



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE BIOCIÊNCIAS

GUSTAVO LOVATTO MICHAELSEN

Análise Integrada de Metilação de DNA e Expressão
Gênica nos Quatro Subgrupos de Meduloblastoma

Porto Alegre

2021

GUSTAVO LOVATTO MICHAELSEN

Análise Integrada de Metilação de DNA e Expressão
Gênica nos Quatro Subgrupos de Meduloblastoma

Trabalho de conclusão de curso de graduação
junto ao Instituto de Biociências da
Universidade Federal do Rio Grande do Sul,
como requisito parcial para a obtenção do
título de Bacharel em Biotecnologia.

Orientadora: Dra. Marialva Sinigaglia.

Porto Alegre

2021

RESUMO

O meduloblastoma (MB) está entre os tumores sólidos pediátricos mais frequentemente diagnosticados e estima-se que um terço dos pacientes com meduloblastoma irão a óbito devido à doença. Este tumor heterogêneo é atualmente classificado em quatro subgrupos moleculares geneticamente distintos: wingless (WNT), sonic hedgehog (SHH), Grupo 3 (G3) e Grupo 4 (G4). Alterações epigenéticas, como a disrupção da metilação do DNA são reconhecidas como uma característica universal da tumorigênese, e para o MB isto não é exceção. Neste estudo foi utilizado um grande conjunto de dados compreendendo 763 amostras pareadas de MB, tanto de microarranjo de expressão gênica como de metilação de DNA. Primeiramente foram identificados os genes diferencialmente metilados (DMs) e diferencialmente expressos (DEs) em cada um dos subgrupos. O subgrupo SHH apresentou uma menor taxa de metilação diferencial com 10,3% dos sítios CpGs DM, enquanto que o subgrupo G3 apresentou a maior taxa com 14,6%. Todos os subgrupos apresentaram uma maior taxa de hipermetilação comparada à hipometilação, em média 1,28 vezes maior. Posteriormente, foi examinada a correlação entre a metilação do DNA e expressão gênica. A correlação negativa está comumente associada a um silenciamento transcricional ao limitar o acesso de fatores de transcrição a regiões promotoras. Neste estudo, a correlação negativa mostrou-se mais comum que a positiva, especialmente em regiões promotoras, contudo a correlação positiva predominou no corpo do gene e na região 3'UTR. A metilação CpG do corpo gênico ainda não é bem compreendida, mas geralmente está associada à maior expressão do gene correspondente. Genes já associados ao MB como *GLI2*, *GAB1* e *PRAME* assim como genes ligados a outros tipos tumorais foram identificados DMs e DEs e possuindo um grande número de sítios CpGs com correlação significativa. Esse conjunto de genes nos ajudam a entender o comportamento da metilação e de que maneira ela influencia a expressão gênica nos diferentes subgrupos de MB.

Palavras-chave: câncer infantil; meduloblastoma; expressão diferencial; metilação diferencial; análise de correlação.

ABSTRACT

Medulloblastoma (MB) is the most frequently childhood solid tumor and it is estimated that one third of patients with medulloblastoma will die from the disease. This heterogeneous tumor is currently distinguished in four molecular subgroups: wingless (WNT), sonic hedgehog (SHH), Group 3 (G3) and Group 4 (G4). Epigenetic changes, such as disruption of DNA methylation, are recognized as a universal feature of tumorigenesis, and for MB this is no exception. In this study a large dataset was evaluated, comprising 763 MB paired samples of microarray gene expression and DNA Methylation. Differentially methylated (DMs) and differentially expressed (DEs) genes were identified in each of the subgroups. The SHH subgroup displayed a lower rate of differential methylation with 10.3% of DM CpGs sites, while the G3 subgroup had the highest rate with 14.6%. All subgroups presented a higher rate of hypermethylation compared to hypomethylation, on average 1.28 times greater. Subsequently, the correlation between DNA methylation and gene expression was performed. The negative correlation is associated with transcriptional silencing by limiting the access of transcription factors to promoter regions. In this study, negative correlation was pointed out to be more common than positive, especially in promoter regions. The positive correlation predominated in the gene body and in the 3'UTR region. The CpG methylation of the gene body is still not well understood, but it is usually associated with higher expression of the corresponding gene. Genes already associated with MB such as *GLI2*, *GAB1* and *PRAME* as well as genes linked to other tumor types have been identified as DMs and DEs and with a large number of CpG sites with good correlations coefficients. This set of genes contributed to a greater understanding of the behavior of methylation and your role in gene expression in different subgroups of MB.

Keywords: childhood cancer; medulloblastoma; differential expression; differential methylation; correlation analysis.

LISTA DE FIGURAS

Figura 1. Distribuição de sítios CpGs DM dentre as diferentes regiões gênicas (TSS1500, TSS200, 5'UTR, 1stExon, Body e 3'UTR) em cada um dos subgrupos de meduloblastoma.....	19
Figura 2. Classificação de sondas DMs de genes DEs em quatro grupos: genes hipermetilados e sobre expressos (hyper-up), hipermetilados e sub-expressos (hypo-down), hipometilados e sobre expressos (hypo-up) e hipometilados e sub-expressos (hypo-down) para cada subgrupo de meduloblastoma.....	20
Figura 3. Incidência da correlação negativa (corr-neg), correlação positiva (corr-pos), ambas as correlações (ambos) e sem correlação (NA) nos genes DEs e DMs e em todos os genes analisados para cada subgrupo de MB.....	21
Figura 4. Distribuição da correlação positiva (corr-pos), correlação negativa (corr-neg) e sem correlação (NA) de sítios CpGs DMs de genes DEs dentre as diferentes regiões gênicas (TSS1500, TSS200, 5'UTR, 1stExon, Body e 3'UTR) em cada um dos subgrupos de meduloblastoma.....	22
Figura 5. Representação dos termos do Gene Ontology (GO) e das rotas metabólicas do Kyoto Encyclopedia of Genes (KEGG) enriquecidas dos genes DMs DEs com no mínimo um sítio CpG com correlação significativa.....	30

LISTA DE TABELAS

Tabela 1. Características clínicas e genéticas dos subgrupos de MB.....	12
Tabela 2. Número de genes DMs e genes DEs por subgrupo comparados com controle (cerebelo normal).....	19
Tabela 3. Sondas dos genes <i>GPR81</i> e <i>ISLR2</i> com correlação negativa do subgrupo SHH.....	24
Tabela 4. Sondas do gene <i>PRAME</i> com correlação negativa dos subgrupos G3 e G4.....	25
Tabela 5. Genes com maior número de sítios CpGs com correlação negativa para cada subgrupo de MB.....	26
Tabela 6. Genes com maior número de sítios CpGs com correlação positiva para cada subgrupo de MB.....	27

LISTA DE ABREVIATURAS

GEO	Gene Expression Omnibus
GO	Gene Ontology
INCA	Instituto Nacional de Câncer
RNA	Ácido ribonucleico
MB	Meduloblastoma
DM	Diferencialmente metilado
DE	Diferencialmente expreso
WNT	Wingless
SHH	Sonic hedgehog
G3	Grupo 3
G4	Grupo 4
CpG	Citosina-fosfato-Guanina
SNC	Sistema nervoso central

SUMÁRIO

1 INTRODUÇÃO	9
1.1 Câncer infantil.....	9
1.2 O Meduloblastoma.....	10
1.3 A Epigenética do Meduloblastoma.....	13
2 JUSTIFICATIVA	14
3 OBJETIVO	15
3.1 Objetivo Geral.....	15
3.2 Objetivos Específicos	15
4 MATERIAL E MÉTODOS	15
4.1 Informações dos dados de microarranjo.....	15
4.2 Análise de metilação e expressão diferencial.....	16
4.3 Análise de correlação e Enriquecimento.....	17
5 RESULTADOS E DISCUSSÃO	18
5.1 Análise de metilação e expressão diferencial.....	18
5.1.1 CpGs DM nos subgrupos de Meduloblastoma.....	18
5.1.2 Distribuição da metilação diferencial nas diferentes regiões gênicas.....	18
5.1.3 Genes DE nos subgrupos de Meduloblastoma.....	19
5.2 Impacto da metilação do DNA na regulação gênica nos subgrupos de Meduloblastoma.....	20
5.2.1 Classificação dos genes baseada na intersecção de genes DMs e DEs.....	20
5.2.2 Comparação da correlação global com a de genes DMs e DEs.....	21
5.2.3 Distribuição da correlação nas diferentes regiões gênicas.....	22
5.2.4 A correlação negativa.....	23
5.2.5 A correlação positiva.....	26
5.2.6 Análise de Enriquecimento.....	28
6 CONCLUSÃO E PERSPECTIVAS	31
7 BIBLIOGRAFIA	32
8 APÊNDICE	39

1 INTRODUÇÃO

1.1 Câncer infantil

O câncer em crianças e adolescentes (entre 0 e 19 anos) se refere a um grande conjunto de doenças que apresentam características próprias, não somente de natureza histológicas, mas também relacionadas ao comportamento clínico (Little, 1999). Taxas de mortalidade eram expressivamente maiores há meio século, contudo o sucesso no tratamento hoje chega a 80% em países desenvolvidos, onde há acesso a abordagens modernas e cuidados à saúde robustos. Salienta-se que apenas 10% das crianças moram nesses países com alta renda (Noone AM et al., 2018). Em países de baixa e média renda onde vivem aproximadamente 90% das crianças com câncer essas estimativas são muito piores com taxas variando de 5% a 60% (Bakta et al., 2019; Ramirez et al., 2018). Devido à melhora no tratamento de doenças infecciosas e não transmissíveis em países de baixa renda, o tratamento do câncer tornou-se foco principal já que é uma das principais causas de morte em crianças e adolescentes (Fadel et al., 2019). Crianças com câncer representam aproximadamente apenas 1% a 3% de todos os cânceres diagnosticados em todo o mundo a cada ano. A medida do impacto global do câncer na infância e adolescência ainda permanece mal quantificada. A criação de estimativas sólidas do impacto global do câncer infantil é um grande desafio devido à escassez de registros de câncer de alta qualidade e dados de estatísticas vitais nos países em desenvolvimento, bem como as diferenças na etiologia, patogênese e apresentação das neoplasias mais comuns entre crianças e adultos (Global Burden of Disease Cancer Collaboration, 2017).

No Brasil, o câncer infantil é a primeira causa de morte por doença em crianças de 1 a 19 anos, superando as doenças infecciosas. O Instituto Nacional do Câncer (INCA) estimou para 2020 a ocorrência de 8.000 novos casos de câncer pediátrico no Brasil com estimativa de sobrevida de 64%, porém com diferenças contrastantes entre as regiões do país (INCA, 2019). Os tumores malignos cerebrais e do sistema nervoso somados as leucemias correspondem a 55% de todos os casos de câncer em crianças e adolescentes (0-14 anos), tornando outros tipos de cânceres pediátricos ainda mais raros (American Cancer Society, 2018).

Já é bem estabelecido que os tumores infantis diferem dos de adultos quanto aos seus aspectos morfológicos, origens histológicas, comportamento clínico, localizações primárias e período de latência, assim, devem ser estudados e tratados separadamente (INCA, 2016). Crianças tratadas em centros de oncologia pediátrica possuem 80% de chances de cura, já quando são tratadas em protocolos e centros de adultos a chance de cura cai para 40%. As

chances de cura efetiva no câncer infantil vão depender ainda do tipo de câncer e estão intimamente relacionadas com: (i) a capacidade de diagnosticar precocemente e com precisão; (ii) tratamento dentro de protocolos clínicos e equipe multidisciplinar especializada em oncologia pediátrica e (iii) eficácia em detectar doença metastática ou recidiva. Ainda segundo dados INCA de 2016, estima-se que 3% do total de casos de câncer ocorrem em crianças e adolescentes (0-19 anos). A perspectiva da sobrevida para todas as neoplasias em crianças e adolescentes foi em torno de 64%, variando de 50% na Região Norte, 60% na Região Nordeste, 65% na Região Centro-Oeste, 70% na Região Sudeste, e 75% na Região Sul. As taxas de sobrevivência de cinco anos para alguns tipos de câncer continuam a melhorar nas últimas décadas (INCA, 2016). Para que o progresso seja sustentado e renovado, é importante que pesquisas sejam conduzidas para identificar novas e inovadoras terapias que se baseiam em uma maior compreensão das vias celulares que promovem o crescimento e a sobrevivência das células tumorais em paciente pediátricos (Smith MA et al., 2010).

O surgimento dos estudos de sequenciamento genômico proporcionou um grande aumento na compreensão das bases genéticas do câncer infantojuvenil e destacou as principais diferenças entre tumores pediátricos e adultos. Enquanto a maioria dos tumores adultos é caracterizada por um número elevado de mutações somáticas, os cânceres pediátricos normalmente têm poucas mutações somáticas, e essas alterações ocorrem com alta frequência em genes que codificam proteínas envolvidas na regulação epigenética (Downin *et al.*, 2012; Huether *et al.*, 2014). Além disso, há uma heterogeneidade nos tipos de alterações genéticas que impulsionam o crescimento dos tumores pediátricos, incluindo alterações nos números de cópia e fusões gênicas (Sweet-Cordero et al., 2019).

1.2 O Meduloblastoma

O meduloblastoma (MB) é um dos tumores cerebrais pediátricos mais frequentes, corresponde a 20% de todos os tumores pediátricos do sistema nervoso central (McNeil et al., 2002; Weil et al., 2017). Estima-se que um terço dos pacientes com meduloblastoma irão a óbito devido a doença (Polkinghorn & Tarbell, 2007). Terapias multimodais agressivas, porém inespecíficas (cirurgia, radioterapia e quimioterapia) melhoraram significativamente a sobrevida de pacientes com MB, mas os sobreviventes apresentam efeitos colaterais neurológicos e cognitivos de início tardio graves incluindo déficit neurocognitivo, endocrinopatias e malignidades secundárias (Crawford et al., 2007; Diller et al., 2009; Packer e Vezina, 2008; Packer et al., 2013). A causa da maioria das mortes relacionadas com MB são

metástases leptomeningeas e a recaída é um evento quase uniformemente fatal, sem taxas de salvamento significativas (Ramaswamy e Taylor, 2017).

O MB é uma classe heterogênea de tumores cerebelares embrionários que apresentam aberrações em vias conhecidas por participarem de papéis críticos no desenvolvimento do cerebelo (Taylor MD et al., 2012). Atualmente a classificação molecular de Tumores do Sistema Nervoso Central da Organização Mundial de Saúde reconhece quatro subgrupos de MB geneticamente distintos: wingless (WNT), sonic hedgehog (SHH), Grupo 3 (G3) e Grupo 4 (G4) (WHO, 2016). Esses subgrupos são caracterizados por apresentarem diferenças transcricionais e genéticas específicas bem como implicações clínicas distintas (Kuzan-Fischer et al., 2017; Northcott et al., 2011). Cada subgrupo de MB responde de forma diferente aos regimes terapêutico tipicamente utilizados que compreendem a ressecção do tumor, radiação cranioespinhal em pacientes com mais de 3 anos e quimioterapia de combinação com um subconjunto de agentes antineoplásicos, incluindo vincristina (antimicrotubular), cisplatina (alquilante), ciclofosfamida (alquilante), carboplatina (baseado em platina) e/ou lomustina (alquilante) (Packer RJ et al., 2006; Needle MN et al., 1997). A adição de terapias direcionadas estão atualmente sendo avaliadas em ensaios clínicos. Os protocolos atuais tiveram um impacto nas taxas de sobrevida de até ~ 70%, embora os pacientes frequentemente sofram de sequelas graves que prejudicam muito sua qualidade de vida adulta (Gajjar A e Finlay JL, 2015). Trabalhos recentes avaliando dados de expressão e metilação sugerem que há uma certa heterogeneidade dentro dos quatro subgrupos, assim, cada subgrupo seria composto por outros subtipos (Cavalli et al., 2017).

O subgrupo WNT representa aproximadamente 10% de todos os MBs sendo caracterizado pela ativação da via de sinalização WNT. Apresenta o melhor prognóstico entre todos os subgrupos, raramente é metastático ao diagnóstico (<10% dos casos), com taxa de sobrevida de cinco anos em pacientes pediátricos superior a 95% (Taylor et al., 2012). Já o subgrupo SHH é definido pela ativação da via SHH e corresponde a aproximadamente 30% dos casos de MB. Mutações somáticas de *PTCH1*, *SMO* e *SUFU* e ampliações de *GLI2* e *MYCN* são características desse subgrupo (Kool M, Jones DT, Jager N, et al., 2014). Semelhante ao WNT, SHH é igualmente distribuído entre os sexos (Taylor et al., 2012). O prognóstico para os pacientes desse subgrupo é intermediário, com taxa de sobrevivência em cinco anos de 75% (Northcott et al., 2012). O MB de G3 representa um quarto de todos os casos de MBs e se caracteriza por ter o pior prognóstico entre todos os subgrupos, com sobrevida inferior a 60%, em cinco anos. A metástase é identificada no diagnóstico de cerca de 45% de todos os pacientes e raramente ocorre em adultos (Taylor et al., 2012). O G3 possui

uma assinatura de expressão gênica gabaminérgica (Northcott et al., 8/2012) e foi proposto que esse tumor pode surgir de células-tronco neurais multipotentes na zona ventricular do cerebelo embrionário que geram neurônios gabaminérgicos cerebelares. Amplificações de *MYC* (*C-MYC*) parecem ser quase exclusivas para o MB de G3, com cerca de 17% de incidência (Shih DJ, Northcott PA, Remke M, et al., 2014). O MB de G4 possui uma assinatura molecular glutamatérgica e alterações citogenéticas recorrentes incluem amplificação de *MYCN*, amplificação de *CDK6*, perda do cromossomo 11 e duplicação do cromossomo cinco onde o gene *SNCAIP* está localizado. O G4 constitui cerca de 35% dos casos, apresentando-se como o subgrupo mais frequente (Northcott et al., 8/2012) com sobrevida em cinco anos de 75% (Northcott et al., 12/2012).

Tabela 1. Características clínicas e genéticas dos subgrupos de MB.

	WNT (10%)	SHH (30%)	Group 3 (25%)	Group 4 (35%)
Grupo de Idade	>4 anos	<3 anos, >16 anos	<4 anos, <16 anos, >16 anos	Todas idades
Propensão a metástase	Baixa	Baixa	Alta	Alta
Prognóstico	Excelente	Intermediário	Ruim	Intermediário
Marcadores	Sem alteração recorrente	<i>MYCN</i> (12%), <i>GLI1</i> (8%)	<i>MYC</i> (17%), <i>PVT</i> (12%), <i>OTX2</i> (8%)	<i>SNCAIP</i> (10%), <i>MYCN</i> (6%), <i>CDK6</i> (5%)
Mutações	<i>CTNNB1</i> (91%), <i>DDX3X</i> (50%), <i>SMARCA4</i> (26%), <i>MLL2</i> (13%), <i>TPS3</i> (13%)	<i>TERT</i> (60%), <i>PTCH1</i> (46%), <i>SUFU</i> (24%), <i>MLL2</i> (16%), <i>SMQ</i> (14%), <i>TPS3</i> (13%)	<i>SMARCA4</i> (11%), <i>MLL2</i> (4%)	<i>KDM6A</i> (13%), <i>MLL2</i> (5%)
Aberrações cromossômicas	Perda do 6	Ganho do 3q. Perda do 9q, 10q e 14q	Ganho do 1q, 7, 17q e 18q. Perda do 8, 10q, 11, 16p, 17p	Ganho do 7, 17q e 18q. Perda do 8, 11p e X
Expressão	Sinalização de WNT	Sinalização de SHH e P13K	Assinatura de MYC/retinal	Assinatura de MYC/neuronal

Adaptado de (Rawal ZD et al., 2020)

1.3 A Epigenética do Meduloblastoma

Comparado aos tumores adultos os tumores pediátricos apresentam menos mutações somáticas sendo as alterações epigenéticas um fator preponderante nesses pacientes (HUETHER et al., 2014). Na última década, a caracterização molecular de vários grupos de tumores pediátricos têm contribuído para a classificação e o prognóstico de grupos de risco. Estas informações revelaram alterações genéticas e epigenéticas importantes que desempenham um papel relevante na iniciação, manutenção e/ou progressão tumoral (La Madrid & Kieran, 2017). As alterações herdáveis que não envolvem perturbações nas sequências de nucleotídeos são chamadas coletivamente de epigenéticas. A regulação epigenética é mediada pela metilação do DNA, modificações de histonas, remodelação da cromatina, microRNAs e LncRNAs (Sharma et al., 2010). Um dos fenômenos epigenéticos mais bem caracterizados é a metilação do DNA, que normalmente ocorre em ilhas CpGs na região promotora dos genes, resultando na repressão transcricional (Esteller M, 2008), entretanto a metilação de apenas um sítio CpG também pode ter um grande impacto na expressão de um gene, dependendo de sua localização (Watt F et al., 1988). A metilação do DNA é um mecanismo epigenético catalisado pela família de DNA metiltransferases que transferem um grupo metil da S-adenil metionina (SAM) para o quinto carbono de um resíduo de citosina para formar a 5mC (Moore LD et al., 2013). Este tipo de metilação regula a expressão gênica recrutando proteínas envolvidas na repressão gênica ou inibindo a ligação de fatores de transcrição ao DNA. Arranjos de metilação de DNA de todo o genoma verificam atualmente 850.000 locais CpG em regiões gênicas e também regiões não traduzidas. Esses arranjos não apenas ajudaram a confirmar os quatro subgrupos principais de MB, mas também identificaram vários subtipos dentro deles (Schwalbe et al., 2013). Essas análises sugerem que a metilação do DNA desempenha um papel importante na patogênese do MB, reprimindo genes para evitar a diferenciação celular e a morte celular, por exemplo (Dubuc et al., 2013).

Estudos iniciais usando conversão de bissulfito de DNA genômico seguido de amplificação por reação em cadeia da polimerase (PCR) revelaram vários genes supressores tumorais silenciados por hipermetilação de promotores CpG-densos incluindo *CDKN2A*, *HIC1* e *RASSF1* em MB (Shahi MH et al., 2010; Shahi MH et al., 2011). Outras abordagens identificaram genes reprimidos, incluindo *PTCH1*, o regulador negativo da sinalização de SHH, a família SFRP, inibidores da via de sinalização WNT e o repressor transcricional *ZIC2*, entre outros (Pfister S et al., 2007).

Os achados relacionados à metilação aberrante de genes participantes da tumorigênese

e de vias relacionadas ao câncer, impulsionaram a utilização da metilação de DNA como estratégia terapêutica no combate ao tumor. Algumas das primeiras drogas epigenéticas propostas como terapia anticâncer foram inibidores da metilação do DNA, incluindo 5-azacitidina (5-aza-CR) e 5-aza-2'-desoxicitidina (5-aza-CdR). Esses agentes citotóxicos se incorporam ao DNA e prendem as metiltransferases, moléculas responsáveis pela metilação do DNA, prevenindo a atividade e restaurando a expressão de genes supressores de tumor (Esteller M, 2008).

Entretanto, o impacto de muitas dessas alterações ainda não é bem compreendido e os mecanismos por trás das alterações epigenéticas e transcricionais não foram profundamente investigados. Análises de correlação têm sido amplamente usadas para examinar a relação entre metilação e expressão gênica (Zhou R, Man Y, 2016). A expressão gênica e a metilação podem estar correlacionadas positiva ou negativamente. Análises integradas anteriores de perfis de metilação de DNA e expressão gênica tornaram-se um método favorável na investigação da patogênese de doenças, particularmente no câncer (Li M et al., 2009). No presente estudo, a expressão gênica e os perfis de metilação do DNA foram investigados em conjunto. Os genes diferencialmente expressos (DE), sondas diferencialmente metiladas (DM) e funções associadas mostraram-se úteis para a elucidação dos mecanismos biológicos presentes no câncer, fundamentando e aperfeiçoando o potencial terapêutico da epigenética no tratamento dessa doença.

2. JUSTIFICATIVA

Estima-se que um terço dos pacientes com meduloblastoma irão a óbito devido à doença, em especial os pacientes dos subgrupos SHH, Grupo 3 e Grupo 4 (Polkinghorn & Tarbell, 2007). A metilação está envolvida na regulação de vários processos biológicos como a regulação da expressão gênica, desenvolvimento celular entre outros (Smith Z. D. et al., 2013; Kulis M. et al.; 2015). Já é bem estabelecido que a metilação aberrante tem um papel fundamental no desenvolvimento do câncer (Thakur C. et al., 2018). A hipermetilação de sítios CpGs em regiões promotoras têm sido amplamente verificada em diversos tumores, e pode levar ao silenciamento de genes supressores tumorais (Xu W, Xu M, Wang L et al., 2019). A identificação dos mecanismos epigenéticos que possuem um grande impacto na expressão gênica permite uma melhor compreensão desses tumores. Assim, a exploração e determinação de seus alvos são de extrema valia para a descoberta de biomarcadores bem como para um

possível direcionamento terapêutico.

3. OBJETIVO

3.1 Objetivo Geral

Avaliar a correlação entre metilação de DNA e expressão gênica dos quatro subgrupos moleculares de meduloblastoma pediátrico.

3.2 Objetivos Específicos

- Identificar os genes diferencialmente expressos nos subgrupos de meduloblastoma;
- Identificar os sítios CpGs diferencialmente metilados nos subgrupos de meduloblastoma;
- Avaliar a correlação entre metilação e expressão de genes que estão diferencialmente expressos e diferencialmente metilados nos subgrupos de meduloblastoma;
- Verificar a ocorrência da correlação metilação-expressão a nível estrutural, ou seja, quais regiões dos genes ocorrem uma maior correlação positiva e/ou negativa;
- Analisar de forma comparativa os quatro subgrupos de meduloblastoma com base nos resultados encontrados.

4 MATERIAL E MÉTODOS

4.1 Informações dos dados de microarranjo

Os dados foram obtidos através do repositório público de dados genômicos Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>)(EDGAR, 2002). Os seguintes datasets com dados tumorais foram analisados: GSE85212 (Illumina HumanMethylation450 BeadChip, GPL13534), GSE85217 (Affymetrix Human Gene 1.1 ST Array, GPL22286), cada um contendo 763 amostras pareadas de meduloblastoma pediátrico com dados de metilação de

DNA e expressão gênica respectivamente, com os subgrupos já definidos. O dataset GSE148388 (Infinium MethylationEPIC, GPL21145) contém 10 amostras controle de cerebelo saudável de metilação de DNA. Por fim, o dataset GSE167447 (Affymetrix Human Gene 1.1 ST Array, GPL11532) contém 8 amostras controle de cerebelo saudável de expressão gênica. As amostras controle não são pareadas.

4.2 Análise de metilação e expressão diferencial

A análise de metilação diferencial entre as amostras de MB e tecido saudável foi realizada para cada um dos subgrupo separadamente, e através dos pacotes R/Bioconductor limma (Ritchie et al., 2015) e minfi (Aryee et al., 2014). O minfi é responsável pelo carregamento dos dados e controle de qualidade das amostras, além de todos os processos de normalização e filtragem dos dados. A normalização foi realizada utilizando a função preprocessFunnorm, apropriada para dados com diferenças globais de metilação como câncer/tecido saudável ou tecidos vastamente diferentes (Fortin JP et al., 2014). As sondas filtradas possuem p-valor de detecção maior que 0,01, pertencem à cromossomos sexuais, possuem alto polimorfismo ou hibridização cruzada. Sondas que não estão mapeadas a genes ou sondas mapeadas a genes em que não há dados de expressão na plataforma Affymetrix Human Gene 1.1 ST Array também não foram utilizadas. Como os dados tumorais e de tecido saudável são de plataformas diferentes, foi utilizada a função combineArrays do pacote minfi, um método que combina diferentes tipos de arranjos de metilação em um arranjo virtual. Especificamente, os microarranjos 450k e EPIC da Illumina compartilham muitas sondas em comum. Esta função combina dados dos dois tipos de arranjos diferentes e produz um único objeto (Aryee et al., 2014). O pacote limma é requerido para a própria análise diferencial. Sondas foram consideradas DMs com base nos seguintes parâmetros de corte: p-valor < 0,05 e média do valor beta de metilação entre caso e controle > 0,2. Genes que apresentam pelo menos uma sonda diferencialmente metilada são definidos como genes DMs (Zhao et al., 2018).

O pacote limma também foi utilizado para análise de expressão diferencial entre as amostras de MB e tecido saudável, além dos pacotes oligo e arrayQualityMetrics (Kauffmann A et al, 2009; Carvalho et al., 2010) para carregamento dos dados, controle de qualidade das amostras e normalização. Foi utilizado o método de normalização RMA do pacote oligo (Bolstad BM et al., 2003; Rafael A et al., 2003 e Irizarry RA et al., 2003). Genes foram consideradas DEs com base nos seguintes parâmetros de corte: p-valor < 0,01 e logFC > 1,5.

Os p-valores, tanto da análise de expressão quanto de metilação diferencial foram ajustados utilizando o método de Benjamini e Hockberg's para controlar as taxas de descobertas falso positivas (Benjamini Y e Hochberg., 1995). Dez amostras não passaram pelo controle de qualidade e foram descartadas. As amostras de metilação pertencentes a esses mesmos dez indivíduos também não foram utilizadas. O número final de amostras pareadas de expressão e metilação por grupo foi de 220 amostras SHH, 141 amostras G3, 323 amostras G4 e 69 amostras WNT, além das 8 amostras de expressão de cerebelo saudável e 10 amostras de metilação de cerebelo saudável.

4.3 Análise de correlação e Enriquecimento

A análise de correlação de Spearman foi realizada entre a expressão gênica e o valor beta de metilação de sondas mapeadas a eles para cada um dos subgrupos de MB. O cálculo foi realizado para todas as sondas que possuem um gene mapeado e com valor de expressão gênica disponível. Foram consideradas correlações positivas aquelas com rho (coeficiente de correlação de Spearman) $> 0,5$ e correlações negativas aquelas com rho $< -0,5$. O p-valor não foi utilizado como parâmetro de corte nesta análise pois não afetaria o resultado final. Os níveis de correlação foram analisados dentro de diferentes regiões do gene previamente definidas pela plataforma, em que as sondas estão mapeadas, sendo elas: TSS1500 (1500 bases antes do sítio de início de transcrição), TSS200, 5'UTR, 1stExon (primeiro exon), Body (corpo do gene) e 3'UTR.

O enriquecimento funcional foi realizado na plataforma GeneTrail 3.0 (<https://genetrail.bioinf.uni-sb.de/>) (Nico G. et al., 2020). Foram identificados os termos do GO (Gene Ontology) e vias do KEGG (Kanehisa, 2020) através do banco de dados do MSigDB (Subramanian A et al., 2005) associados aos genes DMs DEs com pelo menos um sítio CpG com correlação significativa para cada subgrupo de MB. Foram representados os 10 termos e rotas mais significativos com seus p-valores

5 RESULTADOS E DISCUSSÃO

5.1 Análise de metilação e expressão diferencial

5.1.1 CpGs DMs nos subgrupos de Meduloblastoma

Um total de 31325 sítios CpGs estão DMs no subgrupo SHH (17791 hipermetilados e 13534 hipometilados), 44266 no subgrupo G3 (24976 hipermetilados e 19290 hipometilados), 38680 no subgrupo G4 (21935 hipermetilados e 16745 hipometilados) e 44089 no subgrupo WNT (24235 hipermetilados e 19854 hipometilados) de um total de 302766 sondas avaliadas por subgrupo. O subgrupo SHH apresentou uma menor taxa de metilação diferencial com 10,3% dos sítios DM, enquanto que o subgrupo G3 apresentou a maior taxa com 14,6%. Todos os subgrupos apresentaram uma maior frequência hipermetilação comparada à hipometilação, em média 1,28 vezes mais comum. Um padrão maior de metilação em MB comparado ao cerebelo normal adulto já foi reportado em estudos de sequenciamento (Hovestadt V et al., 2014), resultado semelhante visto aqui com cerebelo normal pediátrico.

5.1.2 Distribuição da metilação diferencial nas diferentes regiões gênicas

Examinando o local da metilação diferencial dentre diferentes regiões gênicas a hipermetilação mostrou-se ser mais comum no corpo gênico e na região 3'UTR para todos os subgrupos, chegando a compreender 10,8% das sondas presentes no corpo do gene para as amostras do subgrupo G3 e nunca menor que 4% para qualquer região dos subgrupos (Figura 1). Já a hipometilação mostrou-se mais comum nas regiões TSS1500, corpo do gene e 3'UTR, porém com diferenças pequenas. Em número totais o corpo do gene apresenta uma maior diferença de metilação, contudo ele possui uma maior cobertura de sítios pela plataforma compreendendo 42% das 302766 sondas avaliadas. Posteriormente foi realizado o mapeamento dos sítios DM a seus genes (Tabela 2).

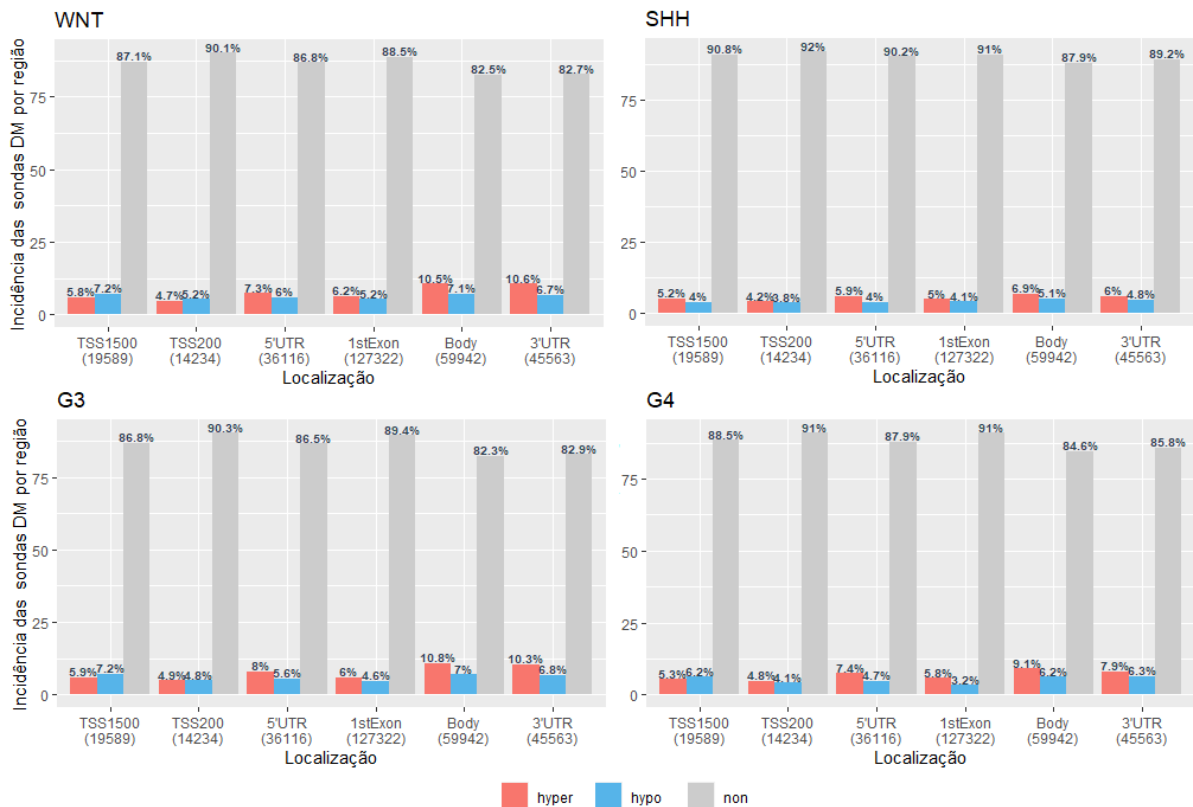


Figura 1. Distribuição de sítios CpGs DM dentre as diferentes regiões gênicas (TSS1500, TSS200, 5'UTR, 1stExon, Body e 3'UTR) em cada um dos subgrupos de meduloblastoma. Em vermelho as sondas hipermetiladas, em azul sondas hipometiladas e em cinza sondas sem diferença de metilação significativa quando comparadas com o controle (cerebelo normal). Entre parênteses o número total de sondas da região.

5.1.3 Genes DE nos subgrupos de Meduloblastoma

A análise de expressão diferencial foi realizada individualmente para cada um dos subgrupos: SHH (220 amostras), G3 (141 amostras), G4 (323 amostras), WNT (69 amostras) e 8 amostras de cerebelo saudável como controle (Tabela 2).

Tabela 2. Número de genes DMs e genes DEs por subgrupo comparados com controle (cerebelo normal).

	SHH	WNT	G3	G4
Genes Hipermetilados	7302	8317	8487	8264
Genes Hipometilados	6273	8152	7864	7280
Total de Genes DM	10902	12671	12724	12017
Genes Sobre Expressos	111	359	194	157
Genes Sub-Expresso	687	882	879	779
Total de Genes DE	798	1241	1073	936

5.2 Impacto da metilação do DNA na regulação gênica nos subgrupos de Meduloblastoma

5.2.1 Classificação dos genes baseada na intersecção de genes DMs e DEs

A análise integrada entre metilação de DNA e expressão gênica teve como objetivo determinar a intersecção entre os genes DMs e DEs. Uma série de estudos mostraram que a metilação do promotor leva à inativação de genes supressores tumorais, sendo um mecanismo importante no desenvolvimento de câncer (Lee, S et al., 2004; Licchesi LD et al., 2008). A metilação CpG do corpo gênico ainda não é bem compreendida, mas geralmente está associada à maior expressão do gene correspondente, ou vice-versa, hipometilação no corpo gênico associada a uma expressão mais baixa (Yang X et al., 2014). Os genes foram classificados em quatro grupos baseado na intersecção de sondas DMs e genes DEs: genes hipermetilados e sobre expressos (hyper-up), hipermetilados e sub-expressos (hypo-down), hipometilados e sobre expressos (hypo-up) e hipometilados e sub-expressos (hypo-down) (Figura 2).

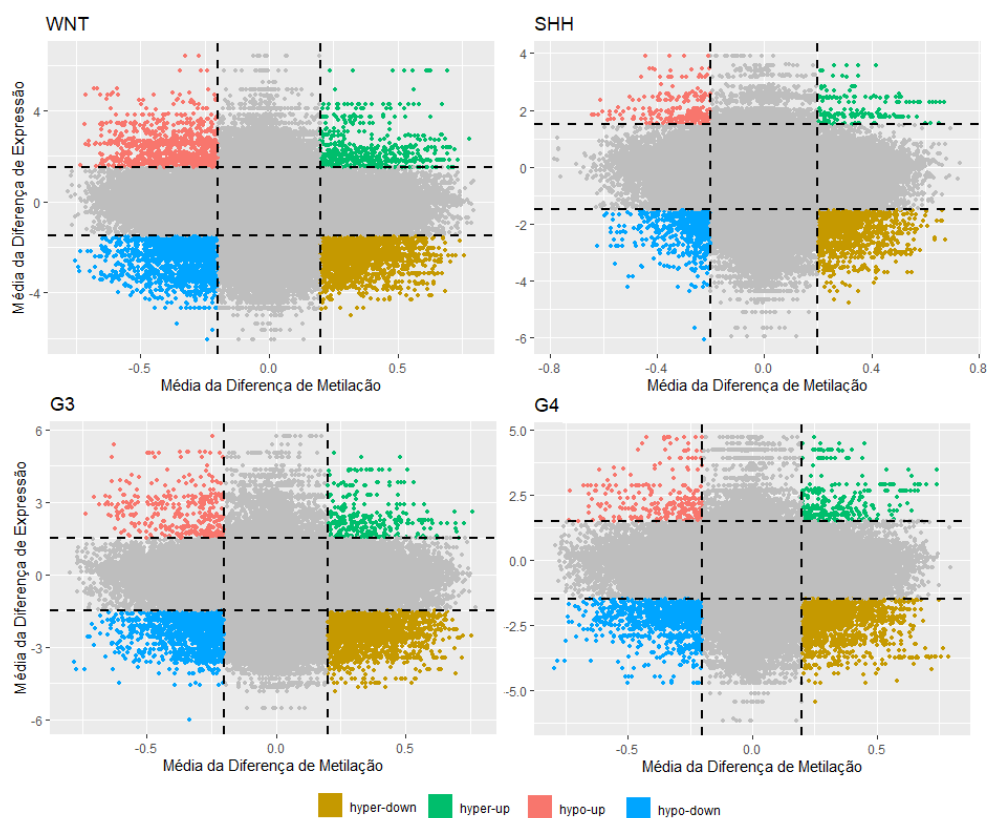


Figura 2. Classificação de sondas DMs de genes DEs em quatro grupos: genes hipermetilados e sobre expressos (hyper-up), hipermetilados e sub-expressos (hypo-down), hipometilados e sobre expressos (hypo-up) e hipometilados e sub-expressos (hypo-down) para cada subgrupo de meduloblastoma.

5.2.2 Comparação da correlação global com a de genes DMs e DEs

A análise de correlação de Spearman foi realizada para todos os sítios CpGs mapeados e genes em que também havia o seu valor de expressão. A correlação positiva é caracterizada em genes com alta metilação e alta expressão ou baixa metilação e baixa expressão. A correlação negativa ocorre quando há alta metilação e baixa expressão ou baixa metilação e alta expressão, classicamente encontrada em regiões promotoras (Lee, S et al., 2004). A correlação global (assumindo todos os 17972 genes com 302766 sondas) foi comparada com a correlação dos genes DM e DE encontrados nas análises anteriores para cada subgrupo (Figura 3). A correlação mostrou-se 2,9 vezes mais comum nos genes DMs e DEs comparada com o conjunto de todos os genes analisados, o que fortalece o papel da metilação na tumorigênese desse tumor, visto o forte impacto da metilação nos genes que diferem o tumor comparado ao tecido saudável.

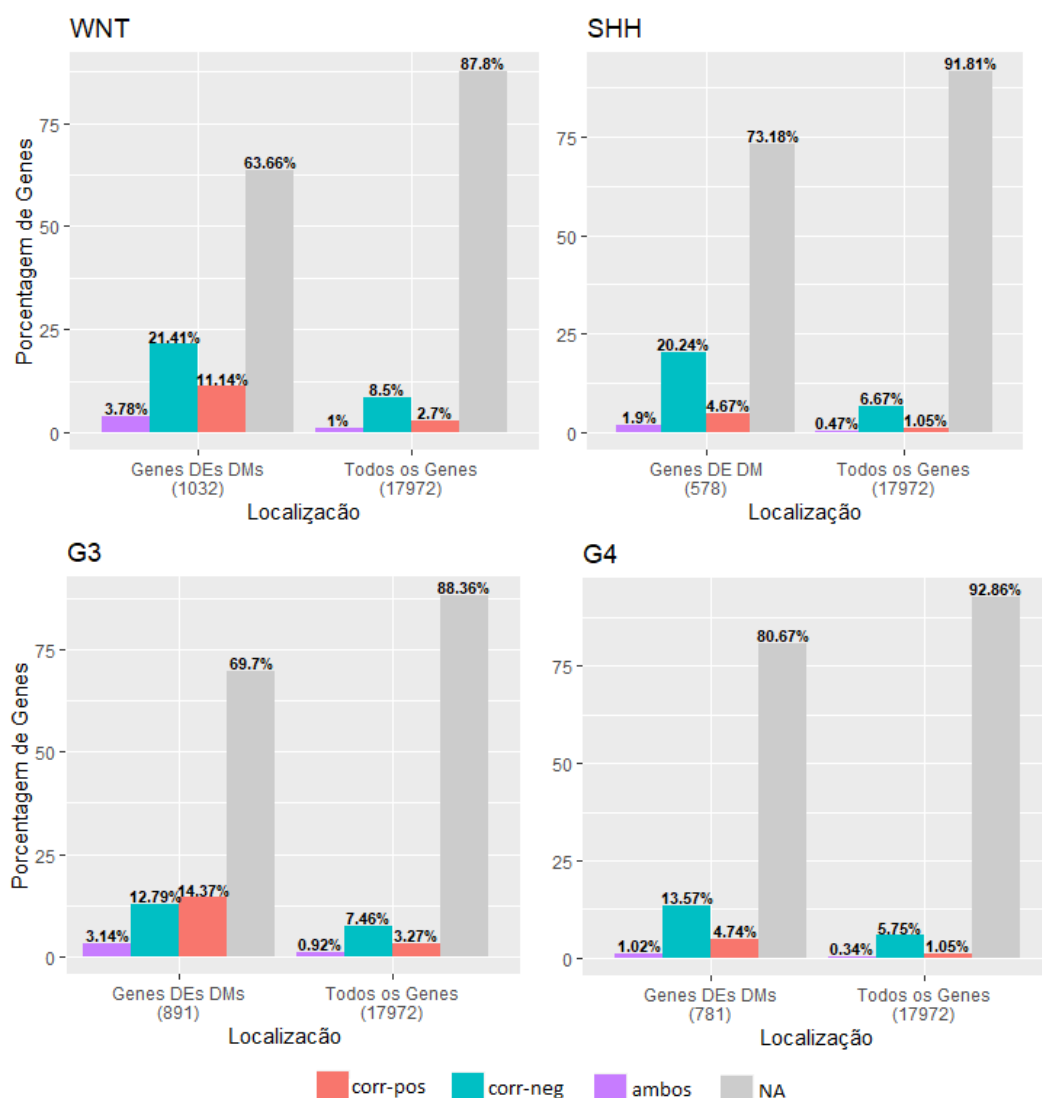


Figura 3. Incidência da correlação negativa (corr-neg), correlação positiva (corr-pos), ambas as correlações (ambos) e sem correlação (NA) nos genes DEs e DMs e em todos os genes analisados para cada subgrupo de MB.

5.2.3 Distribuição da correlação nas diferentes regiões gênicas

Foi avaliada a distribuição da correlação ao longo das diferentes regiões gênicas em cada subgrupo para os genes DMs e DEs. A correlação negativa mostrou-se mais comum nas regiões próximas ao promotor (TSS1500, TSS200, 5'UTR) em proporcionalidade, assim como na região do primeiro exon. Ela também mostrou-se 4,9 vezes mais comum que a correlação positiva nessas regiões. O subgrupo SHH apresentou a maior porcentagem de correlação negativa com média de 13,1% dentre as diferentes regiões com um pico de correlação negativa de 20,8% na região TSS200. A correlação positiva, em proporção, foi encontrada comumente no corpo de gene e na região 3'UTR, em incidência média 5% comparado a 1,7% das demais regiões. O G3 possui a maior porcentagem média de correlação positiva com 4,8%, chegando até 9,6% na região 3'UTR (Figura 4).

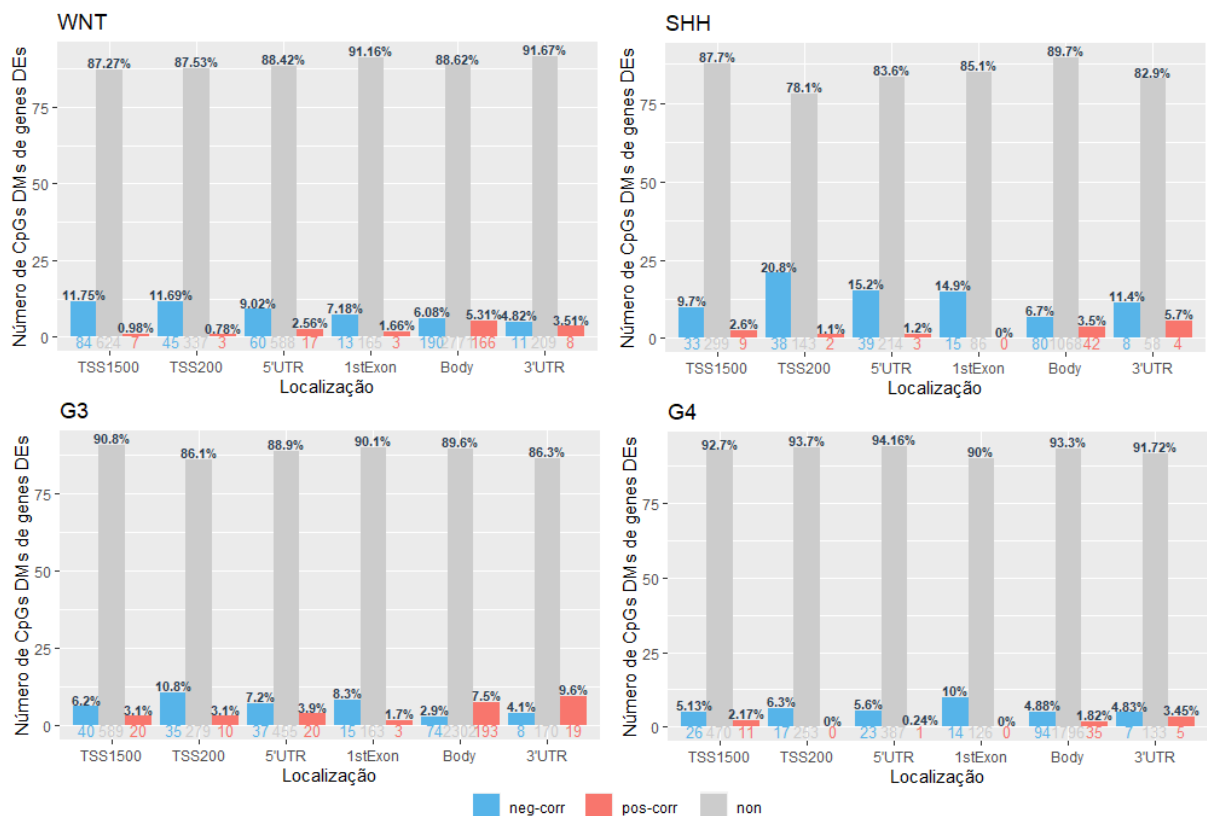


Figura 4. Distribuição da correlação positiva (corr-pos), correlação negativa (corr-neg) e sem correlação (NA) de sítios CpGs DMs de genes DEs dentre as diferentes regiões gênicas (TSS1500, TSS200, 5'UTR, primeiro exon, corpo do gene e 3'UTR) em cada um dos subgrupos de meduloblastoma.

5.2.4 A correlação negativa

Visto que a correlação negativa está associada a um silenciamento transcricional de genes hipermetilados ou um aumento transcricional de genes hipometilados, principalmente em regiões promotoras, e que um conjunto de sítios CpGs potencialmente teriam um maior impacto na transcrição que apenas um único sítio (Esteller M, 2008), foram selecionados os genes hipermetilados sub-expressos e hipometilados sobre expressos (destacados na Figura 2 em amarelo e vermelho, respectivamente) com o maior número de sítios com correlação negativa em cada um dos subgrupos de MB (Tabela 5). Para o subgrupo SHH, há 206 sítios com correlação negativa mapeados a 77 genes. Os genes *GPR81* e *ISLR2* apresentaram o maior número de sítios, 11 cada (Tabela 3), localizados quase exclusivamente em regiões promotoras. O gene *GPR81* é um receptor de lactato pouco estudado em câncer e apresentou-se sobre expresso e hipometilado. Recentemente, foi encontrada sua expressão em linhagens celulares de carcinoma de pulmão, mama, colo, hematocelular, glândula salivar e cervical. Funcionalmente a redução da expressão de *GPR81* em culturas de células pancreáticas leva a uma rápida morte celular quanto suplementadas com baixas concentrações de glicose, porém com altas concentrações de lactato (Roland CL et al., 2014). A reprogramação de vias bioquímicas é uma marca registrada de células tumorais, e a geração de ácido láctico a partir de glicose/glutamina representa uma das consequências de tais alterações metabólicas (Brown TP et al., 2019). In vivo, os níveis de sua expressão estão correlacionados com taxas de crescimento tumoral e metástase em câncer pancreático (Roland CL et al., 2014). O gene *ISLR2* foi encontrado sub-expresso e hipermetilado. Ele é necessário para extensão do axônio durante o desenvolvimento neural. Foi pouco estudado em câncer, contudo é categorizado como um marcador epigenético de glioblastoma devido ao seu status de metilação nesse tipo tumoral (Mock A et al., 2016). O gene *GLI2* também foi encontrado com um grande número de sítios com correlação negativa. Ele é um biomarcador clássico do subgrupo SHH e encontra-se normalmente amplificado (Shih DJ, Northcott PA, Remke M, et al., 2014), assim como o gene *GABI*, também associado ao subgrupo SHH (Shuangshoti S et al., 2020).

Tabela 3. Sondas dos genes *GPR81* e *ISLR2* com correlação negativa do subgrupo SHH.

Sonda	Gene	Local	Rho	Sonda	Gene	Local	Rho
cg00357958	<i>GPR81</i>	1stExon	-0,793	cg00781110	<i>ISLR2</i>	TSS1500	-0,581
cg01306688	<i>GPR81</i>	TSS200	-0,746	cg03689324	<i>ISLR2</i>	Body	-0,519
cg05290737	<i>GPR81</i>	TSS200	-0,729	cg05505745	<i>ISLR2</i>	5'UTR	-0,572
cg13702536	<i>GPR81</i>	TSS1500	-0,790	cg05547902	<i>ISLR2</i>	TSS1500	-0,600
cg15823495	<i>GPR81</i>	1stExon	-0,879	cg07216112	<i>ISLR2</i>	5'UTR	-0,581
cg19975916	<i>GPR81</i>	TSS200	-0,731	cg08036502	<i>ISLR2</i>	TSS1500	-0,566
cg20566840	<i>GPR81</i>	TSS200	-0,781	cg12017911	<i>ISLR2</i>	5'UTR	-0,550
cg22534509	<i>GPR81</i>	1stExon	-0,767	cg23152885	<i>ISLR2</i>	5'UTR	-0,538
cg22972858	<i>GPR81</i>	TSS200	-0,737	cg25828462	<i>ISLR2</i>	TSS1500	-0,587
cg23505823	<i>GPR81</i>	TSS1500	-0,706	cg25833989	<i>ISLR2</i>	Body	-0,554
cg26643476	<i>GPR81</i>	TSS1500	-0,841	cg27200833	<i>ISLR2</i>	5'UTR	-0,572

O subgrupo G3 apresentou 156 sítios mapeados em 82 genes (Tabela 5). O gene *MEIS1* possui o maior número de sítios com correlação negativa, todas as 10 sondas na região do corpo, encontra-se hipermetilado e sub-expresso. *MEIS1* é um fator de transcrição que interage e regula a atividade de fatores de transcrição homeodomínio *HOX* (Johng D et al., 2019). Em câncer de próstata *MEIS1* é regulado negativamente por *MYC* mediando seu desenvolvimento (Whitlock NC et al., 2020). *MYC* está amplificado quase que exclusivamente no subgrupo G3 (Shih DJ, Northcott PA, Remke M, et al., 2014) e mecanismos similares podem estar ocorrendo também neste subgrupo. *MYC* não foi encontrado como DE no subgrupo G3, contudo apresentou um logFC igual a 1,42 (0,08 abaixo do corte estabelecido). O subgrupo G4 apresentou 124 sítios mapeados a 63 genes. Com 8 sítios em regiões próximas ao seu promotor, o gene *HIST1H1A* possuiu o maior número de sondas com correlação negativa para nesse subgrupo, apresentando-se hipermetilado e sub-expresso. Esse gene codifica H1.1, um membro da família das histonas ligantes que está envolvida na organização e compactação da cromatina. *HIST1H1A* possui alguns estudos em câncer, ele foi identificado como um gene que confere resistência à cisplatina em células HeLa (Wu ZZ et al., 2010). Em células de câncer de próstata a sobre expressão de *HIST1H1A* aumentou o crescimento celular, assim como a expressão de marcadores neuroendócrinos e epiteliais/mesenquimais in vitro (Williams KA et al., 2018). Também em câncer de próstata, utilizando modelos animais, a expressão de *HIST1H1A* foi associada como um efeito protetivo contra desenvolvimento agressivo da doença (Williams KA et al., 2014). O gene *PRAME* foi o segundo com o maior número de sítios tanto no subgrupo G3 quanto no subgrupo G4, com 7 e 6 sítios respectivamente,

compartilhando 5 sondas iguais (Tabela 4). Ele encontra-se hipometilado e sobre expresso em ambos subgrupos. *PRAME* é amplamente estudado em câncer, inclusive em MB. *PRAME* foi originalmente identificado como um gene que codifica antígenos para HLA-A*24, capazes de estimular linfócitos citotóxicos tumor-específicos derivados de pacientes com melanoma (Ikeda H et al., 1997). Sua expressão foi demonstrada em tecidos germinativos e em vários tumores (incluindo melanoma, carcinoma de pulmão, carcinoma de mama, carcinoma de células renais, câncer de cabeça e pescoço, tumor de Wilms e linfoma de Hodgkin), porém a sua expressão é limitada em tecidos adultos saudáveis (Wadelin F et al., 2010). Em MB, *PRAME* encontra-se comumente sobre expresso e está correlacionado com uma pior sobrevida. A sua baixa expressão em tecidos adultos e alta expressão em MB tornam esse gene um alvo promissor para imunoterapia. Células T modificadas com receptores específicos para os peptídeos do gene *PRAME* induziram a morte das células de MB. Resultados similares também foram vistos em modelo animal ortotópico (Orlando D et al., 2018).

Tabela 4. Sondas do gene *PRAME* com correlação negativa dos subgrupos G3 e G4.

G3

Sonda	Gene	Local	Rho
cg01966129	<i>PRAME</i>	TSS200	-0,706
cg07969622	<i>PRAME</i>	TSS200	-0,542
cg22871485	<i>PRAME</i>	TSS200	-0,690
cg27303185	<i>PRAME</i>	5'UTR	-0,563
cg20667684	<i>PRAME</i>	TSS200	-0,601
cg17648213	<i>PRAME</i>	Body	-0,627
cg19804859	<i>PRAME</i>	5'UTR	-0,532

G4

Sonda	Gene	Local	Rho
cg01966129	<i>PRAME</i>	TSS200	-0,630
cg07969622	<i>PRAME</i>	TSS200	-0,535
cg22871485	<i>PRAME</i>	TSS200	-0,624
cg27303185	<i>PRAME</i>	5'UTR	-0,568
cg20667684	<i>PRAME</i>	TSS200	-0,571
cg05208878	<i>PRAME</i>	TSS200	-0,500

O subgrupo WNT possui o maior número de sítios com correlação negativa, com 353 sondas mapeadas a 162 genes. Em especial o gene hipermetilado e sub-expresso *NTM* com 12 sítios, participante na adesão celular e no reconhecimento neuronal (Struyk et al., 1995).

Tabela 5. Genes com maior número de sítios CpGs com correlação negativa para cada subgrupo de MB.

WNT		SHH		G3		G4	
Genes	Número de Sítios	Genes	Número de Sítios	Genes	Número de Sítios	Genes	Número de Sítios
<i>NTM</i>	12	<i>GPR81</i>	11	<i>MEIS1</i>	10	<i>HIST1H1A</i>	8
<i>MGC45800</i>	10	<i>ISLR2</i>	11	<i>PRAME</i>	7	<i>PRAME</i>	6
<i>ONECUT2</i>	9	<i>KIF26A</i>	10	<i>SGIP1</i>	7	<i>TEX14</i>	6
<i>DSCR4</i>	8	<i>ALDH1A3</i>	8	<i>ABCA4</i>	5	<i>NR2F1</i>	5
<i>DSCR8</i>	8	<i>C6orf123</i>	8	<i>HIST1H4F</i>	5	<i>C1orf61</i>	4
<i>GNG3</i>	7	<i>GABI</i>	7	<i>RUNX1T1</i>	5	<i>DSCR8</i>	4
<i>HIST1H4D</i>	7	<i>KCNJ3</i>	7	<i>C1orf61</i>	4	<i>FZD7</i>	4
<i>CDKL2</i>	6	<i>PCNT</i>	7	<i>CDKL2</i>	4	<i>KCNKI</i>	4
<i>KHDRBS2</i>	6	<i>TSPYL5</i>	7	<i>ELAVL4</i>	4	<i>LHX2</i>	4
<i>AKRIE2</i>	5	<i>ZNF154</i>	7	<i>GNB3</i>	4	<i>LMX1A</i>	4
<i>CNGA3</i>	5	<i>CPNE8</i>	6	<i>TBR1</i>	4	<i>RASSF8</i>	4
<i>COL6A3</i>	5	<i>GLI2</i>	5	<i>TULP1</i>	4	<i>TBR1</i>	4
<i>FMOD</i>	5	<i>HIST1H4F</i>	5	<i>BMF</i>	3	<i>ZFPM2</i>	3
<i>TCF7</i>	5	<i>ALDOC</i>	4	<i>C3orf64</i>	3	<i>ANK3</i>	2
<i>BAMBI</i>	4	<i>CACNA1E</i>	4	<i>GUCY2D</i>	3	<i>C12orf5</i>	2
<i>CASZ1</i>	4	<i>FGF5</i>	4	<i>MPP4</i>	3	<i>CDH23</i>	2
<i>EBF1</i>	4	<i>KIAA0182</i>	4	<i>PCBP4</i>	3	<i>DDX31</i>	2
<i>FAM174B</i>	4	<i>QPCT</i>	4	<i>B3GALNT2</i>	2	<i>FAM18A</i>	2
<i>LGR5</i>	4	<i>SPRY2</i>	4	<i>C6orf191</i>	2	<i>GABRR2</i>	2
<i>MCTP2</i>	4	<i>C8orf84</i>	3	<i>CACNA2D4</i>	2	<i>KCNA5</i>	2

5.2.5 A correlação positiva

Também foram selecionados os genes hipometilados sub-expressos e hipermetilados sobre expressos com um grande número de sondas com correlação positiva (Tabela 6). A correlação positiva localizada no corpo do gene mostrou-se bem mais comum que nas demais regiões para todos os subgrupos, com 133 do total de 155 no subgrupo WNT, 35 de 44 no SHH, 141 de 191 no G3 e 30 de 41 no G4. A metilação CpG do corpo gênico ainda não é bem compreendida, mas geralmente está associada à maior expressão do gene correspondente, ou vice-versa, hipometilação no corpo gênico associada a uma expressão mais baixa (Yang X et al., 2014). Tal fenômeno ocorre pois a metilação de DNA no corpo do gene o protege contra a entrada da RNA polimerase II em promotores intragênicos normalmente inacessíveis para a

montagem da maquinaria transcricional. Portanto, essa metilação é responsável pela prevenção de eventos de iniciação da transcrição aberrantes e é necessária para garantir a fidelidade da iniciação da transcrição do mRNA original. Tais RNAs parciais gerados pela perda da fidelidade do início de transcrição poderiam afetar múltiplos processos biológicos, prejudicando mecanismos moleculares como regulação da expressão gênica, direcionamento de miRNA, geração de proteína truncada entre outros, favorecendo a heterogeneidade de células tumorais e a predisposição neoplásica (Neri F et al., 2017). Uma série de genes apresentam um conjunto de sondas com correlação positiva quase exclusivamente em seu corpo. Em especial o gene *GABBR1*, que codifica uma subunidade do receptor de GABA tipo B, foi o terceiro com maior número de sondas no subgrupo G3 com 16 sítios CpGs, 15 em seu corpo estando hipometilado e sub-expresso. O subgrupo G3 é conhecido por ter uma assinatura de expressão gabaminérgica, comumente com a alta expressão de *GABRA5*, uma subunidade do receptor de GABA tipo A (Cho YJ et al., 2011). Esse receptor tipo A já foi proposto como alvo terapêutico específico para o subgrupo G3 de MB por aumentar o desequilíbrio iônico induzindo a apoptose (Kallay L et al., 2019). Ambos os tipos de receptores são direcionados por muitos medicamentos clinicamente importantes que afetam a função gabaminérgica e são amplamente usados no tratamento de transtorno de ansiedade, epilepsia, insônia, espasticidade, comportamento agressivo e outras condições e doenças fisiopatológicas (Jembrek et al., 2015), contudo pouco se sabe sobre os papéis do receptor tipo B de GABA em MB.

Tabela 6. Genes com maior número de sítios CpGs com correlação positiva para cada subgrupo de MB.

WNT		SHH		G3		G4	
Genes	Número de Sítios	Genes	Número de Sítios	Genes	Número de Sítios	Genes	Número de Sítios
<i>CACNA1H</i>	16	<i>NBLA00301</i>	9	<i>PYY</i>	20	<i>CDH23</i>	5
<i>ADARB2</i>	13	<i>KIAA1543</i>	7	<i>SDK1</i>	18	<i>PRKAR1B</i>	5
<i>CDH4</i>	7	<i>SHOX2</i>	7	<i>GABBR1</i>	16	<i>HOXA7</i>	4
<i>FAT1</i>	6	<i>ADAP1</i>	5	<i>SHANK2</i>	15	<i>DLX5</i>	3
<i>HCN1</i>	6	<i>NCAN</i>	3	<i>RASA3</i>	13	<i>NCAN</i>	3
<i>MAGI2</i>	5	<i>OR2L13</i>	3	<i>AFF3</i>	12	<i>PALM</i>	3
<i>NRG1</i>	5	<i>KHDRBS2</i>	2	<i>PRK CZ</i>	8	<i>CDH4</i>	2
<i>ADCY2</i>	4	<i>OR2L8</i>	2	<i>DOCK1</i>	7	<i>FAM20C</i>	2
<i>GRIN1</i>	4	<i>SYNJ2</i>	2	<i>HOXA7</i>	7	<i>GRIN1</i>	2
<i>KIAA1543</i>	4	<i>CDH4</i>	1	<i>PDE10A</i>	7	<i>ANO2</i>	1

5.2.6 Análise de Enriquecimento

A análise de enriquecimento foi realizada individualmente para cada subgrupo de MB (Figura 5). Foram utilizados os genes DEs e DMs com pelo menos um sítio CpG com correlação significativa compreendendo 224 genes para o subgrupo WNT, 90 genes para SHH, 139 genes para o G3 e 84 para o G4 (Apêndice). Os processos mais encontrados como neurogênese, desenvolvimento neuronal e diferenciação neuronal são esperados em tumores do sistema nervoso central (SNC). O subgrupo WNT apresentou diversos processos regulatórios envolvidos no desenvolvimento e diferenciação celular e também diferenciação neuronal. Rotas tradicionais em câncer, como a própria rota metabólica de WNT e a rota de sinalização de MAPK mostraram-se enriquecidas (Sinkala M et al., 2021). O subgrupo SHH apresentou duas rotas envolvidas no metabolismo do citocromo P450, uma grande superfamília de proteínas conservadas encontradas na membrana do retículo endoplasmático que catalisam o metabolismo oxidativo de muitas drogas e produtos químicos estruturalmente diversos como colesterol, testosterona, progesterona, ácido retinóico, vitamina D3 e ácido araquidônico (Chang GW et al., 1999). Em câncer renal, a subfamília do citocromo P450 CYP3A potencialmente causa resistência a múltiplas drogas anticâncer (Murray et al., 1999). Em carcinoma de células, o *CYP1B1* é significativamente desregulado e promove a progressão desse câncer (Mitsui et al., 2015). Já o *CYP2E1* contribui para a geração de espécies reativas de oxigênio em células de câncer de mama regulando a autofagia, estimulando o estresse do retículo endoplasmático e suprimindo o potencial metastático das células deste tumor (Leung et al., 2013). Sua sobre expressão pode ser benéfica para o paciente com câncer (Vaclavikova et al., 2007). Pouco se sabe sobre o metabolismo e os papéis das proteínas da citocromo P450 em MB. Recentemente, foi demonstrado que em cultura de células de MB sob hipóxia ocorre a sub-expressão de *CYP2B6*, *CYP3A4* e *CYP3A5*, que por sua vez leva ao aumento de resistência a drogas como ciclofosfamida e ifosfamida impedindo sua conversão para suas formas ativas e, portanto, à diminuição da citotoxicidade (Valencia-Cervantes J et al., 2019). Também verificamos que a gene *ALDH1A3* encontra-se hipometilado e sobre expresso no grupo SHH e o quarto gene com maior número de sondas com correlação negativa (Tabela 5, SHH). Essa proteína faz parte da via do metabolismo de P450 e é um marcador de células-tronco cancerígenas em neoplasias, incluindo glioblastoma, entretanto o seu papel em meduloblastoma ainda não foi explorado. Em glioblastoma, a expressão de *ALDH1A3* foi

associada com uma menor sobrevida global e foi sugerido a sua utilização como marcador prognóstico e um possível alvo para terapias futuras deste tumor (Gan C et al., 2020). Recentemente, demonstrou-se que tanto a inibição farmacológica quanto genética de *ALDH1A3* inibiu a formação de metástase cerebral do câncer de mama em modelo animal e aumentou a sobrevida dos camundongos com tumor, sendo identificado como um alvo terapêutico viável em células iniciadoras de metástase em câncer de mama (Yamashita D et al., 2020).

O G3 é conhecido por ter uma assinatura de expressão retinal, com o oncogene *OTX2* comumente sobre expresso, um fator de transcrição responsável por mediar a embriogênese no SNC, regulando a formação do parênquima cerebral, glândula pineal, cerebelo e olho (Bunt J et al., 2012). Processos biológicos como morfogênese dos olhos, detecção de estímulos luminosos e percepção sensorial de luz foram apresentados para esse subgrupo. O G4 também apresentou diversas rotas metabólicas já relacionadas ao câncer, como as vias de sinalização de insulina, WNT, cálcio e ErbB (Rajaram P et al., 2017; Zhang W., 2000; Vigneri R et al., 2016; Northcott et al., 2011).

WNT

MSigDB GO Biological Processes	P-valor
Neuron differentiation	2,71e-013
Neurogenesis	2,33e-012
Neuron development	2,60e-011
Regulation of neuron differentiation	1,49e-007
Regulation of nervous system development	3,40e-007
Behavior	4,38e-007
Regulation of cell differentiation	5,75e-007
Regulation of cell development	4,61e-006
Animal organ morphogenesis	4,85e-006
Cell cell signaling	4,85e-006

MSigDB Curated KEGG	P-valor
Axon guidance	2,73e-007
Pathways in cancer	3,92e-005
Wnt signaling pathway	1,03e-004
Melanogenesis	1,11e-004
Dilated cardiomyopathy	7,66e-004
Gap junction	7,66e-004
Basal cell carcinoma	1,61e-003
Mapk signaling pathway	1,64e-003
Cardiac muscle contraction	3,69e-003
O glycan biosynthesis	3,69e-003

G3

MSigDB GO Biological Processes	P-valor
Neurogenesis	1,46e-007
Sensory perception of light stimulus	1,46e-007
Neuron differentiation	3,25e-007
Detection of light stimulus	5,32e-007
Neuron development	5,97e-007
Protein localization to synapse	1,23e-005
Nervous system process	2,22e-005
Detection of abiotic stimulus	5,95e-005
Eye morphogenesis	1,05e-004
Regulation of membrane potential	2,14e-004

MSigDB Curated KEGG	P-valor
Mapk signaling pathway	8,68e-003
Neuroactive ligand receptor interaction	8,68e-003
Taste transduction	8,68e-003
Type ii diabetes mellitus	8,68e-003
Calcium signaling pathway	1,26e-002
Focal adhesion	1,63e-002
Axon guidance	3,00e-002
Fructose and mannose metabolism	3,00e-002
Prion diseases	3,00e-002
Tight junction	3,00e-002

SHH

MSigDB GO Biological Processes	P-valor
Neurogenesis	1,12e-002

MSigDB Curated KEGG	P-valor
PPAR signaling pathway	9,01e-003
Gap junction	9,88e-003
Cell adhesion molecules cams	2,05e-002
Basal cell carcinoma	3,38e-002
Drug metabolism cytochrome p450	3,38e-002
Glycolysis gluconeogenesis	3,38e-002
Hedgehog signaling pathway	3,38e-002
Long term depression	3,38e-002
Metabolism of xenobiotics by cytochrome p450	3,38e-002
Neuroactive ligand receptor interaction	4,60e-002

G4

MSigDB GO Biological Processes	P-valor
Cell part morphogenesis	5,10e-005
Neurogenesis	5,10e-005
Neuron differentiation	5,10e-005
Axon development	3,61e-004
Cell morphogenesis involved in differentiation	5,08e-004
Neuron development	5,97e-004
Cell projection organization	6,15e-004
Cellular component morphogenesis	6,15e-004
Cell morphogenesis involved in neuron differentiation	6,31e-004
Regulation of cell differentiation	6,31e-004

MSigDB Curated KEGG	P-valor
Insulin signaling pathway	1,83e-003
Wnt signaling pathway	1,83e-003
Long term potentiation	2,75e-003
ErbB signaling pathway	3,93e-003
Neuroactive ligand receptor interaction	7,11e-003
Axon guidance	7,60e-003
Tight junction	7,60e-003
Pathways in cancer	8,62e-003
Calcium signaling pathway	1,27e-002
Type ii diabetes mellitus	1,27e-002

Figura 5. Representação dos termos do Gene Ontology (GO) e das rotas metabólicas do Kyoto Encyclopedia of Genes and Genomes (KEGG) enriquecidas dos genes DMs DEs com no mínimo um sítio CpG com correlação significativa. A análise foi conduzida na plataforma GeneTrail 3.0, e os 10 maiores processos e rotas com seus p-valores estão reportados.

6 CONCLUSÃO E PERSPECTIVAS

Com este estudo, foi possível identificar o impacto da metilação na regulação da expressão gênica de MB e como ela está estruturada nas diferentes regiões gênicas e nos subgrupos de MB. O entendimento do impacto da metilação na expressão gênica permitiu identificar vias e genes importantes para cada um dos subgrupos de MB.

Como complemento aos resultados encontrados aqui, pretendemos avaliar os dados clínicos em conjunto com os dados de metilação e expressão, através da avaliação do impacto dos genes DM e DE na sobrevida dos pacientes. Outras análises focadas nos padrões de metilação e expressão de oncogenes e genes supressores tumorais serão desenvolvidas com intuito de identificar genes marcadores e genes alvos para o desenvolvimento de novas terapias.

7 BIBLIOGRAFIA

ARYEE MJ, JAFFE AE, CORRADA-BRAVO H, LADD-ACOSTA C, FEINBERG AP, HANSEN KD, IRIZARRY RA (2014). “Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA Methylation microarrays.” **Bioinformatics**, 30(10), 1363–1369.

BENJAMINI Y AND HOCHBERG Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of the Royal Statistical Society Series**. 1995; B 57, 289–300.

BOLSTAD BM, IRIZARRY RA, ASTRAND M, AND SPEED TP. A Comparison of Normalization Methods for High Density Oligonucleotide Array Data Based on Bias and Variance. **Bioinformatics**. 2003; 19(2):185-193.

BROWN TP, GANAPATHY V. Lactate/GPR81 signaling and proton motive force in cancer: Role in angiogenesis, immune escape, nutrition, and Warburg phenomenon. **Pharmacol Ther**. 2020 Feb;206:107451.

BUNT J, HASSELT NE, ZWIJNENBURG DA, HAMDI M, KOSTER J, VERSTEEG R, KOOL M. OTX2 directly activates cell cycle genes and inhibits differentiation in medulloblastoma cells. **Int J Cancer**. 2012 Jul 15;131(2):E21-32.

Cancer Facts & Figures 2018. American Cancer Society. Disponível em: <<https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2018.html>>.

CARVALHO, B.S. AND IRIZARRY, R.A. A framework for oligonucleotide microarray preprocessing. **Bioinformatics**, vol. 26, no. 19, p. 2363–2367, 2010.

CAVALLI, F. M. G. ET AL. Intertumoral Heterogeneity within Medulloblastoma Subgroups. **Cancer Cell** 31 (2017) 737–754.

CHANG GW, KAM PC. The physiological and pharmacological roles of cytochrome P450 isoenzymes. **Anaesthesia**. 1999 Jan;54(1):42-50.

CHO YJ, TSHERNIAK A, TAMAYO P, SANTAGATA S, LIGON A, GREULICH H, ET AL. Integrative genomic analysis of medulloblastoma identifies a molecular subgroup that drives poor clinical outcome. **J Clin Oncol**. 2011;29:1424–1430.

CRAWFORD JR, MACDONALD TJ AND PACKER RJ (2007) Medulloblastoma in childhood: new biological advances. **Lancet Neurol** 6, 1073–1085.

DUBUC AM, ET AL. Aberrant patterns of H3K4 and H3K27 histone lysine methylation occur across subgroups in medulloblastoma. **Acta Neuropathol**. 2013;125(3):373–384.

EDGAR, R. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. **Nucleic Acids Research**, vol. 30, no. 1, p. 207–210, 2002.

- ESTELLER M. Epigenetics in cancer. **N Engl J Med.** 2008;358(11):1148–1159.
- FORTIN, JP., HANSEN, K.D. Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. **Genome Biol** 16, 180 (2015).
- GAJJAR A, FINLAY JL. The management of children and adolescents with medulloblastoma in low and middle income countries. **Pediatr Blood Cancer.** 2015;62(4):549–50.
- GAN C, PIERSCIANEK D, EL HINDY N, ET AL. The predominant expression of cancer stem cell marker ALDH1A3 in tumor infiltrative area is associated with shorter overall survival of human glioblastoma. **BMC Cancer.** 2020;20(1):672. Published 2020 Jul 17.
- Global Burden of Disease Cancer Collaboration.** Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the Global Burden of Disease Study. **JAMA Oncol** 2017; 3: 524–48.
- GREEN DM, KUN LE, MATTHAY KK, ET AL. Relevance of historical therapeutic approaches to the contemporary treatment of pediatric solid tumors. **Pediatr Blood Cancer.** 2013;60:1083-1094.
- HOVESTADT V, JONES DT, PICELLI S, WANG W, KOOL M, ET AL. Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. **Nature.** 2014 Jun 26;510(7506):537-41.
- HUETHER R, DONG L, CHEN X, ET AL. The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. **Nat Commun.** 2014;5:3630
- Ikeda H, Lethé B, Lehmann F, van Baren N, Baurain JF, de Smet C, Chambost H, Vitale M, Moretta A, Boon T, Coulie PG. Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. **Immunity.** 1997 Feb;6(2):199-208.
- Incid. Minist.** Disponível em: <<http://www1.inca.gov.br/wcm/incidencia/2017/>>.
- Incidência, mortalidade e morbidade hospitalar por câncer em crianças, adolescentes e adultos jovens no Brasil: informações dos registros de câncer e do sistema de mortalidade. [s.l.]: INCA, 2016.
- IRIZARRY RA, HOBBS B, COLLIN F, BEAZER-BARCLAY YD, ANTONELLIS KJ, SCHERF U, SPEED TP. Exploration, Normalization, and Summaries of High Density Oligonucleotide Array Probe Level Data. **Biostatistics.** 2003; Vol. 4, Number 2: 249-264.
- JOHNG D, TORGA G, EWING CM, JIN K, NORRIS JD, MCDONNELL DP, ET AL. HOXB13 interaction with MEIS1 modifies proliferation and gene expression in prostate cancer. **Prostate.** 2019;79:414–24.
- KALLAY L, KESKIN H, ROSS A, EL AL. Modulating native GABAA receptors in medulloblastoma with positive allosteric benzodiazepine-derivatives induces cell death. **J Neurooncol.** 2019 May;142(3):411-422.
- KANEHISA, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. **Nucleic Acids Research**, vol. 28, no. 1, p. 27–30, 2000.

- KAUFFMANN A, GENTLEMAN R, HUBER W. arrayQualityMetrics--a bioconductor package for quality assessment of microarray data. **Bioinformatics**. 2009;25(3):415-416.
- KLUTSTEIN M, NEJMAN D, GREENFIELD R, CEDAR H. DNA Methylation in Cancer and Aging. **Cancer Research**. 2016 Jun 15;76(12):3446-50.
- KOOL M, JONES DT, JAGER N, ET AL. Genome sequencing of SHH medulloblastoma predicts genotype-related response to smoothed inhibition. **Cancer Cell**. 2014 Mar 17;25(3):393-405.
- KORSHUNOV A, LICHTER P, TAYLOR MD, PFISTER SM. Medulloblastomics: the end of the beginning. **Nat Rev Cancer**. 2012 Dec;12(12):818-34.
- KULIS, M. ET AL. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. **Nat. Genet.** 47, 746 (2015).
- KUZAN-FISCHER, C. M.; STUCKLIN, A. S. G. E TAYLOR, M. D. Advances in genomics explain medulloblastoma behavior at the bedside. **Neurosurgery** 64 (2017) 21–26.
- LA MADRID, A. M. E KIERAN, M. W. Epigenetics in Clinical Management of Children and Adolescents with Brain Tumors. **Current Cancer Drug Targets** 18 (2017).
- LEE, S., HWANG, K. S., LEE, H. J., KIM, J.-S. & KANG, G. H. Aberrant CpG island hypermethylation of multiple genes in colorectal neoplasia. **Lab. Invest.** 84, 884 (2004)
- LI M, BALCH C, MONTGOMERY JS, JEONG M, CHUNG JH, YAN P, HUANG TH, KIM S, NEPHEW KP. Integrated analysis of DNA methylation and gene expression reveals specific signaling pathways associated with platinum resistance in ovarian cancer. **BMC Med Genomics**. 2009;2:34.
- LICCHESI, J. D., WESTRA, W. H., HOOKER, C. M. & HERMAN, J. G. Promoter hypermethylation of hallmark cancer genes in atypical adenomatous hyperplasia of the lung. **Clin. Cancer Res.** 14, 2570–2578 (2008).
- LITTLE, J., 1999. Epidemiology of Childhood Cancer. IARC Scientific Publications 149. Lyon: International Agency for Research on Cancer.
- LEUNG T, RAJENDRAN R, SINGH S, ET AL. Cytochrome P450 2E1 (CYP2E1) regulates the response to oxidative stress and migration of breast cancer cells. **Breast Cancer Res.** 2013;15:R107.
- MARGOLIN, A. A. ET AL. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. **BMC Bioinformatics** 7 (2006) S7.
- MCNEIL, D.E.; COTE, T.R.; CLEGG, L. E RORKE, L.B. Incidence and trends in pediatric malignancies medulloblastoma/primitive neuroectodermal tumor: A seer update. Surveillance epidemiology and end results. **Med. Pediatr. Oncol.** 39 (2002) 190–194.
- MITSUI Y, CHANG I, FUKUHARA S, ET AL. CYP1B1 promotes tumorigenesis via altered

expression of CDC20 and DAPK1 genes in renal cell carcinoma. **BMC Cancer**. 2015;15:942.

MOCK A, GEISENBERGER C, ORLIK C, ET AL. LOC283731 promoter hypermethylation prognosticates survival after radiochemotherapy in IDH1 wild-type glioblastoma patients. **Int J Cancer**. 2016 Jul 15;139(2):424-32.

MOORE LD, LE T, FAN G. DNA methylation and its basic function. **Neuropsychopharmacology**. 2013 Jan;38(1):23-38.

MURRAY GI, MCFADYEN MC, MITCHELL RT, ET AL. Cytochrome P450 CYP3A in human renal cell cancer. **Br J Cancer**. 1999;79:1836-42.

NEEDLE MN, ET AL. Adjuvant chemotherapy for the treatment of intracranial ependymoma of childhood. **Cancer**. 1997;80(2): 341-7.

NERI F, RAPELLI S, KREPELOVA A, INCARNATO D, PARLATO C, BASILE G, MALDOTTI M, ANSELMINI F, OLIVIERO S. Intragenic DNA methylation prevents spurious transcription initiation. **Nature**. 2017 Mar 2;543(7643):72-77.

NICO GERSTNER, TIM KEHL, ET AL. GeneTrail 3: advanced high-throughput enrichment analysis, **Nucleic Acids Research**, Volume 48, Issue W1, 02 July 2020, Pages W515-W520

NOONE AM, HOWLADER N, KRAPCHO M, ET AL, eds. SEER cancer statistics review, 1975-2015. April 2018. **National Cancer Institute**. https://seer.cancer.gov/csr/1975_2015/ (accessed July 10, 2018).

NORTHCOTT PA, JONES DT, KOOL M, ROBINSON GW, GILBERTSON RJ, CHO YJ, POMEROY SL, KORSHUNOV A, LICHTER P, TAYLOR MD, PFISTER SM. Medulloblastomics: the end of the beginning. **Nat Rev Cancer**. 2012 Dec;12(12):818-34.

NORTHCOTT, P. A. ET AL. Medulloblastoma comprises four distinct molecular variants. **J Clin. Oncol**. 29 (2011) 1408-1414.

NORTHCOTT, P. A. ET AL. Subgroup-specific structural variation across 1,000 medulloblastoma genomes. **Nature**. 2012 Aug 2;488(7409):49-56.

ORLANDO D, MIELE E, DE ANGELIS B, ET AL. Adoptive Immunotherapy Using PRAME-Specific T Cells in Medulloblastoma. **Cancer Res**. 2018 Jun 15;78(12):3337-3349.

PACKER RJ AND VEZINA G (2008) Management of and prognosis with medulloblastoma: therapy at a crossroads. **Arch Neurol** **65**, 1419-1424.

PACKER RJ, ET AL. Phase III study of craniospinal radiation therapy followed by adjuvant chemotherapy for newly diagnosed average-risk medulloblastoma. **J Clin Oncol**. 2006;24(25):4202-8.

PACKER RJ, ZHOU T, HOLMES E, VEZINA G AND GAJJAR A (2013) Survival and secondary tumors in children with medulloblastoma receiving radiotherapy and adjuvant chemotherapy: results of Children's Oncology Group trial A9961. **Neuro Oncol** **15**, 97-103.

PFISTER S, ET AL. Array-based profiling of reference-independent methylation status (aPRIMES) identifies frequent promoter methylation and consecutive downregulation of ZIC2

in pediatric medulloblastoma. **Nucleic Acids Res.** 2007;35(7):e51.

POLKINGHORN, W. R. E. TARSELL, N. J. Medulloblastoma: Tumorigenesis, current clinical paradigm, and efforts to improve risk stratification. **Nature Clin. Pract. Oncol.** 4 (2007) 295–304.

RAFAEL A, IRIZARRY, BENJAMIN M. BOLSTAD, FRANCOIS COLLIN, LESLIE MC, BRIDGET HOBBS AND TERENCE PS. Summaries of Affymetrix GeneChip probe level data. **Nucleic Acids Research.** 2003; 31(4):e15;

RAJARAM P, CHANDRA P, TICKU S, PALLAVI BK, RUDRESH KB, MANSABDAR P. Epidermal growth factor receptor: Role in human cancer. **Indian J Dent Res.** 2017 Nov-Dec;28(6):687-694.

RAMASWAMY V AND TAYLOR MD (2017) Medulloblastoma: from myth to molecular. **J Clin Oncol** 35, 2355–2363.

RAWAL ZD, UPADHYAY VA, PATEL DD, TRIVEDI TI. Medulloblastoma under Siege: Genetic and Molecular Dissection Concerning Recent Advances in Therapeutic Strategies. **J Pediatr Neurosci.** 2020 Jul-Sep;15(3):175-182.

RITCHIE ME, PHIPSON B, WU D, HU Y, LAW CW, SHI W, SMYTH GK (2015). “limma powers differential expression analyses for RNA-sequencing and microarray studies.” **Nucleic Acids Research**, 43(7), e47.

ROLAND CL, ARUMUGAM T, DENG D, LIU SH, PHILIP B, GOMEZ S, BURNS WR, RAMACHANDRAN V, WANG H, CRUZ-MONSERRATE Z, LOGSDON CD. Cell surface lactate receptor GPR81 is crucial for cancer cell survival. **Cancer Res.** 2014 Sep 15;74(18):5301-10.

SCHWALBE EC, ET AL. DNA methylation profiling of medulloblastoma allows robust subclassification and improved outcome prediction using formalin-fixed biopsies. **Acta Neuropathol.** 2013;125(3):359–371.

SHAHI MH, ET AL. Human hedgehog interacting protein expression and promoter methylation in medulloblastoma cell lines and primary tumor samples. **J Neuro-Oncol.** 2011;103(2):287–96.

SHAHI MH, ET AL. REGULATION of sonic hedgehog-GLI1 downstream target genes PTCH1, Cyclin D2, Plakoglobin, PAX6 and NKX2.2 and their epigenetic status in medulloblastoma and astrocytoma. **BMC Cancer.** 2010;10:614.

SHARMA S, KELLY TK, JONES PA. Epigenetics in cancer. **Carcinogenesis.** 2010;31(1):27–36.

SHIH DJ, NORTHCOTT PA, REMKE M, ET AL. Cytogenetic prognostication within medulloblastoma subgroups. **J Clin Oncol.** 2014;32(9):886-896.

SHUANGSHOTI S, TADADONTIP P, TECHAVICHIT P, THORNER PS, SHUANGSHOTI S, TEERAPAKPINYO C. Simplified Molecular Subtyping of Medulloblastoma for Reduced Cost and Improved Turnaround Time. **Appl Immunohistochem Mol Morphol.** 2020 Aug;28(7):538-543.

SINKALA M, NKHOMA P, MULDER N, MARTIN DP. Integrated molecular characterisation of the MAPK pathways in human cancers reveals pharmacologically vulnerable mutations and gene dependencies. **Commun Biol.** 2021;4(1):9. Published 2021 Jan 4.

SMITH MA, SEIBEL NL, ALTEKRUSE SF, ET AL. Outcomes for children and adolescents with cancer: challenges for the twenty-first century. **J Clin Oncol.** 2010;28:2625-2634.

SMITH, Z. D. & MEISSNER, A. DNA methylation: roles in mammalian development. **Nat. Rev. Genet.** 14, 204 (2013)

STRUYK AF, CANOLL PD, WOLFGANG MJ, ROSEN CL, D'EUSTACHIO P, SALZER JL. Cloning of neurotrimin defines a new subfamily of differentially expressed neural cell adhesion molecules. **J Neurosci.** 1995 Mar;15(3 Pt 2):2141-56.

SUBRAMANIAN A, TAMAYO P, MOOTHA VK, MUKHERJEE S, EBERT BL, GILLETTE MA, PAULOVICH A, POMEROY SL, GOLUB TR, LANDER ES, MESIROV JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. **Proc Natl Acad Sci U S A.** 2005 Oct 25;102(43):15545-50.

TAYLOR, M.D. ET AL. Molecular subgroups of medulloblastoma: the current consensus. **Acta Neuropathol.** 2012 Apr;123(4):465-72.

THAKUR, C. ET AL. Loss of mdig expression enhances DNA and histone methylation and metastasis of aggressive breast cancer. **Signal Transduct. Target. Ther.** 3, 25 (2018)

VACLAVIKOVA R, HUBACKOVA M, STRIBRNA-SARMANOVA J, ET AL. RNA expression of cytochrome P450 in breast cancer patients. **Anticancer Res.** 2007;27:4443-50.

VALENCIA-CERVANTES J, HUERTA-YEPEZ S, AQUINO-JARQUÍN G, ET AL. Hypoxia increases chemoresistance in human medulloblastoma DAOY cells via hypoxia-inducible factor 1 α -mediated downregulation of the CYP2B6, CYP3A4 and CYP3A5 enzymes and inhibition of cell proliferation. **Oncol Rep.** 2019;41(1):178-190.

VIGNERI R, GOLDFINE ID, FRITTITTA L. Insulin, insulin receptors, and cancer. **J Endocrinol Invest.** 2016 Dec;39(12):1365-1376.

WADELIN F, FULTON J, MCEWAN PA, SPRIGGS KA, EMSLEY J, HEERY DM. Leucine-rich repeat protein PRAME: expression, potential functions and clinical implications for leukaemia. **Mol Cancer.** 2010 Aug 27;9:226.

WATT F, MOLLOY PL. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. **Genes Dev.** 1988 Sep;2(9):1136-43.

WHITLOCK NC, TROSTEL SY, WILKINSON S, ET AL. MEIS1 down-regulation by MYC mediates prostate cancer development through elevated HOXB13 expression and AR activity. **Oncogene.** 2020;39(34):5663-5674.

WILLIAMS KA, LEE M, HU Y, ANDREAS J, PATEL SJ, ZHANG S, CHINES P, ELKAHLOUN A, CHANDRASEKHARAPPA S, GUTKIND JS, MOLINOLO AA,

CRAWFORD NP. A systems genetics approach identifies CXCL14, ITGAX, and LPCAT2 as novel aggressive prostate cancer susceptibility genes. **PLoS Genet.** 2014 Nov 20;10(11):e1004809.

WILLIAMS KA, LEE M, WINTER JM, GILDEA DE, CALAGUA C, CURRY NL, LICHTENBERG J, YE H, CRAWFORD NPS. Prostate cancer susceptibility gene HIST1H1A is a modulator of androgen receptor signaling and epithelial to mesenchymal transition. **Oncotarget.** 2018 Jun 19;9(47):28532-28546.

XU W, XU M, WANG L, ET AL. Integrative analysis of DNA methylation and gene expression identified cervical cancer-specific diagnostic biomarkers. **Signal Transduct Target Ther.** 2019;4:55. Published 2019 Dec 13.

WU ZZ, LU HP, CHAO CC. Identification and functional analysis of genes which confer resistance to cisplatin in tumor cells. **Biochem Pharmacol.** 2010 Jul 15;80(2):262-76.

YAMASHITA D, MINATA M, IBRAHIM AN, YAMAGUCHI S, COVIELLO V, BERNSTOCK JD, HARADA S, CERIONE RA, TANNOUS BA, LA MOTTA C, NAKANO I. Identification of ALDH1A3 as a Viable Therapeutic Target in Breast Cancer Metastasis-Initiating Cells. **Mol Cancer Ther.** 2020 May;19(5):1134-1147.

YANG X, HAN H, DE CARVALHO DD, LAY FD, JONES PA, LIANG G. Gene body methylation can alter gene expression and is a therapeutic target in cancer. **Cancer Cell.** 2014;26(4):577-590.

ZHANG, W. ET AL. Tamoxifen-induced enhancement of calcium signaling in glioma and MCF-7 breast cancer cells. **Cancer Res.** 60, 5395–5400 (2000).

ZHAO, C., ZOU, H., ZHANG, J., WANG, J., & LIU, H. (2018). An integrated methylation and gene expression microarray analysis reveals significant prognostic biomarkers in oral squamous cell carcinoma. **Oncology reports**, 40(5), 2637–2647.

ZHOU R, MAN Y. Integrated analysis of DNA methylation profiles and gene expression profiles to identify genes associated with pilocytic astrocytomas. **Mol Med Rep.** 2016;13(4):3491-3497.

8 APÊNDICE

Lista de genes diferencialmente metilados e expressos no subgrupo WNT.

NTM	C20orf103	CNPY1	STX1B	NLGN1	CDK6	RAB11FIP2	MTSS1	CNTN4
MGC45800	C8orf84	CROT	SYK	NPAS3	DSE	SCRN1	MYC	DLG2
ONECUT2	CACNA2D1	CRYZ	TGFA	NRG1	EBF3	SERPINF1	MYO7A	DNAH7
DSCR4	CCDC88C	DDR2	TLL1	NTNG1	EPHB6	SFN	NRN1	DPP6
DSCR8	CNR1	DHRS3	TMEM132C	OLFM2	GFPT2	SPOCK2	NTRK1	EPHB1
GNG3	CORO2B	DKK4	TMEM229B	PAX5	GUCY1B3	TBR1	PAPSS1	FAT1
HIST1H4D	CREB3L2	EBF2	TMEM74	PCSK2	HK2	THBS4	PARP12	FBN3
CDKL2	DDR1	EPHA3	TSPAN13	PDE9A	HS3ST3B1	TNFRSF19	PART1	FOXP2
KHDRBS2	DLX3	ERN1	TSPAN9	PEX5L	LHX8	WIF1	PELI2	GNG7
AKR1E2	EMID1	FGF20	TUBB2B	PSD3	LMO3	ZNF154	PLCD4	GRIN1
CNGA3	FBXL7	FREM2	TYW3	SEMA3B	LRIT1	ZNF238	PLEKHF2	GUCY1B3
COL6A3	GABRG3	FZD10	ZNF503	SEMA5A	NET1	ABCA1	PLXDC1	GULP1
FMOD	GAD1	GALNT14	ADARB2	SFMBT2	PDIA5	ACPP	PRAME	HCN1
TCF7	GPR98	GALNTL2	ADCY2	SHANK2	RANBP17	ADAM19	PRELID2	HOXC6
BAMBI	IRX3	GLT8D2	APBB2	SLC22A15	RD3	ADAMTSL1	PRRX1	HOXD10
CASZ1	KIAA1199	GNAI1	ARHGEF4	SORBS2	SIX3	ADHFE1	RASSF3	JPH3
EBF1	LOXL1	HAND2	ARID5B	SYT3	TMEM51	ANK3	RASSF8	KCNK2
FAM174B	LPPR5	IGSF9	BCL11A	SYT5	ZFH3	ASB13	ROBO2	KHDRBS2
LGR5	LRIG3	IMPG2	CACNA1H	TANC2	ZMAT4	ASCL4	ROR2	KIAA1543
MCTP2	MYOT	KCNJ5	CACNA2D1	TCP10L	ABHD12B	ATP1B1	RRAGD	LPHN3
NKD2	NHLH2	KCNK17	CACNG7	TH	ADCY1	C6orf191	RUNX2	LUZP2
ALDH5A1	NRXN2	LAMB1	CDH4	TMTC2	ALK	CALHM2	SH3BP5	MAGI2
C3orf64	OSBPL3	MPP4	CDH8	UNC5D	AMHR2	CBLN1	SHOX2	MED12L
CACNA2D4	PLCH1	MTMR7	CNPY1	XKR4	AXIN2	CERKL	ST3GAL1	MYRIP
C1orf61	CLYBL	STC2	NEGR1	NDST3	BMP4	CHRM4	STC1	

Lista de genes diferencialmente metilados e expressos no subgrupo SHH.

A4GALT	LRP1B	FAM169A	SYNE1	CBLN1	RALYL	HOXA9	KHDRBS2
ACSL6	LSAMP	FEZ1	TLL1	CDKL2	RERG	HSD11B2	KIAA1543
ALDH1A3	MICAL1	FGF5	TMEM130	CLDN1	SCD5	ID2	NBLA00301
ALDOC	MMP11	GAB1	TMEM163	CNTN1	SCN2A	INSM2	NCAN
ANO4	NTM	GLI2	TRH	CNTN6	SLC14A1	ISLR2	OR2L13
ARHGAP26	PCDH9	GPR81	TSPAN13	CPNE8	SLC1A2	KCNJ3	OR2L8
BHLHE41	PCNT	GRM3	TSPYL5	CYYR1	SLC1A4	KIAA0182	
C20orf108	PDE5A	GSTM5	TTC9	DPP6	SLC4A10	KIF26A	
C6orf123	POU6F2	GUCY1B3	ZNF154	DRD2	SPHK1	L3MBTL4	
C8orf84	PRKG1	GULP1	ADAP1	EYA1	SPOCK2	LGI1	
CACNA1E	PRLR	HHIP	CDH4	FABP7	SPRY2	DOCK3	
CADM2	QPCT	HIST1H4F	SYNJ2	SLC1A6	SHOX2	PTPRE	

Lista de genes diferencialmente metilados e expressos no subgrupo G3.

MEIS1	RREB1	C2orf71	GRID1	B3GALNT2	ATP8A2	KCNJ2	PPP2R2C
PRAME	SBK2	CBLN1	GRIN1	C6orf191	C19orf77	KIAA1486	PRKCA
SGIP1	STX3	CERKL	HOXA7	CACNA2D4	CACNA1A	MTMR7	PRKCZ
ABCA4	TAGLN3	CLIP4	IQCA1	CPLX2	CACNA2D4	NCAM2	PSTPIP2
HIST1H4F	TDRD9	EPHA2	IRX6	DAND5	CAP2	NELL2	PTPRT
RUNX1T1	TMEM176A	EPHA7	KCNH6	GFPT2	CBLN2	NEUROG1	PYY
C1orf61	TMEM176B	EYA2	LONRF2	LPPR4	CHD5	NR2E3	RASA3
CDKL2	TSHZ3	FAM19A4	MEG3	NTN5	CLIP2	NRL	RYR3
ELAVL4	UNC5C	FBXL21	NCAN	RCVRN	CNPY1	OAT	SDK1
GNB3	VTN	GABRA5	NKAIN3	SAG	CNTFR	OTX2	SHANK2
TBR1	ZIC5	GABRR2	NOTCH1	SERP2	CPNE9	PDE6H	SHISA6
TULP1	ZNF679	GALNT14	NRXN2	SLC1A7	CSMD2	PELI2	SHOX2
BMF	ABCA4	HK2	NTM	TBC1D8	D4S234E	POTEF	SNORD113-4
C3orf64	ACCN1	IMPG2	ODZ3	AIPL1	DLL3	PTGFRN	SYT7
GUCY2D	AFF3	JAKMIP2	PDE10A	ARL4D	DOCK1	PURG	TRIL
MPP4	APC2	KCNA5	PDE4DIP	ARL6	DSCAM	RASGRF2	UBASH3B
PCBP4	ATCAY	KCNB2	POTEF	BCHE	FLNC	RD3	WNT7A
C12orf60	GABBR1	RERG					

Lista de genes diferencialmente metilados e expressos no subgrupo G4.

NR2F1	MAPK10	RREB1	CDH23	ZFPM2	SCD5	FABP7	NOTCH1
C1orf61	MGAT4C	SLC6A3	CDH4	ANK3	SLC22A15	GRM8	PALM
DSCR8	MPP4	SPRY2	D4S234E	C12orf5	SOX9	IMPG2	PAX5
FZD7	MUDENG	CAMK4	DLX5	CDH23	SRGAP1	JAKMIP2	PPP2R2C
KCNK1	NR2F2	CBLN1	EPHA4	DDX31	TBC1D8	KALRN	PRKAR1B
LHX2	PCDHB7	CERKL	FAM20C	FAM18A	TP53BP2	KCNA2	PRKCB
LMX1A	PTGFRN	CNR1	GRIN1	GABRR2	UNC5D	KCNB2	PRKCZ
RASSF8	RERG	CXXC4	HOXA7	KCNA5	ZNF521	KCNJ2	STAT5A
TBR1	RPH3A	DSEL	NCAN	NUDT4	ZNF679	KIAA1486	WWC1