorrer de seu desenvolvimento. Em sua primeira infância com início aos dois meses de idade, a paciente apresentou recorrentemente pneumonia, bronquite, sinusite, otite média e diarreia crônica, sendo necessário o uso de antibióticos profiláticos. Foi detectada deficiência de IgM.

Em sua primeira consulta com a Unidade de Genética, ao exame físico observaram-se características compatíveis com a SB, incluindo microcefalia, nariz e orelhas proeminentes, hipoplasia malar, manchas hipo e hiperocrômicas (face, braços, abdômen e pernas). Seu desenvolvimento neuropsicomotor estava adequado.

Até o presente estudo, a paciente evoluiu sem intercorrências.

3.1.2.2 Paciente M1

Paciente M1, sexo feminino, 12 anos, brasileira, é a segunda filha de um casal não consanguíneo e sem ancestral Judeu de origem Ashkenazi. Nasceu prematura com 28 semanas de uma gestação de alto risco (hemorragia e perda de líquido amniótico). Ao nascimento apresentou peso de 870 g (< 5 percentil), e comprimento de 33 cm (< 5 percentil). Apesar de apresentar desenvolvimento cognitivo normal durante o primeiro ano de vida, foi observado um discreto atraso no desenvolvimento neuropsicomotor no decorrer de seu desenvolvimento.

Em sua primeira consulta com a Unidade de Genética, ao exame físico observaram-se características compatíveis com a SB, incluindo microcefalia, nariz e orelhas proeminentes, hipoplasia malar, manchas hipo e hiperocrômicas (face, tórax, braços, pernas), bom como clinodactilia de mão. As medidas antropométricas estavam todas abaixo do percentil 5. Além disso, observou-se que a paciente apresentava dificuldade de aprendizado.

Seu desenvolvimento foi marcado por infecções recorrentes do trato respiratório (pneumonia e bronquite), sendo necessário o uso de antibióticos profiláticos com frequência. Além disso, a paciente apresenta deficiência de IgM e IgE. Ademais, desenvolveu-se sem intercorrências.

3.2 Métodos

O primeiro passo da realização deste estudo foi confirmar a hipótese clínica para a SB. Para tanto realizou-se o estudo citogenético para pesquisa de TCI, seguido pela pesquisa de mutações no gene *BLM* pela metodologia de sequenciamento Sanger. O teste citogenético é considerado padrão ouro para o diagnóstico da SB, pois a frequência aumentada de TCI é uma característica patognomônica da síndrome.

Após a confirmação do diagnóstico da doença, foi realizado o estudo do transcriptoma pela metodologia de RNA-seq, com duas abordagens diferentes de análise

de dados: a análise de expressão diferencial (A1) e a análise para verificação de variantes genômicas (A2). Por fim, fez-se a validação da expressão do gene *BLM* pela técnica de PCR quantitativo em tempo real (RT-qPCR). É possível visualizar o fluxograma da metodologia realizada neste trabalho na Figura 9.

Os testes citogenéticos, os ensaios moleculares e a análise de dados foram realizados no Laboratório de Citogenômica do LIM03 do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo. O sequenciamento do transcriptoma foi inicialmente realizado no Laboratório Multiusuário de Genômica Funcional da Escola Superior de Agricultura Luiz de Queiroz (ESALQ – USP), com a colaboração do Prof. Dr. Luiz Lehmann Coutinho e apoio do Departamento de Patologia da FMUSP.

Figura 9 – Fluxograma da metodologia utilizada no presente estudo.

3.2.1 Estudo Citogenético

O primeiro passo para confirmar a hipótese diagnóstica de SB foi o estudo citogenético para pesquisa de TCI, utilizando 5'-bromo-2'-deoxiuridina (BrdU) em culturas linfocitárias.

Para este procedimento, foram coletados 5 mL de sangue periférico em tubo heparinizado (Heparina sódica). Em fluxo laminar, com auxílio de uma pipeta Pasteur estéril, 13 gotas do sangue total de cada amostra foram semeadas em tubo estéril do tipo Falcon de 15 mL (TPP[®]), contendo 5 mL de meio PBmax (Gibco[®]) e 80 µL de BrdU (1,6 mg de BrdU diluída em 100 mL). As culturas foram então incubadas em posição inclinada e com a tampa semiaberta em estufa de CO₂ a 37°C por aproximadamente 72 horas. Ao final do período de incubação, 65 µL de colcemid (Gibco[®]) foram adicionados às culturas e essas novamente incubadas em estufa de CO₂ a 37°C por mais uma hora. O material cultivado foi submetido à centrifugação a 1500 rpm durante 5 minutos. Com auxílio de uma pipeta Pasteur, o sobrenadante foi desprezado e o sedimento ressuspendido em 8 mL de solução hipotônica (KCl 0,075M) aquecida à 37°C. Os tubos de cultura foram então incubados em estufa a 37°C na posição vertical por 20 minutos e em seguida foram adicionados 0,5 mL de solução fixadora *Carnoy* (3 metanol:1 ácido acético) (Synth[®]) com o intuito de interromper a hipotonia. Após esse período, o material foi novamente centrifugado e o sobrenadante desprezado. Ao pellet foram adicionados 6 mL da solução fixadora e o material foi homogeneizado. Esse processo foi repetido por três vezes. O material final obtido foi gotejado em três lâminas foscas lapidadas, previamente lavadas, sobre banho Maria a 60°C.

Em ambiente escuro, as lâminas confeccionadas com o material proveniente da cultura celular foram cobertas por solução de KCl a 2% (20 g de KCl (Merck[®]) diluído em 1.000 mL de água destilada) por 10 minutos. As lâminas foram escorridas e cobertas por corante Hoechst 33258 (7,5 mg de corante Hoeschst 33258 (Sigma[®]) diluído em 100 mL de água destilada). Após esse período, as lâminas foram enxaguadas em solução de KCl a 2%, cobertas por uma lamínula e expostas a luz UV (254 nm) por 40 minutos. As lâminas foram então lavadas em água destilada, onde foi possível retirar as lamínulas sem raspar a lâmina. Em seguida, as lâminas foram coradas com corante

Giemsa a 4% (4 mL de corante Giemsa (Merck®) diluído em 100 mL de tampão fosfato por 4 minutos e enxaguadas em água corrente.

A análise foi realizada em um microscópio óptico (AxioVision A2 – Zeiss, Alemanha) utilizando a objetiva de imersão de 100x e as metáfases foram fotografadas pelo sistema de captura de imagem Ikaros versão 4.5 (Zeiss, Alemanha). Para cada paciente foram analisadas dez metáfases. O valor de TCI/metáfase foi estabelecido utilizando a média entre a soma dos pontos de trocas em cada metáfase observada. O valor de referência utilizado foi o mesmo descrito pela literatura, >10-15 TCI/metáfase (Sanz e German, 2006), e o resultado descrito seguindo o padrão do Sistema Internacional de Nomenclatura para Citogenética de 2016 (ISCN 206 – Shaffer et al., 2016)

3.2.2 Pesquisa de mutações do gene *BLM* por sequenciamento Sanger

O DNA genômico foi extraído de 3 mL de sangue periférico coletado em tubo EDTA, por meio do Kit QIAamp DNA Mini Kit (Qiagen®), de acordo com o protocolo descrito pelo fabricante. A qualidade e a quantidade de DNA das amostras foram determinadas pelo espectrofotômetro NanoVue Plus (GE Healthcare Life Sciences). Foram consideradas adequadas as amostras com razão de absorbância 260/280 nm entre 1.8 e 2.0.

Primers específicos para as regiões codificadoras (éxons) do gene *BLM* foram desenhados com auxílio do *software* primer 3 out (www.primer3out.com) com base na sequência de referência NM_000057.3, uc002brp.5 (CRCh38/hg38 Genome Reference Consortium Human Reference 38 build, Dec/2013) e Go Taq Hot Start Polymerase (Promega®). Também foram incluídas regiões intrônicas entre os limites exon-intron para cada éxon. Todas as regiões desenhadas foram amplificadas por reação de cadeia em polimerase (**PCR** – do inglês *Polimerase Chain Reaction*).

A PCR pré sequenciamento foi realizada com 6 µl de polimerase GoTaq, 1 µl dos *primers forward* e *reverse* desenhados para o gene *BLM* e 5 µl das amostras do DNA genômico das pacientes (1 – 200 ng). A PCR foi realizada no StepOneplus (Applied

Biosystems®) com uso de TaqMan universal PCR Master Mix (Applied Biosystems®). As condições do ciclo foram: 2 minutos a 94 °C, 30 segundos 94 °C, 30 segundos a 60 °C, 72 °C por 1:30 minutos por 35 ciclos e temperatura final de 4 °C. Não foram inclusos controles.

Os produtos da reação de PCR foram verificados por eletroforese em gel de agarose a 2% (2 g de agarose diluídos em 100 mL de buffer TAE 1x) e purificados pelo kit ExoSAP-IT (Affymetrix®), adicionando 3 µl de EXOSAP para cada 13 µl do produto de PCR e um ciclo de 15 minutos a 37 °C, 80 °C a 15 minutos e temperatura final de 4°C.

Em seguida foi realizada a reação de sequenciamento usando 2 µl de Big-Dye Terminator versão 3.1, 1 µl de primer (2.5 µl) e 1 µl de água estéril e ciclo térmico com início de incubação de 30 segundos a 96 °C seguido por 10 segundos a 96 °C, 4 minutos a 60 °C por 30 ciclos e temperatura final de 4 °C. A reação foi purificada por XTerminator.

O sequenciamento do *BLM* foi realizado no sequenciador ABI 3500xL Genetic Analyzer (Applied Biosystems®) e a análise dos dados foi realizada com auxílio do *software* SeqMan NGen® versão 12.0 (DNASTAR, Madison, WI) em comparação com o genoma de referencia (uc002bpr.5 (CRCh38/hg38) e NM_000057.3).

As variantes encontradas foram submetidas aos bancos de dados dbSNP (*Single Nucleotide Database* – <http://www.ncbi.nlm.nih.gov/SNP/>), ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>), LOVD (<https://oi.gene.le.ac.uk/home.php>) e HGMD (<http://www.hgmd.cf.ac.uk/ac/index.php>) que associam variantes e fenótipos (Landrum et al., 2016). As variantes foram classificadas de acordo com as recomendações do *American College of Medical Genetics* (ACMG) (www.acmg.net), cujos critérios incluem:

- **Benignas:** quando não há evidências de que sejam responsáveis pela condição sob investigação e podem ser observadas na população geral;
- **Patogênicas:** quando há evidências de serem causadoras da doença investigada, segundo os bancos de dados genômicos e a literatura;
- **Provavelmente patogênicas:** quando nunca foram descritas como causadora da doença, não são vistas na população geral e afetam substancialmente a função da

proteína, como *stop* códons, deleções e inserções em *frameshift*;

- **Possivelmente patogênicas:** quando nunca foram descritas como causadoras da doença, não são vistas na população geral e são trocas simples de bases que potencialmente afetam a função da proteína;
- **Significado incerto:** quando não há evidências de que sejam responsáveis pela condição e não é possível descartar que sejam patogênicas nem afirmar que sejam benignas.

Após a classificação das variantes, foi realizada uma predição de proteína, que possibilita avaliar se a alteração encontrada pode afetar substancialmente a mesma, e uma predição de *splicing*, que possibilita predizer se a alteração encontrada poderá afetar um sítio de *splicing* ou se poderá criar um novo sítio de *splicing*. Para tanto foram utilizados os programas Mutation Taster (www.mutationtaster.org) e NNSplice (http://fruitfly.org/seq_tools/splice.html).

3.2.3 Análise do transcriptoma pela metodologia de RNA-seq

Os pacientes e controles foram submetidos à coleta de aproximadamente 2,5 mL de sangue periférico em tubo PAXgene™ (PAXgene RNA collection tubes – PreAnalytiX, BD Alemanha), por punção venosa, para extração de RNA. Estes tubos foram incubados em temperatura ambiente por 2 horas para estabilização do RNA e então foram armazenados a -80°C até o momento da extração. Após a extração, seguiram-se o sequenciamento do RNA mensageiro pela tecnologia de RNA-Seq e análise dos dados gerados.

A extração do RNA foi realizada na estação de trabalho semiautomatizada denominada QIAcube® (QIAGEN®) utilizando o kit de extração QIAGEN® PAXgene™ Blood RNA System Kit, seguindo o protocolo descrito pelo fabricante. Concisamente, após o descongelamento seguido por aproximadamente 2 h de repouso em posição vertical, os tubos de amostras foram centrifugados a 4.000 rpm por 10 minutos. O sobrenadante foi descartado vertendo o tubo, e ao pellet foram adicionados 4 mL de água livre de RNase (RNFw - RNase free water). O precipitado foi dissolvido por homogeneização em vórtex. Os tubos foram novamente centrifugados a 4.000 rpm por

10 minutos, o sobrenadante desprezado e ao pellet foram adicionados 350 µL de Buffer 1 (solução BR1 contida no kit de extração). O precipitado foi dissolvido por homogeneização em vórtex. Esse material foi transferido para tubos de 1,5 mL PT (tubo fornecido no kit de extração). Os tubos foram dispostos no QIAcube[®] conforme orientação do fabricante, onde foram extraídos automaticamente. Ao final da extração, os materiais foram eluídos em 40 µL de Buffer 5 (solução BR5 contida no kit de extração).

Após a extração, o RNA total obtido foi quantificado por espectrofotometria pelo Nanodrop[®] ND-1000 (NanoDrop Thermo Technologies, Wilmington, DE) e sua integridade foi determinada pelo Agilent[®] 2100 Bioanalyzer chip (Agilent Technologies, Palo Alto, CA), seguindo os protocolos descritos pelos fabricantes. Foram consideradas adequadas as amostras com razão de absorbância 260/280 nm próxima a 2,0 e integridade com RIN >8 (RNA *Integrity Number*).

As amostras adequadas para a realização do estudo foram processadas para que do RNA total fossem capturados apenas os RNAs mensageiros (mRNA), que por sua vez foram convertidos em DNA complementar (cDNA), para que pudessem ser sequenciados. Para tanto, utilizou-se o protocolo mRNA TruSeq RNA *sample preparation* versão 2 LS *protocol* Illumina[®], constituído por nove etapas principais: Purificação e fragmentação de mRNA; Síntese da primeira fita de cDNA, Síntese da segunda fita de cDNA; *End repair*; adenilação de 3'; Ligação de adaptadores; Enriquecimento; Validação e Normalização, seguindo o protocolo descrito pelo fabricante.

A clusterização (processo de amplificação isotérmica do cDNA em uma *flowcell*) e o sequenciamento, foram realizados na plataforma Illumina[®] HiSeq 2500 (Illumina, Inc., San Diego, CA).

A análise dos dados foi conduzida com auxílio de bioinformatas colaboradores. Na etapa de obtenção dos dados brutos, utilizou-se o *software* CASAVA 1.8.2 fornecido pela Illumina[®], que faz a identificação das bases dos dados brutos (*base call*) e as transforma em *reads* (sequências provenientes da plataforma de sequenciamento) no formato fastq., acompanhados dos scores de qualidade Phred. Os *reads* foram visualizados utilizando o programa FastQC (www.bioinformatics.bbsrc.ac.uk/projects/).

Por se tratar de uma tecnologia nova, ainda não há um referencial teórico ou diretrizes que descrevam uma forma única e precisa de análise dos dados gerados pelo RNA-seq. Nesse sentido, os dados obtidos pelo Illumina® HiSeq 2500 (Illumina, Inc., San Diego, CA) foram analisados utilizando duas abordagens, com metodologias distintas, uma baseada nos programas Bowtie2 e EdgeR com filtragem de *reads* para detecção de expressão diferencial (**Abordagem de análise A1**) e a outra baseada nos programas Rsubread e DESeq2, sem a filtragem de *reads* para detecção de variações genômicas (**Abordagem de análise A2**).

3.2.3.1 Abordagem de análise A1 – Expressão Diferencial (programas Bowtie2 e EdgeR)

Para a análise de expressão diferencial foi utilizada uma abordagem por meio dos programas Bowtie2 e EdgeR, realizada em quatro etapas principais: filtragem dos dados, mapeamento, análise de expressão diferencial e análise de enriquecimento funcional dos genes considerados diferencialmente expressos.

A filtragem dos *reads* de baixa qualidade, sequências de adaptadores e vetores foi realizada pelo programa Seqclean (<https://bitbucket.org/izhbannikov/seqclean>), utilizando como *cutoff* bases com qualidade inferior a 24 QScore. A base de dados de contaminantes usada foi a Univec (<http://www.ncbi.nlm.nih.gov/VecScreen/UniVec.html>). Após a filtragem, *reads* com comprimento inferior a 65 pb foram removidos.

O mapeamento das amostras foi feito contra o genoma referência de *Homo sapiens* (hg19, *Genome Reference Consortium GRCh37 build*, Fev/2009) pelo programa Bowtie2 v2.1.0 (Langmead e Salzberg, 2012), utilizando como opção o mapeamento local e alta sensibilidade, o que maximiza o score do alinhamento e permite uma maior acurácia. Em seguida a qualidade do mapeamento foi checada para cada amostra utilizando o pacote Samtools v.0.1.18 (ferramenta *flagstat*, Li *et al.*, 2009). Por meio do mesmo programa, o mapeamento foi organizado pelo nome dos *reads* (*ferrament sort*).

Para a análise de expressão diferencial utilizou-se a ferramenta EdgeR baseada em linguagem computacional R. Análise de expressão diferencial significa identificar genes que mudaram seus níveis de expressão em diferentes condições biológicas. Neste trabalho, as condições biológicas utilizadas para comparação foram portadores de síndrome de Bloom *versus* controles que não apresentavam a doença, considerados normais. Nessa etapa, como guia, foi dado o arquivo de anotação .gtf do genoma de referência previamente utilizado para extração das contagens brutas pelo *script* do HTSeq-count v.05.4.p2 (<http://www-huber.embl.de/users/anders/HTSeq/doc/index.html>). O HTSeq-count identifica *reads* com mapeamento ambíguo e *reads* que mapearam em múltiplas regiões do genoma, removendo-os.

Os genes considerados diferencialmente expressos foram submetidos a uma análise de enriquecimento funcional baseada em ontologia gênica, que exhibe os termos mais frequentes presentes em uma lista de genes, possibilitando a identificação de processos biológicos significativos que despontam em evidência. Para tanto, foram utilizados os programas DAVID (*Database for Annotation, Visualization and Integrated Discovery*) (Huang et al., 2009) que integra anotações funcionais com sumários gráficos intuitivos e o *Gene Ontology consortium* (www.geneontology.org). Foram consideradas enriquecidas, categorias com valor de $p > 0,05$.

3.2.3.2 Abordagem de análise A2 – Análise de Variantes Genômicas (programas Rsubread e DESeq2)

Na literatura existe uma controvérsia sobre a filtragem de *reads* ainda em estado bruto. Por conta disso, decidiu-se testar o mapeamento de *reads* por sua qualidade, sendo esse o único processo de seleção para identificação de variantes. Há evidências de que esse processo produz melhores resultados em termos de aproveitamento dos *reads*, sem aumento de falsos positivos (Minoche et al., 2011).

O mapeamento das *reads* obtidas foi feito contra o genoma de referência CRCh38/hg38 pelo pacote Subread versão 1.5.0-p1 (Liao et al., 2013). Foram

consideradas para este estudo, apenas *reads* mapeadas uma única vez. Para tanto, foi utilizado o pacote SAMtools versão 1.3 (Li et al., 2009) para remoções de *reads* duplicadas. A cobertura de genes e transcritos foi calculada pelo *software* BEDtools versão 2.17.0 (Aron e Hall, 2010).

Durante o mapeamento, foi possível obter variações genômicas de um único (SNVs, do inglês – *Single Nucleotide Variants*) e Indels (**I**nserções e **D**eleções) pela função exactSNP do pacote Rsubread (Liao et al., 2013), com parâmetros ajustados para detectar quaisquer alterações, e sensibilidade para Indels de até 17 bases de tamanho.

Foram consideradas neste trabalho, apenas SNVs com uma profundidade mínima de 5 *reads* e qualidade de 10. Indels com menos de dois *reads* de suporte, ou razão entre qualidade da sequência e profundidade de leitura menor do que dois foram descartadas. Todas as outras variantes foram descartadas por não apresentarem confiabilidade suficiente de *reads*.

A anotação das variantes e a predição de seus efeitos foram obtidas utilizando o programa VEP (*Ensembl's Variant Effect Predictor*, release 83 – <http://www.ncbi.nlm.nih.gov/pubmed/20562413>), fornecido pelo banco de dados Ensembl.

3.2.4 Validação da expressão do gene *BLM* pela metodologia de RT-qPCR

A validação da expressão do gene *BLM* foi realizada por meio da técnica de reação em cadeia de polimerase quantitativa em tempo real (RT-PCR). Para esta reação, o RNA extraído de cada paciente foi submetido a transcrição reversa para conversão em c-DNA, que foi utilizado para o estabelecimento da técnica de RT-PCR.

A síntese do cDNA utilizado foi realizada a partir de 1 µg de RNA total por meio do kit *High Capacity cDNA Reverse Transcription* (Applied Biosystem, Foster, Estados Unidos), seguindo as orientações do fabricante. Imediatamente antes deste processo, todas as amostras foram tratadas com o kit *Deoxyribonuclease I Amplification Grade* (Invitrogen, Carlsbad, CA, Estados Unidos), para eliminação de qualquer resquício de DNA genômico.

As amplificações por PCR foram realizadas em duplicatas e usando 50 ng de cDNA por reação. As reações foram preparadas com reagentes padronizados para PCR em tempo real (TaqMan Universal PCR Master Mix, Applied Biosystems) adicionados a sonda específica para cada gene. As condições da reação foram: 50°C durante 2 minutos, 95°C durante 10 minutos e 40 repetições de 95°C por 15 segundos, finalizando a 60°C por 1 minuto.

As leituras da fluorescência foram realizadas pelo equipamento StepOne Real Time-PCR (Applied Biosystems) a cada ciclo de amplificação e, posteriormente, analisadas pelo StepOne *Software* v2.3 (Applied Biosystems).

Todas as reações foram submetidas às mesmas condições de análise e normalizadas pelo sinal do corante de referência passiva ROX para correção de flutuações na leitura decorrentes das variações de volume e evaporação ao longo da reação. O resultado em número de CT, refere-se ao número de ciclos de PCR necessários para que o sinal fluorescente atinja o limiar de detecção.

Os níveis de expressão relativa foram calculados segundo o método de quantificação relativa $2^{-\Delta\Delta C_t}$ proposto por Pfaffl (2001), utilizando o *GAPDH* como gene de referência.

Os dados foram verificados quanto à normalidade pelo teste K-S (Teste de Kolmogorov-Smirnov). Os resultados foram expressos em termos de média e erro padrão e analisados por meio do programa GraphPad prism 5 (GraphPad, EUA).

4 RESULTADOS

4.1 Análise Citogenética

Os resultados do estudo citogenético das pacientes K1 e M1 encontram-se resumidos na Tabela 2.

Tabela 2 – Estudo citogenético para confirmação diagnóstica da síndrome de Bloom, baseado em 20 células analisadas para coloração convencional e bandamento G, e em 10 células analisadas para trocas entre cromátides irmãs

Pacientes	CV	BG		CHG
	Nº de quebras espontâneas/metáfase	Bandamento G	Nº de quebras espontâneas/metáfase	Nº de TCI/metáfase
K1	0,55	46,XX	0,25	60,9
M1	0,8	46,XX	0,2	61,3

CV – Coloração convencional; BG – Bandamento G; CHG – Coloração Hoechst-Giemsa para observação de TCI (Trocias entre cromátides irmãs)

A paciente K1, apresentou 11 alterações cromossômicas em análise sob coloração convencional, sendo principalmente, cromossomos dicêntricos e gaps cromatídicos (Figura 10). O cariótipo por bandamento G foi normal para o sexo feminino, porém, com alterações em cinco células, sendo estas fragmentos e quebras cromossômicas. A análise para TCI apresentou frequência aumentada, com média de 60,9 TCI/célula (Figura 11).

Figura 10 – Exemplo de alterações cromossômicas observadas sob análise citogenética convencional da paciente K1. (a) Gap cromatídico – chtg e (b) Cromossômo dicêntrico.

Figura 11 - Exemplo de uma metáfase sob coloração diferencial de cromátides irmãs da paciente K1. A figura demonstra a frequência aumentada de trocas entre cromátides irmãs.

A paciente M1, exibiu 16 alterações cromossômicas em análise sob coloração convencional, sendo que as principais foram gaps cromatídicos e fragmentos (Figura

12). O cariótipo banda G foi normal para o sexo feminino, porém com alterações em quatro células, sendo estas translocações e cromossomos marcadores. A análise para TCI apresentou frequência aumentada, com média de 61,3 TCI/célula (Figura 13).

Figura 12 – Exemplo de alterações cromossômicas observadas sob análise citogenética convencional da paciente M1. (a) Gap cromatídico – chtg e (b) Cromossômo dicêntrico.

Figura 13 - Exemplo de uma metáfase sob coloração diferencial de cromátides irmãs da paciente M1. A figura demonstra a frequência aumentada de trocas entre cromátides irmãs.

4.2 Análise mutacional pela metodologia de Sanger

Com o intuito de confirmar a SB, além da análise citogenética, foi realizada a pesquisa de mutações no gene *BLM* pela metodologia de Sanger. Todas as variantes estão descritas em detalhes na Tabela 3.

O sequenciamento da paciente K1 identificou a presença de três variantes sinônimas do tipo SNV já descritas no banco de dados dbSNP: c.3102G>A:p.Thr1034=, c.3531C>A:p.Ala1177=, c.3945C>T:p.Leu1315=, como observado nas figuras 14, 15 e 16 respectivamente. Além disso, foi observada uma variante intrônica c.3358+32T>G, também já descrita pelo banco de dados dbSNP (Figura 17). Em relação à patogenicidade, as três variantes exônicas foram consideradas benignas de acordo com o banco de dados ClinVar e o programa Mutation Taster, assim como a variante intrônica (ANEXO C). Além disso, a frequência populacional das quatro mutações é de aproximadamente 15%, sendo, portanto, todas variantes consideradas comuns na população em geral e não causadoras da SB.

Figura 14 – Eletroferograma da análise do sequenciamento Sanger, demonstrando uma troca de uma Adenina (A) por Guanina (G) na posição 3102 do éxon 16 do gene *BLM* (c.3102G>A:p.Thr1034=), observada na paciente K1.

Figura 15 – Eletroferograma da análise do sequenciamento Sanger, demonstrando uma troca de uma Citosina (C) por Adenina (A) na posição 3531 do éxon 18 do gene *BLM* (c.3531C>A;p.Ala1177=), observada na paciente K1.

Figura 16 – Eletroferograma da análise do sequenciamento Sanger, demonstrando uma troca de uma Citosina (C) por Timina (T) na posição 3945 do éxon 21 do gene *BLM* c.3945C>T;p.Leu1315=), observada na paciente K1.

Figura 17 – Eletroferograma da análise do sequenciamento Sanger, demonstrando uma troca de uma Timina (T) por Guanina (G) na posição 3358 no íntron entre os éxons 17 e 18 do gene *BLM* (c.3358+32T>G), observada na paciente K1.

A análise da paciente M1 mostrou duas variantes do tipo *missense* já descritas no banco de dados dbSNP: c.3164G>C:p.Cys1055Ser e c.3625T>A:p.Ser1209Thr (Figuras 18 E 19, respectivamente). A primeira variante é descrita como patogênica, de acordo com o banco de dados ClinVar e pelo programa Mutation Taster (ANEXO D). Já a segunda variante, apresenta um conflito de interpretações entre ser patogênica e de significado incerto, mas de acordo com o programa Mutation Taster, essa alteração pode ser lesiva à proteína BLM (ANEXO D). Além disso, essa alteração pode gerar um novo sítio de *splicing* e está em uma região altamente conservada.

Figura 18 – Eletroferograma da análise do sequenciamento Sanger, demonstrando uma variação do tipo *missense* na posição 3164 do éxon 16 do gene *BLM* (c.3164G>C:p.Cys1055Ser), observada na paciente M1.

Figura 19 – Eletroferograma da análise do sequenciamento Sanger, demonstrando uma variação do tipo *missense* na posição 3625 do éxon 19 do gene *BLM* (c.3625T>A:p.Ser1209Thr), observada na paciente M1.

Tabela 3 – Descrição das variantes observadas no gene BLM pela metodologia de sequenciamento Sanger para as pacientes K1 e K2

	PACIENTES					
	K1				K2	
	16	17	18	21	16	19
Éxon						
Mudança do nucleotídeo	c.3102G>A	c.3358+32T>G :c.17IVS(+32)	c.3531C>A	c.3945C>T	c.3164G>C	c.3265T>A
Mudança de aminoácido	p.T1034T	---	p.A1177A	p.L1315L	p.C1055S	p.S1209R
Tipo de Variante	SNV	SNV	SNV	SNV	SNV - não sinônima	SNV - não sinônima
Consequência Molecular	Sinônima	---	Sinônima	Sinônima	Missense	Missense
Significância Clínica	Benigna	Incerto	Benigna	Benigna	Patogênica	Patogênica
ID dbSNP	rs2227933	rs17273842	rs2227934	rs1063147	rs367543029	rs1801256
Posição (HG38)	Chr15:90794249	---	Chr15:90803693	Chr15:90811275	Chr15:90794311	Chr15:90804233

ID – Identificação; dbSNP – Banco de dados de polimorfismos de um único nucleotídeo; SNV – variação de um único nucleotídeo

4.3 Análise do transcriptoma pela metodologia de RNA-seq

Foram sequenciados cinco indivíduos pela metodologia de RNA-seq utilizando-se a plataforma de sequenciamento Illumina® HiSeq 2500, sendo três controles considerados saudáveis e duas pacientes portadoras da SB.

Para ambas abordagens de análise utilizadas neste estudo (A1 e A2), os dados brutos gerados pelo sequenciamento foram processados, obtendo-se um cenário transcricional altamente complexo, no qual foi possível observar a localização precisa dos limites da transcrição, com resolução de um único nucleotídeo e alto nível de eficiência (Tabelas 4 e 5).

Tabela 4 – Dados brutos gerados pela Abordagem de análise A1

Amostras	RT	RAP	RM	RAU	%
K1	66999012	56304018	56256763	43682607	77,6
M1	90798306	75173246	74916578	44412281	59,1
C1	99847166	83403588	83205738	60497992	72,5
C2	126006700	106798830	106591710	68462498	64,1
C3	108081294	90500518	90340118	61279600	67,7

RT – Reads total; RAP – Reads após filtragem (QS>24 e retirando contaminantes); RM – Reads mapeados; RAU – Reads alinhados apenas uma vez; % - Porcentagem de reads alinhados.

Tabela 5 – Dados brutos gerados pela Abordagem de análise A2

Amostra	Quantidade Bruta de <i>reads</i>	<i>Reads</i> mapeadas	Junções	Fusões	Indels
K1	66999012	50374631 (75.19%)	173885	87393	189873
M2	90798306	59144607 (65.14%)	186607	124885	237798
C1	99847166	77476882 (77.60%)	194831	148681	260750
C2	108081294	82783805 (76.59%)	209448	189990	346924
C3	126006700	98927134 (78.5%)	229297	238556	433309

A partir da obtenção dos dados brutos, procedeu-se às análises de expressão diferencial, pela abordagem A1 e de variantes, pela abordagem A2.

4.3.1 Expressão diferencial (abordagem A1 de análise)

Nesta abordagem de análise, ao todo foram mapeados 22.334 genes ao genoma de referência. Desses, um total de 11.938 apresentou o mínimo de cobertura necessária para análise de expressão diferencial. Dentre esses 11.938 genes, 399 foram considerados diferencialmente expressos (DE) entre os portadores de SB e os controles considerados normais, conforme observado na Figura 20. Destes 399 genes DE, 216 foram determinados como *down-regulated* (Hipo-expressos) nos portadores de SB e 183 *up-regulated* (Hiper-expressos) no grupo controle.

Figura 20 – *Maplot* demonstrando a distribuição de *foldchange* para os 11.939 genes adequados para análise de expressão diferencial. Cada ponto representa um gene. Genes qualificados como diferencialmente expressos estão em vermelho.

Todos os 216 genes DE *down regulated* (hipo-expressos) encontrados no grupo com SB, foram submetidos à análise de enriquecimento funcional, onde foi possível observar o envolvimento com importantes processos biológicos e classes de doenças

Foram identificadas três importantes classes de doenças relacionadas às respostas imune, infecciosa e ao câncer, seguindo o DAVID. E, foram identificados sete processos biológicos enriquecidos, seguindo o *Gene Ontology*, sendo todos os processos relacionados à vias imunológicas: replicação negativa da regulação da replicação do genoma viral, regulação positiva da proliferação de células B, via de sinalização mediada por *interferon* gama, ativação de células B, resposta a vírus, resposta imune adaptativa e processo efetor imune.

Dentre esses genes, apresentando maior importância funcional, destacam-se os relacionados a processos imunes (GO:0002376), considerados genes envolvidos em qualquer processo do desenvolvimento ou funcionamento do sistema imune; genes

relacionados à resposta imune, considerados genes relacionados a qualquer processo ligado ao sistema imune que atua no funcionamento da resposta de um organismo em relação a um potencial invasor (GO:0006955), e genes relacionados à resposta de defesa, considerados genes relacionados a reações provocadas em resposta à presença de um corpo estranho ou em decorrência de uma lesão, resultando na restrição de danos para o organismo atacado ou, prevenção/recuperação de uma infecção causada pelo ataque (GO:0006952). Alguns dos genes envolvidos se repetem entre os diferentes processos observados, incluindo os genes *FCGR1B*, *KLRC2*, *CIITA*, *MS4A1* e *FCGR2C*.

Expandindo a análise ontológica ($p < 0,05$), observamos também genes DE associados a vias de controle da apoptose, como processo apoptótico na medula óssea (GO:0071839), processo apoptótico (GO:0006915), morte celular programada (GO:0012501) e morte celular (GO:0008219). Alguns destes genes são *BCL2L1*, *CASP7*, *CDKN1A*, *E2F2* e *ITPR1*.

Surpreendentemente não foram detectados genes DE relacionados a vias de reparo de DNA como era esperado, e o gene *BLM*, considerado na literatura como responsável pela síndrome de Bloom, apresentou níveis de expressão diferencial normal em relação ao grupo controle em ambas as pacientes, por esta metodologia.

4.3.2 Análise de variantes (abordagem A2 de análise)

4.3.2.1 Paciente K1

Para a paciente K1, foram geradas 66999012 *reads*, sendo 50.374.631 (75,1%) mapeadas ao genoma de referência. Após a redução do volume de dados do backup, para evitar redundância nos dados (deduplicação), foram obtidas 32.440.448 *reads* com 15.816.067 (48.75%) mapeadas ao genoma de referência e utilizadas neste estudo.

Foram detectadas um total de 626.008 variantes, classificadas em SNPs, Indels, e em função do impacto exercido, resultando em: 36.219 Indels e 5.513 SNPs de alto

impacto; 1095 Indels e 11.670 SNPs de impacto moderado; 711 Indels e 16.189 SNPs de baixo impacto; e 306.338 Indels e 248.273 SNPs de impacto modificado. Todas as alterações podem ser observadas na Figura 21, exceto as de impacto modificado, que foram excluídas de nossas análises por serem consideradas alterações duvidosas e de difícil predição.

A análise de enriquecimento funcional dos genes relacionados aos transcritos que apresentam as alterações de alto impacto demonstra processos biológicos relacionados às vias de regulação transcricional, atividades catalíticas e vias imunológicas.

No gene *BLM*, foram encontradas 41 variantes, distribuídas em dez transcritos: BLM-001, BLM-002, BLM-003, BLM-004, BLM-005, BLM-006, BLM-007, BLM-009, BLM-010 e BLM-011, como descrito na Tabela 6 e observado na Figura 23.

Figura 21 – Mapa do tipo *circus plot* demonstrando uma visão geral das variações genômicas observadas na paciente K1, pelo estudo do transcriptoma, pela metodologia de RNA-seq pela abordagem de análise A2. Lendo-se de fora para dentro: Cada um dos cromossomos 1-22, X e Y estão representados pelo ideograma disposto circularmente, demarcado por uma escala de megabase na borda exterior da figura; logo em seguida, em vermelho, estão as Indels de alto impacto, em laranja as de médio impacto e em amarelo as de baixo impacto; em seguida encontram-se os tipos de variantes encontradas (círculo cinza), seguido pelos SNPs (círculo azul), impacto dos SNPs (círculo amarelo) e os tipos das variantes encontradas (círculo cinza mais interno). A legenda desta figura está descrita por completo na **Figura 22**.

Figura 22 – Legenda do *circus plot* da Figura 21.

Figura 23 – Mapa do tipo *circus plot* demonstrando uma visão geral das variações genômicas observadas nas pacientes K1 e M1, apenas no gene *BLM*, pelo estudo do transcriptoma, pela metodologia de RNA-seq pela abordagem de análise A2. A legenda desta figura está descrita por completo na **Figura 24**.

Figura 24 – Legenda do *circus plot* da Figura 23.

Tabela 6 – Distribuição das variantes encontradas nos transcritos do gene *BLM* na paciente K1

Localização	Consequência	Impacto	Transcrito	
chr15:90743726_C/T	IV	M	ENST00000355112	BLM-001
chr15:90743726_C/T	IV e NMD	M	ENST00000559724	BLM-003
chr15:90743726_C/T	IV	M	ENST00000560509	BLM-002
chr15:90743726_C/T	UGV	M	ENST00000559282	BLM-006
chr15:90743845_C/T	IV e NMD	M	ENST00000559724	BLM-003
chr15:90743845_C/T	IV	M	ENST00000355112	BLM-001
chr15:90743845_C/T	IV	M	ENST00000560509	BLM-002
chr15:90743845_C/T	UGV	M	ENST00000559282	BLM-006
chr15:90749920_AT/AAT	FSH	H	ENST00000355112	BLM-001
chr15:90749920_AT/AAT	FSH e NMD	H	ENST00000559724	BLM-003
chr15:90749920_AT/AAT	UGV	M	ENST00000558599	BLM-007
chr15:90749920_AT/AAT	FSH	H	ENST00000560509	BLM-002
chr15:90749920_AT/AAT	NCTE e NCTV	M	ENST00000559282	BLM-006
chr15:90766984_AG/AAG	NCTE e NCTV	M	ENST00000559426	BLM-009
chr15:90766984_AG/AAG	NCTE e NCTV	M	ENST00000560136	BLM-004
chr15:90766984_AG/AAG	3PUTR e NMD	M	ENST00000559724	BLM-003
chr15:90766984_AG/AAG	FSH	H	ENST00000355112	BLM-001
chr15:90766984_AG/AAG	FSH	H	ENST00000560509	BLM-002
chr15:90794249_G/A	SV	L	ENST00000560509	BLM-002
chr15:90794249_G/A	SV	L	ENST00000355112	BLM-001
chr15:90794249_G/A	3PUTR e NMD	M	ENST00000559724	BLM-003
chr15:90794249_G/A	NCTE e NCTV	M	ENST00000558825	BLM-005
chr15:90794249_G/A	NCTE e NCTV	M	ENST00000560136	BLM-004
chr15:90794249_G/A	NCTE e NCTV	M	ENST00000560559	BLM-011
chr15:90803693_C/A	SV	L	ENST00000355112	BLM-001
chr15:90803693_C/A	3PUTR e NMD	M	ENST00000559724	BLM-003
chr15:90803693_C/A	IV	M	ENST00000560509	BLM-002
chr15:90803693_C/A	NCTE e NCTV	M	ENST00000560136	BLM-004
chr15:90803693_C/A	NCTE e NCTV	M	ENST00000558825	BLM-005
chr15:90804208_AC/AAAC	3PUTR e NMD	M	ENST00000559724	BLM-003
chr15:90804208_AC/AAAC	FSH	H	ENST00000355112	BLM-001
chr15:90804208_AC/AAAC	IV	M	ENST00000560509	BLM-002
chr15:90804208_AC/AAAC	NCTE e NCTV	M	ENST00000560136	BLM-004
chr15:90804208_AC/AAAC	UGV	M	ENST00000560821	BLM-010
chr15:90804208_AC/AAAC	NCTE e NCTV	M	ENST00000558825	BLM-005
chr15:90811275_C/T	SV	L	ENST00000560509	BLM-002
chr15:90811275_C/T	SV	L	ENST00000355112	BLM-001
chr15:90811275_C/T	3PUTR e NMD	M	ENST00000559724	BLM-003
chr15:90811275_C/T	NCTE e NCTV	M	ENST00000558825	BLM-005
chr15:90811275_C/T	NCTE e NCTV	M	ENST00000560821	BLM-010
chr15:90811275_C/T	NCTE e NCTV	M	ENST00000560136	BLM-004

3PUTR e NMD – 3_prime_UTR_variant&NMD_transcript_variant; FSH – frameshift_variant;
 FSH e NMD – frameshift_variant&NMD_transcript_variant; IV – Intron_variant; IV e NMD –
 intron_variant&NMD_transcript_variant; NCTE e NCTV –
 non_coding_transcript_exon_variant&non_coding_transcript_variant; SV – synonymous_variant;
 UGV - upstream_gene_variant

4.3.2.2 Paciente M1

Para a paciente M1 foram geradas 90.798.306 *reads*, sendo que 59.144.607 (65,1%) foram mapeadas ao genoma de referência. Após deduplicação, foram obtidas 51328076 *reads*, sendo que 19674377 (38,3%) mapearam ao genoma de referência.

Foram detectadas um total de 837.931 variantes que foram classificadas em SNPs, Indels, e em função do impacto exercido, resultando em: 50.893 Indels e 6.584 SNPs de alto impacto, 2.076 Indels e 11.943 SNPs de impacto moderado, 1.098 Indels e 16.638 SNPs de baixo impacto, e 451.413 Indels e 297.286 SNPs de impacto modificado. Todas as alterações podem ser observadas na Figura 25, exceto as de impacto modificado, que foram excluídas de nossas análises por serem consideradas alterações duvidosas e de difícil predição.

Assim como na paciente K1, a análise de enriquecimento funcional dos genes relacionados aos transcritos que apresentam as alterações de alto impacto demonstra processos biológicos relacionados à vias de regulação transcricional, atividades catalíticas e vias imunológicas.

No gene *BLM*, foram detectadas 58 variantes distribuídas em nove transcritos: BLM-001, BLM-002, BLM-003, BLM-004, BLM-005, BLM-006, BLM-007, BLM-009 e BLM-010. Dentre estas, destacam-se duas variantes de alto impacto no transcrito BLM-001, uma do tipo *frameshift* e a outra do tipo *frameshift* e *stop gained*. Uma compilação das variantes encontradas pode ser observada na Tabela 7 e na Figura 23.

Figura 25 – Mapa do tipo *circus plot* demonstrando uma visão geral das variações genômicas observadas na paciente M1, pelo estudo do transcriptoma, pela metodologia de RNA-seq pela abordagem de análise A2. Lendo-se de fora para dentro: Cada um dos cromossomos 1-22, X e Y estão representados pelo ideograma disposto circularmente, demarcado por uma escala de megabase na borda exterior da figura; logo em seguida, em vermelho, estão as Indels de alto impacto, em laranja as de médio impacto e em amarelo as de baixo impacto; em seguida encontram-se os tipos de variantes encontradas (círculo cinza), seguido pelos SNPs (círculo azul), impacto dos SNPs (círculo amarelo) e os tipos das variantes encontradas (círculo cinza mais interno). A legenda desta figura está descrita por completo na **Figura 26**.

Figura 26 – Legenda do *circus plot* da Figura 25.

Tabela 7 – Distribuição das variantes encontradas nos transcritos do gene *BLM* da paciente M1

Nome	Consequência	Impacto	Transcrito	
chr15:90716988_A/G	UGV	M	ENST00000560509	BLM-002
chr15:90716988_A/G	UGV	M	ENST00000559724	BLM-003
chr15:90716988_A/G	UGV	M	ENST00000355112	BLM-001
chr15:90751877_T/A	DGV	M	ENST00000559282	BLM-006
chr15:90751877_T/A	MV e NMD	MOD	ENST00000559724	BLM-003
chr15:90751877_T/A	NCTE e NCTV	M	ENST00000558599	BLM-007
chr15:90751877_T/A	MV	MOD	ENST00000560509	BLM-002
chr15:90751877_T/A	MV	MOD	ENST00000355112	BLM-001
chr15:90760917_AT/AAT	3PUTR e NMD	M	ENST00000559724	BLM-003
chr15:90760917_AT/AAT	FSH	H	ENST00000560509	BLM-002
chr15:90760917_AT/AAT	DGV	M	ENST00000558599	BLM-007
chr15:90760917_AT/AAT	FSH	H	ENST00000355112	BLM-001
chr15:90760917_AT/AAT	UGV	M	ENST00000559426	BLM-009
chr15:90760917_AT/AAT	UGV	M	ENST00000560136	BLM-004
chr15:90761070_T/A	MV	MOD	ENST00000355112	BLM-001
chr15:90761070_T/A	UGV	M	ENST00000559426	BLM-009
chr15:90761070_T/A	UGV	M	ENST00000560136	BLM-004
chr15:90761070_T/A	3PUTR e NMD	M	ENST00000559724	BLM-003
chr15:90761070_T/A	MV	MOD	ENST00000560509	BLM-002
chr15:90761070_T/A	DGV	M	ENST00000558599	BLM-007
chr15:90766984_AG/AAG	FSH	H	ENST00000355112	BLM-001
chr15:90766984_AG/AAG	NCTE e NCTV	M	ENST00000560136	BLM-004
chr15:90766984_AG/AAG	NCTE e NCTV	M	ENST00000559426	BLM-009
chr15:90766984_AG/AAG	3PUTR e NMD	M	ENST00000559724	BLM-003
chr15:90766984_AG/AAG	FSH	H	ENST00000560509	BLM-002
chr15:90785084_A/T	SRV e IV	L	ENST00000560509	BLM-002
chr15:90785084_A/T	SRV e IV e NMD	L	ENST00000559724	BLM-003
chr15:90785084_A/T	SRV e IV e NCTV	L	ENST00000560136	BLM-004
chr15:90785084_A/T	SRV e IV	L	ENST00000355112	BLM-001
chr15:90803600_TA/TTA	IV	M	ENST00000560509	BLM-002
chr15:90803600_TA/TTA	NCTE e NCTV	M	ENST00000558825	BLM-005
chr15:90803600_TA/TTA	3PUTR e NMD	M	ENST00000559724	BLM-003
chr15:90803600_TA/TTA	NCTE e NCTV	M	ENST00000560136	BLM-004
chr15:90803600_TA/TTA	SG e FSH	H	ENST00000355112	BLM-001
chr15:90804208_AC/AAC	3PUTR e NMD	M	ENST00000559724	BLM-003
chr15:90804208_AC/AAC	NCTE e NCTV	M	ENST00000558825	BLM-005
chr15:90804208_AC/AAC	IV	M	ENST00000560509	BLM-002
chr15:90804208_AC/AAC	FSH	H	ENST00000355112	BLM-001
chr15:90804208_AC/AAC	NCTE e NCTV	M	ENST00000560136	BLM-004
chr15:90804208_AC/AAC	UGV	M	ENST00000560821	BLM-010
chr15:90804233_T/A	3PUTR e NMD	M	ENST00000559724	BLM-003
chr15:90804233_T/A	NCTE e NCTV	M	ENST00000558825	BLM-005
chr15:90804233_T/A	IV	M	ENST00000560509	BLM-002
chr15:90804233_T/A	MV	MOD	ENST00000355112	BLM-001
chr15:90804233_T/A	NCTE e NCTV	M	ENST00000560136	BLM-004

Continuação Tabela 7 – Distribuição das variantes encontradas nos transcritos do gene *BLM* da paciente M1

chr15:90804233_T/A	UGV	M	ENST00000560821	BLM-010
chr15:90808786_T/A	IV e NCTV	M	ENST00000558825	BLM-005
chr15:90808786_T/A	IV	M	ENST00000560509	BLM-002
chr15:90808786_T/A	FSH e NMD	M	ENST00000559724	BLM-003
chr15:90808786_T/A	IV e NCTV	M	ENST00000560136	BLM-004
chr15:90808786_T/A	NCTE e NCTV	M	ENST00000560821	BLM-010
chr15:90808786_T/A	IV	M	ENST00000355112	BLM-001
chr15:90812726_AC/AAC	IV	M	ENST00000355112	BLM-001
chr15:90812726_AC/AAC	IV e NCTV	M	ENST00000560821	BLM-010
chr15:90812726_AC/AAC	DGV	M	ENST00000560136	BLM-004
chr15:90812726_AC/AAC	FSH e NMD	M	ENST00000559724	BLM-003
chr15:90812726_AC/AAC	IV	M	ENST00000560509	BLM-002
chr15:90812726_AC/AAC	IV e NCTV	M	ENST00000558825	BLM-005

3'UTR e NMD – 3_prime_UTR_variant&NMD_transcript_variant; FSH – frameshift_variant; FSH e NMD – frameshift_variant&NMD_transcript_variant; IV – Intron_variant; IV e NMD – intron_variant&NMD_transcript_variant; NCTE e NCTV – non_coding_transcript_exon_variant&non_coding_transcript_variant; SV – synonymous_variant; UGV – upstream_gene_variant

4.4 Validação do gene *BLM* pela metodologia de RT-qPCR

Para validar a expressão do gene *BLM* nas amostras das pacientes submetidas ao estudo do transcriptoma, foi utilizada a técnica de quantificação relativa de PCR em tempo real (RT-qPCR).

Foi detectado um aumento do nível de expressão do gene *BLM* em ambas as pacientes (Figura 27), em relação ao controle normal. Os resultados desta análise foram normalizados para o nível de *GAPDH*.

Figura 27 – Gráfico demonstrando a análise da quantificação relativa da expressão do gene *BLM* observada nas pacientes K1 e M1. Os níveis relativos estão representados por barras.

5 DISCUSSÃO

5.1 Análise Citogenética

O estudo citogenético demonstrou uma frequência aumentada de TCI nas duas pacientes estudadas, em relação ao controle normal e instabilidade cromossômica, confirmando a SB nesses pacientes.

Pacientes com SB apresentam instabilidade cromossômica, com achados citogenéticos característicos e exclusivos em suas células somáticas, permitindo assim a confirmação inequívoca no diagnóstico. German (1965) descreveu a alta frequência de alterações cromossômicas na SB, sendo estas de dois tipos: (1) quebras que levam a fragmentos ou a translocações recíprocas entre cromossomos não homólogos e (2) característica exclusiva da SB que consiste na tendência a trocas entre segmentos homólogos, que se expressa em um aumento da frequência de TCI (aproximadamente 10 vezes maior que em pessoas normais).

Recentemente (Tubio e Estivill, 2011) foi descrito, em células tumorais, um fenômeno de instabilidade genômica chamado *chromotripsis*. Trata-se da ocorrência de uma explosão cromossômica, com subsequente reunião de todos os fragmentos (Figura 28). Na tentativa de manter a estabilidade celular, os mecanismos de reparo poderiam ser ineficientes para uma reordenação efetiva das partes quebradas, levando à formação de rearranjos complexos e acúmulo de alterações. Trata-se de um único evento, em uma única célula, que leva a uma catástrofe cromossômica que pode gerar células cancerígenas. Contudo, esse mesmo mecanismo também é sugerido para a formação de rearranjos complexos, observados em diferentes doenças genéticas, não apenas em câncer (Liu et al., 2011). Na SB, observamos diferentes tipos de rearranjos e instabilidades cromossômicas em diferentes células de um mesmo paciente.

Podemos inferir que, além de um defeito nos mecanismos de reparo, estes pacientes são mais suscetíveis a sofrerem eventos de *chromotripsis*, uma das hipóteses

que explicaria a exacerbada instabilidade e a predisposição aumentada para o desenvolvimento de neoplasias nesses indivíduos. Estudos futuros devem completar os dados faltantes para a elucidação de tal questão.

Figura 28 – Demonstração esquemática de chromotripsis: quando um evento de catástrofe cromossômica (como uma explosão) causa a fragmentação do DNA que é subsequentemente reparado levando a um rearranjo do cromossomo. Esse rearranjo pode levar a perda de algum segmento (Tubio e Estivill, 2011).

5.2 Análise mutacional pela metodologia de Sanger

As três variantes exônicas sinônimas encontradas na paciente K1 já foram descritas isoladamente na literatura em portadores da SB (German, 2007). No entanto, foram consideradas não causadoras da doença, uma vez que, além destas variantes, os pacientes já apresentavam mutações causadoras da SB. Além disso, na literatura foi descrita uma variante exônica diferente por paciente, sendo que, não há relatos de que essas três variantes tenham sido encontradas todas em um mesmo paciente, como observado em nosso estudo, fato que aumenta a relevância da descrição e discussão do estudo molecular nessas pacientes.

Aparentemente, mutações sinônimas não levariam à alteração na proteína (Griffiths et al., 2000). Contudo, à luz dos avanços científicos e novas técnicas moleculares (Sauna e Kimchi-Sarfaty, 2011), observamos que as variantes sinônimas podem comprometer a conformação, expressão e função da proteína final, uma vez que afetam: (1) os sítios de *splicings*, levando à formação de um mRNA mais curto e subsequente uma proteína não funcional; (2) a estabilidade do mRNA levando à uma degradação rápida e redução da expressão e (3) resultar em elongação da tradução influenciando os níveis proteicos. Segundo a predição obtida via *mutation taster*, nenhuma das variantes observadas na paciente K1 afeta diretamente a proteína BLM final. Contudo, é recomendada a realização de ensaios funcionais para confirmar tal predição, uma vez que, outros diferentes fatores não avaliados pelo ensaio do transcriptoma como por exemplo, os fatores epigenômicos e a conformação da cromatina, poderiam comprometer a transcrição proteica (Sauna e Kimchi-Sarfaty, 2011; Cao et al., 2015).

Mesmo não sendo consideradas causa direta da doença, aparentemente as variações exônicas observadas na paciente K1, podem, juntas, estar interferindo na expressão do gene *BLM*, e, portanto, causando a SB na paciente em questão.

Em relação a variante intrônica, não é possível predizer se a mesma pode ser causadora da SB, uma vez que temos ainda três variantes exônicas a considerar. No entanto, existem algumas possibilidades de interferência no equilíbrio da regulação gênica, uma vez que tais variantes poderiam gerar um *splicing* anormal de mRNA, quando envolvidas sequências doadoras ou aceptoras, alterar o nível de expressão do gene ou ainda gerar isoformas alternativas da proteína final, afetando a ligação de fatores de spliceossomos próximos a um sítio de *splicing* e levando à utilização de sítios alternativos. Além disso, podem ser relevantes como *enhancers*, atuando como reguladores do próprio gene, intensificando ou diminuindo a expressão de muitos outros genes. Dessa forma, afetariam a regulação e sinalização gênica, levando a modificações pós-transcricionais e resultando em defeitos como a parada do ciclo celular (Cappi et al., 2016). Todas essas possibilidades necessitam ser analisadas em maior profundidade.

No momento, já existem estudos que estão sendo aplicados, principalmente após resultados encontrados por sequenciamento do exoma em diferentes doenças, para detectar se uma SNV é funcional. Em breve espera-se comprovar se essas variantes têm

envolvimento com a produção proteica e se afetam diretamente o fenótipo clínico (Chami et al., 2016).

A variante c.3164G>C(p.Cys1055Ser), observada na paciente M1, já foi descrita como causadora da SB por German (2007), contribuindo assim, para confirmar o diagnóstico clínico dessa paciente.

Já a variante c.3625T>A(p.Ser1209Thr), descrita como sendo de significado incerto, poderá futuramente ser reclassificada como uma nova variante patogênica, já que está localizada em uma região altamente conservada e, de acordo com o *mutation taster*, gera um novo sítio de splicing, capaz de causar um desequilíbrio de haploinsuficiência. Essa potencial reclassificação ainda depende de evidências do papel fisiopatológico da variante no distúrbio clínico.

5.3 Análise do transcriptoma

5.3.1 Análise de expressão diferencial entre SB x controles

Para análise de expressão diferencial recomenda-se o uso de replicatas biológicas para o mesmo grupo. Essa repetição parece não contribuir, necessariamente, para o incremento da precisão do experimento, mas pode ser importante para o aumento da precisão das estimativas de médias e de outras funções das variáveis respostas, contribuindo para o aumento da confiabilidade dessas estimativas e da sensibilidade do experimento para detectar pequenas, mas importantes, diferenças de expressão (Kvam e Liu, 2012). No entanto, devido ao fato de que as pacientes deste estudo são portadoras de uma doença rara, só foi possível realizar essa análise em duas pacientes acometidas pela SB (pacientes SB1 e SB2) *versus* o grupo controle (composto por três indivíduos considerados saudáveis, C1, C2 e C3). Apesar do número restrito de amostras, obtivemos resultados extremamente valiosos, que muito podem esclarecer a cerca do comportamento da doença.

Foi observado um total de 216 genes diferencialmente expressos no grupo com SB. Os mesmos foram anotados para obtenção da associação com vias biologicamente relevantes associadas à SB. A anotação do transcriptoma promove uma visão da estrutura, função e processo biológico dos genes montados. A ontologia gênica promove uma linguagem sistemática em 3 domínios biológicos chave: Função molecular, processo biológico e componente celular para unificar a representação de características dos genes. Os termos GO (*Gene Ontology*) são estruturados como um gráfico e podem ser distribuídos em níveis: nível 1 representa a maioria das categorias gerais e promove maior cobertura, enquanto níveis mais altos promovem informações mais específicas e menor cobertura (Gramazio et al., 2016). Brevemente, o nível 1 é uma descrição geral de um processo, enquanto os outros níveis promovem uma descrição mais específica.

Por meio dessa análise, notamos que o perfil transcricional dos 216 genes DE *down-regulated* (hipo-expressos), encontrados no grupo com SB em relação aos controles, estão relacionados principalmente a processos de vias imunológicas como o reconhecimento e apresentação de antígenos, ativação e regulação do sistema complemento, entre outros. Esse achado corrobora com a função imunológica anormal apresentada pelos pacientes acometidos pela SB, como a redução das concentrações de imunoglobulinas plasmáticas. Nesses pacientes, são observados graus variados de imunodeficiência celular e humoral de etiologia desconhecida (Sanz e German, 2006; Resende et al., 2007, Cuniff et al., 2016). Nesse sentido, o perfil transcricional descrito em nosso trabalho pode explicar a desregulação do sistema imune apresentada pelos portadores de SB.

Além disso, foram observados genes DE relacionados à ativação e regulação de apoptose em medula óssea e linfócitos, um mecanismo de morte celular programada. Esse processo geralmente é caracterizado por mudanças morfológicas da célula e ativações de processos bioquímicos, sendo considerado um mecanismo vital que envolve vários processos biológicos normais, como o desenvolvimento celular, o funcionamento do sistema imune e o desenvolvimento embrionário, entre outros (Elmore, 2007). Uma célula que tenha seu mecanismo de apoptose desregulado, não sofrerá a morte programada e, portanto, poderá acumular mutações e repassá-las para a próxima geração celular. Nesse sentido, consideramos que a alteração dos mecanismos regulatórios da apoptose poderia gerar uma instabilidade genômica afetando o sistema

imune e ainda levando a quadros neoplásicos nesses pacientes, uma vez que, as células tendem a acumular danos (mutações) que podem alterar a sinalização de progressão do ciclo celular. Além disso, a resposta a apoptose ocorre em resposta a proteólise da proteína BLM, estando ligada portanto a expressão do gene *BLM* (Wang e Hu, 2008). Corroborando com nossos achados, na literatura também consta que mecanismos de apoptose estão fortemente relacionados à SB, sendo os responsáveis pela baixa estatura, inclusive intrauterino, como observado em nossas duas pacientes (Kaneko et al., 2011).

Do mesmo modo, um mecanismo inadequado de apoptose pode estar associado aos diferentes tipos de câncer observados entre os pacientes com SB. Sabemos que neoplasia é a principal causa de morte nesses pacientes (Amor-Gueret, 2006; Liu e West, 2008) e cerca de 42% dos afetados desenvolvem um câncer em idade precoce e, aproximadamente 10% destes desenvolvem mais de um tipo de neoplasia - principalmente linfomas, carcinomas epiteliais, hematopoiéticos e em alguns casos tumor de *Wilms* e retinoblastomas (German, 1997; Ellis et al., 1998; Ma et al., 2001; Bhisitkul e Rizem, 2004; Sanz e German, 2006; Resende et al., 2007; Moreira et al., 2013; Cuniff et al., 2016).

Embora nenhuma das duas pacientes estudadas nesse projeto tenha desenvolvido algum tipo de neoplasia até o momento, sabemos que o surgimento da doença está fortemente associado a um mecanismo inapropriado de apoptose e alterações imunológicas. Portanto, a SB requer uma vigilância constante no manejo dessas pacientes.

Outro achado inesperado foi o nível de expressão do gene *BLM* que, pela técnica de RNA-seq, foi considerado normal. Esse fato pode também ser decorrente da falta de replicatas biológicas. Contudo, mesmo não observando diferença na expressão do gene *BLM*, nosso resultado sugere que pode haver um desequilíbrio pós transcricional, gerado por um provável defeito na proteína BLM final, que pode estar atuando de maneira incorreta, gerando a instabilidade genômica observada nas pacientes deste estudo.

5.3.2 Análise de variações genômicas pela metodologia de RNA-seq

A enorme quantidade de dados gerados pelo sequenciamento nos leva a um novo e interessante desafio, que é identificar quais Indels e/ou SNVs, detectados ao longo de todo o genoma, estão de fato relacionadas com os achados clínicos observados em nossas pacientes.

Conforme esperado, observou-se uma quantidade exorbitante de variações genômicas em ambas as pacientes estudadas. Dessa forma, nosso enfoque sobressaiu em todas as variantes de alto impacto e nas variantes observadas nos diferentes transcritos do gene *BLM*. O estudo destas variantes revelou alterações em vias biológicas importantes que causam doenças, portanto sua identificação foi essencial para nosso estudo.

A maioria das variações obtidas, assim como os genes diferencialmente expressos, também está associada à vias imunológicas, demonstrando uma ligação entre essas alterações e a imunodeficiência observada nas pacientes deste estudo.

Identificar variações genômicas (Indels e SNPs) é um passo crucial para desvendar a correlação genético-fenótipo, podendo revelar mecanismos importantes sobre a doença estudada (Piskol et al., 2013).

5.3.3 Transcriptoma aplicado à SB

A análise do transcriptoma é uma abordagem nova para o esclarecimento de variáveis funcionais do genoma. É uma técnica que agrega capacidades para responder a diversas questões, tais como a quantificação dos níveis de expressão gênica, a detecção de splicing alternativo, fusões gênicas e edição de RNA, entre outras, sendo fundamental uma análise computacional adequada e eficiente (Piskol et al., 2013).

No entanto, ainda não existe um consenso sobre quais são as melhores ferramentas estatísticas de análise (*pipe-lines*) ou quantas réplicas biológicas são

necessárias para assegurar uma interpretação biológica válida dos resultados (Schurch et al., 2016).

Nos experimentos com RNA-seq, o tamanho do efeito biológico também é um determinante importante para ajudar a validar um *insight* relevante no nível individual do gene ou dos sistemas biológicos (Khang e Lau, 2015).

Em nosso estudo, a análise dos pacientes chamou a atenção para mecanismos importantes que envolvem diferentes vias de resposta imunológica que, nesse caso, têm um peso relevante no aspecto clínico dos pacientes que apresentam a SB.

Embora bastante interessantes, os resultados obtidos podem ser considerados subjetivos, em razão da falta de replicatas das amostras, o que poderia resultar em vies nas diferenças de expressão, devido à variação biológica inerente a cada amostra (Fang et al., 2011; Li et al., 2013).

Apesar do baixo ruído e da alta sensibilidade da plataforma de RNA-Seq utilizada, projetar experimentos com repetição biológica é ainda fundamental para identificar alterações na abundância de RNA.

Até o momento, a literatura recomenda que experimentos que envolvam a técnica de RNA-Seq incluam pelo menos três replicatas biológicas (Auer & Doerge, 2010; Oshlack et al., 2010). Nesse caso, um ponto importante a considerar no presente estudo seria o tamanho da casuística, uma vez que a raridade da doença estudada proporciona grande dificuldade para recrutamento de pacientes em um estudo prospectivo, principalmente quando os critérios de inclusão são bastante restritivos.

Adicionalmente, a relevância de efeito biológico na desregulação imune e o ineditismo do estudo podem conferir um valor significativo aos achados.

Ainda, a sutil discrepância observada entre os resultados obtidos pelas diferentes técnicas utilizadas neste trabalho, revela que uma análise personalizada para cada paciente, talvez seja mais benéfica para o manejo da doença.

A análise conjunta pode nos indicar um perfil geral de como a doença se comporta. No entanto, o transcriptoma é tempo dependente, o que o torna diferente em

cada estágio da doença, fazendo-se necessário, também, um estudo individual para verificar como ele se comporta em cada fase. Assim, sugerimos que o estudo do transcriptoma seja realizado de duas maneiras: conjunta (análise de expressão diferencial) e individual (análise de variações genômicas), principalmente nos casos de falta de replicatas biológicas.

5.4 Validação do gene *BLM* por RT-qPCR

O experimento para a validação dos transcritos do gene *BLM* pela metodologia de RT-qPCR demonstrou uma expressão diferencialmente aumentada para ambas as pacientes do estudo, não corroborando de forma esperada com os resultados obtidos pelo transcriptoma, cuja a expressão do gene *BLM* foi considerada normal. Esse fato pode ser explicado pela possibilidade do resultado aumentado por RT-qPCR refletir a presença de isoformas dos transcritos, já que existem 11 isoformas descritas e apenas duas codificadoras de proteínas.

Não há nada descrito na literatura sugerindo níveis elevados do gene *BLM* em casos de SB. Contudo, sua expressão aumentada é descrita em casos muito agressivos de câncer de mama e cólon retal (Cunniff et al., 2016).

Em nossas pacientes, observamos o contrário, ou seja, quadros clínicos menos agressivos de manifestação da doença. Na paciente K1, em que a expressão do gene *BLM* foi considerada superior em relação à paciente M1 e o controle normal, observamos que a doença se manifesta quase sem intercorrências, exceto as infecções recorrentes observadas na infância.

A literatura prediz que a SB é decorrente de mutações que causam uma hipo-expressão do gene *BLM* (Killen, 2011). Contudo, é possível que a hiper-expressão observada em nosso estudo esteja afetando de forma negativa os mecanismos antirecombinantes levando à manifestação (mesmo que de forma mais branda) do quadro clínico de SB.

6 CONSIDERAÇÕES FINAIS

Os resultados obtidos neste trabalho mostram que a técnica de sequenciamento do transcriptoma de RNAs mensageiros (mRNAs) sanguíneo, por meio da técnica de RNA-Seq utilizando a plataforma Illumina® HiSeq 2500 (Illumina, Inc., San Diego, CA), em pacientes com a síndrome de Bloom, foi extremamente útil para obter um perfil de expressão gênica diferencial inicial, que sugere principalmente interferência na regulação de processos imunológicos, como os relacionados à apoptose, ativação e regulação do sistema complemento e diferenciação de células B, entre outros.

Foi realizado o estudo de enriquecimento funcional dos genes diferencialmente expressos obtidos para a síndrome de Bloom e foram detectadas vias biológicas relevantes. Não foram encontrados, na literatura, estudos que descrevam tais vias.

O presente trabalho, também permitiu detectar alterações genômicas em portadores da síndrome de Bloom, que podem ser relevantes para o fenótipo, demonstrando que, possivelmente, uma análise personalizada do transcriptoma seja cada vez mais necessária para um manejo adequado da doença. Contudo, é necessário o desenvolvimento de novas ferramentas que permitam uma análise mais robusta, excluindo a necessidade de replicatas biológicas.

Observamos por fim, que o grande desafio está em interpretar os significados biológicos consistentes, associados à enorme quantidade de dados gerados pelas plataformas de sequenciamento, tornando imperativo o incremento de métodos eficazes nas análises e interpretação funcional das listas de genes derivadas destes estudos.

7 CONCLUSÕES

Durante este estudo:

- ✓ Realizamos o sequenciamento do transcriptoma sanguíneo de RNAs mensageiros (mRNAs), por meio da técnica de RNA-seq em pacientes com a síndrome de Bloom.
- ✓ Implantamos a técnica de RNA-seq no Laboratório de Citogenômica do LIM03 (HC-FMUSP);
- ✓ Obtivemos o perfil de expressão gênica da síndrome de Bloom, detectando vias biológicas relevantes
- ✓ Estudamos e anotamos com detalhes, pela metodologia de RNA-seq, as variantes genômicas de alto impacto relevantes para o fenótipo e consequente manejo da doença, em ambas as pacientes.

ANEXO A

Parecer aprovado da Comissão de Ética para Análise de Projetos de Pesquisa (Cappesq)

Hospital das Clínicas da FMUSP
Comissão de Ética para Análise de Projetos de Pesquisa - CAPPesq

PROJETO DE PESQUISA

Título: INVESTIGAÇÃO DO TRANSCRIPTOMA EM PORTADORES DE SÍNDROMES DE INSTABILIDADE CROMOSSÔMICA

Pesquisador Responsável: Leslie Domenici Kulikowski **Versão:** 3

Pesquisador Executante: Marília Borges Moreira **CAAE:** 04082512.5.0000.0068

Finalidade Acadêmica: Doutorado

Instituição: HCFMUSP

Departamento: PATOLOGIA

PARECER CONSUBSTANCIADO DO CEP

Número do Parecer: 63954

Data da Relatoria: 31/07/12 – *ad-referendum*

Apresentação do Projeto: O Projeto está apresentado de maneira clara, com objetivos pertinentes e será desenvolvido como Tese de Doutorado. O trabalho proposto é bastante interessante e de extrema importância para a área de diagnóstico de doenças genéticas, oferecendo uma metodologia de ponta na tentativa de elucidação de mecanismos gênicos envolvidos em síndromes de instabilidade cromossômica.

Objetivo da Pesquisa: O objetivo da pesquisa é investigar o transcriptoma total sanguíneo de pacientes com diferentes síndromes de instabilidade cromossômica e avaliar a expressão de RNAs para construir uma possível rede de interações gênicas observadas nas síndromes de instabilidade cromossômica. Serão avaliados 15 pacientes em acompanhamento no ambulatório de Genética do Instituto da Criança do HCFMUSP portadores das seguintes síndromes: Anemia de Fanconi, Ataxia Telangiectasia, síndrome de Bloom, síndrome de Rothmund-Thomson, síndrome de Nijmegen Breakage, síndrome de Cockayne e Xeroderma Pigmentoso, que já tenham os respectivos diagnósticos moleculares.

Avaliação dos Riscos e Benefícios: Não há riscos para os pacientes, uma vez que o projeto contempla apenas coleta de sangue. Os benefícios seriam melhor diagnóstico dessas síndromes, com possibilidade de descoberta de mecanismos que ajudariam no melhor direcionamento de tratamento.

Comentários e Considerações sobre a Pesquisa: A pesquisa proposta é de extrema importância para o entendimento dessas síndromes de instabilidade cromossômica, na medida em que o estudo do transcriptoma poderá gerar dados inéditos que não são obtidos por meio de sequenciamento direto de DNA. A metodologia é de ponta e inovadora.

Considerações sobre os Termos de apresentação obrigatória: A pesquisadora realizou todas as modificações e correções sugeridas anteriormente, enviando novas versões do projeto e do TCLE.

Recomendações: Não há.

Conclusões ou Pendências e Lista de Inadequações: A pesquisadora realizou as modificações sugeridas. Não há mais nenhuma pendência ética.

Situação do Parecer: Aprovado

Necessita Apreciação da CONEP: Não

Considerações Finais a critério do CEP: Aprovado

São Paulo, 31 de Julho de 2012

Prof. Dr. Luiz Eugênio Garcez Leme
Coordenador
Comissão de Ética para Análise de
Projetos de Pesquisa-CAPPesq

Rua Dr. Ovídio Pires de Campos, 225 - Prédio da Administração - 5º andar
CEP 05403-010 - São Paulo - SP.
55 11 2661-6442 ramais: 16, 17, 18 e 20 | cappesq@hcnet.usp.br

ANEXO B

Termo de Consentimento Livre e Esclarecido

1

**HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DA UNIVERSIDADE
DE SÃO PAULO-HCFMUSP**
TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO**DADOS DE IDENTIFICAÇÃO DO SUJEITO DA PESQUISA OU RESPONSÁVEL LEGAL**

1. NOME DO PACIENTE:
DOCUMENTO DE IDENTIDADE Nº: SEXO: M ☐ F ☐
DATA NASCIMENTO:/...../.....
ENDEREÇO Nº APTO:
BAIRRO: CIDADE
CEP: TELEFONE: DDD (.....)
2. RESPONSÁVEL LEGAL
NATUREZA (grau de parentesco, tutor, curador etc.)
DOCUMENTO DE IDENTIDADE: SEXO: M ☐ F ☐
DATA NASCIMENTO:/...../.....
ENDEREÇO: Nº APTO:
BAIRRO: CIDADE:
CEP: TELEFONE: DDD (.....)

DADOS SOBRE A PESQUISA

1. TÍTULO DO PROTOCOLO DE PESQUISA: *“Investigação do transcriptoma em portadores síndromes de instabilidade cromossômica”*

PESQUISADOR: **Dra. Leslie Domenici Kulikowski**

CARGO/FUNÇÃO: **Pesquisador**

INSCRIÇÃO CONSELHO REGIONAL Nº **56833/01D**

UNIDADE DO HCFMUSP: **Departamento de Patologia**

2. AVALIAÇÃO DO RISCO DA PESQUISA:

RISCO MÍNIMO ☒ RISCO MÉDIO ☐
RISCO BAIXO ☐ RISCO MAIOR ☐

3. DURAÇÃO DA PESQUISA : **4 anos**

Rubrica do sujeito de pesquisa ou responsável _____

Rubrica do pesquisador _____

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DA UNIVERSIDADE DE SÃO PAULO-HCFMUSP**1 – Desenho do estudo e objetivo(s)**

O seu filho está sendo convidado a participar desse estudo, pois, ele apresenta uma doença genética e cromossômica. Essa participação é voluntária e se em qualquer momento você não quiser mais que seu filho participe do estudo, não haverá nenhum prejuízo de atendimento ao seu filho e/ou aos seus familiares na instituição.

2 – Descrições dos procedimentos que serão realizados, com seus propósitos e identificação dos que forem experimentais e não rotineiros

Será realizada uma coleta de 5ml de sangue periférico.

3 – Relação dos procedimentos rotineiro e como serão realizados

Coleta de sangue periférico: será feita uma punção venosa ("colher o sangue da veia") de 5ml de sangue periférico de veias de um dos braços, o local poderá ficar um pouco dolorido. Também pode ocorrer a formação de um pequeno hematoma ("roxo") que pode persistir por alguns dias; porém se desfaz sem a necessidade de medicamentos.

4 – Descrição dos desconfortos e riscos esperados nos procedimentos do item 2

O desconforto associado ao exame consiste apenas naquele relacionado à coleta de sangue, o que não confere riscos à saúde.

5 – Benefícios para o participante

Trata-se de um estudo experimental no qual esperamos encontrar pistas para o tratamento das doenças cromossômicas e para buscar melhora no prognóstico. No futuro, o estudo poderá ajudar a detectar mais precocemente complicações que possam vir a surgir com a doença, para garantir um tratamento mais efetivo para muitos doentes.

Rubrica do sujeito de pesquisa ou responsável_____

Rubrica do pesquisador_____

6 – Relação de procedimentos alternativos que possam ser vantajosos, pelos quais o paciente pode optar

Não se aplica.

7 – Garantia de acesso: em qualquer etapa do estudo

Você terá acesso aos profissionais responsáveis pela pesquisa para esclarecimento de eventuais dúvidas. O principal investigador é a **Dra. Leslie Domenici Kulikowisk** que pode ser encontrada no endereço **Av. Dr. Enéas de Carvalho Aguiar, 255**, Telefone(s) **2661.9506**. Se você tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com o Comitê de Ética em Pesquisa (CEP) – Rua Ovídio Pires de Campos, 225 – 5º andar – tel: 2661-6442 ramais 16, 17, 18 –

e-mail: cappesq@hcnet.usp.br

8 – Você tem a garantia da liberdade de retirada do consentimento a qualquer momento e deixar de participar do estudo, sem qualquer prejuízo à continuidade de seu tratamento na Instituição

9 – Você tem direito de confidencialidade – As informações obtidas serão analisadas em conjunto com outros pacientes, não sendo divulgada a identificação de nenhum paciente.

10 – Você tem direito de ser mantido atualizado sobre os resultados parciais das pesquisas, quando em estudos abertos, ou de resultados que sejam do conhecimento dos pesquisadores

11 – Despesas e compensações: não há despesas pessoais para o participante em qualquer fase do estudo, incluindo exames e consultas. Também não há compensação financeira relacionada à sua participação.

12 - Você quer ser consultado se o material coletado do seu filho for usado em uma próxima etapa da pesquisa? () Sim () Não

Mesmo que não queira ser informado, o material coletado de seu filho só será utilizado com autorização do comitê de ética em pesquisa (Cappesq – HCFMUSP).

Rubrica do sujeito de pesquisa ou responsável_____

Rubrica do pesquisador_____

**HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DA UNIVERSIDADE
DE SÃO PAULO-HCFMUSP**

Acredito ter sido suficientemente informado a respeito das informações que li ou que foram lidas para mim, descrevendo a participação do meu filho no estudo "*Investigação do transcriptoma em portadores síndromes de instabilidade cromossômica*".

Eu discuti com a **Dra. Leslie Domenici Kulikowisk** sobre a minha decisão em deixar meu filho participar desse estudo. Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, seus desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro também que a participação do meu filho é isenta de despesas e que ele terá garantia do acesso a tratamento hospitalar quando necessário. Concordo voluntariamente em deixar meu filho participar deste estudo e poderei retirar o meu consentimento a qualquer momento, antes ou durante o mesmo, sem penalidades ou prejuízo ou perda de qualquer benefício que meu filho possa ter adquirido, ou no meu atendimento neste Serviço.

Assinatura do representante legal Data ____ / ____ / ____

Assinatura da testemunha Data ____ / ____ / ____

(para casos de pacientes menores de 18 anos, analfabetos, semi-analfabetos ou portadores de deficiência auditiva ou visual.)

(Somente para o responsável do projeto)

Declaro que obtive de forma apropriada e voluntária o Consentimento Livre e Esclarecido deste paciente ou representante legal para a participação neste estudo.

Assinatura do responsável pelo estudo Data ____ / ____ / ____

Rubrica do sujeito de pesquisa ou responsável _____

Rubrica do pesquisador _____

ANEXO C

Análises de predição de proteína para as mutações encontradas na Paciente K1

03/11/2016

no title [MT]

[documentation](#)

mutation t@sting

Prediction polymorphism

Model: *without_aae*, prob: 0.99999925140095
 (classification due to TGP/ExAC, [real probability](#) is shown anyway) ([explain](#))
[hyperlink](#)

Summary

- homozygous in TGP or ExAC
- protein features (might be) affected
- splice site changes

analysed issue analysis result

name of alteration no title
 alteration (phys. location) chr15:91341599T>G [show variant in all transcripts](#) [IGV](#)
 HGNC symbol [BLM](#)
 Ensembl transcript ID [ENST00000355112](#)
 Genbank transcript ID [NM_000057](#)
 UniProt peptide [P54132](#)
 alteration type single base exchange
 alteration region intron
 DNA changes g.81042T>G
 AA changes N/A
 position(s) of altered AA N/A
 if AA alteration in CDS
 frameshift N/A
 known variant Reference ID: [rs17273842](#)

database	homozygous (G/G)	heterozygous	allele carriers
1000G	54	617	671
ExAC	1636	15976	17612

regulatory features H3K36me3, Histone, Histone 3 Lysine 36 Tri-Methylation

phyloP / phastCons

PhyloP PhastCons
 (flanking) -2.562 0
 -0.288 0
 (flanking) -0.069 0

[explain score\(s\)](#) and/or inspect your position(s) in [in UCSC Genome Browser](#)

splice sites effect gDNA position score wt detection sequence exon-intron border
 Acc 81034 wt: 0.25 / wt: 0.43 AGTCATCTGTTTGAATGTTGAGTTACTTCAATTGAAATT gttt|GAGT
 increased mu: AGTCATCTGTTTGAATGTTGAGTTACGTCATTGAAATT
 Acc 81036 wt: 0.46 / wt: ttgt|GTTA

03/11/2016

no title [MT]

increased mu: 0.58 TCATCTGTTTTGAATGTTTGAGTTACTTCAATTGAAATTGA
 mu: TCATCTGTTTTGAATGTTTGAGTTACGTCAATTGAAATTGA

Donor 81046 wt: 0.54 / wt: CTTCAATTGAAATTG TCAA|ttga
 increased mu: 0.85 mu: CGTCAATTGAAATTG

Donor 81041 wt: 0.47 / wt: AGTTACTTCAATTGA TTAC|ttca
 increased mu: 0.98 mu: AGTTACGTCAATTGA

Donor 81036 wt: GTTTGAGTTACTTCA TTGA|gtta
 marginally 0.9557 / mu: GTTTGAGTTACGTCA
 increased mu: 0.9987
 (marginal change - not scored)

distance from splice site 32

Kozak consensus sequence altered? N/A

conservation protein level for non-synonymous changes N/A

protein features

start (aa)	end (aa)	feature	details
1207	1209	STRAND	might get lost (downstream of altered splice site)
1211	1233	HELIX	might get lost (downstream of altered splice site)
1212	1292	DOMAIN	HRDC. might get lost (downstream of altered splice site)
1237	1240	HELIX	might get lost (downstream of altered splice site)
1243	1252	HELIX	might get lost (downstream of altered splice site)
1257	1261	HELIX	might get lost (downstream of altered splice site)
1263	1265	STRAND	might get lost (downstream of altered splice site)
1268	1273	HELIX	might get lost (downstream of altered splice site)
1275	1288	HELIX	might get lost (downstream of altered splice site)
1295	1295	MOD_RES	Phosphoserine. might get lost (downstream of altered splice site)
1296	1296	MOD_RES	Phosphoserine. might get lost (downstream of altered splice site)
1303	1303	MOD_RES	Phosphoserine. might get lost (downstream of altered splice site)
1310	1310	MOD_RES	Phosphoserine. might get lost (downstream of altered splice site)
1334	1349	MOTIF	Nuclear localization signal (Potential). might get lost (downstream of altered splice site)
1411	1411	MOD_RES	N6-acetyllysine. might get lost (downstream of altered splice site)

length of protein N/A

AA sequence altered N/A

position of stopcodon in wt / mu CDS N/A

03/11/2016	no title [MT]				
position (AA) of stopcodon in wt / mu AA sequence	N/A				
position of stopcodon in wt / mu cDNA	N/A				
poly(A) signal	N/A				
conservation nucleotide level for all changes - no scoring up to now	N/A				
position of start ATG in wt / mu cDNA	119 / 119				
chromosome	15				
strand	1				
last intron/exon boundary	4195				
theoretical NMD boundary in CDS	4026				
length of CDS	4254				
coding sequence (CDS) position	N/A				
cDNA position (for ins/del: last normal base / first normal base)	N/A				
gDNA position (for ins/del: last normal base / first normal base)	81042				
chromosomal position (for ins/del: last normal base / first normal base)	91341599				
original gDNA sequence snippet	GTTTGAATGTTTGAGTTAC T TCAATTGAAATTGAACATCT				
altered gDNA sequence snippet	GTTTGAATGTTTGAGTTAC G TCAATTGAAATTGAACATCT				
original cDNA sequence snippet	N/A				
altered cDNA sequence snippet	N/A				
wildtype AA sequence	MAAVPQNNLQ EQLERHSART LNNKLSLSKP KFSGFTFKKK TSSDNNVSVT NVSVAKTPVL RNKDVNVTED FSFSEPLPNT TNQQRVKDFF KNAPAGQETQ RGGSKSLPD FLQTPKEVVC TTQNTPTVKK SRDTALKKLE FSSSPDSLST INDWDDMDDF DTSETSKSFV TPPQSHFVRV STAQKSKKGK RNFFKAQLYT TNTVKTDLPP PSSESEQIDL TEEQKDDSEW LSSDVICIDD GPiAEVHINE DAQESDSLKT HLEDERDNSE KKKNLEEAEH HSTEKVPCIE FDDDDYDTDF VPPSPPEEII ASSSSSKCLS TLKDLTSDR KEDVLSTSKD LLSKPEKMSM QELNPETSTD CDARQISLQQ QLIHVMERIC KLIDTIPDDK LKLDCGNEL LQQRNIRRLK LTEVDFNKSD ASLLGSLWRY RPDSDLGPME GDSCPTGNSM KELNFSHLP NSVSPGDCLL TTTLGKTGFS ATRKNLFERP LFNTHLQKSF VSSNWAETPR LGKKNESSYF PGNVLTSTAV KDQNKHTASI NDLERETQPS YDIDNFDIDD FDDDDWEDI MHNLAASKSS TAAYQPIKEG RPIKSVSERL SSAKTDCLPV SSTAQNINFS ESIQNYTDKS AQNLASRLK HERFQSLSFP HTKEMMKIFH KKFGLHNFRT NQLEAINAAL LGEDCFILMP TGGGKSLCYQ LPACVSPGVT VVISPLRSLI VDQVQKLTSL DIPATYLTGD KTDSEATNIY LQLSKKDPPI KLLYVTPEKI CASNRLISTL ENLYERKLLA RFVIDEAHCV SQWGHDFRQD YKRMNMLRQK FPSVPMALT ATANPRVQKD ILTQLKILRP QVFSMSFNRH NLKYYVLPKK PKKVAFDCLC WIRKHHPYDS GIIYCLSRRE CDTMADTLQR DGLAALAYHA GLSDSARDEV QQKWINQDGC QVICATIAFG MGIDKPDVRF VIHASLPKSV EGYYESGRA GRDGEISHCL LFYTYHDVTR LKRLIMMEKD GNHHTRETHF NNLYSMVHYC ENITECRRIQ LLAYFGENG FNPDFCKKHPD VSCDNCKCTK DYKTRDVTDD VKSIVRVQVE HSSSQGMRNI KHVGPGRFT MNMLVDIFLG SKSAKIQSGI FGKGSAYSRRH NAERLFKKLI LDKILDEDLY INANDQAIAY VMLGNKAQTV LKRLIMMEKD GNHHTRETHF QKALVAKVSQ REEMVKKCLG ELTEVCKSLG KVFGVHYFNI FNTVTLKKLA ESLSSDPEVL				

03/11/2016	no title [MT]
	LQIDGVTEDK LEKYGAEVIS VLQKYSEWTS PAEDSSPGIS LSSSRGPGRS AAEEELDEEIP VSSHVFASKT RNERKRKKMP ASQRSKRRKT ASSGSKAKGG SATCRKISSK TKSSSIIGSS SASHTSQATS GANSKLGIMA PPKPINRPFL KPSYAFS*
mutated AA sequence	N/A
speed	0.14 s
	Report bugs / help us to improve MutationTaster!

All positions are in basepairs (bp) if not explicitly stated differently.
AA/aa: amino acid; CDS: coding sequence; mu: mutated; NMD: nonsense-mediated mRNA decay; nt: nucleotide;
wt: wildtype; TGP: 1000 Genomes Project

03/11/2016

no title [MT]

[documentation](#)

Prediction polymorphism

Model: *without_aae*, prob:
 3.81656130189341e-14
 (classification due to TGP/ExAC,
real probability is shown
 anyway) ([explain](#))
[hyperlink](#)

Summary

- homozygous in TGP or ExAC

analysed issue analysis result

name of alteration no title
 alteration (phys. location) chr15:91346923C>A [show variant in all transcripts](#) [IGV](#)
 HGNC symbol [BLM](#)
 Ensembl transcript ID [ENST00000355112](#)
 Genbank transcript ID [NM_000057](#)
 UniProt peptide N/A
 alteration type single base exchange
 alteration region CDS
 DNA changes c.3531C>A
 cDNA.3649C>A
 g.86366C>A
 AA changes no AA changes
 position(s) of altered AA N/A
 if AA alteration in CDS
 frameshift no
 known variant Reference ID: [rs2227934](#)

database	homozygous (A/A)	heterozygous	allele carriers
1000G	44	579	623
ExAC	1569	15687	17256

regulatory features H3K36me3, Histone, Histone 3 Lysine 36 Tri-Methylation

phyloP / phastCons PhyloP PhastCons

(flanking) 5.193 1

0.105 0.573

(flanking) 0.981 0.549

[explain score\(s\)](#) and/or inspect your position(s) in [in UCSC Genome Browser](#)

splice sites effect gDNA score wt detection sequence exon-intron border
 position
 Donor 86363 wt: 0.4547 / mu: 0.4581 wt: TAAA|gccc

03/11/2016		no title [MT]	
	marginally increased	(marginal change - not scored)	AATAAGCCCAAAC mu : AATAAGCACAAAC
distance from splice site	28		
Kozak consensus sequence altered?	N/A		
conservation protein level for non-synonymous changes	N/A		
protein features	N/A		
length of protein	normal		
AA sequence altered	no		
position of stopcodon in wt / mu CDS	4254 / 4254		
position (AA) of stopcodon in wt / mu AA sequence	1418 / 1418		
position of stopcodon in wt / mu cDNA	4372 / 4372		
poly(A) signal	N/A		
conservation nucleotide level for all changes - no scoring up to now	N/A		
position of start ATG in wt / mu cDNA	119 / 119		
chromosome	15		
strand	1		
last intron/exon boundary	4195		
theoretical NMD boundary in CDS	4026		
length of CDS	4254		
coding sequence (CDS) position	3531		
cDNA position (for ins/del: last normal base / first normal base)	3649		
gDNA position (for ins/del: last normal base / first normal base)	86366		
chromosomal position (for ins/del: last normal base / first normal base)	91346923		
original gDNA sequence snippet	GTGATGCTCGGAAATAAGCCCAAACGTACTAAATGGCAA		
altered gDNA sequence snippet	GTGATGCTCGGAAATAAGCACAAACGTACTAAATGGCAA		
original cDNA sequence snippet	GTGATGCTCGGAAATAAGCCCAAACGTACTAAATGGCAA		
altered cDNA sequence snippet	GTGATGCTCGGAAATAAGCACAAACGTACTAAATGGCAA		

03/11/2016

no title [MT]

wildtype AA
sequence

```

MAAVPQNNLQ EQLERHSART LNNKLSLSKP KFSGFTFKKK TSSDNNVSVT NVSVAKTPVL
RNKDVNVTED FSFSEPLPNT TNQQRVKDFF KNAPAGQETQ RGGSKSLLPD FLQTPKEVVC
TTQNTPTVKK SRDTALKKLE FSSSPDSLST INDWDDMDDF DTSETSKSFV TPPQSHFVRV
STAQSKKKGK RNFFKAQLYT TNTVKTDLPP PSSESEQIDL TEEQKDDSEW LSSDVICIDD
GPIAEVHINE DAQESDSLKT HLEDERDNSE KKKNLEEAEL HSTKVPCIE FDDDDYDTDF
VPPSPREEIIS ASSSSSKCLS TLKDLTSDR KEDVLSTSKD LLSKPEKMSM QELNPETSTD
CDARQISLQQ QLIHVMHIC KLIDTIPDDK LKLLDCGNEL LQQRNIRRLK LTEVDFNKSD
ASLLGSLWRY RPDSDLGPME GDSCPTGNSM KELNFSHLPs NSVSPGDCLL TTTLGKTGFS
ATRKNLFRP LFNTHLQKSF VSSNWAETPR LGKKNESSYF PGNVLTSTAV KDQNKHTASI
NDLERETQPS YDIDNFDIDD FDDDDWEDI MHNLAASKSS TAAYQPIKEG RPIKSVSERL
SSAKTDCLPV SSTAQNIIFS ESIQNYTDKS AQNLASRNK HERFQSLSFP HTKEMMKIFH
KKFGLHNFRT NQLEAINAAL LGEDCFILMP TGGGKSLCYQ LPACVSPGVT VVISPLRSI
VDQVQKLTSL DIPATYLTGD KTDSEATNIY LQLSKKDPPII KLLYVTPEKI CASNRLISTL
ENLYERKLLA RFVIDEAHCV SQWGHDFRQD YKRMNMLRQK FPSVPVMALT ATANPRVQKD
ILTQLKILRP QVFSMSFNRH NLKYYVLPKK PKKVAFDCL E WIRKHHYPYDS GIIYCLSRRE
CDTMADTLQR DGLAALAYHA GLSDSARDEV QQKWINQDGC QVICATIAFG MGIDKPDVRF
VIHASLPKSV EGYQESGRA GRDGEISHCL LFYTYHDVTR LKRLIMMEKD GNHHTRETHF
NNLYSMVHYC ENITECRRIQ LLAYFGENG NPDFCKKHPD VSCDNCCKTK DYKTRDVTDD
VKSIVRFVQE HSSSQGMRNI KHVGPSGRFT MNMLVDIFLG SKSAKIQSGI FGKGSAYSRH
NAERLFKKLI LDKILDEDLY INANDQAIAY VMLGNKAQTV LNGNLKVDFM ETENSSSVKK
QKALVAKVSQ REEMVKKCLG ELTEVCKSLG KVFVGHYFNI FNTVTLKKLA ESLSSDPEVL
LQIDGVTEDEK LEKYGAEVIS VLQKYEWS PAEDSSPGIS LSSSRGPGRS AAELDEEIP
VSSHVFASKT RNERKRKKMP ASQSKRRKT ASSGSKAKGG SATCRKISSK TKSSSIIGSS
SASHTSQATS GANSKLGIMA PPKPINRPFL KPSYAFS*

```

mutated AA
sequence

```

MAAVPQNNLQ EQLERHSART LNNKLSLSKP KFSGFTFKKK TSSDNNVSVT NVSVAKTPVL
RNKDVNVTED FSFSEPLPNT TNQQRVKDFF KNAPAGQETQ RGGSKSLLPD FLQTPKEVVC
TTQNTPTVKK SRDTALKKLE FSSSPDSLST INDWDDMDDF DTSETSKSFV TPPQSHFVRV
STAQSKKKGK RNFFKAQLYT TNTVKTDLPP PSSESEQIDL TEEQKDDSEW LSSDVICIDD
GPIAEVHINE DAQESDSLKT HLEDERDNSE KKKNLEEAEL HSTKVPCIE FDDDDYDTDF
VPPSPREEIIS ASSSSSKCLS TLKDLTSDR KEDVLSTSKD LLSKPEKMSM QELNPETSTD
CDARQISLQQ QLIHVMHIC KLIDTIPDDK LKLLDCGNEL LQQRNIRRLK LTEVDFNKSD
ASLLGSLWRY RPDSDLGPME GDSCPTGNSM KELNFSHLPs NSVSPGDCLL TTTLGKTGFS
ATRKNLFRP LFNTHLQKSF VSSNWAETPR LGKKNESSYF PGNVLTSTAV KDQNKHTASI
NDLERETQPS YDIDNFDIDD FDDDDWEDI MHNLAASKSS TAAYQPIKEG RPIKSVSERL
SSAKTDCLPV SSTAQNIIFS ESIQNYTDKS AQNLASRNK HERFQSLSFP HTKEMMKIFH
KKFGLHNFRT NQLEAINAAL LGEDCFILMP TGGGKSLCYQ LPACVSPGVT VVISPLRSI
VDQVQKLTSL DIPATYLTGD KTDSEATNIY LQLSKKDPPII KLLYVTPEKI CASNRLISTL
ENLYERKLLA RFVIDEAHCV SQWGHDFRQD YKRMNMLRQK FPSVPVMALT ATANPRVQKD
ILTQLKILRP QVFSMSFNRH NLKYYVLPKK PKKVAFDCL E WIRKHHYPYDS GIIYCLSRRE
CDTMADTLQR DGLAALAYHA GLSDSARDEV QQKWINQDGC QVICATIAFG MGIDKPDVRF
VIHASLPKSV EGYQESGRA GRDGEISHCL LFYTYHDVTR LKRLIMMEKD GNHHTRETHF
NNLYSMVHYC ENITECRRIQ LLAYFGENG NPDFCKKHPD VSCDNCCKTK DYKTRDVTDD
VKSIVRFVQE HSSSQGMRNI KHVGPSGRFT MNMLVDIFLG SKSAKIQSGI FGKGSAYSRH
NAERLFKKLI LDKILDEDLY INANDQAIAY VMLGNKAQTV LNGNLKVDFM ETENSSSVKK
QKALVAKVSQ REEMVKKCLG ELTEVCKSLG KVFVGHYFNI FNTVTLKKLA ESLSSDPEVL
LQIDGVTEDEK LEKYGAEVIS VLQKYEWS PAEDSSPGIS LSSSRGPGRS AAELDEEIP
VSSHVFASKT RNERKRKKMP ASQSKRRKT ASSGSKAKGG SATCRKISSK TKSSSIIGSS
SASHTSQATS GANSKLGIMA PPKPINRPFL KPSYAFS*

```

speed

0.15 s

[Report bugs / help us to improve MutationTaster!](#)

All positions are in basepairs (bp) if not explicitly stated differently.

AA/aa: amino acid; CDS: coding sequence; mu: mutated; NMD: nonsense-mediated mRNA decay; nt: nucleotide; wt: wildtype; TGP: 1000 Genomes Project

03/11/2016

no title [MT]

mutation
t@sting

[documentation](#)

Prediction polymorphism

Model: *without_aae*,
prob: 0.999999878656444
(classification due to
TGP/ExAC, **real**
probability is shown
anyway) ([explain](#))
[hyperlink](#)

Summary

- homozygous in TGP or ExAC
- protein features (might be) affected
- splice site changes

analysed issue analysis result

name of alteration no title
 alteration (phys. location) chr15:91354505C>T [show variant in all transcripts](#) [IGV](#)
 HGNC symbol [BLM](#)
 Ensembl transcript ID [ENST00000355112](#)
 Genbank transcript ID [NM_000057](#)
 UniProt peptide [P54132](#)
 alteration type single base exchange
 alteration region CDS
 DNA changes c.3945C>T
 cDNA.4063C>T
 g.93948C>T
 AA changes no AA changes
 position(s) of altered AA N/A
 if AA alteration in CDS
 frameshift no
 known variant Reference ID: [rs1063147](#)

database	homozygous (T/T)	heterozygous	allele carriers
1000G	49	567	616
ExAC	1589	15716	17305

regulatory features H3K36me3, Histone, Histone 3 Lysine 36 Tri-Methylation

phyloP / phastCons PhyloP PhastCons

(flanking) -1.961 0

-1.059 0

(flanking) 0.58 0.004

[explain score\(s\)](#) and/or inspect your position(s) in [in UCSC Genome Browser](#)

03/11/2016	no title [MT]				
splice sites	effect	gDNA position	score	wt detection sequence	exon-intron border
	Donor gained	93941	0.35	mu: CGCTGAGGAGCTTGA	CTGA ggag
distance from splice site	71				
Kozak consensus sequence altered?	N/A				
conservation protein level for non-synonymous changes	N/A				
protein features	start (aa)	end (aa)	feature	details	
	1334	1349	MOTIF	Nuclear localization signal (Potential).	might get lost (downstream of altered splice site)
	1411	1411	MOD_RES	N6-acetyllysine.	might get lost (downstream of altered splice site)
length of protein	normal				
AA sequence altered	no				
position of stopcodon in wt / mu CDS	4254 / 4254				
position (AA) of stopcodon in wt / mu AA sequence	1418 / 1418				
position of stopcodon in wt / mu cDNA	4372 / 4372				
poly(A) signal	N/A				
conservation nucleotide level for all changes - no scoring up to now	N/A				
position of start ATG in wt / mu cDNA	119 / 119				
chromosome	15				
strand	1				
last intron/exon boundary	4195				
theoretical NMD boundary in CDS	4026				
length of CDS	4254				
coding sequence (CDS) position	3945				
cDNA position (for ins/del: last normal base / first normal base)	4063				
gDNA position (for ins/del: last normal base / first normal base)	93948				
chromosomal position (for ins/del: last normal base / first normal base)	91354505				
original gDNA sequence snippet	AGAAGTGCCGCTGAGGAGCTCGACGAGGAAATACCCGTATC				
altered gDNA	AGAAGTGCCGCTGAGGAGCTTGACGAGGAAATACCCGTATC				

03/11/2016

no title [MT]

sequence snippet

original cDNA AGAAGTGCCGCTGAGGAGCTGACGAGGAAATACCCGTATC

sequence snippet

altered cDNA AGAAGTGCCGCTGAGGAGCTGACGAGGAAATACCCGTATC

sequence snippet

wildtype AA

sequence

```

MAAVPQNNLQ EQLERHSART LNNKLSLSKP KFSGFTFKKK TSSDNNVSVT NVSVAKTPVL
RNKDVNVTED FSFSEPLPNT TNQQRVKDFF KNAPAGQETQ RGGSKSLLPD FLQTPKEVVC
TTQNTPTVKK SRDTALKKLE FSSSPDSLST INDWDDMDDF DTSETSKSFV TPPQSHFVRV
STAQSKKKGK RNFFKAQLYT TNTVKTDLPP PSSESEQIDL TEEQKDDSEW LSSDVICIDD
GPPIAEVHINE DAQESDSLKT HLEDERDNSE KKKNLEEAEL HSTKVPCIE FDDDDYDITDF
VPPSPPEEII ASSSSSKCLS TLKDLTSDR KEDVLSTSKD LLSKPEKMSM QELNPETSTD
CDARQISLQQ QLIHVMHEIC KLIDTIPDDK LKLLDCGNEL LQQRNIRRLK LTEVDFNKSD
ASLLGSLWRY RPDSDLGPME GDSCPTGNSM KELNFSHLPS NSVSPGDCLL TTTLGKTGFS
ATRKNLFERP LFNTHLQKSF VSSNWAETPR LGKKNESSYF PGNVLTSTAV KDQNKHTASI
NDLERETQPS YDIDNFDIDD FDDDDWEDI MHNLAASKSS TAAYQPIKEG RPIKSVSERL
SSAKTDCLPV SSTAQNINFS ESIQNYTDKS AQNLASRNK HERFQSLSPF HTKEMMKIFH
KKFGLHNFRT NQLEAINAAL LGEDCFILMP TGGGKSLCYQ LPACVSPGVT VVISPLRSLI
VDQVQKLTSL DIPATYLTGD KTDSEATNIY LQLSKKDPII KLLYVTPEKI CASNRLISTL
ENLYERKLLA RFVIDEAHCV SQWGHDFRQD YKRMNMLRQK FPSVPVMALT ATANPRVQKD
ILTQLKILRP QVFSMSFNRH NLKYYVLPKK PKKVAFDCL E WIRKHHYPYDS GIIYCLSRRE
CDTMADTLQR DGLAALAYHA GLSDSARDEV QQKWINQDGC QVICATIAFG MGIDKPDVRF
VIHASLPKSV EGYQESGRA GRDGEISHCL LFYTYHDVTR LKRLIMMEKD GNHHTRETHF
NNLYSMVHYC ENITECRRIQ LLAYFGENG NPFCCKKHPD VSCDNCCKTK DYKTRDVTDD
VKSIVRFVQE HSSSQGMRNI KHVGPSGRFT MNMLVDIFLG SKSAKIQSGI FGKGSAYSRH
NAERLFKKLI LDKILDEDLY INANDQAIAY VMLGNKAQTV LNGNLKVDFM ETENSSSVKK
QKALVAKVSQ REEMVKKCLG ELTEVCKSLG KVFGVHYFNI FNTVTLKKLA ESLSSDPEVL
LQIDGVTEDE LEKYGAEVIS VLQKYEWS PAEDSSPGIS LSSSRGPGRS AAELDEEIP
VSSHVFASKT RNERKRKKMP ASQSKRRKT ASSGSKAKGG SATCRKISSK TKSSSIIGSS
SASHTSQATS GANSKLGIMA PPKPINRPFL KPSYAFS*

```

mutated AA

sequence

```

MAAVPQNNLQ EQLERHSART LNNKLSLSKP KFSGFTFKKK TSSDNNVSVT NVSVAKTPVL
RNKDVNVTED FSFSEPLPNT TNQQRVKDFF KNAPAGQETQ RGGSKSLLPD FLQTPKEVVC
TTQNTPTVKK SRDTALKKLE FSSSPDSLST INDWDDMDDF DTSETSKSFV TPPQSHFVRV
STAQSKKKGK RNFFKAQLYT TNTVKTDLPP PSSESEQIDL TEEQKDDSEW LSSDVICIDD
GPPIAEVHINE DAQESDSLKT HLEDERDNSE KKKNLEEAEL HSTKVPCIE FDDDDYDITDF
VPPSPPEEII ASSSSSKCLS TLKDLTSDR KEDVLSTSKD LLSKPEKMSM QELNPETSTD
CDARQISLQQ QLIHVMHEIC KLIDTIPDDK LKLLDCGNEL LQQRNIRRLK LTEVDFNKSD
ASLLGSLWRY RPDSDLGPME GDSCPTGNSM KELNFSHLPS NSVSPGDCLL TTTLGKTGFS
ATRKNLFERP LFNTHLQKSF VSSNWAETPR LGKKNESSYF PGNVLTSTAV KDQNKHTASI
NDLERETQPS YDIDNFDIDD FDDDDWEDI MHNLAASKSS TAAYQPIKEG RPIKSVSERL
SSAKTDCLPV SSTAQNINFS ESIQNYTDKS AQNLASRNK HERFQSLSPF HTKEMMKIFH
KKFGLHNFRT NQLEAINAAL LGEDCFILMP TGGGKSLCYQ LPACVSPGVT VVISPLRSLI
VDQVQKLTSL DIPATYLTGD KTDSEATNIY LQLSKKDPII KLLYVTPEKI CASNRLISTL
ENLYERKLLA RFVIDEAHCV SQWGHDFRQD YKRMNMLRQK FPSVPVMALT ATANPRVQKD
ILTQLKILRP QVFSMSFNRH NLKYYVLPKK PKKVAFDCL E WIRKHHYPYDS GIIYCLSRRE
CDTMADTLQR DGLAALAYHA GLSDSARDEV QQKWINQDGC QVICATIAFG MGIDKPDVRF
VIHASLPKSV EGYQESGRA GRDGEISHCL LFYTYHDVTR LKRLIMMEKD GNHHTRETHF
NNLYSMVHYC ENITECRRIQ LLAYFGENG NPFCCKKHPD VSCDNCCKTK DYKTRDVTDD
VKSIVRFVQE HSSSQGMRNI KHVGPSGRFT MNMLVDIFLG SKSAKIQSGI FGKGSAYSRH
NAERLFKKLI LDKILDEDLY INANDQAIAY VMLGNKAQTV LNGNLKVDFM ETENSSSVKK
QKALVAKVSQ REEMVKKCLG ELTEVCKSLG KVFGVHYFNI FNTVTLKKLA ESLSSDPEVL
LQIDGVTEDE LEKYGAEVIS VLQKYEWS PAEDSSPGIS LSSSRGPGRS AAELDEEIP
VSSHVFASKT RNERKRKKMP ASQSKRRKT ASSGSKAKGG SATCRKISSK TKSSSIIGSS
SASHTSQATS GANSKLGIMA PPKPINRPFL KPSYAFS*

```

speed

0.16 s

[Report bugs / help us to improve MutationTaster!](#)

03/11/2016

no title [MT]

All positions are in basepairs (bp) if not explicitly stated differently.

AA/aa: amino acid; CDS: coding sequence; mu: mutated; NMD: nonsense-mediated mRNA decay; nt: nucleotide; wt: wildtype; TGP: 1000 Genomes Project

03/11/2016

no title [MT]

[documentation](#)

Prediction polymorphism

Model: *without_aae*, prob:
8.24762523640819e-15
(classification due to
TGP/ExAC, [real probability](#)
is shown anyway) ([explain](#))
[hyperlink](#)

Summary

- homozygous in TGP or ExAC
- protein features (might be) affected
- splice site changes

analysed issue analysis result

name of alteration no title
 alteration (phys. location) chr15:91337479G>A [show variant in all transcripts](#) [IGV](#)
 HGNC symbol [BLM](#)
 Ensembl transcript ID [ENST00000355112](#)
 Genbank transcript ID [NM_000057](#)
 UniProt peptide [P54132](#)
 alteration type single base exchange
 alteration region CDS
 DNA changes c.3102G>A
 cDNA.3220G>A
 g.76922G>A
 AA changes no AA changes
 position(s) of altered AA N/A
 if AA alteration in CDS
 frameshift no
 known variant Reference ID: [rs2227933](#)

database	homozygous (A/A)	heterozygous	allele carriers
1000G	60	637	697
ExAC	1682	16160	17842

regulatory features H3K36me3, Histone, Histone 3 Lysine 36 Tri-Methylation

phyloP / phastCons PhyloP PhastCons

(flanking) 1.361 0.72

-0.871 0.444

(flanking) 5.563 1

[explain score\(s\)](#) and/or inspect your position(s) in [in UCSC Genome Browser](#)

splice sites effect gDNA score wt detection sequence exon-intron

<http://www.mutationtaster.org/cgi-bin/MutationTaster/MutationTaster69.cgi>

1/4

03/11/2016

no title [MT]

	Donor increased	position 76920	wt: 0.37 / mu: 0.68	wt : ATATAACGGAATGCA mu : ATATAACAGAATGCA	border ATAA cgga
	Donor increased	76927	wt: 0.42 / mu: 0.50	wt : GGAATGCAGGAGAAT mu : AGAATGCAGGAGAAT	AATG cagg
distance from splice site	83				
Kozak consensus sequence altered?	N/A				
conservation protein level for non-synonymous changes	N/A				
protein features	start (aa)	end (aa)	feature	details	
	1207	1209	STRAND		might get lost (downstream of altered splice site)
	1211	1233	HELIX		might get lost (downstream of altered splice site)
	1212	1292	DOMAIN	HRDC.	might get lost (downstream of altered splice site)
	1237	1240	HELIX		might get lost (downstream of altered splice site)
	1243	1252	HELIX		might get lost (downstream of altered splice site)
	1257	1261	HELIX		might get lost (downstream of altered splice site)
	1263	1265	STRAND		might get lost (downstream of altered splice site)
	1268	1273	HELIX		might get lost (downstream of altered splice site)
	1275	1288	HELIX		might get lost (downstream of altered splice site)
	1295	1295	MOD_RES	Phosphoserine.	might get lost (downstream of altered splice site)
	1296	1296	MOD_RES	Phosphoserine.	might get lost (downstream of altered splice site)
	1303	1303	MOD_RES	Phosphoserine.	might get lost (downstream of altered splice site)
	1310	1310	MOD_RES	Phosphoserine.	might get lost (downstream of altered splice site)
	1334	1349	MOTIF	Nuclear localization signal (Potential).	might get lost (downstream of altered splice site)
	1411	1411	MOD_RES	N6-acetyllysine.	might get lost (downstream of altered splice site)
length of protein	normal				
AA sequence altered	no				
position of stopcodon in wt / mu CDS	4254 / 4254				
position (AA) of stopcodon in wt / mu AA sequence	1418 / 1418				
position of	4372 / 4372				

03/11/2016

no title [MT]

stopcodon in wt /
mu cDNA

poly(A) signal N/A

conservation N/A

nucleotide level for all
changes - no scoring up
to now

position of start 119 / 119

ATG in wt / mu
cDNA

chromosome 15

strand 1

last intron/exon
boundary 4195

theoretical NMD
boundary in CDS 4026

length of CDS 4254

coding sequence
(CDS) position 3102

cDNA position 3220
(for ins/del: last
normal base / first
normal base)

gDNA position 76922
(for ins/del: last
normal base / first
normal base)

chromosomal
position 91337479
(for ins/del: last
normal base / first
normal base)

original gDNA
sequence snippet CATTACTGTGAAATATAACGAATGCAGGAGAATACAGCT

altered gDNA
sequence snippet CATTACTGTGAAATATAACGAATGCAGGAGAATACAGCT

original cDNA
sequence snippet CATTACTGTGAAATATAACGAATGCAGGAGAATACAGCT

altered cDNA
sequence snippet CATTACTGTGAAATATAACGAATGCAGGAGAATACAGCT

wildtype AA
sequence

MAAVPQNNLQ	EQLERHSART	LNNKLSLSKP	KFSGFTFKKK	TSSDNVSVT	NVSVAKTPVL
RNKDVNVTE	FSFSEPLPNT	TNQQRVKDFF	KNAPAGQETQ	RGGSKSLPD	FLQTPKEVVC
TTQNTPTVKK	SRDTALKKLE	FSSSPDSLST	INDWDDMDDF	DTSETSKSFV	TPPQSHFVRV
STAQKSKGKG	RNFFKAQLYT	TNTVKTDLPP	PSSSESEQIDL	TEEQKDDSEW	LSSDVICIDD
GPIAEVHINE	DAQESDSLKT	HLEDERDNSE	KKKNLEEAEL	HSTKVPCIE	FDDDDYDTDF
VPPSPPEEIS	ASSSSSKCLS	TLKDLTSDR	KEDVLSTSKD	LLSKPEKMSM	QELNPETSTD
CDARQISLQQ	QLIHVMEHIC	KLIDTIPDDK	LKLLDCGNEL	LQQRNIRRLK	LTEVDNFKSD
ASLLGSLWRY	RPDSLDPGME	GDSCPTGNSM	KELNFSHLPS	NSVSPGDCLL	TTTLGKTGFS
ATRKNLFERP	LFNTHLQKSF	VSSNWAETPR	LGKKNESSYF	PGNVLSTAV	KDQNKHTASI
NDLERETQPS	YDIDNFDIDD	FDDDDWEDI	MHNLAASKSS	TAAYQPIKEG	RPIKSVSERL
SSAKTDCLPV	SSTAQNINFS	ESIQNYTDKS	AQNLASRNK	HERFQSLSFP	HTKEMMKIFH
KKFGLHNFRT	NQLEAINAAL	LGEDCFILMP	TGGGKSLCYQ	LPACVSPGVT	VVISPLRSLI
VDQVQKLTS	LIPATYLTGD	KTDSEATNIY	LQLSKKDPPI	KLLVVTPEKI	CASNRLISTL
ENLYERKLLA	RFVIDEAHCV	SQWGHDFRQD	YKRMNMLRQK	FPSVPVMALT	ATANPRVQKD
ILTQLKILRP	QVFSMSFNH	NLKYYVLPKK	PKKVAFDCL	EIRKHPYDS	GIIYCLSRRE
CDTMADTLQR	DGLAALAYHA	GLSDSARDEV	QQKWINQDGC	QVICATIAFG	MGIDKPDVRF
VIHASLPKSV	EGYYQESGRA	GRDGEISHCL	LFYTYHDVTR	LKRLIMMEKD	GNHHTRETHF
NNLYSMVHYC	ENITECRRIQ	LLAYFGENG	NPDFCKKHPD	VSCDNCKCTK	DYKTRDVTDD
VKSIVRFVQE	HSSSQGMRNI	KHVGPSGRFT	MNMLVDIFLG	SKSAKIQSGI	FGKGSAYSRR
NAERLFKKLI	LDKILDEDLY	INANDQAIAY	VMLGNKAQTV	LNGNLKVDPM	ETENSSSVKK
QKALVAKVSQ	REEMVKKCLG	ELTEVCKSLG	KVFGVHYFNI	FNTVTLLKLA	ESLSSDPEVL

03/11/2016

no title [MT]

mutated AA
sequence

LQIDGVTEDEK LEKYGAEEVIS VLQKYSEWTS PAEDSSPGIS LSSSRGPGRS AAEEELDEEIP
 VSSHYFASKT RNERKRKKMP ASQSKRRKT ASSGSKAKGG SATCRKISSK TKSSSIIGSS
 SASHTSQATS GANSKLGIMA PPKPINRPFL KPSYAFS*
 MAAVPQNNLQ EQLERHSART LNNKLSLSKP KFSGFTFKKK TSSDNNVSVT NVSVAKTPVL
 RNKDVNVTED FSFSEPLPNT TNQQRVKDFF KNAPAGQETQ RGGSKSLLPD FLQTPKEVVC
 TTQNTPTVKK SRDTALKKLE FSSSPDSLST INDWDDMDDF DTSETSKSFV TPPQSHFVRV
 STAQKSKKKG RNFFKAQLYT TNTVKTDLPP PSSESEQIDL TEEQKDDSEW LSSDVICIDD
 GPIAEVHINE DAQESDSLKT HLEDERDNSE KKKNLEEAEI HSTKVPCIE FDDDDYDTDF
 VPPSPPEEII ASSSSSKCLS TLKDLDTSDR KEDVLSTSKD LLSKPEKMSM QELNPETSTD
 CDARQISLQQ QLIHVMFHIC KLIDTIPDDK LKLLDCGNEL LQQRNIRRLK LTEVDFNKSD
 ASLLGSLWRY RPDSDLGPME GDSCPTGNSM KELNFSHLPS NSVSPGDCLL TTTLGKTGFS
 ATRKNLFERP LFNTHLQKSF VSSNWAETPR LGKKNESSYF PGNVLTSTAV KDQNKHTASI
 NDLERETQPS YDIDNFDIDD FDDDDWEDI MHNLAASKSS TAAYQPIKEG RPIKSVSERL
 SSAKTDCLPV SSTAQNINFS ESIQNYTDKS AQNLASRNK HERFQSLFHP HTKEMMKIFH
 KKFGLHNFRT NQLEAINAAL LGEDCFILMP TGGGKSLCYQ LPACVSPGVT VVISPLRSLI
 VDQVQKLTSL DIPATYLTGD KTDSEATNIY LQLSKKDPII KLLYVTPKEI CASNRLISTL
 ENLYERKLLA RFVIDEAHCV SQWGHDFRQD YKRMNMLRQK FPSVPVMALT ATANPRVQKD
 ILTQLKILRP QVFSMSFNH NLKYYVLPKK PKKVAFDCLC WIRKHPYDS GIIYCLSRRE
 CDTMADTLQR DGLAALAYHA GLSDSARDEV QQKWINQDGC QVICATIAFG MGIDKPDVRF
 VIHASLPKSV EGYQESGRA GRDGEISHCL LFYTYHDVTR LKRLIMMEKD GNHHTRETHF
 NNLYSMVHYC ENITECRRIQ LLAYFGENGF NPDFCKKHPD VSCDNCKCTK DYKTRDVTDD
 VKSIVRFVQE HSSSQGMRNI KHVGPSGRFT MNMLVDIFLG SKSAKIQSGI FGKGSAYSRH
 NAERLFKKLI LDKILDEDLY INANDQAIAY VMLGNKAQTV LNGNLKVDFM ETENSSSVKK
 QKALVAKVSQ REEMVKKCLG ELTEVCKSLG KVFVHYFNI FNTVTLKKLA ESLSSDPEVL
 LQIDGVTEDEK LEKYGAEEVIS VLQKYSEWTS PAEDSSPGIS LSSSRGPGRS AAEEELDEEIP
 VSSHYFASKT RNERKRKKMP ASQSKRRKT ASSGSKAKGG SATCRKISSK TKSSSIIGSS
 SASHTSQATS GANSKLGIMA PPKPINRPFL KPSYAFS*

speed

0.17 s

[Report bugs / help us to improve MutationTaster!](#)

All positions are in basepairs (bp) if not explicitly stated differently.

AA/aa: amino acid; CDS: coding sequence; mu: mutated; NMD: nonsense-mediated mRNA decay; nt: nucleotide; wt: wildtype; TGP: 1000 Genomes Project

ANEXO D

Análises de predição de proteína para as mutações encontradas na Paciente M1

05/11/2016

no title [MT]

[documentation](#)

mutation t@sting

Prediction disease causing

Model: *simple_aae*, prob:
0.999999999963319 (classification due to
ClinVar, **real probability** is shown anyway)
([explain](#))

[hyperlink](#)

Summary

- amino acid sequence changed
- known disease mutation at this position (HGMD CM950139)
- known disease mutation: rs367543029 (pathogenic)
- protein features (might be) affected
- splice site changes

analysed issue analysis result

name of alteration no title
 alteration (phys. location) chr15:91337541G>C [show variant in all transcripts](#) [IGV](#)
 HGNC symbol [BLM](#)
 Ensembl transcript ID [ENST00000355112](#)
 Genbank transcript ID [NM_000057](#)
 UniProt peptide [P54132](#)
 alteration type single base exchange
 alteration region CDS
 DNA changes c.3164G>C
 cDNA.3282G>C
 g.76984G>C
 AA changes C1055S Score: 112 [explain score\(s\)](#)
 position(s) of altered AA 1055
 if AA alteration in CDS
 frameshift no
 known variant Reference ID: [rs367543029](#)

database	homozygous (C/C)	heterozygous	allele carriers
1000G	-	-	-
ExAC	0	2	2

known disease mutation: rs367543029 (pathogenic for Bloom_syndrome|Hereditary_cancer-predisposing_syndrome) [dbSNP](#) [NCBI variation viewer](#)

known disease mutation at this position, [please check HGMD for details](#) (HGMD ID CM950139)

regulatory features H3K36me3, Histone, Histone 3 Lysine 36 Tri-Methylation

phyloP / phastCons PhyloP PhastCons

(flanking) 4.704 1

5.695 1

(flanking) 1.4 1

[explain score\(s\)](#) and/or inspect your position(s) in [in UCSC Genome Browser](#)

splice sites effect gDNA position score wt detection sequence exon-intron border

Acc marginally increased 76987 wt: 0.7275 / mu: 0.7616 (marginal change - not scored) wt: ATTTAATCCTGATTTTGTAAAGAAACACCCAGATGTTTCTT mu: ATTTAATCCTGATTTTCTAAGAAACACCCAGATGTTTCTT gtaa | GAAA

05/11/2016

no title [MT]

Donor increased	76982	wt: 0.21 / mu: 0.33	wt: GATTTTGTAGAAA mu: GATTTTCTAGAAA	TTTT tgta
Acc gained	76988	0.55	mu: TTTAATCCTGATTTTCTAAGAAACACCCAGATGTTTCTTG	taag AAAC
Acc gained	76989	0.68	mu: TTAATCCTGATTTTCTAAGAAACACCCAGATGTTTCTTG	aaga AACA
Acc gained	76986	0.52	mu: GATTTAATCCTGATTTTCTAAGAAACACCCAGATGTTTCT	tcta AGAA
distance from splice site	47			
Kozak consensus sequence altered?	N/A			
conservation protein level for non-synonymous changes	species	match	gene	aa alignment
	Human			1055 Y F G E N G F N P D F C K K H P D V S C D N C C
	mutated	not conserved		1055 Y F G E N G F N P D F S K K H P D V S C D N C
	Ptroglydotes	all identical	ENSPTRG0000007460	1055 Y F G E N G F N P D F C K K H P D V S C D N C
	Mmulatta	all identical	ENSMUG00000019526	1054 Y F G E N G F N P D F C K K H P D V S C D N C
	Fcatus	all identical	ENSFCAG0000006087	1054 Y F G E N G F N P D F C K K Y P D V S C D N C
	Mmusculus	all identical	ENSMUSG00000030528	1066 Y F G E K G F N P D F C K K Y P D V S C D N C
	Ggallus	all identical	ENSGALG0000008256	1023 Y F G E T D F N P N F C K D H P E V I C D N C
	Trubripes	no homologue		
	Drerio	all identical	ENSDARG00000077089	1047 Y F G E H T F N T S F C K E H P E V I C D N C
	Dmelanogaster	all identical	FBgn0002906	1123 S E Q C L E N R E T A C D N C
	Celegans	all identical	T04A11.6	632 H F G E - V Y D E Q S C - R N S K T P C D I C
	Xtropicalis	all identical	ENSXETG0000007847	1063 Y F G E N N F N P N F C K E N T R V S C D N C
protein features	start (aa)	end (aa)	feature details	
	1207	1209	STRAND	might get lost (downstream of altered splice site)
	1211	1233	HELIX	might get lost (downstream of altered splice site)
	1212	1292	DOMAIN HRDC.	might get lost (downstream of altered splice site)
	1237	1240	HELIX	might get lost (downstream of altered splice site)
	1243	1252	HELIX	might get lost (downstream of altered splice site)
	1257	1261	HELIX	might get lost (downstream of altered splice site)
	1263	1265	STRAND	might get lost (downstream of altered splice site)
	1268	1273	HELIX	might get lost (downstream of altered splice site)
	1275	1288	HELIX	might get lost (downstream of altered splice site)
	1295	1295	MOD_RES Phosphoserine.	might get lost (downstream of altered splice site)
	1296	1296	MOD_RES Phosphoserine.	might get lost (downstream of altered splice site)
	1303	1303	MOD_RES Phosphoserine.	might get lost (downstream of altered splice site)
	1310	1310	MOD_RES Phosphoserine.	might get lost (downstream of altered splice site)
	1334	1349	MOTIF Nuclear localization signal	might get lost (downstream of altered splice site)

05/11/2016			no title [MT]	
	1411	1411	(Potential). MOD_RES N6-acetyllysine.	site) might get lost (downstream of altered splice site)
length of protein	normal			
AA sequence altered	yes			
position of stopcodon in wt / mu CDS	4254 / 4254			
position (AA) of stopcodon in wt / mu AA sequence	1418 / 1418			
position of stopcodon in wt / mu cDNA	4372 / 4372			
poly(A) signal	N/A			
conservation	N/A			
nucleotide level for all changes - no scoring up to now				
position of start ATG in wt / mu cDNA	119 / 119			
chromosome	15			
strand	1			
last intron/exon boundary	4195			
theoretical NMD boundary in CDS	4026			
length of CDS	4254			
coding sequence (CDS) position	3164			
cDNA position (for ins/del: last normal base / first normal base)	3282			
gDNA position (for ins/del: last normal base / first normal base)	76984			
chromosomal position (for ins/del: last normal base / first normal base)	91337541			
original gDNA sequence snippet	TGGATTTAATCCTGATTTTTGTAAGAAACACCCAGATGTTT			
altered gDNA sequence snippet	TGGATTTAATCCTGATTTTTCTAAGAAACACCCAGATGTTT			
original cDNA sequence snippet	TGGATTTAATCCTGATTTTTGTAAGAAACACCCAGATGTTT			
altered cDNA sequence snippet	TGGATTTAATCCTGATTTTTCTAAGAAACACCCAGATGTTT			
wildtype AA sequence	MAAVPQNNLQ EQLERHSART LNNKLSLSKP KFSGFTFKKK TSSDNNVSVT NVSVAKTPVL RNKDVNVTED FSFSEPLPNT TNQQRVKDFF KNPAGQETQ RGGSKSLLPD FLQTPKEVVC TTQNTPTVKK SRDTALKKLE FSSSPDSLST INDWDDMDDF DTSETSKSFV TPPQSHFVRV STAQSKKGGK RNFFKAQLYT TNTVKTDLPP PSSESEQIDL TEEQKDDSEW LSSDVICIDD GPJAEVHINE DAQESDSLKT HLEDERDNSE KKKNLEEAEEL HSTEKVPCIE FDDDDYDTDF VPPSPPEEIISSSSSSSKCLS TLKDLDTSDR KEDVLSTSKD LLSKPEKMSM QELNPETSTD CDARQISLQQ QLIHVMEHIC KLIDTIPDDK LKLDCGNEL LQQRNIRRLK LTEVDFNKSD ASLLGSLWRY RPDSDLGPMG GDSCPTGNM KELNFSHLP NSVSPGDCLL TTTLGKTGFS ATRKNLFERP LFNTHLQKSF VSSNWAETPR LGKKNESSYF PGNVLTSTAV KDQNKHTASI NDLERETQPS YDIDNFDIDD FDDDDDWEDI MHNLAASKSS TAAYQPIKEG RPIKSVSERL SSAKTDCLPV SSTAQNINFS ESIGNYTDKS AQNLASRNK HERFQSLSFP HTKEMMKIFH KKFGLHNFRT NQLEAINAAL LGEDCFILMP TGGGKSLCYQ LPACVSPGVT VVISPLRSIL VDQVQKLTSL DIPATYLTGD KTDSEATNIY LQLSKKDPII KLLYVTPEKI CASNRLISTL ENLYERKLLA RFVIDEAHCV SQWGHDFRQD YKRMNMLRQK FPSVPVMAIT ATANPRVQKD ILTQLKILRP QVFSMSFNRH NLKYYVLPKK PKKVAFDCLE WIRKHHYPYDS GIITYCLSRRE CDTMADTLQR DGLAALAYHA GLSDSARDEV QQKWINDGCG QVICATIAFG MGIDKPDVRF VIHASLPKSV EGYYESGRA GRDGEISHCL LFYTYHDVTR LKRLIMMEKD GNHHTRETHF NNLYSMVHYC ENITECRRIQ LLAYFGENG FNPFCCKHPD VSCDNCKCTK DYKTRDVTDD			

05/11/2016	no title [MT]									
mutated AA sequence	VKSIVRFVQE	HSSSQGMNRI	KHVGPSGRFT	MNMLVDIFLG	SKSAKIQSGI	FGKGSAYSRRH				
	NAERLFKKLI	LDKILDEDLY	INANDQAIAY	VMLGNKAQTV	LNGNLKVDFM	ETENSSSVKK				
	QKALVAKVSQ	REEMVKKCLG	ELTEVCKSLG	KVFGVHYFNI	FNTVTLKKLA	ESLSSDPEVL				
	LQIDGVTEDEK	LEKYGAEVIS	VLQKYSEWTS	PAEDSSPGIS	LSSSRGPGRS	AAEELDEEIP				
	VSSHVFASKT	RNERKRKKMP	ASQRSKRRKT	ASSGSKAKGG	SATCRKISSK	TKSSSIIGSS				
	SASHTSQATS	GANSKLGIMA	PPKPINRPFL	KPSYAFS*						
	MAAVPQNNLQ	EQLERHSART	LNNKLSLSKP	KFSGFTFKKK	TSSDNNVSVT	NVSVAKTPVL				
	RNKDVNVTED	FSFSEPLPNT	TNQQRVKDFF	KNAPAGQETQ	RGGSKSLLPD	FLQTPKEVVC				
	TTQNTPTVKK	SRDTALKKLE	FSSSPDSLST	INDWDDMDDF	DTSETSKSFV	TPPQSHFVRV				
	STAQKSKKGG	RNFFKAQLYT	TNTVKTDLPP	PSSSESEQIDL	TEEQKDDSEW	LSSDVICIDD				
	GPIAEVHINE	DAQESDSLKT	HLEDERDNSE	KKKNLEEAEI	HSTEKVPICIE	FDDDDYDTDF				
	VPPSPREEIIS	ASSSSSKCLS	TLKDLDTSDR	KEDVLSTSKD	LSSKPEKMSM	QELNPETSTD				
	CDARQISLQQ	QLIHVMEHIC	KLIDTIPDDK	LKLLDCGNEL	LQQRNIRRLK	LTEVDFNKSD				
	ASLLGSLWRY	RPDSLDGPME	GDSCPTGNSM	KELNFSHLP	NSVSPGDCLL	TTTLGKTGFS				
	ATRKNLFERP	LFNTHLQKFS	VSSNWAETPR	LGKKNESYF	PGNVLSTAV	KDQNKHTASI				
	NDLERETQPS	YDIDNFDIDD	FDDDDDWEDI	MHNLAASKSS	TAAYQPIKEG	RPIKSVSERL				
	SSAKTDCLPV	SSTAQNINFS	ESIQNYTDKS	AQNLASRNK	HERFQSLSFP	HTKEMMKIFH				
	KKFGLHNFRT	NQLEAINAAL	LGEDCFILMP	TGGGKSLCYQ	LPACVSPGVT	VVISPLRSLI				
	VDQVQKLTSL	DIPATYLTGD	KTDSEATNIY	LQLSKKDPPI	KLLVYVTEKI	CASNRLISTL				
	ENLYERKLLA	RFVIDEAHCV	SQWGHDFRQD	YKRMNMLRQK	FPSVPVMALT	ATANPRVQKD				
	ILTLKILRP	QVFSMSFNRH	NLKYVVLPPK	PKKVAFDCL	WIRKHHYPDS	GIIYCLSRRE				
	CDTMADTLQR	DGLAALAYHA	GLSDSARDEV	QKWINQDGC	QVICATIAFG	MGIDKPDVRF				
	VIHASLPKSV	EGYYQESGRA	GRDGEISHCL	LFYTYHDVTR	LKRLIMMEKD	GNHHTRETHF				
	NNLYSMVHYC	ENITECRRIQ	LLAYFGENG	NPDFSKKHPD	VSCDNCKCTK	DYKTRDVTDD				
	VKSIVRFVQE	HSSSQGMNRI	KHVGPSGRFT	MNMLVDIFLG	SKSAKIQSGI	FGKGSAYSRRH				
	NAERLFKKLI	LDKILDEDLY	INANDQAIAY	VMLGNKAQTV	LNGNLKVDFM	ETENSSSVKK				
	QKALVAKVSQ	REEMVKKCLG	ELTEVCKSLG	KVFGVHYFNI	FNTVTLKKLA	ESLSSDPEVL				
	LQIDGVTEDEK	LEKYGAEVIS	VLQKYSEWTS	PAEDSSPGIS	LSSSRGPGRS	AAEELDEEIP				
	VSSHVFASKT	RNERKRKKMP	ASQRSKRRKT	ASSGSKAKGG	SATCRKISSK	TKSSSIIGSS				
	SASHTSQATS	GANSKLGIMA	PPKPINRPFL	KPSYAFS*						
speed	0.18 s									
	Report bugs / help us to improve MutationTaster!									

All positions are in basepairs (bp) if not explicitly stated differently.

AA/aa: amino acid; CDS: coding sequence; mu: mutated; NMD: nonsense-mediated mRNA decay; nt: nucleotide; wt: wildtype;

TGP: 1000 Genomes Project

05/11/2016	no title [MT]				
	Trubripes	identical			
		no			
		homologue			
	Drerio	all	ENSDARG00000077089	1201	K V S K R E E M V K K C L G
		identical			
	Dmelanogaster	all	FBgn0002906	1271 - K E A K A A V G S V S D G A T S S T A D G Q S G M R E	
		identical			
	Celegans	not	T04A11.6	787 V S E A Q A L K E R H M V K	H G D V F
		conserved			
	Xtropicalis	all	ENSXETG00000007847	1214 L R K Q K A S V V A N T S Q	R E E M V K K C Q
		identical			
protein features	start (aa)	end (aa)	feature	details	
	1207	1209	STRAND		lost
	1211	1233	HELIX		might get lost (downstream of altered splice site)
	1212	1292	DOMAIN	HRDC.	might get lost (downstream of altered splice site)
	1237	1240	HELIX		might get lost (downstream of altered splice site)
	1243	1252	HELIX		might get lost (downstream of altered splice site)
	1257	1261	HELIX		might get lost (downstream of altered splice site)
	1263	1265	STRAND		might get lost (downstream of altered splice site)
	1268	1273	HELIX		might get lost (downstream of altered splice site)
	1275	1288	HELIX		might get lost (downstream of altered splice site)
	1295	1295	MOD_RES Phosphoserine.		might get lost (downstream of altered splice site)
	1296	1296	MOD_RES Phosphoserine.		might get lost (downstream of altered splice site)
	1303	1303	MOD_RES Phosphoserine.		might get lost (downstream of altered splice site)
	1310	1310	MOD_RES Phosphoserine.		might get lost (downstream of altered splice site)
	1334	1349	MOTIF Nuclear localization signal (Potential).		might get lost (downstream of altered splice site)
	1411	1411	MOD_RES N6-acetylysine.		might get lost (downstream of altered splice site)
length of protein	normal				
AA sequence altered	yes				
position of stopcodon in wt / mu CDS	4254 / 4254				
position (AA) of stopcodon in wt / mu AA sequence	1418 / 1418				
position of stopcodon in wt / mu cDNA	4372 / 4372				
poly(A) signal	N/A				
conservation	N/A				
nucleotide level for all changes - no scoring up to now					
position of start ATG in wt / mu cDNA	119 / 119				
chromosome	15				
strand	1				
last intron/exon boundary	4195				
theoretical NMD boundary in CDS	4026				
length of CDS	4254				
coding sequence (CDS) position	3625				
cDNA position (for ins/del: last normal base / first normal base)	3743				
gDNA position (for ins/del: last normal base / first normal base)	86906				
chromosomal position (for ins/del: last normal base / first normal base)	91347463				
original gDNA sequence snippet	AAGCGTTAGTAGCAAAAAGTGTCTCAGAGGGAAGAGATGGTT				
altered gDNA sequence snippet	AAGCGTTAGTAGCAAAAAGTGA CT CAGAGGGAAGAGATGGTT				
original cDNA sequence snippet	AAGCGTTAGTAGCAAAAAGTGTCTCAGAGGGAAGAGATGGTT				
altered cDNA sequence snippet	AAGCGTTAGTAGCAAAAAGTGA CT CAGAGGGAAGAGATGGTT				

05/11/2016

no title [MT]

mutation t@sting

[documentation](#)

Prediction disease causing

 Model: *simple_aae*, prob: 0.9999988851353 ([explain](#))

Summary

- amino acid sequence changed
- protein features (might be) affected
- splice site changes

[hyperlink](#)

analysed issue analysis result

name of alteration no title
 alteration (phys. location) chr15:91347463T>A [show variant in all transcripts](#) [IGV](#)
 HGNC symbol [BLM](#)
 Ensembl transcript ID [ENST00000355112](#)
 Genbank transcript ID [NM_000057](#)
 UniProt peptide [P54132](#)
 alteration type single base exchange
 alteration region CDS
 DNA changes c.3625T>A
 cDNA.3743T>A
 g.86906T>A
 AA changes S1209T Score: 58 [explain score\(s\)](#)
 position(s) of altered AA 1209
 if AA alteration in CDS
 frameshift no
 known variant Reference ID: [rs1801256](#)

database	homozygous (A/A)	heterozygous	allele carriers
1000G	-	-	-
ExAC	0	37	37

regulatory features H3K36me3, Histone, Histone 3 Lysine 36 Tri-Methylation

phyloP / phastCons PhyoP PhastCons

(flanking) -1.815 0.068

4.311 0.996

(flanking) 3.637 1

[explain score\(s\)](#) and/or inspect your position(s) in [in UCSC Genome Browser](#)

splice sites effect gDNA position score wt detection sequence exon-intron border
 Donor gained 86907 0.37 mu: AGTGACTCAGAGGGA TGAC|tcag
 Donor gained 86902 0.95 mu: GCAAAAGTGACTCAG AAAA|gtga

distance from splice site 67

Kozak consensus sequence altered? N/A

conservation protein level for non-synonymous changes	species	match	gene	aa	alignment	
	Human			1209	V K K Q K A L V A K V S Q	R E E M V K K C L G E
	mutated	all conserved		1209	Q K A L V A K V T Q	R E E M V K K C L G
	Ptroglydotes	all identical	ENSPTRG0000007460	1209	Q K A L V A K V S Q	R E E M V K K C L G
	Mmulatta	all identical	ENSMUG00000019526	1208	Q K A L V A K V S Q	R E E M V K K C L G
	Fcatus	all identical	ENSFCAG0000006087	1208	V K K Q K A L V A K M S Q	R
	Mmusculus	all identical	ENSMUSG00000030528	1217	I K K Q K A L V A K V S Q	R E E V V K K C L G
	Ggallus	all	ENSALG00000008256	1177	I K K Q R A S V T K M S Q	R E

<http://www.mutationtaster.org/cgi-bin/MutationTaster/MutationTaster69.cgi>

1/3

05/11/2016	no title [MT]				
	Trubripes	identical			
		no			
		homologue			
	Drerio	all	ENSDARG00000077089	1201	K V S K R E E M V K K C L G
		identical			
	Dmelanogaster	all	FBgn0002906	1271 - K E A K A A V G S V S	DGATSSSTADGQSG M R E
		identical			
	Celegans	not	T04A11.6	787 V S E A Q A L K E R H M V K	H G D V F
		conserved			
	Xtropicalis	all	ENSXETG00000007847	1214 L R K Q K A S V V A N T S Q	R E E M V K K C Q
		identical			
protein features	start (aa)	end (aa)	feature	details	
	1207	1209	STRAND		lost
	1211	1233	HELIX		might get lost (downstream of altered splice site)
	1212	1292	DOMAIN	HRDC.	might get lost (downstream of altered splice site)
	1237	1240	HELIX		might get lost (downstream of altered splice site)
	1243	1252	HELIX		might get lost (downstream of altered splice site)
	1257	1261	HELIX		might get lost (downstream of altered splice site)
	1263	1265	STRAND		might get lost (downstream of altered splice site)
	1268	1273	HELIX		might get lost (downstream of altered splice site)
	1275	1288	HELIX		might get lost (downstream of altered splice site)
	1295	1295	MOD_RES	Phosphoserine.	might get lost (downstream of altered splice site)
	1296	1296	MOD_RES	Phosphoserine.	might get lost (downstream of altered splice site)
	1303	1303	MOD_RES	Phosphoserine.	might get lost (downstream of altered splice site)
	1310	1310	MOD_RES	Phosphoserine.	might get lost (downstream of altered splice site)
	1334	1349	MOTIF	Nuclear localization signal (Potential).	might get lost (downstream of altered splice site)
	1411	1411	MOD_RES	N6-acetylysine.	might get lost (downstream of altered splice site)
length of protein	normal				
AA sequence altered	yes				
position of stopcodon in wt / mu CDS	4254 / 4254				
position (AA) of stopcodon in wt / mu AA sequence	1418 / 1418				
position of stopcodon in wt / mu cDNA	4372 / 4372				
poly(A) signal	N/A				
conservation	N/A				
nucleotide level for all changes - no scoring up to now					
position of start ATG in wt / mu cDNA	119 / 119				
chromosome	15				
strand	1				
last intron/exon boundary	4195				
theoretical NMD boundary in CDS	4026				
length of CDS	4254				
coding sequence (CDS) position	3625				
cDNA position (for ins/del: last normal base / first normal base)	3743				
gDNA position (for ins/del: last normal base / first normal base)	86906				
chromosomal position (for ins/del: last normal base / first normal base)	91347463				
original gDNA sequence snippet	AAGCGTTAGTAGCAAAAGTGTCTCAGAGGGAAGAGATGGTT				
altered gDNA sequence snippet	AAGCGTTAGTAGCAAAAGTGA CT CAGAGGGAAGAGATGGTT				
original cDNA sequence snippet	AAGCGTTAGTAGCAAAAGTGTCTCAGAGGGAAGAGATGGTT				
altered cDNA sequence snippet	AAGCGTTAGTAGCAAAAGTGA CT CAGAGGGAAGAGATGGTT				

05/11/2016	no title [MT]
sequence snippet	
wildtype AA sequence	MAAVPQNNLQ EQLERHSART LNNKLSLSKP KFSGFTFKKK TSSDNNVSVT NVSVAKTPVL RNKDVNVTED FSFSEPLPNT TNQQRVKDFF KNAPAGQETQ RGGSKSLLPD FLQTPKEVVC TTQNTPTVKK SRDTALKKLE FSSSPDSLST INDDWDDMDDF DTSETSKSFV TPPQSHFVRV STAQSKKGG RNFFKAQLYT TNTVKTDLPP PSSESEQIDL TEEQKDDSEW LSSDVCIDDD GPFAEVHINE DAQESDSLKT HLEDERDNSE KKKNLEEAEEL HSTEKVPDIE FDDDDVDTDF VPPSPPEEIISS ASSSSSKCLS TLKDLTSDR KEDVLSTSKD LLSKPEKMSM QELNPETSTD CDARQSLSLQQ QLIHVMHEIC KLIDTIPDDK LKLLDCGNEL LQQRNIRRLK LTEVDFNKS ASLLGSLWRY RPDSDLGPMG GDSCPTGNSM KELNFSHLPS NSVSPGDCLL TTTLGKTGFS ATRNKLFERP LFNTLQKSF VSSNWAETPR LGKKNESYF PGNVLTSTAV KDQNKHTASI NDLERETQPS YDIDNFDIDD FDDDDWEDI MHNLAASKSS TAAYQPIKEG RPIKSVSERL SSAKTDCLPV SSTAQNINFS ESIGNYTDKS AQNLASRNK HERFQSLFPP HTKEMMKIFH KKFGLHNFRT NQLEAINAAL LGEDCFILMP TGGGKSLCYQ LPACVSPGVV VVISPLRLSI VDQVQKLTSL DIPATYLTGD KTDSEATNIY LQLSKKDPII KLLYVTPEKI CASNRLISTL ENLYERKLLA RFVIDEAHCV SQWGHDFRQD YKRMNMLRQK FPSVPVMALT ATANPRVQKD ILTQLKILRP QVFSMSFNRH NLKYVYLPKK PKKVAFCLE WIRKHHYPDS GIIVCLSRRE CDTMADTLQR DGLAALAYHA GLSDSARDEV QKQWINDQGC QVICATIAFG MGIDKPDVRF VIHASLPKSV EGYQESGRA GRDGEISHCL LFYTYHDVTR LKRLIMMEKD GNHHTRETHF NNLYSMVHYC ENITECRRIQ LLAYFGENG NPDCCKKHPD VSCDNCCKTK DYKTRDVTDD VKSIVRFVQE HSSSQGMRI KHVGPGRFT MNMLVDIFLG SKSAKIQSGI FGKGSAYSRIH NAERLFKKLI LDKILDEDLY INANDQAIAY VMLGNKAQTV LNGNLKVDPM ETENSSSVKK QKALVAKVQ REEMVKCLG ELTEVCKSLG KVFVGHYFNI FNTVTLKLLA ESLSDPEVL LQIDGVTEDK LEKYGAEVIS VLQKYEWS PAEDSSPGIS LSSSRGPGRS AAEELDEEIP VSSHVFASKT RNERKRRKMP ASQSKRRKT ASSGSKAKGG SATCRKISSK TKSSSIIGSS SASHTSQATS GANSKLGIMA PPKPINRPFL KPSYAFS*
mutated AA sequence	MAAVPQNNLQ EQLERHSART LNNKLSLSKP KFSGFTFKKK TSSDNNVSVT NVSVAKTPVL RNKDVNVTED FSFSEPLPNT TNQQRVKDFF KNAPAGQETQ RGGSKSLLPD FLQTPKEVVC TTQNTPTVKK SRDTALKKLE FSSSPDSLST INDDWDDMDDF DTSETSKSFV TPPQSHFVRV STAQSKKGG RNFFKAQLYT TNTVKTDLPP PSSESEQIDL TEEQKDDSEW LSSDVCIDDD GPFAEVHINE DAQESDSLKT HLEDERDNSE KKKNLEEAEEL HSTEKVPDIE FDDDDVDTDF VPPSPPEEIISS ASSSSSKCLS TLKDLTSDR KEDVLSTSKD LLSKPEKMSM QELNPETSTD CDARQSLSLQQ QLIHVMHEIC KLIDTIPDDK LKLLDCGNEL LQQRNIRRLK LTEVDFNKS ASLLGSLWRY RPDSDLGPMG GDSCPTGNSM KELNFSHLPS NSVSPGDCLL TTTLGKTGFS ATRNKLFERP LFNTLQKSF VSSNWAETPR LGKKNESYF PGNVLTSTAV KDQNKHTASI NDLERETQPS YDIDNFDIDD FDDDDWEDI MHNLAASKSS TAAYQPIKEG RPIKSVSERL SSAKTDCLPV SSTAQNINFS ESIGNYTDKS AQNLASRNK HERFQSLFPP HTKEMMKIFH KKFGLHNFRT NQLEAINAAL LGEDCFILMP TGGGKSLCYQ LPACVSPGVV VVISPLRLSI VDQVQKLTSL DIPATYLTGD KTDSEATNIY LQLSKKDPII KLLYVTPEKI CASNRLISTL ENLYERKLLA RFVIDEAHCV SQWGHDFRQD YKRMNMLRQK FPSVPVMALT ATANPRVQKD ILTQLKILRP QVFSMSFNRH NLKYVYLPKK PKKVAFCLE WIRKHHYPDS GIIVCLSRRE CDTMADTLQR DGLAALAYHA GLSDSARDEV QKQWINDQGC QVICATIAFG MGIDKPDVRF VIHASLPKSV EGYQESGRA GRDGEISHCL LFYTYHDVTR LKRLIMMEKD GNHHTRETHF NNLYSMVHYC ENITECRRIQ LLAYFGENG NPDCCKKHPD VSCDNCCKTK DYKTRDVTDD VKSIVRFVQE HSSSQGMRI KHVGPGRFT MNMLVDIFLG SKSAKIQSGI FGKGSAYSRIH NAERLFKKLI LDKILDEDLY INANDQAIAY VMLGNKAQTV LNGNLKVDPM ETENSSSVKK QKALVAKVQ REEMVKCLG ELTEVCKSLG KVFVGHYFNI FNTVTLKLLA ESLSDPEVL LQIDGVTEDK LEKYGAEVIS VLQKYEWS PAEDSSPGIS LSSSRGPGRS AAEELDEEIP VSSHVFASKT RNERKRRKMP ASQSKRRKT ASSGSKAKGG SATCRKISSK TKSSSIIGSS SASHTSQATS GANSKLGIMA PPKPINRPFL KPSYAFS*
speed	0.18 s
Report bugs / help us to improve MutationTaster!	

All positions are in basepairs (bp) if not explicitly stated differently.

AA/aa: amino acid; CDS: coding sequence; mu: mutated; NMD: nonsense-mediated mRNA decay; nt: nucleotide; wt: wildtype; TGP: 1000 Genomes Project

9 REFERÊNCIAS

American College of Medical Genetics (ACMG): www.acmg.net

Amor-Guéret, M. Bloom's syndrome. Orphanet Encyclopedia. 2004. <http://www.orpha.net/data/patho/gb/uk-Bloomsyndrome.pdf>. Bhisitkul RB, Rizen M. Bloom syndrome: multiple retinopathies in a chromosome breakage disorder. *Br J Ophthalmol*. 2004;88:354-357.

Auer PL, Doerge RW. Statistical Design and Analysis of RNA Sequencing Data. *Genetics*. 2010;185:405-416.

Bachrati CZ, Hickson ID. RecQ helicases: suppressors of tumorigenesis and premature aging. *Biochem J*. 2003 Sep;374(Pt 3):577-606.

Banco de dados de variantes genômicas clínicas (ClinVar). Disponível em: <https://www.ncbi.nlm.nih.gov/clinvar/>

Banco de dados de variantes genômicas HGMD. Disponível em: <http://www.hgmd.cf.ac.uk/ac/index.php>

Banco de dados de variantes genômicas LOVD. Disponível em: <https://oi.gene.le.ac.uk/home.php>

Bhisitkul RB, Rizen M. Bloom syndrome: multiple retinopathies in a chromosome breakage disorder. 2004.

Bloom D. Congenital telangiectatic erythema resembling lupus erythematosus in dwarfs; probably a syndrome entity. *Am J Dis Child*. 1954 Dec;88(6):754-8.

Cao QF, Yamamoto J, Isobe T, Tateno S, Murase Y, Chen Y, et al. Characterization of the Human Transcription Elongation Factor Rtf1: Evidence for Nonoverlapping Functions of Rtf1 and the Paf1 Complex. *Mol Cell Biol*. 352015. p. 3459-70.

Chaganti RS, Schonberg S, German J. A manyfold increase in sister chromatid exchanges in Bloom's syndrome lymphocytes. *Proc Natl Acad Sci U S A*. 1974 Nov;71(11):4508-12.

Cunniff C, Bassetti JA, Ellis NA. Bloom's Syndrome: Clinical Spectrum, Molecular Pathogenesis, and Cancer Predisposition. *Molecular Syndromology*. 2016;8(1):4-23.

Deckbar D, Jeggo PA, Lobrich M. Understanding the limitations of radiation-induced cell cycle checkpoints. *Crit Rev Biochem Mol Biol*. 2011 Aug;46(4):271-83

Ellis NA, Groden J, Ye TZ, Straughen J, Lennon DJ, Ciocchi S, et al. The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell*. 1995 Nov;83(4):655-66.

Ellis NA, Ciocchi S, Proytcheva M, Lennon D, Groden J, German J. The Ashkenazic Jewish Bloom syndrome mutation blmAsh is present in non-Jewish Americans of Spanish ancestry. *Am J Hum Genet*. 1998 Dec;63(6):1685-93.

Elmore S. Apoptosis: A review of programmed cell death. *Toxicol Pathol*. 2007; 35(4):495-516.

Ensembl's Variant Effect Predictor, release 83. Disponível em <http://www.ncbi.nlm.nih.gov/pubmed/20562413>

Fang Z, Cui X. Design and validation issues in RNA-seq experiments. *Briefing in Bioinformatics*. 2011;12(3):280-287.

FastQC. Disponível em: www.bioinformatics.bbsrc.ac.uk/projects/.

Filiatrault MJ, Stodghill PV, Myers CR, Bronstein PA, Butcher BG, Lam H, et al. Genome-wide identification of transcriptional start sites in the plant pathogen *Pseudomonas syringae* pv. tomato str. DC3000. *PLoS One*. 2011;6(12):e29335.

Gene Ontology consortium. Disponível em www.geneontology.org.

German J. Bloom's syndrome. I. Genetical and clinical observations in the first twenty-seven patients. *Am J Hum Genet*. 1969 Mar;21(2):196-227.

German, J. and Passarge, E. (1989), Bloom's Syndrome. XII. Report from the Registry for 1987. *Clinical Genetics*, 35: 57–69.

German J, Takebe H. Bloom's syndrome. XIV. The disorder in Japan. *Clin Genet*. 1989 Feb;35(2):93-110.

German J, Passarge E. Bloom's syndrome. XII. Report from the Registry for 1987. *Clin Genet*. 1989 Jan;35(1):57-69.

German J, Crippa LP, Bloom D. Bloom's syndrome. III. Analysis of the chromosome aberration characteristic of this disorder. *Chromosoma*. 1974;48(4):361-6.

German J, Bloom D, Passarge E, Fried K, Goodman RM, Katzenellenbogen I, et al. Bloom's syndrome. VI. The disorder in Israel and an estimation of the gene frequency in the Ashkenazim. *Am J Hum Genet*. 1977 Nov;29(6):553-62.

German J, Schonberg S, Louie E, Chaganti RS. Bloom's syndrome. IV. Sister-chromatid exchanges in lymphocytes. *Am J Hum Genet*. 1977 May;29(3):248-55.

German J. Cytological evidence for crossing-over in vitro in human lymphoid cells. *Science*. 1964 Apr;144(3616):298-301.

German J. Bloom syndrome: a mendelian prototype of somatic mutational disease. *Medicine (Baltimore)*. 1993 Nov;72(6):393-406.

German J, Roe AM, Leppert MF, Ellis NA. Bloom syndrome: an analysis of consanguineous families assigns the locus mutated to chromosome band 15q26.1. *Proc Natl Acad Sci U S A*. 1994 Jul;91(14):6669-73.

German J. Bloom's syndrome. XX. The first 100 cancers. *Cancer Genet Cytogenet*. 1997 Jan;93(1):100-6.

German J, Sanz MM, Ciocchi S, Ye TZ, Ellis NA. Syndrome-causing mutations of the BLM gene in persons in the Bloom's Syndrome Registry. *Hum Mutat*. 2007 Aug;28(8):743-53.

Goya R, Sun MG, Morin RD, Leung G, Ha G, Wiegand KC, et al. SNVMix: predicting single nucleotide variants from next-generation sequencing of tumors. *Bioinformatics*. 2010 Mar;26(6):730-6.

Gramazio et al., 2016

Gretzula JC, Hevia O, Weber PJ. Bloom's syndrome. *J Am Acad Dermatol*. 1987 Sep;17(3):479-88.

Griffiths AJF, Miller JH, Suzuki DT, et al. New York: W. H. Freeman; 2000.

Helleday T. Pathways for mitotic homologous recombination in mammalian cells. *Mutat Res*. 2003 Nov 27;532(1-2):103-15.

Hickson ID, Davies SL, Li J-L, Levitt NC, Mohaghegh P, North PS, et al. Role of the Bloom's syndrome helicase in maintenance of genome stability. 2001. 29.

Hoeijmakers JH. Genome maintenance mechanisms for preventing cancer. *Nature*. 2001 May 17;411(6835):366-74.

Hoeijmarkers, 2007

HTSeq-count v.05.4.p2. Disponível em:<http://www-huber.embl.de/users/anders/HTSeq/doc/index.html>

http://www.med.cornell.edu/bsr/data_from_registry/

<http://www.genome.jp/kegg/>

Huang WD, Sherman BT, Lempicki RA. System and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*. 2009; (4):44-57.

Hurley D, Araki H, Tamada Y, Dummore B, Sanders D, et al. Gene network inference and visualization tools for biologists: application to new human transcriptome datasets. *Nucleic Acids Research*. 2011;1-22.

Ju YS, Kim JI, Kim S, et al. Extensive genomic and transcriptional diversity indentified throug massively parallel DNA and RNA sequencing of eighteen Korean individuals. *Nature Genetics*. 43(8):745-754.

Kaneko H, Fukao T, Kasahara K, Yamada T, Kondo N. Augmented cell death with Bloom syndrome helicase deficiency. *Mol Med Rep*. 2011 Jul-Aug;4(4):607-9.

Kang TF, Lau CY. Getting the most out of RNA-seq data analysis. *PeerJ PrePri nts*. [Internet}. 2015. Disponível em; <https://dx.doi.org/10.7287/peerj.preprints.1198v1>

Killen, MW. Loss of Bloom syndrome protein causes destabilization of genomic architecture and is complemented by ectopic expression of Escherichia coli RECG in human cells. (2011). University of Kentucky Doctoral Dissertations. Paper 201.[Internet]. Disponível em:http://uknowledge.uky.edu/gradschool_diss/201

Kuhn EM, Therman E. Cytogenetics of Bloom 's syndrome. *Cancer Genet Cytogenet*.1986;22:1-18.

Kvan VM, Liu P, Si Y. A comparison of statistical methods for detecting differentially expressed genes from RNA-seq data. *Am J Bot*. 2012 Feb;99(2):248-56.

Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*. 2016 Jan;44(D1):D862-8.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009 Aug 15;25(16):2078-9.

Li CI, Su PS, Guo Y, SHyr Y. Sample size calculation for differential expression analysis of RNA-seq data under Poisson distribution. *Int J Comput Biol Drugs Des*. 2013;6(4):1- 19.

Li L, Eng C, Desnick RJ, German J, Ellis NA. Carrier frequency of the Bloom syndrome blmAsh mutation in the Ashkenazi Jewish population. *Mol Genet Metab*. 1998 Aug;64(4):286-90.

Liu P, Erez A, Nagamani SCS, Dhar SU, Kolodziejska KE et al. Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. *Cell*. 2011;(146)889-903.

Liu Y, West SC. More complexity to the Bloom's syndrome complex. *Genes Dev*. 2008;22:2737-42.

Ma B, Corry J, Rischin D, Leong T, Peters LJ. Combined modality treatment for locally advanced squamous-cell carcinoma of the oropharynx in a woman with Bloom's syndrome: A case report and review of the literature. *Annals of oncology*. 2001;12:1015-1017.

Martin SA, Hewish M, Lord CJ, Ashworth A. Genomic instability and the selection of treatments for cancer. *J Pathol*. 2010 Jan;220(2):281-9.

McGettigan PA. Transcriptomics in the RNA-seq era. *Curr Opin Chem Biol*. 2013 Feb;17(1):4-11.

Minoche AE, Dohm JC, Himmelbauer H. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. *Genome Biol*. 2011 Nov 08;12(11):R112.

Mladenov E, Iliakis G. Induction and repair of DNA double strand breaks: the increasing spectrum of npn-homologous end joining pathways. *Mutat Res*. 2011;711(1-2):61-72.

Moreira MB, Quaió CRDC, Zandon -Teixeira AC, Novo-Filho GM, Zanardo EA, Kulikowski LD, Kim CA. Discrepant outcomes in two Brazilian patients with Bloom syndrome and Wilms tumor: two case reports. *Journal of Medical Case Reports*. 2013;7:284.

Mulcahy MT, French M. Pregnancy in Bloom's syndrome. *Clin Genet*. 1981 Mar;19(3):156-8.

Mutation Taster. Disponível em: www.mutationtaster.org

NNSplice. Disponível em: http://fruitfly.org/seq_tools/splice.html

Oshlack A, Robinson MD, YOUNG MD. From RNA-seq reads to differential expression results. *Genome Biol*. 2010;11:220.

Paffl MW. Relative Quantification. *Real-time PCR*. 2001;63-82.

Payne M, Hickson ID. Genomic instability and cancer: lessons from analysis of Bloom's syndrome. *Biochem Soc Trans*. 2009 Jun;37(Pt 3):553-9.

Perry P & Wolff S. New Giemsa method for the differential staining of sister chromatids. *Nature* 251:156-8. 1974.

Piskol R, Ramaswami G, Li JB. Reliable identification of genomic variants from RNA-seq data. *Am J Hum Genet*. 2013 Oct;93(4):641-51.

Protein expression of BLM gene and its apoptosis sensitivity in hematopoietic tumor cell strains | SpringerLink. 2016.

Ray JH, German J. Bloom's syndrome and EM9 cells in BrdU-containing medium exhibit similarly elevated frequencies of sister chromatid exchange but dissimilar amounts of cellular proliferation and chromosome disruption. *Chromosoma*. 1984;90(5):383-8.

Resende Ana Cristina Brandão de, Pereira Luciana Baptista, Melo Batielle Moreira Franco de, Santos Helena Hollanda, Aguiar Marcos José Burle de. Você conhece esta síndrome? *An. Bras. Dermatol.* [Internet]. 2007 Aug [cited 2016 Dec 08];82(4): 363-365. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S036505962007000400011&lng=en. <http://dx.doi.org/10.1590/S0365-05962007000400011>.

Sanz MM, German J, Cunniff C. Bloom's Syndrome. 2006 Mar 22 [Updated 2016 Apr 7]. In: Pagon RA, Adam MP, Ardinger HH, et al., editors. *GeneReviews®* [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2016. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK1398/>

Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. *Nat Rev Genet*. 2011 Aug 31;12(10):683-91.

Seki M, Otsuki M, Ishii Y, Tada S, Enomoto T. RecQ family helicases in genome stability: Lessons from gene disruption studies in DT40 cells. [Internet]. Disponível em: <http://dxdoiorg/104161/cc7166462>. 2008 4 Aug 2008.

Seqclean. Disponível em: <https://bitbucket.org/izhbannikov/seqclean>. Shaffer et al., 2016

Schurch NJ, Schofiel P, Gierlinski M, Cole C, Sherstnev A, Singh V, Wrobel N, Gharbi K, Simpson GG, Owen-Hughes T, Blaxter M, Barton GJ. How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA*. 2016 Jun;22(6):839-51.

Shastri VM, Schmidt KH. Cellular defects caused by hypomorphic variants of the Bloom syndrome helicase gene BLM. *Mol Genet Genomic Med*. 4 2016. p. 106-19.

Singh DK, Ahn B, Bohr VA. Roles of RECQ helicases in recombination based DNA repair, genomic stability and aging. *Biogerontology*. 2009 Jun;10(3):235-52.

Single Nucleotide Database – <http://www.ncbi.nlm.nih.gov/SNP/>

Taniguchi N, Mukai M, Nagaoki T, Miyawaki T, Moriya N, Takahashi H, et al. Impaired B-cell differentiation and T-cell regulatory function in four patients with Bloom's syndrome. *Clin Immunol Immunopathol*. 1982 Feb;22(2):247-58.

Transcriptome: Connecting the Genome to Gene Function. 2016.

Tubio JMC, Estivill X. Cancer: When catastrophe strikes a cell. *Nature*. 2011 2011-02-23;470:476-7.

Univec. Disponível em: <http://www.ncbi.nlm.nih.gov/VecScreen/UniVec.html>.

Wilson DM, Thompson LH. Molecular mechanisms of sister-chromatid exchange. *Mutation research*. 2007;616:11-23.

Wu L, Hickson ID. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature*. 2003 Dec 18;426(6968):870-4.

APÊNDICE A

Artigo submetido para publicação

Genome Medicine

Immunological insights in Bloom's syndrome patients: the transcriptome study of BLM gene in two Brazilian by RNA-seq. --Manuscript Draft--

Manuscript Number:	
Full Title:	Immunological insights in Bloom's syndrome patients: the transcriptome study of BLM gene in two Brazilian by RNA-seq.
Article Type:	Research
Section/Category:	Genomics & epigenomics of disease
Funding Information:	
Abstract:	<p>Background: Bloom's Syndrome (BS) is a rare chromosome instability syndrome, with recessive autosomal inheritance. The main clinical manifestations are pre and postnatal growth deficiency, microcephaly, malar hypoplasia, telangiectatic facial erythema and compromised immune system, among others. Patients with BS present increased risk to the development of neoplasias at an early age, which is the main cause of death. Cytogenetic test is used as a diagnostic marker for BS since the patient's cells present increase in spontaneous chromosomal breaks and sister chromatid exchange (SCE). In addition, the literature reveals that most patients also present mutations in the BLM gene, which are related to defects in the DNA repair mechanism; however, it is still not completely understood. Methods: We studied the transcriptome of two (2) patients with Bloom's syndrome and three (3) controls using the RNA-seq methodology (Illumina, Inc., San Diego, CA). Result: Differential expression analysis revealed 216 differentially expressed genes related to immunological pathways such as: negative regulation of the regulation of the viral genome replication, positive regulation of B cells proliferation, gamma-interferon mediated signalization pathway, B cells activation, virus response, adaptive immune response and immune effector process, and absence of difference of DNA repair genes expression. At the same time, we observed the hyperexpression of the BLM gene for both patients contributing for the destabilization of genes involved in immunological pathways, a phenomenon also observed in some tumors. Conclusion: We suggest that the combination of lymphoid proliferation defects and cell signaling defects added to others such as cell loss and altered expression of the BLM gene may contribute directly to the main characteristics observed in Bloom's syndrome, such as growth failure and high risk of cancer. In the future, the study of the transcriptome applied to other BS carriers and other instability syndromes, will allow a more accurate analysis of the relevant gene interactions to the destabilization of the genome.</p>
Corresponding Author:	Marília Moreira Montenegro Universidade de São Paulo Faculdade de Medicina BRAZIL
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Universidade de São Paulo Faculdade de Medicina
Corresponding Author's Secondary Institution:	
First Author:	Marília Moreira Montenegro

APÊNDICE B

Artigo publicado durante o desenvolvimento deste doutorado

Moreira et al. *Journal of Medical Case Reports* 2013, **7**:284
http://www.jmedicalcasereports.com/content/7/1/284

CASE REPORT

Open Access

Discrepant outcomes in two Brazilian patients with Bloom syndrome and Wilms' tumor: two case reports

Marilia Borges Moreira^{1,2,3*}, Caio Robledo DC Quao¹, Aline Cristina Zandoná-Teixeira¹, Gil Monteiro Novo-Filho¹, Evelin Aline Zanardo¹, Leslie Domenici Kulikowski² and Chong Ae Kim¹

Abstract

Introduction: Bloom syndrome is a rare, autosomal recessive, chromosomal instability disorder caused by mutations in the *BLM* gene that increase the risk of developing neoplasias, particularly lymphomas and leukemias, at an early age.

Case presentation: Case 1 was a 10-year-old Brazilian girl, the third child of a non-consanguineous non-Jewish family, who was born at 36 weeks of gestation and presented with severe intrauterine growth restriction. She had Bloom syndrome and was diagnosed with a unilateral Wilms' tumor at the age of 3.5 years. She responded well to oncological treatment and has remained disease-free for the last 17 years. Case 2 was a 2-year-old Brazilian girl born to non-Jewish first-degree cousins. Her gestation was marked by intrauterine growth restriction. She had Bloom syndrome; a unilateral stage II Wilms' tumor was diagnosed at the age of 4 years after the evaluation of a sudden onset abdominal mass. Surgical removal, neoadjuvant chemotherapy and radiotherapy were not sufficient to control the neoplasia. The tumor recurred after 8 months and she died from clinical complications.

Conclusion: Our study reports the importance of rapid diagnostics and clinical follow-up of these patients.

Keywords: Bloom syndrome, Cancer risk, Chromosomal instability, Sister chromatid exchange, Wilms' tumor

Introduction

Bloom syndrome (BS; Online Mendelian Inheritance in Man database, number 210900) [1] is a rare, autosomal recessive, chromosomal instability disorder [2] caused by mutations in the *BLM* gene, which encodes a product necessary for the maintenance of genomic stability [3]. The prominent clinical features associated with BS include severe growth deficiency (pre- and postnatal), sun-sensitive facial erythema, immunological deficiency and a remarkably increased risk of developing neoplasias of various types at a younger age than expected in the general population; the neoplasias are the main cause of death among affected individuals [4,5].

Among the 265 cases of BS reported in the Bloom's Syndrome Registry (which includes 222 families), 122 developed some type of neoplasia during their lives; leukemias, lymphomas and carcinomas were common, but several other cancers have been reported [4,5]. These patients all present a remarkably increased frequency of sister chromatid exchange (SCE). The cancer predisposition in patients with BS can be attributed to excessive chromosomal breakage and homologous recombination events that lead to spontaneous mutations in the somatic cells and defective damage response functionality [6].

Wilms' tumor (WT) is the most common pediatric solid cancer; it has been estimated to occur in 1:10,000 children below the age of 15 years and was once considered a rare event among patients with BS [7]. Nevertheless, six patients with BS who developed this tumor have been described previously in the literature [8-11].

Here we report the cases of two new unrelated Brazilian patients diagnosed with BS who developed WT.

* Correspondence: mariliamoreira@usp.br

¹Genetics Unit, Instituto da Criança do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil

²Department of Pathology, Cytogenomic Laboratory – LIM03 – Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil

Full list of author information is available at the end of the article

© 2013 Moreira et al.; licensee BioMed Central Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Case presentation

Case 1

A 10-year-old Brazilian girl, the third child of a non-consanguineous, non-Jewish family, was born at 36 weeks of gestation and was marked by severe intrauterine growth restriction. After resolving the clinical complications of her extreme low birth weight (BW) of 1500g, she was discharged from the hospital; however, despite presenting normal cognitive development during the first year of her life, a remarkable, refractory failure to thrive was noted. She had recurrent diarrhea and upper and lower respiratory tract infections by the age of 1 year and required prophylactic antibiotics. A WT was diagnosed in her left kidney after evaluating an abdominal mass at the age of 3.5 years. She underwent surgical removal of the mass and neoadjuvant chemotherapy (unknown regimen).

The first genetic evaluation occurred at age 10 when a physical examination revealed features characteristic of BS, including facial features (elongated face, prominent nose, prominent ears, malar hypoplasia, and thin upper vermillion), microcephaly, nose telangiectasias, hypomelanotic macules in her upper limbs and clinodactyly of her bilateral fifth fingers. Her anthropometric measurements were all below the fifth percentile (weight (W) 15.9kg, height (H) 116cm, and occipital-frontal circumference (OFC) 46cm; Centers for Disease Control and Prevention growth charts, National Center for Health Statistics, USA), and she had developed learning disabilities.

A cytogenetic test was performed to examine the frequency of SCE using bromodeoxyuridine (BrdU) in lymphocyte cultures. The results demonstrated an increased frequency of SCE, with an average of 48.5 SCE per cell (Figure 1A). The G-band karyotype was normal (46, XX).

She responded well to oncological treatment and has remained disease-free for the last 17 years. Currently, at the age of 21, all of her anthropometric measurements continue to be below the standards for the normal population (H: 140cm; W: 26.4kg; and OFC: 47cm).

Case 2

The gestation of a 2-year-old Brazilian girl, who was born to non-Jewish first-degree cousins, was marked by intrauterine growth restriction. Vaginal delivery occurred at term with no complications. The neonate had a low BW of 2060g, microcephaly (OFC: 30.5cm) and a normal length (45.5cm); her anthropometric measurements were all below the fifth percentile. A physical examination at the age of 2 years revealed a short stature (78cm), low W (8.1kg), and microcephaly (OFC: 45cm); her anthropometric measurements were all below the fifth percentile. Features characteristic of BS, including facial features (long, narrow face; telangiectatic erythema involving her nose, malar and oral regions; prominent nose; and retrognathia),

café-au-lait macules throughout her body, diffuse hypomelanotic macules, and bilateral fifth finger clinodactyly, were identified. Recurrent upper respiratory tract infections were common, but she had not experienced any serious infection or required prophylactic antibiotic use.

A stage II unilateral WT was diagnosed at the age of 4 years after the evaluation of a sudden onset abdominal mass. Surgical removal, neoadjuvant chemotherapy and radiotherapy were not sufficient to control the neoplasia. The tumor recurred after 8 months, and the patient died from clinical complications.

A cytogenetic test was performed to examine the frequency of SCE using BrdU in lymphocyte cultures. The results demonstrated an increased frequency of SCE, with an average of 49.5 SCE per cell (Figure 1B). The G-band karyotype was normal (46, XX).

Discussion

We described two Brazilian patients affected by BS who each developed a WT at a young age. These two cases, along with the other six described in the literature (Table 1) [8-11], increase the estimated frequency of WT in individuals with BS to at least 3%, a 300-fold increase in the risk relative to the general pediatric population.

BS is caused by mutations in the *BLM* gene, whose product is a 1417-amino acid protein that belongs to the RecQL helicase group and plays important roles in replication, recombination and cellular repair. Upon alteration, *BLM* loses its function and causes genomic instability and an increased rate of spontaneous mutations in somatic cells, primarily by SCE [3]. Cytogenetic analyses of cases 1 and 2 have demonstrated an increase in the frequency of exchange between sister chromatids, which is pathognomonic for BS and confirmed the diagnosis of BS [12]. Unequal crossing-over has probably played a major role in the evolution of various genes and heterochromatin. Hypermutability and the hyperrecombinability of somatic cells leads to an increased chance of homozygous tumor suppressor genes and/or oncogenes being affected and, consequently, increases the rate of the development of a wide variety of neoplasias at an early age [5,8]. In some cases, retinoblastoma and WT are associated with the homozygosity of a chromosome segment resulting from a mitotic crossover. Similarly, the high incidence of cancer in BS may be caused by a mitotic crossover that leads to homozygosity or the amplification of oncogenes.

The literature [13] describes that two events lead to the formation of WT. First, the child inherits one genomic alteration, and a second event occurs in the child that would be caused easily in BS because of the high rate of genomic exchanges. Nevertheless, if recombination in BS occurs as a defect in the repair machinery, maybe there is another mechanism that may contribute to the formation of tumors and cancer.

New technological advances in array-based genomics revealed the contribution of structural alterations in the human genome to several different diseases, including cancer (both solid and hematologic tumors) [14]. In fact, certain copy number variations (CNVs) potentially compromise fundamental processes controlling genomic stability, including deoxyribonucleic acid (DNA) replication and the DNA damage response, and have been reported to be associated with the response to chemotherapy,

which affects the disease prognosis [12]. From a clinical perspective, CNVs might interfere in the DNA damage response and create a permissive environment for the acquisition of additional pathogenic alterations, such as an individual's predisposition to cancer [15].

Furthermore, recently, Stephens and colleagues recently described a novel mechanism of genomic rearrangement in cancer cells, termed 'chromothripsis', that associates specific CNVs and their contribution to cancer development.

Table 1 Patients affected by Bloom syndrome associated with Wilms' tumor

	Ashkenazim	Consanguinity	Age at diagnosis of WT	References
1	No	No	3.5y	Our patient
2	No	Yes	4y	Our patient
3	Yes	No	8y	[8]
4	No	No	5m	[8]
5	No	No	22m	[8]
6	No	No	4y	[9]
7	Not mentioned	Not mentioned	3y	[10]
8	No	No	3.5y	[11]

Abbreviations: m months, WT Wilms' tumor, y years.

Chromothripsis arises through chromosome breakage and inaccurate reassembly, produces highly complex derivative chromosomes, and causes oncogene amplification. A disaster of this magnitude could possibly affect both WT alleles because patients with BS demonstrate susceptibility for breaks in the genome [16].

In our study, both patients developed WT before 4 years of age; however, despite both patients undergoing surgical removal of the mass and chemotherapy, the treatment results were different. The patient described in case 2 (stage II WT) had a fatal outcome after the recurrence of the disease, but the patient described in case 1 has remained disease-free after 17 years of follow-up. All reported cases of WT in patients with BS in the literature and the two cases described here occurred in patients aged 8 years or less.

WT has been previously reported in six cases of BS in the literature [8-11]. The age at which WT was diagnosed ranged from 5 months to 8 years. Two of these cases were diagnosed at early stages (stages I and II) and had apparently been cured by nephrectomy and chemotherapy. Although patients with BS develop cancer at early ages, the age of onset of WT is the same in literature reports of patients without previous medical histories who are under the age of 15 years and in patients with BS [17].

Thus, the frequency of WT in patients with BS is considerable. Clinicians need to be more aware of this fact, particularly because the occurrence of WT in patients with BS was once presumed to be rare. This frequency summed with the frequencies of other solid tumors in patients with BS increases the overall frequency of solid tumors to approximately 12% [5]. Unfortunately, insufficient attention is given to this group of tumors in patients with BS.

Beckwith-Wiedemann syndrome is a prototype genetic disease with an increased risk of the development of early onset solid cancers, mainly WT and hepatoblastoma, with a total estimated lifelong risk for solid tumors of

7.5% [18]. This increased risk has prompted specialists to seriously consider solid tumor surveillance because surveillance has been demonstrated to reduce treatment-related morbidity. Renal ultrasonography is currently the optimal surveillance modality and is accessible, is non-invasive and has minor risks [19]. However, this method may have unfavorable consequences because false positive results may lead to unnecessary investigations and surgical procedures.

Although the evidence does not show benefits for leukemia screening in patients with BS because early treatment does not improve clinical outcomes, surveillance for solid tumors may have significant advantages and improve survival [9,10]. In addition, the frequency of solid tumors in patients with BS is approximately the same as that for patients with Beckwith-Wiedemann syndrome, and a screening program has been demonstrated to be feasible in the latter patients. Thus, we suggest that ultrasound be performed as a regular method of surveillance for the early detection of solid tumors in individuals with BS because the early detection of these tumors may have clinical benefits. The frequency of this type of screening must be individualized, but the performance of an ultrasound examination at least every 6 months is advisable based on the vast experience with Beckwith-Wiedemann syndrome reported in the literature.

Cytogenetic analyses for SCE are the gold standard method for the diagnosis of BS. Cytogenetic analysis is considered a fast and low-cost method to confirm the diagnosis of BS, and this methodology can be implemented in routine and diagnostic laboratories. Regardless, other molecular studies will also be conducted in living patients.

Conclusions

BS is a rare, autosomal recessive, chromosomal instability disorder with remarkably increased risk for developing neoplasias at a young age. The neoplasias represent the main cause of death among affected individuals. Awareness of the high frequency of solid tumors among patients affected by this disorder must be raised because these individuals may benefit from individualized screening for solid tumors, including WT.

Consent

Written informed consent was obtained from the patients' legal guardians for the publication of this manuscript and the accompanying images. Copies of the written consents are available for review by the Editor-in-Chief of this journal.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MBM performed cytogenetic studies and wrote the paper. CRDQ and CAK collected the patients' clinical data. ACZT, GMNF and EAZ performed the cytogenetic studies. LDK and CAK coordinated the study and helped draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors thank the patients and their families, as well as Coordenação de Aperfeiçoamento de Pessoal de Nível Superior for financial support.

Author details

¹Genetics Unit, Instituto da Criança do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil. ²Department of Pathology, Cytogenomic Laboratory – LIM03 – Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil. ³Unidade de Genética do Instituto da Criança, HC – FMUSP, Av. Dr. Enéas Carvalho de Aguiar, 647, CEP: 05403-000, São Paulo, SP, Brazil.

Received: 12 March 2013 Accepted: 21 October 2013
Published: 30 December 2013

References

- Bloom D: Congenital telangiectatic erythema resembling lupus erythematosus in dwarfs. *Amer J Dis Child* 1954, **88**:754–758.
- German J: Bloom's syndrome I. Genetical and clinical observations in the first twenty-seven patients. *Am J Genet* 1969, **21**:196–227.
- Ellis NA, Groden J, Ye TZ, Straughen J, Lennon DJ, Ciochi S, Proytcheva M, German J: The Bloom syndrome gene product is homologous to RecQ helicases. *Cell* 1995, **83**:655–666.
- German J: Bloom syndrome: a mendelian prototype of somatic mutational disease. *Medicine* 1993, **72**:393–406.
- Sanz MM, German J: Bloom's syndrome [2006, updated 2013]. In *GeneReviews*. Edited by Pagon RA, Adam MP, Bird TD, Dolan CR, Fong CT, Stephens K. Seattle (WA): University of Washington, Seattle; 1993–2013. <http://www.ncbi.nlm.nih.gov/books/NBK1398/>.
- Skorski T: BCR/ABL, DNA damage and DNA repair: implications for new treatment concepts. *Leukemia Lymphoma* 2008, **49**:610–614.
- Dome JS, Coppes MJ: Recent advances in Wilms tumor genetics. *Curr Opin Pediatr* 2002, **14**(1):5–11.
- Cairney AEL, Andrews M, Greenberg M, Smith D, Weksberg R: Wilms tumor in three patients with Bloom syndrome. *J Pediatr* 1987, **111**(3):414–416.
- Berger C, Frappaz D, Leroux D, Blez F, Vercherat M, Bouffet E, Jalbert P, Brunat-Mentigny M: Tumeur de Wilms et syndrome de Bloom. *Arch Pediatr* 1996, **3**:802–805.
- German J: Bloom's syndrome. XX. The first 100 cancers. *Cancer Genet Cytogenet* 1997, **93**:100–106.
- Jain D, Hui P, McNamara J, Schwartz D, German J, Reyes-Múgica M: Bloom syndrome in sibs: first reports of hepatocellular carcinoma and Wilms tumor with documented anaplasia and nephrogenic rests. *Pediatr Dev Pathol* 2001, **4**(6):585–589.
- Chaganti RS, Schonberg S, German J: A manyfold increase in sister chromosome exchanges in Bloom syndrome lymphocytes. *Proc Natl Acad Sci* 1974, **71**:4508.
- Coppes MJ, Haber DA, Grundy PE: Genetic events in the development of Wilms' tumor. *N Engl J Med* 1994, **331**(9):586–593.
- Kim KI, Kim TK, Kim IW, Ahn KS, Yoon SS, Shin WG, Oh JM: Copy number variations in normal karyotype acute myeloid leukaemia and their association with treatment response. *Basic Clin Pharmacol Toxicol* 2012, **111**(5):317–324. doi:10.1111/j.1742-7843.2012.00904.x.
- Liu P, Carvalho CMB, Hastings PJ, Lupski JR: Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev* 2012, **22**(3):211–220.
- Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Beare D, Stebbings LA, McLaren S, Lin ML, McBride DJ, Varela I, Nik-Zainal S, Leroy C, Jia M, Menzies A, Butler AP, Teague JW, Quail MA, Burton J, Swerdlow H, Carter NP, Morsberger LA, Iacobuzio-Donahue C, Follows GA, Green AR, Flanagan AM, Stratton MR, et al: Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 2011, **144**:27–40.
- Sadak KT, Ritchey ML, Dome JS: Paediatric genitourinary cancers and late effects of treatment. *Nat Rev Urol* 2013, **10**:15–25.
- Shuman C, Beckwith JB, Smith AC, Weksberg R: Beckwith-Wiedemann syndrome [2000, updated 2010]. In *GeneReviews*. Edited by Pagon RA, Adam MP, Bird TD, Dolan CR, Fong CT, Stephens K. Seattle (WA): University of Washington, Seattle; 1993–2013. <http://www.ncbi.nlm.nih.gov/books/NBK1394/>.
- Tan TY, Amor DJ: Tumor surveillance in Beckwith-Wiedemann syndrome and hemihyperplasia: a critical review of the evidence and suggested guidelines for local practice. *J Paediatr Child Health* 2006, **42**:486–490.

doi:10.1186/1752-1947-7-284

Cite this article as: Moreira et al.: Discrepant outcomes in two Brazilian patients with Bloom syndrome and Wilms' tumor: two case reports. *Journal of Medical Case Reports* 2013 **7**:284.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

APÊNDICE C

Abstracts de artigos publicados durante este doutorado, que não estão diretamente relacionados a temática do presente estudo

Mol Genet Genomics (2014) 289:1037–1043
DOI 10.1007/s00438-014-0876-7

REVIEW

Complex structural rearrangement features suggesting *chromoanagenesis* mechanism in a case of 1p36 deletion syndrome

Évelin Aline Zanardo · Flavia Balbo Piazzon · Roberta Lelis Dutra · Alexandre Torchio Dias ·
Marília Moreira Montenegro · Gil Monteiro Novo-Filho · Thaís Virgínia Moura Machado Costa ·
Amom Mendes Nascimento · Chong Ae Kim · Leslie Domenici Kulikowski

Received: 8 October 2013 / Accepted: 2 June 2014 / Published online: 2 July 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract Genome rearrangements are caused by the erroneous repair of DNA double-strand breaks, leading to several alterations that result in loss or gain of the structural genomic of a dosage-sensitive genes. However, the mechanisms that promote the complexity of rearrangements of congenital or developmental defects in human disease are unclear. The investigation of complex genomic abnormalities could help to elucidate the mechanisms and causes for the formation and facilitate the understanding of congenital or developmental defects in human disease. We here report one case of a patient with atypical clinical features of the 1p36 syndrome and the use of cytogenomic techniques to characterize the genomic alterations. Analysis by multiplex ligation-dependent probe amplification and array revealed a complex rearrangement in the 1p36.3 region with deletions and

duplication interspaced by normal sequences. We also suggest that *chromoanagenesis* could be a possible mechanism involved in the repair and stabilization of this rearrangement.

Keywords Complex structural rearrangement · *Chromoanagenesis* · 1p36 syndrome · Copy number variation · Cytogenomic techniques

Introduction

The human genome can be modulated by several mutational mechanisms that occasionally result in structural variation and thus congenital alterations and disease (Kloosterman et al. 2011; Sobreira et al. 2011). Some genome rearrangements are caused by the erroneous repair of DNA double-strand breaks (DSBs), generating deletions,

Communicated by S. Hohmann.

É. A. Zanardo · F. B. Piazzon · R. L. Dutra · A. T. Dias ·
M. M. Montenegro · G. M. Novo-Filho · T. V. M. M. Costa ·
A. M. Nascimento · L. D. Kulikowski (✉)
Department of Pathology, Laboratório de Citogenômica,
LIM 03, Universidade de São Paulo, Av. Dr. Enéas de Carvalho
Aguar, 155, 2° floor, block 12, Cerqueira César, São Paulo,
SP CEP: 05403-000, Brazil
e-mail: leslielikul@ gmail.com

É. A. Zanardo
e-mail: evelinzanardo@yahoo.com.br

F. B. Piazzon
e-mail: flapiazzon@gmail.com

R. L. Dutra
e-mail: lelis.roberta@gmail.com

A. T. Dias
e-mail: dias.torchio@gmail.com

M. M. Montenegro
e-mail: mariliammontenegro@gmail.com

G. M. Novo-Filho
e-mail: gilvictorio@yahoo.com.br

T. V. M. M. Costa
e-mail: thaisvmmc@hotmail.com

A. M. Nascimento
e-mail: amomnm@hotmail.com

É. A. Zanardo · R. L. Dutra · M. M. Montenegro ·
G. M. Novo-Filho · A. M. Nascimento · C. A. Kim ·
L. D. Kulikowski
Department of Pediatrics, Instituto da Criança, LIM 36,
Universidade de São Paulo, Av. Dr. Enéas de Carvalho Aguiar,
647, 5° floor, Cerqueira César, São Paulo, SP CEP: 05403-000,
Brazil
e-mail: chong.kim@hc.fm.usp.br

CLINICAL REPORT

AMERICAN JOURNAL OF
medical genetics PART A

Rare Genomic Rearrangement in a Boy with Williams–Beuren Syndrome Associated to XYY Syndrome and Intriguing Behavior

Roberta L. Dutra,^{1,2*} Flavia B. Piazzon,² Évelin A. Zanardo,² Thais Virginia Moura Machado Costa,² Marília M. Montenegro,^{1,2} Gil M. Novo-Filho,^{1,2} Alexandre T. Dias,² Amom M. Nascimento,^{1,2} Chong Ae Kim,¹ and Leslie D. Kulikowski^{2,3}

¹Genetics Unit, Instituto da Criança, Faculdade de Medicina da Universidade de São Paulo (USP), São Paulo, Brazil

²Department of Pathology, Cytogenomics Lab - LIM03, Faculdade de Medicina da Universidade de São Paulo (USP), São Paulo, Brazil

³Department of Collective Health - Human Reproduction and Genetics Center, Faculdade de Medicina do ABC, Santo André, São Paulo, Brazil

Manuscript Received: 19 August 2014; Manuscript Accepted: 17 August 2015

Williams–Beuren syndrome (WBS) is caused by a hemizygous contiguous gene microdeletion of 1.55–1.84 Mb at 7q11.23 region. Approximately, 28 genes have been shown to contribute to classical phenotype of SWB with presence of dysmorphic facial features, supravalvular aortic stenosis (SVAS), intellectual disability, and overfriendliness. With the use of Microarray-based comparative genomic hybridization and other molecular cytogenetic techniques, is possible define with more accuracy partial or atypical deletion and refine the genotype-phenotype correlation. Here, we report on a rare genomic structural rearrangement in a boy with atypical deletion in 7q11.23 and XYY syndrome with characteristic clinical signs, but not sufficient for the diagnosis of WBS. Cytogenetic analysis of G-banding showed a karyotype 47,XYY. Analysis of DNA with the technique of MLPA (Multiplex Ligation-dependent Probe Amplification) using kits a combination of kits (P064, P036, P070, and P029) identified an atypical deletion on 7q11.23. In addition, high resolution SNP Oligonucleotide Microarray Analysis (SNP-array) confirmed the alterations found by MLPA and revealed others pathogenic CNVs, in the chromosomes 7 and X. The present report demonstrates an association not yet described in literature, between Williams–Beuren syndrome and 47,XYY. The identification of atypical deletion in 7q11.23 concomitant to additional pathogenic CNVs in others genomic regions allows a better comprehension of clinical consequences of atypical genomic rearrangements. © 2015 Wiley Periodicals, Inc.

Key words: atypical deletion; 7q11.23; Williams syndrome; 47, XYY syndrome; genomic rearrangement; behavioral phenotype

INTRODUCTION

Copy Number variants or CNVs are the most prevalent types of structural variations in the human genome and result in the

How to Cite this Article:

Dutra RL, Piazzon FB, Zanardo ÉA, Costa TVMM, Montenegro MM, Novo-Filho GM, Dias AT, Nascimento AM, Kim CA, Kulikowski LD. 2015. Rare genomic rearrangement in a boy with Williams–Beuren syndrome associated to XYY syndrome and intriguing behavior. Am J Med Genet Part A 167A:3197–3203.

alteration of a variable number of genes causing changes in gene dosages (deletions and duplications), which ultimately may lead to population diversity, disease predisposition or specific clinical phenotypes [Merla et al., 2010].

Williams–Beuren syndrome (WBS; OMIM #194050) is a neurodevelopmental disorder described independently [Williams et al., 1961; Beuren et al., 1962] as a syndrome involving peculiar facial appearance, supravalvular aortic stenosis (SVAS), and intellectual disability. In fact, WBS presents a wide collection of symptoms affecting blood vessels, growth, intelligence, and behavior. Children with this condition have

Conflict of interest: None.

Roberta Lelis Dutra and Flavia Balbo Piazzon contributed equally to this manuscript.

Grant sponsor: Fundação de Amparo a Pesquisa do Estado de São Paulo; Grant numbers: 2010/52710-3, 2009/53105-9.

*Correspondence to:

Roberta L. Dutra, Ph.D., Cytogenetics Lab, Department of Pathology, Faculdade de Medicina da Universidade de São Paulo, 155, Dr. Enéas de Carvalho Aguiar Avenue, 2° floor, block 12, room 7, São Paulo, Brazil. E-mail: lelis.roberta@gmail.com

Article first published online in Wiley Online Library

(wileyonlinelibrary.com): 30 September 2015

DOI 10.1002/ajmg.a.37360

RESEARCH

Open Access

Cytogenomic delineation and clinical follow-up of 10 Brazilian patients with Pallister-Killian syndrome

Larissa Sampaio de Athayde Costa^{2*}, Aline C. Zandoná-Teixeira^{1,2}, Marília M. Montenegro^{1,2}, Alexandre T. Dias¹, Roberta L. Dutra^{1,2}, Rachel S. Honjo², Debora R. Bertola², Leslie D. Kulikowski^{1,2} and Chong A. Kim²

Abstract

Background: Pallister-Killian syndrome (PKS) is a sporadic genetic disorder caused by the presence of a tissue-specific mosaicism for isochromosome 12p - i(12) (p10) and is characterized by facial dysmorphism including coarse facies, upslanting palpebral fissures, bitemporal alopecia, pigmentary skin anomalies, developmental delay, hypotonia and seizures. Although typical clinical features of PKS commonly exist, clinicians often do not raise the possibility of this diagnosis.

Results: We reviewed the medical records of 10 patients with confirmed PKS followed in our service (since 1990 to 2015). Age at diagnosis varied from prenatal to 3 years and clinical features were consistent with those described in the literature. In all patients, peripheral blood karyotypes were normal and cytogenomic study was performed in order to confirm the diagnosis. Three of these patients had PKS diagnosis confirmed by buccal smear MLPA.

Conclusion: An early conclusion from our results demonstrated that MLPA on buccal smears is a good and non-invasive method to detect extra copies of 12p and should be considered as the first exam, before a skin biopsy for a fibroblast karyotype is performed.

Keywords: MLPA, Buccal smear, Pallister-Killian syndrome, isochromosome 12p

Background

Pallister-Killian syndrome (PKS) is a sporadic and rare chromosomal disorder caused by the mosaicism for isochromosome 12p - i(12)(p10). There are just over 200 cases described in the literature [1]. The syndrome is known by several names, such as Pallister mosaic syndrome, Pallister-Killian syndrome, Pallister-Killian-Teschler-Nicola syndrome, tetrasomy 12p and isochromosome 12p [2, 3].

Clinical findings of PKS are distinctive, especially the combination of coarse facies, pigmentary skin anomalies, developmental delay, hypotonia and seizures. The facial appearance includes coarse facies (coarseness becomes more pronounced with age), prominent forehead, hypertelorism, upslanting palpebral fissures with epicanthal folds, small nose with upturned nares, high arched palate, macrostomia, long philtrum, micrognathia, bitemporal

alopecia and an extension of the philtral skin into the vermilion border of the upper lip, which is termed "Pallister lip". Other associated findings include the following: congenital heart defects, diaphragmatic hernia, cryptorchidism, renal malformations, imperforate or anteriorly placed anus [2–4]. Patients have hypotonia, mental and motor retardation from early infancy, which can vary from mild to severe and profound [2, 5–7].

PKS is caused by a tissue-limited mosaicism for supernumerary i(12)(p10) [4, 8]. Although different levels of mosaicism and genetic variation in isochromosome composition have already been described in the literature, there is currently no evidence for genotype-phenotype correlation [3].

A fibroblast karyotype from skin biopsy is commonly used to achieve PKS diagnosis, but other tissues such as lung, tissue from a buccal smear, bone marrow and even amniotic fluid yield higher percentages of mosaicism i(12)(p10), which can help diagnose PKS [9–11].

* Correspondence: larissasathayde@gmail.com

²Unidade de Genética, Departamento de Pediatria, Instituto da Criança-HCFMUSP, Universidade de São Paulo, São Paulo, Brasil
Full list of author information is available at the end of the article

© 2015 Costa et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Accepted Manuscript

Post-mortem cytogenomic investigations in patients with congenital malformations

Alexandre Torchio Dias, Évelin Aline Zanardo, Roberta Leis Dutra, Flavia Balbo Piazzon, Gil Monteiro Novo-Filho, Marília Moreira Montenegro, Amom Mendes Nascimento, Mariana Rocha, Fabricia Andreia Rosa Madia, Thais Virgínia Moura Machado Costa, Cintia Milani, Regina Schultz, Fernanda Toledo Gonçalves, Cintia Fridman, Guilherme Lopes Yamamoto, Débora Romeo Bertola, Chong Ae Kim, Leslie Domenici Kulikowski

PII: S0014-4800(16)30155-1
DOI: doi: [10.1016/j.yexmp.2016.07.003](https://doi.org/10.1016/j.yexmp.2016.07.003)
Reference: YEXMP 3946

To appear in: *Experimental and Molecular Pathology*

Received date: 28 September 2015
Revised date: 16 June 2016
Accepted date: 19 July 2016

Please cite this article as: Dias, Alexandre Torchio, Zanardo, Évelin Aline, Dutra, Roberta Leis, Piazzon, Flavia Balbo, Novo-Filho, Gil Monteiro, Montenegro, Marília Moreira, Nascimento, Amom Mendes, Rocha, Mariana, Madia, Fabricia Andreia Rosa, Costa, Thais Virgínia Moura Machado, Milani, Cintia, Schultz, Regina, Gonçalves, Fernanda Toledo, Fridman, Cintia, Yamamoto, Guilherme Lopes, Bertola, Débora Romeo, Kim, Chong Ae, Kulikowski, Leslie Domenici, Post-mortem cytogenomic investigations in patients with congenital malformations, *Experimental and Molecular Pathology* (2016), doi: [10.1016/j.yexmp.2016.07.003](https://doi.org/10.1016/j.yexmp.2016.07.003)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Subtelomeric Copy Number Variations: The Importance of 4p/4q Deletions in Patients with Congenital Anomalies and Developmental Disability

Gil M. Novo-Filho^a Marília M. Montenegro^a Évelin A. Zanardo^a
 Roberta L. Dutra^b Alexandre T. Dias^a Flavia B. Piazzon^a Taís V.M.M. Costa^a
 Amom M. Nascimento^{a,b} Rachel S. Honjo^b Chong A. Kim^b Leslie D. Kulikowski^a

^aDepartment of Pathology, Cytogenomics Laboratory and ^bDepartment of Pediatrics, Instituto da Criança, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil

Key Words

4p/4q deletion · Arrays · Developmental disability · MLPA · Subtelomeric copy number variations

Abstract

The most prevalent structural variations in the human genome are copy number variations (CNVs), which appear predominantly in the subtelomeric regions. Variable sizes of 4p/4q CNVs have been associated with several different psychiatric findings and developmental disability (DD). We analyzed 105 patients with congenital anomalies (CA) and developmental and/or intellectual disabilities (DD/ID) using MLPA subtelomeric specific kits (P036/P070) and 4 of them using microarrays. We found abnormal subtelomeric CNVs in 15 patients (14.3%), including 8 patients with subtelomeric deletions at 4p/4q (53.3%). Additional genomic changes were observed at 1p36, 2q37.3, 5p15.3, 5q35.3, 8p23.3, 13q11, 14q32.3, 15q11.2, and Xq28/Yq12. This indicates the prevalence of independent deletions at 4p/4q, involving *PIGG*, *TRIML2*, and *FRG1*. Furthermore, we identified 15 genes with changes in copy number that contribute to neurological development and/or function, among them *CRMP1*, *SORCS2*, *SLC25A4*, and *HELT*. Our results highlight the association of genes with changes in copy number at 4p and 4q

subtelomeric regions and the DD phenotype. Cytogenomic characterization of additional cases with distal deletions should help clarifying the role of subtelomeric CNVs in neurological diseases.

© 2016 S. Karger AG, Basel

Genomic variations are a frequent cause of miscarriage, congenital anomalies (CA), and developmental disability (DD) [Nielsen and Wohler, 1991; Shaffer and Lupski, 2000]. Copy number variations (CNVs) are the most prevalent type of structural variation in the human genome and can affect the transcription rate, sequence, structure, and function of genes. They appear predominantly in the subtelomeric regions, which include repetitive sequences and are therefore more susceptible to chromosomal rearrangements. Previous studies have shown that CNVs associated with intellectual disability (ID) and CA are significantly enriched in genes whose mouse orthologous produce a nervous system phenotype when disrupted [Ballif et al., 2007; De Vries et al., 2003; Fernandez et al., 2005; Leonard and Wen, 2002; Ravnan et al., 2006; Schouten et al., 2002; Shaffer et al., 2005; Shao et al., 2008].

KARGER

E-Mail karger@karger.com
www.karger.com/cgr

© 2016 S. Karger AG, Basel
1424-8581/16/0000-0000\$39.50/0

Gil M. Novo-Filho
Department of Pathology, Cytogenomics Lab, Universidade de São Paulo
Av. Dr. Enéas de Carvalho Aguiar, 155
São Paulo, SP (Brazil)
E-Mail gilmonteiro@usp.br