Maria Prates Rivas

Mecanismos epigenéticos em hepatoblastomas: análise do transcriptoma e sua regulação por metilação de DNA.

Epigenetic mechanisms in hepatoblastomas: transcriptome analysis and regulation by DNA methylation.

São Paulo

2021

Maria Prates Rivas

Mecanismos epigenéticos em hepatoblastomas: análise do transcriptoma e sua regulação por metilação de DNA.

Epigenetic mechanisms in hepatoblastomas: transcriptome analysis and regulation by DNA methylation.

Tese apresentada ao Instituto de Biociências da Universidade de São Paulo, para a obtenção de Título de Doutora em Ciências, na Área de Biologia (Genética).

São Paulo

2021

Ficha Catalográfica

Prates Rivas, Maria.

Mecanismos epigenéticos em hepatoblastomas: análise do transcriptoma e sua regulação por metilação de DNA / Maria Prates Rivas ; orientadora Ana Cristina Victorino Krepischi – São Paulo, 2021. 171p.

Tese (Doutorado) -- Instituto de Biociências da Universidade de São Paulo. Programa de Pós-Graduação em Genética e Biologia Evolutiva.

1. Hepatoblastoma 2. Epigenética 3. Metilação

I. Victorino Krepischi, Ana Cristina, orient. II. Título.

Comissão Julgadora:

Prof. Dr. Eduardo M. R. Reis	Dr(a). Cecília M. L. da Costa
Prof. Dr. Oswaldo K. Okamoto	Prof(a). Dr(a). Ana C. V. Krepischi Orientadora

Dedicatória

A todos que acompanham e apoiam

à ciência básica.

Agradecimentos

Chegando ao fim dos 5 anos de doutorado, certamente tenho muitas pessoas a agradecer. Gostaria de iniciar pela minha orientadora, Dra. Ana Cristina Krepischi, a qual me recebeu de portas abertas como sua primeira aluna vinculada a Universidade de São Paulo. Obrigada por dividir e embarcar comigo nesta jornada da epigenética, cultura celular, ensaios moleculares e muitas colaborações. Não conseguimos realizar todos os planos do projeto, mas certamente tivemos bastante jogo de cintura e sintonia, construindo uma parceria exitosa, com muito conhecimento publicado e repartido com a comunidade. O seu apoio e pensamento à frente foram propulsores para as realizações deste trabalho. Agradeço também a todos os membros do laboratório de Genética Humana, tanto o corpo técnico como os alunos de iniciação científica, mestrado, doutorado e pós-doutorado, tão importantes nas discussões e compartilhamentos. Em especial, agradeço a minha primeira parceira do laboratório Dra. Talita Aguiar, que me apresentou os "caminhos" do Hepatoblastoma. Desde a obtenção das amostras e apresentações aos médicos até discussões de resultados e planos de trabalho, nossa troca culminou em diversos artigos conjuntos. De fato, ter alguém para compartilhar as dores e delícias do dia a dia na pesquisa básica fez muito diferença para este trabalho. Ainda na USP, agradeço a colaboração de todos os membros do CEPID-Centro de Pesquisas sobre o Genoma Humano e células-tronco, em especial Dr. Luiz Caires, Dr. Ernesto Goulart, Kayque Silva e Dra. Mayana Zatz, pelo constante interesse e disponibilidade em contribuir com o avanço científico, assim como os colegas do Departamento de Genética e Biologia Evolutiva, pelos encontros de café da tarde os quais renderam desabafos, risadas e boas colaborações. Ainda no âmbito científico, agradeço as agências de fomento CAPES, CNPq e em especial a FAPESP (Projeto número: 2016/23462-8), as quais foram imprescindíveis para realização deste trabalho e de tão importante apoio a ciência Brasileira, assim como todos os colaboradores da rede de hospitais pediátricos envolvidos neste projeto.

Por fim, agradeço aos meus queridos pais pelo apoio imensurável e diário nesta jornada acadêmica, sempre acreditando e permitindo que este desejo do Doutorado se torne uma realidade. Também aos meus irmãos e cunhados, os quais estimulam e compartilham o meu caminho. Aos amigos Paulistas, pela construção de uma família e rede de apoio, assim como os amigos Baianos que mesmo de longe, se mantêm tão perto. A Deus, Iemanjá e Pacha Mama, por abrir os meus caminhos.

Índice

Capítulo I. Introdução Geral

I.1. Tumores pediátricos

Em países desenvolvidos, o óbito por neoplasia é a maior causa de morte não violenta na infância[1] e o mesmo já ocorre no Brasil. De acordo com o Instituto Nacional do Câncer (INCA), a estimativa para novos casos de câncer em crianças e adolescentes foi de 8.460em 2020[2]. Segundo o Ministério da Saúde, as neoplasias ocuparam a segunda posição (8%) de óbitos de crianças e adolescentes em 2014, ultrapassadas somente pelos óbitos por causas externas, assim se configurando como a doença que mais mata nessa faixa etária[3].

Apesar de apresentarem alta taxa de proliferação e invasão, os tumores pediátricos respondem bem a tratamentos, apresentando bom prognóstico e sobrevida global média de 80%. No entanto, a despeito de grandes avanços no tratamento, uma em cada cinco crianças com câncer morre em decorrência da doença e, dentre as sobreviventes, três em cinco irão sofrer efeitos colaterais graves do tratamento[1].

É importante ressaltar que tumores pediátricos são intrinsecamente diferentes de tumores que se desenvolvem em adultos e, de maneira geral, os tipos de cânceres mais frequentes em crianças são diferentes daqueles de adultos. Em crianças, são mais frequentes as neoplasias hematológicas e de sistema nervoso central, enquanto em adultos, carcinomas são predominantes[4]. Adicionalmente, tumores pediátricos não são fortemente associados a fatores de risco ambientais e, portanto, não há prevenção efetiva. Devido ao curto período de desenvolvimento ou menor período de latência de tumores pediátricos em comparação a tumores em adultos, o acúmulo de mutações é menor, em adição à menor carga de fatores ambientais em sua gênese[4–6]. Entretanto, a diferença primordial é que o câncer em crianças está vinculado ao período de crescimento do organismo; portanto, o câncer na infância é indissociável dos processos de diferenciação celular e organogênese.

Uma classe especial de tumor pediátrico é a de tumores embrionários, definida por idade muito precoce de acometimento - em geral diagnosticados antes dos 3 anos de idade - e histologia característica, que se assemelha a estágios do desenvolvimento do órgão de origem[7]. Tumores embrionários constituem um grupo de neoplasias intrinsecamente associadas ao processo do desenvolvimento embrionário, que envolve a interação entre diversos fatores genéticos e mecanismos epigenéticos[5]. Dentre os tumores embrionários sólidos estão meduloblastoma, neuroblastoma, retinoblastoma, tumor de Wilms e

hepatoblastoma[8]. Poucos fatores ambientais foram associados de maneira significativa à gênese de tumores embrionários e, dessa forma, o modelo preponderante de origem pressupõe a ruptura do processo normal do programa de desenvolvimento embrionário do tecido/órgão de origem, que se daria no período fetal e/ou perinatal[7] (**Figura 1**). Dessa forma, em geral estes tumores apresentam características de células não totalmente diferenciadas e alguns podem mimetizar a atividade de células embrionárias; como exemplo, produção de α-fetoproteína nos hepatoblastomas, uma proteína normalmente produzida apenas pelo fígado fetal[8].

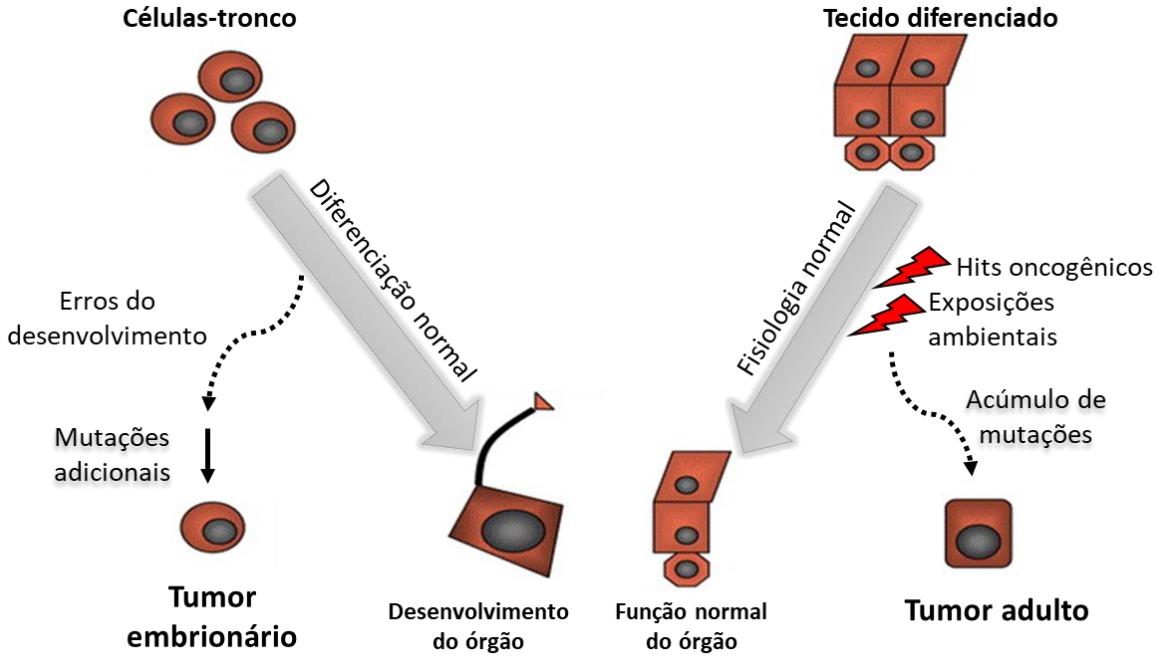


Figura 1. Tumores embrionários: ruptura do processo normal de diferenciação celular. Auto-renovação, proliferação e migração celular são aspectos comuns à organogênese e à tumorigênese. Nos tumores de adultos, há um acúmulo de mutações ao longo do tempo, o que não é necessariamente verdadeiro para tumores pediátricos. A classe dos tumores embrionários, que apresenta idade muito precoce de acometimento, tem como principal hipótese de origem a ruptura do processo normal do programa de desenvolvimento embrionário (Adaptado de: *Pahlman et al.*[9]).

I.2. Desenvolvimento embrionário do fígado

O fígado é o maior órgão interno e realiza funções imprescindíveis como metabolismo de compostos, regulação do nível de glicose pelo glicogênio, controle da homeostase sanguínea, detoxificação e produção de bile[10]. Hepatócitos compõem 70% da massa do

fígado adulto e são derivados da endoderme embrionária. A endoderme é um dos três folhetos embrionários, juntamente com a ectoderme e mesoderme, originando o intestino primitivo, o qual é subdividido em três segmentos: intestino anterior, médio e posterior[11]. O segmento do intestino anterior dará origem a pâncreas, vesícula biliar, pulmão e fígado, neste último pelos hepatoblastos. Já as células estromais, células de kuppfer, estreladas, e vasos sanguíneos possuem origem mesodérmica[12].

Estudos em modelos celulares e animais auxiliaram no entendimento do desenvolvimento do fígado, demonstrando que a hepatogênese é evolutivamente conservada e ocorre por meio de uma série de interações recíprocas entre a endoderme embrionária e a mesoderme próxima[11,12]. O primeiro sinal morfológico do fígado embrionário é a formação do divertículo hepático, cuja porção anterior origina o fígado e a via biliar intra-hepática, enquanto a porção posterior forma a vesícula biliar e os ductos biliares extra-hepáticos. Posteriormente, as células endodérmicas hepáticas, conhecidas como hepatoblastos, delaminam (desprendem-se) do epitélio e invadem o mesênquima do septo transverso adjacente para formar o broto do fígado, juntamente com fibroblastos e células estreladas[10] . Uma vez formado o broto do fígado, é iniciado um processo de grande crescimento, vascularização e colonização por células hematopoiéticas, se tornando o principal órgão hematopoiético fetal. Este crescimento é mediado por sinalizações do mesênquima hepático, assim como por expressão gênica intrínseca a hepatoblastos[13,14].

Posteriormente, ocorre a diferenciação dos hepatoblastos em colangiócitos ou hepatócitos. Neste processo, a localização dos hepatoblastos é preponderante para determinação da célula a ser diferenciada. Do ponto de vista molecular, os hepatoblastos expressam marcadores de fígado fetal, como a alfa-feto proteína (*AFP*), marcadores de hepatócito maduro como albumina (*ALB*), e marcadores de colangiócitos como citoqueratina-19 (*CK19*), o que confere seu caráter bipotencial[10]. Os hepatoblastos em contato com a veia porta são expostos a gradientes de sinalização TGF-β, WNT/β-catenina e fator de crescimento epidérmico (EGF), que induzem a diferenciação destas células em precursores de colangiócitos, assim como regulam negativamente a expressão de genes hepáticos, gerando ao final a diferenciação em colangiócitos maduros[15–17]. Em contrapartida, os hepatoblastos localizados no parênquima hepático não sofrem ação das vias sinalizadoras descritas acima e se diferenciam gradualmente em hepatócitos maduros[10]. Para isto, os hepatoblastos no parênquima são expostos a citocinas secretadas por células hematopoiéticas, como a oncostatina M que, em combinação com fator de crescimento de

hepatócito (HGF) e da via WNT, direcionam a diferenciação de hepatócitos imaturos, por meio da regulação positiva de fatores de transcrição associados à maturação hepática (HNF4α, C/EBPα) e posterior diferenciação em hepatócito maduro[17,18]. É importante ressaltar que a maturação dos hepatócitos funcionais e a formação de uma rede biliar são processos graduais, que prosseguem após o nascimento para a formação completa do órgão e desenvolvimentos de suas funções. A **Figura 2** resume a ação de vias de sinalização e marcadores envolvidos na diferenciação dos hepatoblastos.

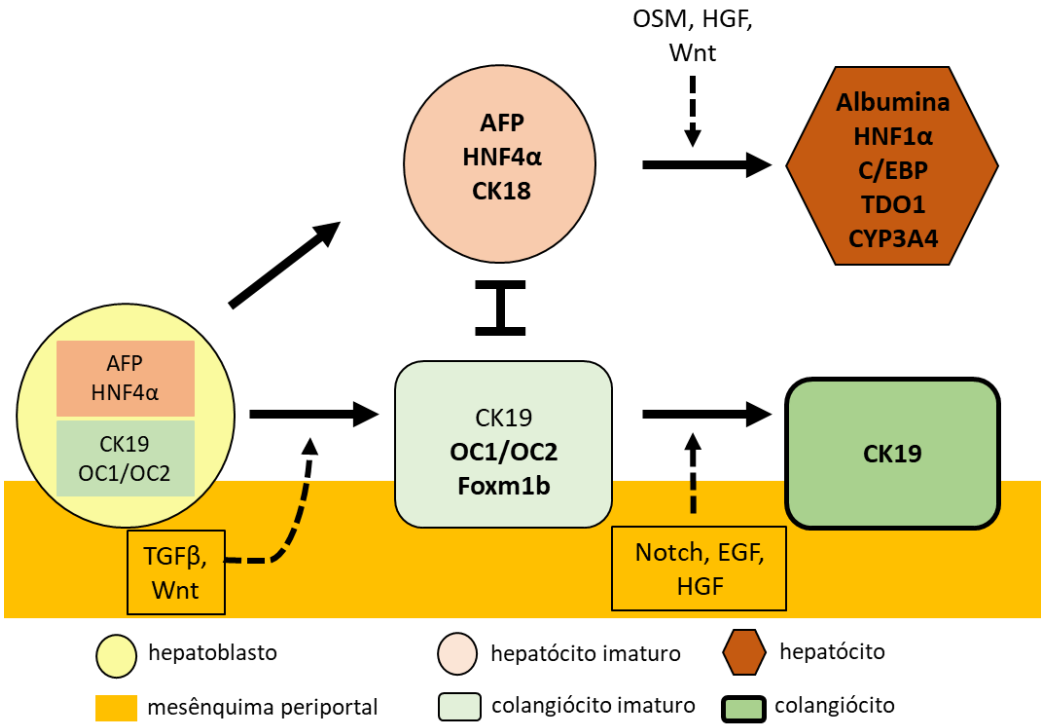


Figura 2: Diferenciação de hepatoblastos em linhagens de hepatócitos e colangiócitos. Os hepatoblastos são bipotenciais e expressam genes hepáticos fetais (*AFP*; *HNF4α*), e marcadores de colangiócitos (*CK19*; *OC1/OC2*). Evidências sugerem que vias de sinalização (TGFβ e Wnt) do mesênquima periportal aumentam a expressão de marcadores de colangiócitos, promovendo a ação de fatores de transcrição (*CK19, OC1, OC2*) nos hepatoblastos adjacentes, enquanto ao mesmo tempo estes sinais reprimem a expressão de fatores de transcrição hepatogênica (HNF4α e C/EBP). Em contraste, os hepatoblastos no parênquima, que não sofrem a atuação da sinalização do mesênquima periportal, apresentam expressão de fatores hepatogênicos (*AFP*; *HNF4α, CK18*). Neste processo, o antagonismo mútuo entre os dois grupos de fatores de transcrição reforça essa segregação das linhagens. Por fim, a sinalização pelas vias Notch, EGF e HGF do mesênquima periportal são essenciais para a diferenciação dos colangiócitos maduros, enquanto outros fatores (OSM, HGF e Wnt) promovem a maturação dos hepatócitos, com expressão de marcadores característicos como albumina e CYP3A4. OSM: Oncostatina M (Figura adaptada de Zorn, 2008[10]).

Durante o desenvolvimento embrionário, mecanismos epigenéticos são reguladores cruciais da diferenciação celular. Desta forma, no início do desenvolvimento, as células-tronco embrionárias, caracterizadas pela totipotência, alta capacidade de proliferação e auto-renovação, apresentam em geral nível global de metilação de DNA (DNAm) reduzido, o que permite a expressão de genes associados à pluripotência, como *NANOG*, *POU5F1* e *SOX2*[19]. No decorrer do desenvolvimento, os genes responsáveis por manter a pluripotência, bem como aqueles que não serão expressos no tipo celular determinado, são reprimidos, parte como resultado de um aumento da DNAm, ocorrendo concomitantemente aumento da expressão de genes específicos para o desenvolvimento daquele tipo celular, como *AFP* e *CK19* em hepatoblastos. Em consequência, as células-tronco iniciam o processo de diferenciação celular, que culmina no acúmulo progressivo de diversas marcas epigenéticas, incluindo DNAm, que resultarão num padrão específico de identidade celular de expressão gênica, característico para cada tipo de célula[20]. Desta forma, no decorrer do desenvolvimento de um órgão, quanto mais diferenciada a célula, maior será o nível global de DNAm.

I.3. Hepatoblastoma

Tumores hepáticos em crianças são raros, representando 1-4% de todos os tumores sólidos pediátricos[8]. Massas hepáticas primárias constituem o terceiro grupo mais comum de tumores abdominais sólidos da infância[21], com incidência de 0,4 a 1,9 por milhão de crianças a cada ano[22]. Ao contrário dos tumores de fígado em adultos, nos quais o carcinoma hepatocelular é predominante, em crianças o câncer hepático prevalente é o hepatoblastoma[22,23].

O diagnóstico de hepatoblastomas ocorre em média aos 18 meses de vida, sendo que a maioria dos casos surge até os quatro anos de idade[24]. A principal hipótese assume que tais tumores teriam origem em células precursoras de hepatócitos, os hepatoblastos, que teriam sofrido um bloqueio do processo normal de diferenciação celular[22,23].

Assim, hepatoblastomas provavelmente se desenvolvem a partir de hepatoblastos e são geralmente compostos por combinações de tipos diferentes de células epiteliais, mesenquimais, indiferenciadas ou outros tipos histológicos[23,25]. Quanto à classificação histopatológica, o padrão mais comum é o tipo embrionário, no qual o tumor se assemelha ao fígado entre 6-8 semanas de gestação (**Figura 3A**). Em 20 a 30% dos casos, o

hepatoblastoma também contêm células derivadas do estroma, como tecido cartilaginoso, esquelético, músculo, entre outros, sendo designado de hepatoblastoma "misto". Quando há uma mistura heterogênea de componentes do tumor, como derivados de ectoderme, neuroectoderme ou células produtoras de melanina, este tumor é classificado como hepatoblastoma "teratóide"[23,26]. Dentre os subtipos desta neoplasia destacam-se o "hepatoblastoma fetal bem diferenciado", caracterizado por células uniformes, variando de 10 a 20 μ de diâmetro e com baixa atividade mitótica; as células deste padrão se assemelham a hepatócitos fetais e estes tumores geralmente contêm células precursoras hematopoiéticas, componentes epiteliais, ou uma mistura destes com grupos de células mesenquimais (**Figura 3B**). O "hepatoblastoma fetal com alta atividade mitótica", constituído por células com núcleo proeminente e alta taxa de mitose, em geral é composto com outras áreas tumorais formadas por células de padrão fetal bem diferenciado (**Figura 3C**). O "hepatoblastoma de componente pleomórfico" é pouco comum, sendo derivado de tumores primários ou de metástases pós-quimioterapia; este subtipo, quando assume um crescimento macrotubular, é de difícil diferenciação do carcinoma hepatocelular (**Figura 3D**). Já o "hepatoblastoma de crescimento macrotubular" possui baixa incidência, representando menos de 5% dos casos, e possui padrão de crescimento caracterizado por placas de células epiteliais (**Figura 3E**) Em um subconjunto de hepatoblastomas, algumas células neoplásicas se diferenciam em colangiócitos formando pequenos ductos. Estes componentes expressam marcadores de linhagem de colangiócitos (citoqueratinas 7 e 19) e pode estar situado dentro ou em torno do componente hepatocelular do tumor[23,26] (**Figura 3F**).

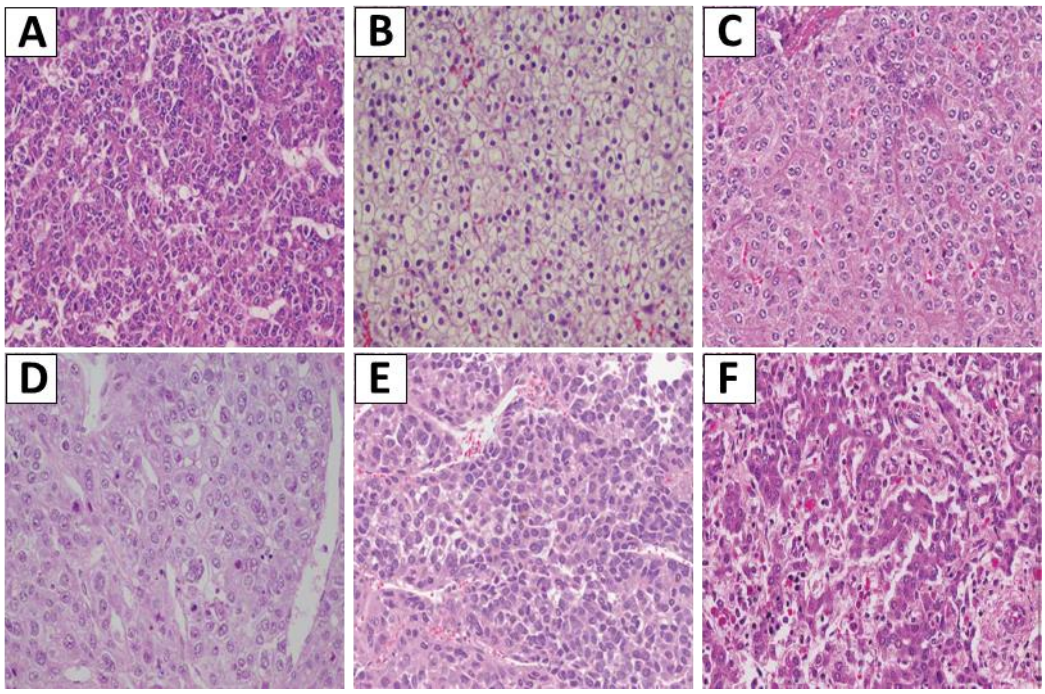


Figura 3: Fotomicrografias de diferentes subtipos histológicos de hepatoblastomas. (a) embrionário; **(b)** fetal bem diferenciado; **(c)** fetal com alta atividade mitótica; **(d)** componente pleomórfico; **(e)** macrotubular; **(f)** colangioblástico (Adaptado de: LOPÉZ-TERRADA, D., *et al.*, 2014[23])

I.3.1 Estratificação de risco do hepatoblastoma

No início de 1970, surgiram evidências de que o hepatoblastoma era um tumor responsivo a quimioterapia; contudo, com os regimes terapêuticos utilizados no período, a sobrevida global dos pacientes apresentava taxas baixas, de 20% a 30% [27]. O sistema de segmentação hepática, descrito por Couinaud, é baseado na identificação das três veias hepáticas e no estabelecimento de um plano imaginário transverso passando pela bifurcação da veia porta, sendo considerado um marco para os procedimentos cirúrgicos hepáticos. Atualmente, este modelo ainda é amplamente utilizado, servindo como base para a criação de sistemas de estratificação de risco para tumores de fígado[28,29].

Desenvolvido pela SIOPEL (sigla para Sociedade Internacional de Oncologia Pediátrica), o sistema PRETEXT (*PRE-Treatment EXTent of disease*) é o modelo principal de estadiamento e estratificação de risco para o diagnóstico do hepatoblastoma[30]. A fim de facilitar o diálogo entre grupos de diferentes instituições, o PRETEXT é usado para descrever a extensão e localização do tumor antes de intervenções terapêuticas[31]. Em resumo, este sistema divide o fígado virtualmente em quatro setores e, considerando a quantidade de setores acometidas pelo tumor, proximidade entre eles, envolvimento de veias hepáticas,

assim como acometimento de metástases, o tumor é classificado em quatro categorias (**Figura 4**).

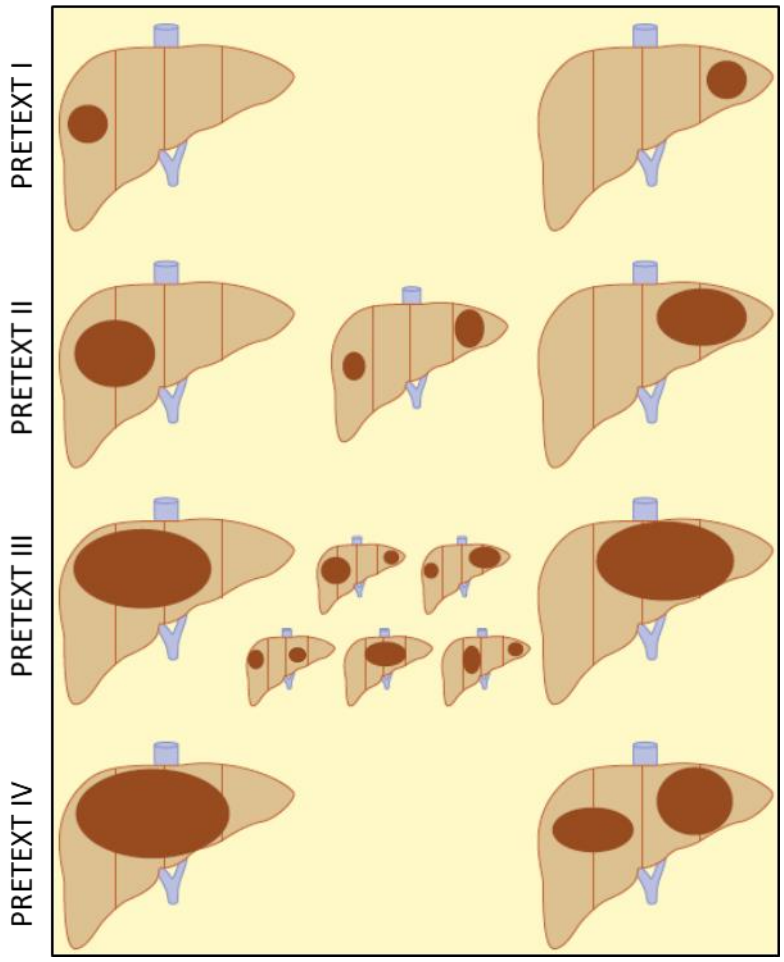


Figura 4. Classificação do hepatoblastoma pelo sistema PRETEXT. Neste modelo, o fígado é dividido em quatro setores, sendo o número PRETEXT derivado a partir da subtração do número de seções contíguas de fígado que não estavam envolvidas pelo tumor. Desta forma, o PRETEXT I indica que um setor está envolvido e três setores adjacentes estão livres da doença; PRETEXT II indica que um ou dois setores estão envolvidos, porém dois setores adjacentes estão livres da doença; PRETEXT III indica que dois ou três setores estão envolvidos e não há dois setores adjacentes livres; PRETEXT IV indica que todos os quatro setores estão comprometidos (figura adaptada de SIDEBOTHAM, 2019[32]).

Em 2016 foram publicados resultados da colaboração entre os maiores centros de tratamento para hepatoblastoma (COG, SIOPEL, GPOH e JPLT), reunindo dados clínicos de pacientes com hepatoblastomas a fim de estabelecer um critério internacional para estratificação e tratamento deste tipo tumoral, denominado CHIC (*The Children's Hepatic Tumors International Collaboration*)[33]. Esta análise foi baseada em 1.605 casos de hepatoblastoma tratados entre 1988 e 2010; os fatores prognósticos considerados foram

PRETEXT; concentração de AFP igual ou inferior a 100 ng/mL, entre 101-1000 ng/mL e acima de 1000 ng/mL; faixas etárias diferentes (idade menor que 3 anos, idade entre 3-7 anos e idade superior a 8 anos); os fatores de anotação PRETEXT, como doença metastática (M), envolvimento macrovascular de todas as veias hepáticas (V) ou bifurcação da veia porta (P), tumor extra-hepático (E), tumor multifocal (F) e ruptura espontânea (R). Para o PRETEXT I e II foi avaliada a possibilidade de ressecabilidade cirúrgica ao momento do diagnóstico; aspectos como PRETEXT IV, positividade dos fatores de anotação do PRETEXT (incluindo metástase), concentração de AFP ≤ 100 ng/mL ao diagnóstico e idade superior a 8 anos de idade foram definidos como clinicamente relevantes. De acordo com todos estes critérios, foram estabelecidos quatro grupos de estratificação de risco: muito baixo, baixo, intermediário e alto, os quais são apresentados no fluxograma da **Figura 5**. Estes avanços nos atuais protocolos, combinando quimioterapia adjuvante e abordagens cirúrgicas, resultaram em sobrevida livre de eventos em três anos de 80% [34].

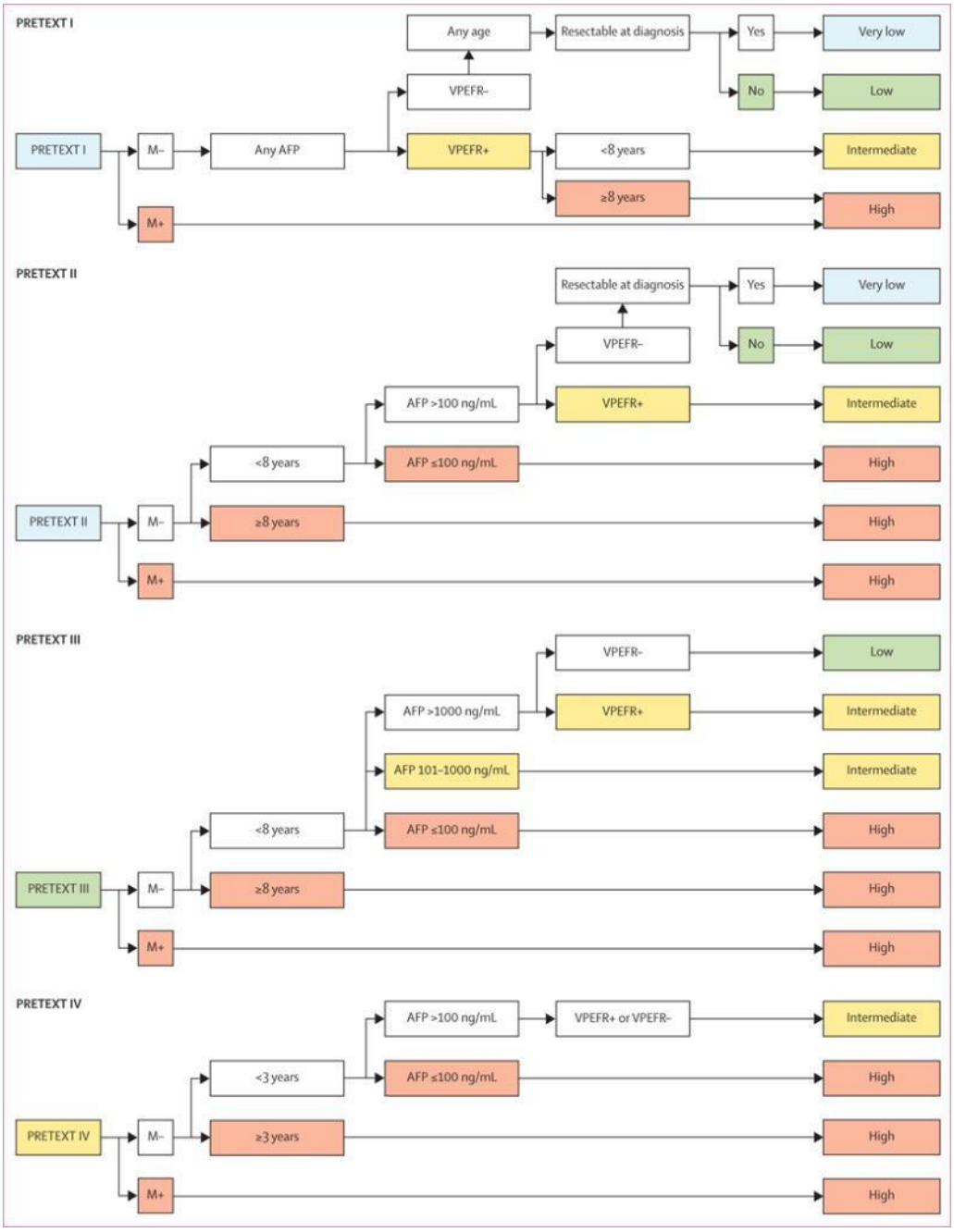


Figura 5. Fluxograma da estratificação de risco para tumores hepáticos infantis, segundo critérios CHIC. A categoria PRETEXT é utilizada para iniciar a estratificação, em seguida a presença de metástase (M+ e M-); para PRETEXT I é avaliado a dosagem de AFP e, para os demais grupos, a idade do paciente. Posteriormente, são considerados fatores adicionais (presentes ou ausentes) como envolvimento macrovascular de todas as veias hepáticas (V) ou bifurcação da veia porta (P), tumor extra-hepático (E), tumor multifocal (F) e ruptura espontânea (R) (MEYERS et al., 2017[35]).

I.3.2 Alterações genômicas descritas em hepatoblastomas

O câncer é uma doença causada por inúmeras alterações em diferentes níveis de complexidade, principalmente mutações genéticas. Alguns desses fatores podem já estar presentes de maneira constitutiva nos indivíduos (predisposição genética), enquanto na maioria dos casos tais alterações são adquiridas somaticamente no decorrer da vida[4]. A combinação de algumas das alterações envolvidas na tumorigênese fornece vantagem competitiva, aumentando a proliferação celular e aumento da massa tumoral e, em estágios mais avançados, promovendo a capacidade de invasão de tecidos vizinhos e causando metástases[36].

No processo de desenvolvimento tumoral, a heterogeneidade genotípica e fenotípica é grande e raramente se estabiliza por completo[37]. Cada subtipo tumoral apresenta ao mesmo tempo um padrão de anomalias que é comum ao grupo (assinatura tumoral) e um grande conjunto de alterações individuais (privadas), variáveis entre tumores do mesmo grupo[38,39]. As alterações podem estar causalmente relacionadas ao início da proliferação neoplásica e/ou à sua progressão (*driver mutation*), ou ainda terem sido adquiridas como consequência do processo tumoral (*passenger mutation*), sendo que apenas uma minoria destas últimas deve ser primordial para a oncogênese.

Tumores embrionários apresentam um forte componente de predisposição genética, evidenciado não apenas pela manifestação muito precoce, mas também pela ocorrência em associação a diversas síndromes genéticas[40]. Os hepatoblastomas não fogem à regra e, embora a maioria ocorra como casos isolados e seja esporádica, aproximadamente 15% dos casos estão associados à presença de uma mutação germinativa que eleva substancialmente o risco de desenvolvimento. Dentre elas destacam-se a polipose adenomatosa familial e a síndrome de Beckwith-Wiedemann[41,42], causadas pela presença de mutações germinativas no gene *APC* e na região de imprinting genômico 11p15, respectivamente.

Hepatoblastomas apresentam genomas relativamente estáveis, com poucas alterações genéticas detectáveis. Recente análise pan-cancer demonstrou que hepatoblastomas são os tumores com menor taxa de mutação somática dentre os tumores sólidos, apresentando de 1-7 mutações por genoma tumoral[43]. As alterações cromossômicas mais comumente descritas neste tumor são aneuploidias, em especial ganhos dos cromossomos 2, 8 e 20, bem como perdas dos cromossomos 4 e 18 [42,44], embora dados de nosso grupo e de revisão de literatura indicam que a perda do cromossomo 18 não seria de fato tão frequente[45]. Poucos rearranjos

focais foram identificados, em especial uma amplificação em 2q24, que nosso grupo restringiu a um segmento de 10Mb contendo cinco genes com aumento de número de cópias e expressão exacerbada, que seriam oncogenes em potencial para este tipo tumoral, como *DAPL1* [44].

O gene da β-catenina (*CTNNB1*) é afetado de maneira recorrente por mutações ativadoras em grande parte dos hepatoblastomas (aproximadamente 70%), enquanto outros genes da via Wnt, como *AXIN1*, *AXIN2*, *AXIN3* e *APC*, também podem estar mutados, porém em frequência muito menor[41,42,46]. A via de sinalização Wnt é central no desenvolvimento embrionário e pós-natal do fígado e sua ativação é descrita em neoplasias hepáticas em geral[42,47], o que indica sua associação com a gênese do câncer de fígado tanto em crianças como em adultos. Mais recentemente, estudo de exomas de alguns hepatoblastomas demonstrou que o gene *NFE2L2*, que codifica fator de transcrição regulador de genes associados a processos antioxidantes e inflamatórios, seria o segundo gene com maior taxa de mutação no hepatoblastoma, estando presente em aproximadamente 10% dos casos[41]. É importante notar que em casuística de pacientes brasileiros, em estudo de nosso grupo, não foram identificadas mutações em *NFE2L2* [48,49]. Mutações na região promotora no gene *TERT* também foram descritas em trabalhos envolvendo hepatoblastomas, associadas a maior agressividade tumoral [41,43,50].

Ferramentas de análise gênica demonstram que, em comparação com hepatocarcinoma celular, o hepatoblastoma apresenta taxa de mutação 10 vezes menor [51], o que sugere que o processo de tumorigênese de fígado na infância é guiado por outros mecanismos além de mutações do DNA, como alterações de expressão gênica e modificações epigenéticas.

Mudanças no padrão de expressão gênica também já foram identificadas nos hepatoblastomas, mais recentemente a partir de estudos de RNA-seq. O sequenciamento massivo de transcriptomas (RNA-seq), desenvolvido há mais de uma década, tornou-se uma ferramenta onipresente que está moldando diversos aspectos de nossa compreensão da função genômica[52]. O RNA-seq é mais frequentemente usado para analisar a expressão diferencial de genes (*diferencial gene expression - DGE*) nas amostras alvo em relação a um conjunto de controles. O fluxo de trabalho geral se inicia com extração de RNA, que pode ser realizada por várias técnicas, seguida de enriquecimento de RNAm e/ou depleção de RNA ribossomal, síntese de cDNA e preparação de bibliotecas de sequenciamento[53]. As

bibliotecas são então sequenciada buscando a obtenção de coberturas de 10-30 milhões de *reads* por amostra em uma plataforma de alto rendimento[52]. As etapas finais envolvem análises computacionais complexas: alinhamento e montagem das *reads* sequenciadas com base em um banco de dados da espécie alvo, quantificação de *reads* que se sobrepõem aos transcritos, normalização e aplicação de ferramentas de análises comparativas entre os grupos de amostras sequenciados[54].

O primeiro grupo a publicar dados de RNA-seq de hepatoblastomas foi Hooks *et al.*, realizando a classificação de amostras tumorais de acordo com a expressão gênica e dados clínicos, o que gerou a hipótese de estratificação deste tipo tumoral em 3 subgrupos utilizando o padrão de expressão de quatro genes (hidroxiesteroide 17-β desidrogenase 6, integrina α 6, topoisomerase 2-α, e vimentina) [55]. Apesar desta técnica ser amplamente utilizada em pesquisa do câncer [56], apenas dois outros trabalhos a utilizaram no estudo do hepatoblastomas [57,58], sempre com o objetivo de buscar aperfeiçoar a estratificação tumoral para permitir avanços no diagnóstico da doença. Entretanto, até o momento ainda não há uma classificação amplamente utilizada para hepatoblastomas, dado que os determinantes genéticos da heterogeneidade clínica deste câncer ainda não estão claros e os biomarcadores prognósticos sugeridos nestes estudos apresentaram limitações na reprodutibilidade de sua correlação clínica, além de não serem claramente compreendidas suas implicações biológicas [58].

I.4. Epigenética e Câncer

A epigenética engloba mecanismos moleculares pelos quais a expressão gênica é controlada e modificada, com marcas sendo transmitidas de maneira estável através de sucessivas gerações celulares, contudo sem alterar a sequência de DNA[59] e por este motivo potencialmente reversíveis. Tais mecanismos epigenéticos determinam o nível de compactação da cromatina, afetando diretamente a acessibilidade da maquinaria de transcrição e, portanto, o estado de atividade dos genes. Os principais mecanismos epigenéticos descritos são metilação e hidroximetilação de DNA; modificações covalentes na cauda N-terminal de histonas por processos de acetilação, fosforilação, sumoilação e metilação; atuação de RNAs não-codificantes[60,61]. Atualmente, sabe-se que os mecanismos epigenéticos são essenciais para a diferenciação e manutenção da identidade celular, específicos para o tipo de tecido, e vinculados a uma contingência temporal[62].

Tradicionalmente, o modelo de desenvolvimento do câncer preconiza o acúmulo de mutações genéticas ao longo de um período extenso de tempo; contudo, este paradigma foi expandido para acomodar um quadro ainda mais complexo, no qual mecanismos de regulação epigenética foram incorporados ao processo de tumorigênese[63]. Atualmente, a interação entre processos genéticos e epigenéticos é considerada a chave para o desenvolvimento tumoral, resultando em expressão anormal de vias biológicas, seja quantitativa ou qualitativamente. Se considerarmos a via estritamente genética, oncogenes e genes supressores de tumor apresentam expressão gênica anormal por meio de mutações de ganho ou perda de função, respectivamente, ao passo que mutações em genes de reparo de DNA amplificam a ocorrência de mutações somáticas adicionais e aumentam a instabilidade genômica. Por outro lado, pela via epigenética, a tumorigênese é modulada por modificações na compactação e acessibilidade da cromatina, estados definidos pelo intercâmbio de mecanismos dinâmicos como modificações de histonas, DNAm, remodelamento de nucleossomos e regulação de RNAs não-codificantes[64]. Mutações e mecanismos epigenéticos estão fortemente ligados ao câncer e, adicionalmente, mutações em genes reguladores de processos epigenéticos podem orquestrar alterações globais no epigenoma[63,64], recentemente descritas para diversas doenças humanas, incluindo câncer.

I.4.1. Metilação do DNA (DNAm)

Em 1975, dois grupos independentes sugeriram que a metilação teria papel central como marca epigenética em animais, realizando regulação da expressão gênica[65]. A DNAm é uma modificação covalente na qual um grupamento metil (CH_3) oriundo da S-adenosilmetionina (AdoMet) é transferido para o carbono na posição 5 de uma citosina, resultando em uma citosina metilada (5mC)[66] (**Figura 6**). Este processo ocorre predominantemente em citosinas que precedem uma guanina (dinucleotídeo CpG), sendo 5mC considerado o quinto nucleotídeo do DNA[67]. Os dinucleotídeos CpG ocorrem esparsos pelo genoma ou mais frequentemente agrupados em regiões denominadas ilhas CpG (*CpG Island - CGI*), definidas como regiões com mais de 200 pares de base (bp), com alto conteúdo de guanina e citosina e presença de ≥60% de CpGs[68].

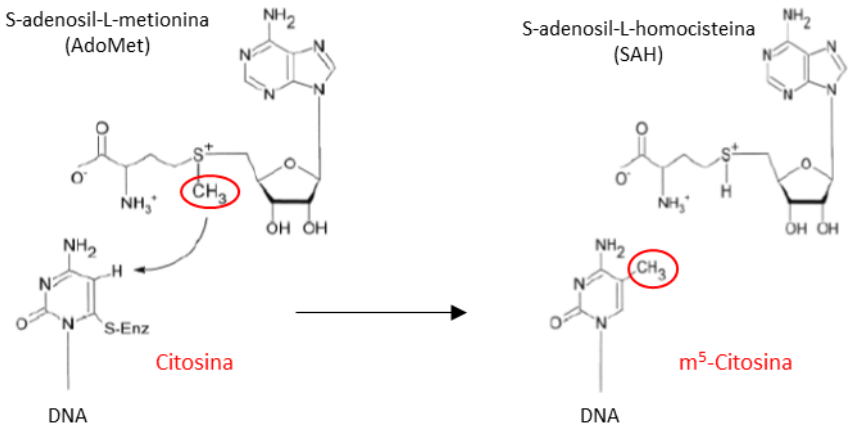


Figura 6. Metilação do DNA. Modificação covalente do DNA com a transferência do grupamento metila (CH_3) oriundo da S-adenosil-L-metionina (AdoMet) para o carbono na posição 5 da citosina do genoma, resultando em uma citosina metilada (5mC) (Adaptado de: Plewa & Jagodzinski, 2005 [68])

O processo de DNAm ocorre pela ação das enzimas DNA metiltransferases (DNMTs), que catalisam a adição do grupo metil à citosina[69]. Os membros da família de enzimas DNMT agem de duas maneiras principais: adicionando grupos metil na sequência de DNA, onde antes não havia metilação (metilação *de novo*), ou copiando a metilação da fita de DNA molde para uma nova fita complementar durante a replicação (metilação de manutenção), o que possibilita a transmissão do padrão de metilação durante as divisões celulares. As DNMTs descritas são DNMT1, DNMT2, DNMT3A, DNMT3B, DNMT3L e DNMT3C[68,70]. A DNMT1 está envolvida na metilação de fitas de DNA recém-sintetizadas durante o processo de replicação celular, sendo classificada como metilase de manutenção, em sítios CpG hemimetilados; as enzimas DNMT3A e DNMT3B estão associadas à metilação *de novo*, responsáveis pelo estabelecimento do padrão de metilação durante a embriogênese e adição do grupo metil em sítios sem a presença de metilação prévia[69]; DNMT2 possui apenas um domínio catalítico e parece desempenhar papel na metilação de tRNA[71,72]; DNMT3L está associada à regulação da atividade das metiltransferases *de novo*, atuando como co-fator ou inibidor da atividade de DNMT3A e DNMT3B [68,73]; recentemente foi descoberta uma nova DNA metiltransferase *de novo*, denominada DNMT3C. Descrita em camundongos, DNMT3C é considerada enzima responsável pela metilação de regiões promotoras de retrotransposons em linhagem germinativa masculina, sendo necessária para a fertilidade destes mamíferos[70].

Apesar de ser observada principalmente em dinucleotídeos CpG, a metilação também foi descrita em citosinas seguidas por timina (CpT), adenina (CpA), ou outra citosina (CpC), sendo denominada metilação não-CpG[74]. Assim como a metilação CpG, a metilação não-CpG é encontrada em todo o genoma, incluindo sequências repetitivas, *enhancers*, promotores e corpos gênicos[75]. A metilação não-CpG, catalisada por DNMT3A e DNMT3B, ainda pouca explorada funcionalmente, é considerada restrita a tipos celulares específicos, como células-tronco pluripotentes, neurônios, oócitos e células da glia[67,74,76].

A maioria dos dinucleotídeos CpG que ocorrem dispersos e isolados pelo genoma está metilada (70-80%), ao contrário das CGIs, que estão em sua maioria não metiladas, particularmente quando localizadas em sítios de início de transcrição/promotores de genes[75]. Embora os padrões de metilação de promotores gênicos estejam sujeitos a um processo de regulação muito estrito em mamíferos, os mecanismos que especificariam tais padrões ainda são investigados. Estudo recente indica a existência de elementos na própria sequência de DNA proximal ao gene (denominados de *small methylation-determining regions* – MDR), atuando com um papel regulatório na determinação ou especificação do perfil de metilação local[77]. Tais sequências podem mediar tanto a hipometilação do promotor gênico quanto a metilação *de novo*, atuando como um fator *in cis*.

A metilação é associada de forma geral à repressão da expressão dos genes-alvo, como exemplo, a inativação de genes do cromossomo X e também de genes expressos predominantemente em linhagens germinativas[78]. Contudo, observou-se mais recentemente que a metilação no corpo gênico pode conduzir à ativação de sua transcrição, tanto em animais como no genoma de plantas[67,79]. Destas evidências surge o paradoxo de que a metilação, quando ocorre na região promotora, está inversamente correlacionada com a expressão gênica, enquanto que no corpo do gene, a metilação está positivamente correlacionada à sua expressão, podendo até ser um marcador mais fiel do padrão de expressão daquele gene. Assim, a DNAm é um mecanismo complexo de controle da expressão gênica, que altera a interação entre proteínas e a sequência de DNA, modificando a estrutura da cromatina e, consequentemente, a acessibilidade da maquinaria transcricional, o que leva ao aumento ou diminuição da atividade gênica[68,80].

O padrão de DNAm é aberrante em células de câncer, quando comparado ao de células oriundas de tecidos normais. Este fenômeno é bastante característico e bem documentado, sendo descrito de forma sumária como hipermetilação de CGIs em promotores

de supressores tumorais, levando à repressão gênica, e hipometilação de sequências repetitivas (**Figura 7**), ocasionando instabilidade genômica. Alterações na metilação podem mimetizar o efeito de mutações em genes supressores de tumor ou proto-oncogenes. Supressores tumorais como *RB1*, *MLH1*, *CDKN2A* e *BRCA1* contêm sequências promotoras que se encontram sobrepostas a CGIs e, em diversos tipos de tumores, estas regiões se encontram extensivamente metiladas, produzindo repressão transcricional[78]. Se a hipermetilação de promotor de genes supressores de tumor conduz ao seu silenciamento transcricional[81], a hipometilação de proto-oncogenes induziria ao aumento de sua expressão[78,82]. Concomitantemente, o genoma das células tumorais perde metilação nas abundantes sequências repetitivas do genoma humano (hipometilação global do genoma), levando à reativação de retrotransposons e consequente instabilidade genômica[59,82]. Assim, o estudo de alterações do perfil de DNAm nas células neoplásicas e a sua associação com a gênese e progressão tumoral tornaram-se de grande relevância para a compreensão da biologia tumoral.

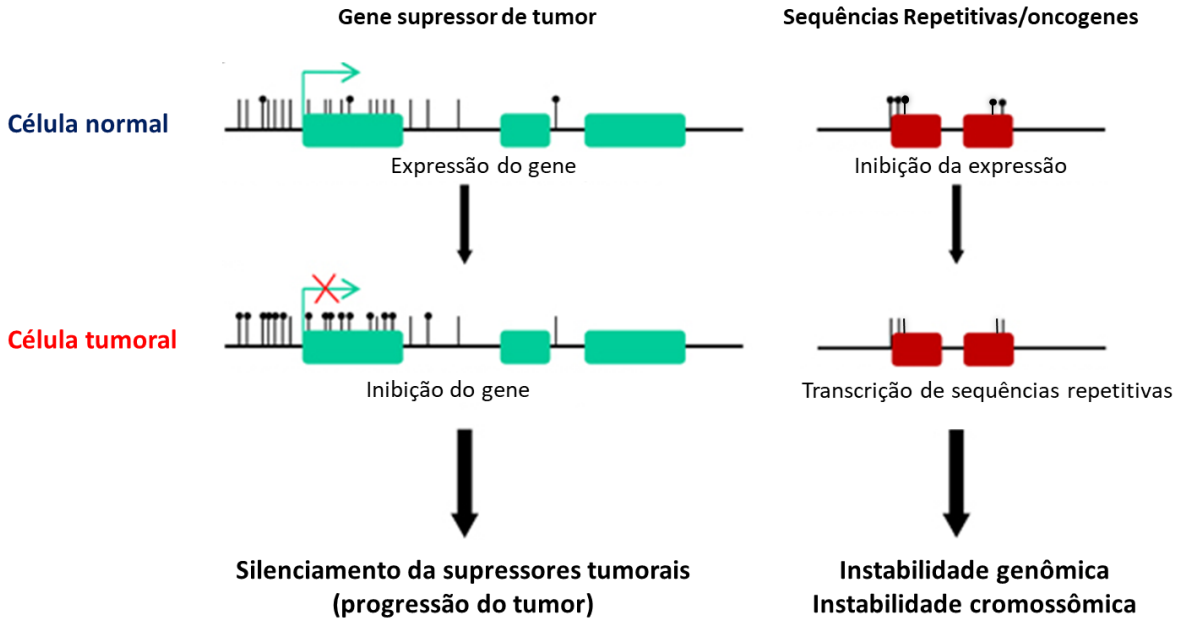


Figura 7. Comparação entre os padrões de metilação de DNA encontrados em células normais e em células tumorais e caracterização de suas consequências para a transcrição genômica. Em células normais (acima), há pouca metilação em sítios de início de transcrição e em CGIs, permitindo a transcrição gênica (à esquerda, associada à metilação de CpGs no corpo do gene, que inibe a ativação de sítios espúrios de início de transcrição; por outro lado, sequências repetitivas apresentam grande metilação de CpGs, inibindo sua expressão (à direita). No câncer (abaixo) ocorre aumento de metilação de CpGs nas CGIs das regiões gênicas promotoras e perda no corpo do gene, causando repressão transcricional e ativação de sítios de início de transcrição incorretos, respectivamente (à esquerda); enquanto em sequências repetitivas, os tumores

apresentam hipometilação, ativando sítios de início de transcrição incorretos e retrotransposons, com consequente aumento de instabilidade genômica (à direita) (Adaptado de: Lopez-Serra & Esteller, 2012[83]).

I.4.2. Hidroximetilação do DNA

A desmetilação do DNA pode ocorrer de forma passiva ou ativa. Na primeira situação, a 5mC é perdida através de consecutivas replicações do DNA. Em contrapartida, a desmetilação ativa implica a modificação da base de citosina via desaminação e/ou oxidação, esta última produzindo a 5-hidroximetilcitosina (5hmC)[84]. A descoberta da via de hidroximetilação do DNA, em que há produção de uma base chamada 5hmC, tem grande valor na epigenética, pois demonstrou que a metilação não é o único mecanismo de regulação da transcrição gênica atuante no DNA[85]. A função da 5hmC ainda não foi completamente elucidada; contudo, além de se tratar de um resíduo intermediário do processo de desmetilação ativa do DNA, estudos apontam a relação da hidroximetilação com a manutenção da pluripotência em células-tronco, a ativação da transcrição gênica e remodelamento da cromatina[86–88]. As proteínas TET são membros de uma família de DNA hidroxilases composta por três enzimas TET1, TET2 e TET3, que realizam reações de oxidação que fazem a conversão de 5mC em 5hmC. Tais proteínas compartilham um domínio catalítico, tendo como cofator α-cetoglutarato e Fe^{2+} e o domínio CXXC, encontrado em muitas proteínas que interagem com a cromatina[87].

A **Figura 8** resume os tipos de DNAm e seus principais efetores em mamíferos.

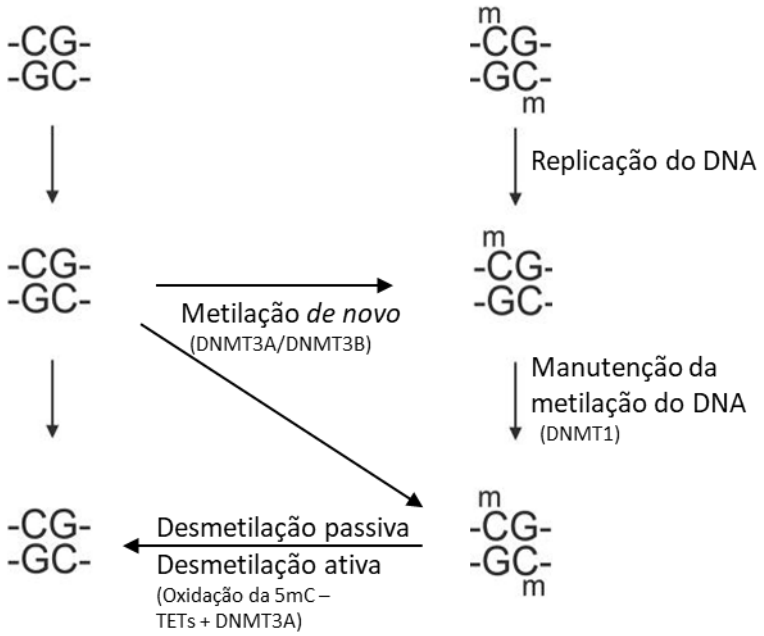


Figura 8. Mecanismos de metilação do DNA e enzimas envolvidas. A coluna do lado esquerdo representa o padrão de CpG não metilado, o qual é mantido durante a replicação do DNA. A coluna direita representa sítios CpG metilados, os quais retém a metilação durante a replicação do DNA na fita molde, enquanto a fita nova será posteriormente metilada (DNMT1), mantendo assim o padrão da metilação desta região do DNA (metilação de manutenção). O processo de metilação *de novo* ocorre pela metilação de CpGs previamente não metilados (DNMT3A/DNMT3B), ilustrado pela linha do meio. O processo de desmetilação pode ocorrer de forma passiva, havendo perda da metilação da citosina, ou de forma ativa, onde há oxidação da 5mC pela ação de TETs e associação com DNMT3A (terceira linha).

O nível de 5hmC está relacionado com o status de diferenciação celular; estudos apontam a existência de uma maior quantidade de 5mC e 5hmC em tecidos adultos, enquanto células-tronco de órgãos adultos teriam redução de ambas as modificações da citosina, de modo que em tecidos saudáveis um maior nível de 5hmC é indicação de maior diferenciação celular[89,90]. Por outro lado, já houve descrição em células-tronco embrionárias de altos níveis de 5hmC em associação à manutenção do estado de pluripotência e continuidade da expressão do gene *NANOG* [91]. Apesar de sua função ainda ser não ser profundamente conhecida, revisão mais recente sugere um quadro global no qual, em ação conjunta com fatores de transcrição e *enhancers*, a marca epigenética 5hmC estaria diretamente associada à regulação de genes importantes para a diferenciação celular, sendo que seu conteúdo e localização no genoma poderia ser considerado um marcador do subtipo celular [92]. Neste trabalho é ressaltado que, na maioria dos casos, o nível de 5hmC global parece ser progressivamente reduzido da célula progenitora para o estado diferenciado, mas permanece (ou é adquirido) em *loci* específicos de cada linhagem ou em genes relacionados à função celular adulta, funcionando como um marcador da "identidade" celular [92].

Diversos tipos celulares tumorais apresentam níveis significativamente mais baixos de 5hmC em comparação ao tecido normal circundante[91,93]. Contudo, trabalhos que modularam a expressão das proteínas TET demonstraram que o nível de 5hmC *per se* não seria suficiente para iniciar a tumorigênese, mas seria um fator aditivo, colaborando com outros mecanismos moleculares[94]. Em células tumorais, mutações em DNA metiltransferases, além das enzimas que modificam 5hmC, como as pertencentes à família TET, foram associadas à transformação oncogênica. Em linfoma de células-T, Couronné e colaboradores sugeriram a cooperação entre mutações em TET2 e DNMT3A, que envolveria a desregulação da metilação de citosinas e processos de desmetilação[95]. Mutações de TET2 também foram identificadas em mieloma e leucemia[87]. Em meduloblastoma, tumor

embrionário do sistema nervoso, foi observado a redução do nível global de hidroximetilação e associação de baixo nível de 5hmC com metástase; os níveis de expressão das *TET* também foram específicos a cada subtipo tumoral, sendo associadas à presença de eventos como recidiva e morte, indicando um possível alvo terapêutico [96]. Recente trabalho envolvendo tumores pediátricos do sistema nervoso central demonstrou que a marca 5hmC seria perdida nos tumores em um padrão *locus* específico e esses locais, por sua vez, terminam hipermetilados; além disso, a redução de 5hmC foi associada à diminuição da sobrevida e tempo livre de recorrência nesta classe tumoral [97].

I.4.3. Epigenética e tumores hepáticos

Hepatoblastomas ainda são pouco estudados, em especial por causa de sua raridade e, até recentemente, trabalhos realizados eram restritos a poucos *loci*, apontando hipermetilação também em regiões promotoras dos genes *CASP8, HHIP* e *RASSF1A*, além de *APC, SFRP1* e *SOCS1* [42,98].

No carcinoma hepatocelular, há descrição de hipermetilação de alguns promotores gênicos, tais como os genes supressores tumorais *CASP8, HHIP, RASSF1A*, entre outros; entretanto, destaca-se a identificação neste tipo tumoral de um padrão não usual de hipometilação global afetando sequências não-repetitivas [47,99]. Foi observado também aumento da expressão de genes que codificam proteínas reguladoras da maquinaria epigenética como DNMT1, DNMT3A e DNMT3B, EZH2 e HDAC1, HDAC2 e HDAC3 [99]. Adicionalmente, trabalho de 2015 associa o aumento de expressão do gene *UHFR1*, um regulador essencial da DNAm, com este padrão não usual de hipometilação global, atuando por meio de desestabilização da proteína DNMT1 [100].

Em trabalho de nosso grupo[101], o metiloma de hepatoblastomas foi investigado em comparação a amostras de fígado diferenciado. Esta análise utilizou duas coortes independentes de hepatoblastomas, examinando o status de DNAm de > 450.000 sítios CpG distribuídos por todo o genoma, bem como sequências LINE-1. Como resultado, foram encontradas 1359 regiões diferencialmente metiladas em hepatoblastomas, sendo hipermetilação observada principalmente em corpos gênicos e ilhas CpG, enquanto a hipometilação foi detectada em regiões intergênicas. Vale ressaltar que, de modo geral, foi evidenciando um padrão de hipometilação em sequências não-repetitivas, resultado similar ao descrito para carcinomas hepatocelulares. A análise pareada ainda revelou que as

diferenças de metilação afetam primordialmente genes associados a vias de câncer, diferenciação de células do fígado, metabolismo de lipídios e drogas, glicólise, dentre outras, e, deste modo, a metilação se apresenta como um mecanismo regulatório importante da expressão gênicas. Este número de sítios CpG diferencialmente metilados contrasta fortemente com a paucidade de mutações genéticas recorrentes em hepatoblastomas, reforçando que a epigenética está no eixo central da gênese deste tipo tumoral.

Recentemente, análise com 113 tumores corroborou nossos dados de hipometilação global em amostras de hepatoblastoma e hipermetilação em CpG específicos[57]. Em especial, este estudo mostrou que hepatoblastomas podem ser clusterizados em dois grupos de acordo com o grau de hipometilação global e hipermetilação em CpG específicos, denominados *Epigenetic-Cluster A* (Epi-CA) e *Epigenetic-Cluster B* (Epi-CB); ao integrar tais dados epigenômicos com análises de genoma, transcriptoma e aspectos clínicos, foi observado que os tumores com a assinatura Epi-CB apresentam maior hipometilação global, ativação da via Wnt/B-catenina, estando associados a pior prognóstico. Além disto, foi destacado a superexpressão de região de 300 kb localizada no *locus* 14q32 *DLK1/DIO3*, sendo observado nesta região a predominância de hipometilação de DNA, o que indica um papel direto da regulação da expressão dos genes desta região pelo controle da metilação[57]. O *locus* 14q32 *DLK1/DIO3* ganhou maior evidencia em recente estudo em hepatoblastomas, que corroborou grande aumento de expressão de um *cluster* de microRNAs e snoRNAs localizados nesta região, especialmente em tumores metastáticos[102]. Utilizando análises multiômicas integradas, Sekiguchi *et al*. prôpos a subdivisão do hepatoblastomas em três grupos (F, E1 e E2) [58]. A análise de *cluster* baseada no perfil de metilação dos tumores demonstrou que o grupo F tem padrão de expressão mais semelhante ao de fígado adulto, com expressão aumentada de genes associados a fígado diferenciado (*ALB*, *CYP3A4* e família *UGT*). Já os grupos E1/E2 apresentam maior nível de expressão de marcadores de células progenitoras do fígado (*AFP*) e padrão de metilação semelhante a fígado fetal, diferindo principalmente na idade de acometimento tumoral entre os dois grupos. Desta forma, o grupo sugere que os tumores do grupo E1/E2 originam-se de células hepáticas em um estado mais imaturo, enquanto que tumores do grupo F surgem de hepatoblastos mais diferenciados[58].

I.5. Proposta e justificativa do estudo

Como apresentado nos itens anteriores, há uma forte conexão entre epigenética, desenvolvimento fetal e tumores pediátricos. Os eventos de diferenciação celular e organização de tecidos durante o desenvolvimento embrionário decorrem do equilíbrio entre vias de regulação e expressão gênicas, e distúrbios podem resultar em letalidade, malformações congênitas e doenças, como câncer. Trabalhos em meduloblastomas[103,104] e em tumores de Wilms, que são tumores embrionários renais[105,106] evidenciaram mecanismos comuns ao desenvolvimento do tumor e à diferenciação do órgão originário. Adicionalmente, tumores pediátricos reconhecidamente carregam uma carga mutacional menor, sugerindo que um número menor de mutações seria requerido para criar um fenótipo tumoral[6]. Eventualmente, fatores ambientais poderiam modular modificações da atuação da maquinaria epigenética no embrião e, portanto, ser a chave da tumorigênese em crianças.

A faixa etária em que os hepatoblastomas ocorrem evidencia forte associação de sua gênese com a diferenciação hepática; contudo, o mecanismo gerador deste tumor ainda é pouco conhecido[58]. Diante da escassez de mutações já identificadas em tumores hepáticos em geral, e em hepatoblastomas em particular, como também recentemente demonstramos em pacientes brasileiros [49], alterações em transcriptoma e em mecanismos epigenéticos surgem como objeto de estudo relevante para elucidação dessas neoplasias.

Considerando a diminuição que havíamos previamente observado no nível de metilação, não havia sido possível discernir entre o conteúdo de 5mC e 5hmC. Estes dados, analisados em conjunto com a literatura científica na área, originaram quatro questionamentos principais:

I. o padrão identificado de metilação é predominantemente 5mC ou há enriquecimento de 5hmC nestes tumores?
II. a hipometilação observada seria decorrente de um processo ativo de desmetilação (atuação de proteínas TET) ou de desmetilação passiva (DNMT1/DNMT3A)?
III. considerando a hipótese de ruptura do processo normal de desenvolvimento do fígado como origem do hepatoblastoma, em que fase da diferenciação hepática se encontrariam as células precursoras deste tipo tumoral?
IV. dado o papel regulatório reconhecido da metilação de DNA na expressão gênica, qual o impacto do metiloma alterado de HBs em seu transcriptoma?

Capítulo II. Objetivos

O presente trabalho teve como objetivo investigar o papel e mecanismos de atuação da metilação de DNA em hepatoblastomas.

II.1. Objetivos específicos

1. Determinar o nível de 5hmC em amostras de hepatoblastomas e tecidos hepáticos não-tumorais.

2. Avaliar a expressão de genes codificadores de proteínas reguladoras da metilação (*DNMT1*, *DNMT3A*, *DNMT3B*, *DNMT3L* e *UHFR1*) e hidroximetilação (*TET1*, *TET2* e *TET3*) em hepatoblastomas, tecidos hepáticos não-tumorais e linhagens celulares de tumores hepáticos.

3. Caracterizar as amostras de hepatoblastomas de acordo com as diferentes fases da diferenciação hepática, para proposição de estratificação dos tumores com base em metilação, expressão gênica e dados clínicos.

4. Avaliar o impacto regulatório da alteração de metilação do gene *NNMT*, descrito em trabalho anterior do grupo[101], em sua expressão gênica e proteica, bem como impacto no metaboloma em hepatoblastomas.

5. Delinear as vias biológicas implicadas em hepatoblastomas por meio da análise de transcriptoma (RNA-Seq).

6. Correlacionar os resultados de transcriptoma obtidos com dados já existentes de metilação de DNA, identificando genes potencialmente regulados por este mecanismo epigenético.

Capítulo III. Casuística e Metodologia

III.1. Casuística

III.1.1. Amostras de pacientes

Este estudo utilizou um total de 33 amostras de tecidos tumorais de hepatoblastomas primários e 11 amostras de tecido de fígado não-tumoral (amostras derivadas de procedimento cirúrgico, congeladas a fresco). As amostras foram provenientes de colaboração previamente estabelecida com o A.C. Camargo Câncer Center, Grupo de Apoio ao Adolescente e Criança com Câncer (GRAACC) e Instituto de Tratamento do Câncer Infantil (ITACI). O material já havia sido previamente obtido para projetos anteriores do grupo [45,48,107], tendo aprovação do Comitê de Ética em Pesquisa das respectivas instituições. Todas as amostras foram coletadas após consentimento informado assinado por pais ou responsáveis legais.

III.1.2. Linhagens celulares

III.1.2.1. Modelo *Hepatocyte-like a partir de iPSC*

Wei Cu e colaboradores desenvolveram um protocolo composto por três estágios celulares, que se baseiam na troca de meio de cultura com adição de componentes distintos, conduzindo a diferenciação de células-tronco em células *hepatocyte-like*[108]. Estas células não atingem a completa maturação hepática, condição que as põe como candidatas para estudos em tumores embrionários hepáticos. Este modelo já foi estabelecido no Centro de Pesquisa sobre o Genoma Humano e Células Tronco – USP pelo Dr. Luiz Caires, MSc. Ernesto Goulart e Kayque Silva (supervisão Dra. Mayana Zatz), grupo com o qual estabelecemos colaboração.

Foram obtidas amostras biológicas (DNA e RNA) para análises subsequentes de três linhagens de iPSC humanas, assim como das fases intermediárias da diferenciação hepática, sendo estas: endoderme definitiva, hepatoblasto e hepatócito precursor *(hepatocyte-like*). Em resumo, foram utilizadas três linhagens de iPSC, obtidas a partir de sangue periférico humano (F3, F8799 e F7007). No dia 0 as células mononucleadas do sangue periférico foram isoladas a partir da centrifugação de gradiente de densidade, transfectadas com mistura de vetor

epissomal e plaqueadas em placa de 6 poços recobertos com monocamada de fibroblastos[109]. Estas células foram cultivadas com meio específicos para crescimento de células-T, no dia 02 diluído com meio específico para célula-tronco, suplementado com fator de crescimento para fibroblastos e inibidor Rock (inibidor de apopotose), sendo completamente substituído no dia 04. As células foram então mantidas com meio específico para célula-tronco suplementado com fator de crescimento para fibroblastos e inibidor Rock (*StemCell Technologies, USA*), e entre os dias 16 e 25 as colônias de iPSC foram contadas e replaqueadas para expansão e caracterização[109]. Uma vez caracterizadas, as linhagens de iPSC foram submetidas ao protocolo de diferenciação hepática[108]. Neste processo, ao atingir confluência de 50 a 70%, o meio de cultivo das iPSC foi substituído pelo meio de iniciação A (RPMI1640 suplementado com 1X B27, ambos *Invitrogen, U.K.*), 1mM de butirato de sódio e 100ng/mL de activina A (*Sigma, USA*). Após 24 horas, houve substituição para meio de cultivo B, o qual é semelhante ao meio de iniciação A, exceto que a concentração do butirato de sódio foi reduzida para 0,5 mM e as células cultivadas por 72 horas, atingindo a fase de endoderme definitiva. As células foram então coletadas e plaqueadas em nova placa revestida com matrigel e cultivadas em meio de diferenciação (*KO*-DMEM suplementado com 20% de *serum replacement*, ambos *Invitrogen, U.K*), e 1% de dimetilsulfóxido (DMSO - *Sigma, USA*) por sete dias, quando chegam à fase de hepatoblastos. Posteriormente o meio de cultivo é substituído por meio de maturação L5 suplementado com 8,3% de soro bovino fetal (SFB - *ThermoFisher, USA*) contendo 10ng/mL de fator de crescimento de hepatócitos e 20ng/mL de oncostatina M (*Invitrogen, U.K.*), para obtenção dos hepatócitos maduros após sete dias[108].

III.1.2.2. Linhagens celulares de câncer de fígado

As linhagens celulares de câncer de fígado foram adquiridas da empresa ATCC (USA). As linhagens de hepatoblastoma HEPG2 (ATCC® HB-8065™) e C3A (ATCC® CRL-10741™) foram cultivadas em garrafas de 100 mm em meio *Minimum Essential Medium* – MEM (*ThermoFisher, USA*) suplementado com 10% de SFB (*ThermoFisher, USA*), em estufa a 37° e 5% de CO_2. O meio de cultura das células foi substituído duas vezes por semana e, ao atingir 70% de confluência, as células foram tripsinizadas para novas garrafas de cultura. Já as linhagens de carcinoma hepatocelular SNU-387(ATCC® CRL-2237™), SNU-423 (ATCC® CRL-2238™), SNU-449 (ATCC® CRL-2234™), e SNU-475

(ATCC® CRL-2236™), foram cultivadas em garrafas de 100 mm em meio RPMI 1640 (*ThermoFisher, USA*) suplementado com 10% de SFB, sendo mantidas em estufa a 37° e 5% de CO_2. O meio de cultura destas células foi substituído a cada 2 dias e, ao atingir 70% de confluência, as células foram tripsinizadas para novas garrafas de cultura. Amostras das células foram guardadas por criopreservação em nitrogênio líquido em meio de cultura específico de cada linhagem, suplementado com 10% SFB, com adição de 5% de DMSO.

III.2. Metodologia

III.2.1. Extração de biomoléculas

Para extração de RNA dos tecidos e células utilizados no projeto foi utilizado o *RNeasy Mini Kit* (*QIAGEN, Alemanha*), seguindo recomendações dos fabricantes. Em resumo, o processo de extração e purificação é baseado em colunas cromatográficas que utilizam uma resina como matriz de separação, que permite a ligação de RNA de maneira dependente da concentração iônica. A qualidade das amostras foi verificada por eletroforese microfluídica (*Bioanalyzer, Agilent Technologies, USA*) e apenas amostras com RNA Integrity Number (RIN) > 7.0 foram usadas.

A extração de DNA das amostras foi realizada com o *QIASymphony DNA Mini kit* (*QIAGEN, Alemanha*), seguindo recomendações dos fabricantes. Em resumo, o kit utiliza partículas magnéticas que se ligam ao DNA para realizar a extração e posterior purificação do lisado celular.

A extração de proteínas seguiu protocolo de duas partes, para obtenção de extratos enriquecidos com material citoplasmático e extratos enriquecidos com proteínas do núcleo. Para isto, as amostras foram lisadas com auxílio de triturador de vidro em tampão de solução A (Tris-HCL (*Merck, Alemanha*), KCl (*Merck, Alemanha*), MgCl2 (*Merck, Alemanha*), EDTA (*Merck, Alemanha*), β-mercaptoethanol (*Merck, Alemanha*)), seguido de centrifugação por 15 minutos a 12000 rpm a 4°C, o sobrenadante contendo o extrato protéico citoplasmático foi coletado em novo tubo de 1,5 mL e armazenado a -80°C. O pellet foi ressuspendido em solução B (Tris-HCL, NaCl (*Merck, Alemanha*), MgCl2, β-mercaptoethanol) e vortexado 4x durante 40 minutos, sendo mantido em gelo. Por fim, a solução foi centrifugada por 15 minutos a 12000 rpm a 4°C, o sobrenadante contendo o extrato protéico nuclear foi coletado em novo tubo de 1,5 mL e armazenado a -80°C.

III.2.2. Quantificação do nível global de 5hmC

A determinação do nível global de 5hmC foi realizada em amostras de DNA genômico (50 ng a 200 ng) de hepatoblastomas e das linhagens celulares, utilizando o kit *MethylFlash Hydroxymethylated DNA 5-hmC Quantification Kit* (*Epigentek, USA*), seguindo as recomendações do fabricante. Em resumo, o kit baseia-se em uma reação de ELISA e utiliza placa de 96 *wells* pré-tratada pelo fabricante, capaz de detectar especificamente a 5hmC. Para quantificar a quantidade DNA hidroximetilado foi utilizada curva padrão sugerida pelo fabricante (0 ng; 0,03 ng; 0,1 ng; 0,2 ng e 0,5ng); as amostras da curva-padrão foram avaliadasem absorbância de 450 nm juntamente com as amostras alvo. Estes valores de densidade ótica/ng da curva padrão obtidos foram submetidos a regressão linear com as amostras alvo. Em seguida, a porcentagem de 5hmC foi calculada usando a fórmula sugerida pelo fabricante:

$$5hmC\ \% = \frac{5hmC\ Amount\ (ng)}{S} \times 100\%$$

S: quantidade de DNA em ng da amostra alvo.

III.2.3. Análise de expressão de transcritos (RNAm) por real-time PCR

III.2.3.1 Síntese de cDNA

O cDNA foi sintetizado com o kit *High Capacity RNA-to-cDNA* (*ThermoFisher, USA*). Aproximadamente 1 μg de RNA foi utilizado para montagem da reação, com incubação a 37°C por 1 h, seguido de 95°C por 5 minutos. As amostras de cDNAs foram estocada s a -20°C.

III.2.3.2 Reação de PCR em tempo real para quantificação relativa de expressão gênica

O cDNA sintetizado foi quantificado por PCR em tempo real usando o sistema de detecção QuantStudio 5 (*ThermoFisher, USA*), com tecnologias TaqMan e *SYBR green* (**Tabela 2**). As condições de ciclagem foram: 5 minutos a 95ºC, seguido por 40 ciclos de 5 segundos a 95ºC e 10 segundos a 60ºC. Após análise com o algoritmo geNorm [110], utilizando os valores de expressão dos genes candidatos: *ACTB*, *B2M*, 18S e *GAPDH*, o valor de

expressão do gene de interesse foi quantificado em relação aos genes de referências endógenos *B2M* e 18S, usando-se o método 2-ΔΔCq. O experimento foi desenhado e interpretado de acordo com as recomendações contidas no *guideline* internacional MIQE (*The Minimum Information for Publication of Quantitative Real-Time PCR Experiments*)[111]. As representações gráficas foram realizadas com o auxílio do software *GraphPad Prism 7* (USA) e para as análises estatísticas foi utilizado o teste não paramétrico *Mann-Whitney*.

TABELA 2 – **(A) Genes avaliados pelo sistema TaqMan na determinação da expressão gênica. (B) Genes avaliados pelo sistema de SYBR green e respectivos pares de *primers* utilizados para determinação da expressão gênica.**

(A)

	Nome do gene	*probe ID*
Technologia Taqman	***TET1***	Hs00286756_m1
	TET2	Hs00325999_m1
	TET3	Hs00379125_m1
	DNMT1	Hs00154749_m1
	DNMT3A	Hs01027166_m1
	DNMT3B	Hs00171876_m1
	DNMT3L	Hs01081364_m1
	UHRF1	Hs01086727_m1
	NNMT	Hs00196287_m1
	18S	Hs03928992_g1
	B2M	Hs00187842_m1
	GAPDH	Hs02758991_g1
	ACTIN	Hs03023943_g1

(B)

	Nome do gene	Sequencia *forward*	Sequencia *reverse*
Tecnologia SYBR green	***AFP***	TGCAGCCAAAGTGAAGAGGGAAGA	CATAGCGAGCAGCCCAAAGAAGAA
	ALB	AATGTGCTGATGACAGGGCG	CGGCAATGCAGTGGGATTTT
	CK18	ACATCCGGGCCCAATATGAC	GGTGCTCTCCTCAATCTGCT
	CK19	AGAGCTGGCCTACCTGAAGA	GCCCCTCAGCGTACTGATTT
	CXCR4	CAAGGCCCTCAAGACCACAG	TGTAGTAAGGCAGCCAACAGG
	CYP3A4	ATGGAAAAGTGTGGGGCTT	TCATGTCAGGATCTGTGATAGC
	FOXA2	TGCACTCGGCTTCCAGTATG	CATGTTGCTCACGGAGGAGT
	G6PC	CTACTACAGCAACACTTCCGTG	GGTCGGCTTTATCTTTCCCTGA
	HNF1A	GCAGCAGTTCACCCATGCAG	CGCCCCTTCTTGGTTGGTAG
	HNF4A	GGTGTCCATACGCATCCTTGAC	AGCCGCTTGATCTTCCCTGGAT
	NANOG	TGGACACTGGCTGAATCCTTC	CGTTGATTAGGCTCCAACCAT
	POU5F1	TCCCATGCATTCAAACTGAGG	CCAAAAACCCTGGCACAAACT
	SOX2	CCATCCACACTCACGCAAAA	TATACAAGGTCCATTCCCCCG
	SOX17	CAAGGGCGAGTCCCGTATC	GTCCTTAGCCCACACCATGA
	TDO2	AACTCCAGGTTTAGAGCCACAT	TGAATTCCTCTTCCAGGCCTCT
	UGTA1	TGCAACCCTTGCCTCAGAAT	ATTCCTGGGATAGTGGATTTTGGT
	B2M	CCACTGAAAAAGATGAGTATGCCT	CCAATCCAAATGCGGCATCTTCA

III.2.4. Análise de proteína

III.2.4.1. *Western Blot*

A concentração de proteína dos extratos foi determinada com o Pierce 660-nm Protein Assay (*ThermoFisher, USA*). Para isto foi realizada uma curva padrão com BSA 1 mg/mL (*ThermoFisher, USA*) com os pontos: branco (sem BSA), 6,3 μg , 12,5 μg , 25 μg , 50 μg e 100 μg. Posteriormente 10 μL de cada amostra da curva, assim como das amostras-alvo, foram transferidos para uma placa de 96 *wells* seguido da adição de 150 μL do reagente Pierce, a absorbância foi determinada após leitura em espectrofotômetro em comprimento de onda de 660 nm. Os valores foram analisados no programa de estatística *GraphPad Prism 7* (*GraphPad Software*, Inc. – EUA) para determinar a concentração protéica.

Para verificar a expressão das proteínas foram realizados ensaios de *Western Blot*, em que há a separação de proteínas de uma mistura complexa em gel desnaturante pela diferença de suas massas e posterior detecção com o uso de anticorpos específicos. Para isto, após determinação da concentração dos extratos proteicos, foi adicionado tampão de amostra reduzido em 15 μg de proteína. As proteínas foram separadas em gel de gradiente 4-20% (*BioRad, USA*), sob corrente elétrica de 110 V, em tampão de corrida; em seguida, transferidas para membrana de nitrocelulose (*BioRad, USA*), durante 1 hora, sob tensão elétrica de 70 V em tampão de transferência. Após 1 hora de bloqueio de sítios inespecíficos com 5% de leite em pó desnatado em PBS-Tween 0,01%, a membrana foi incubada com os anticorpos primários específicos diluídos em 5% de leite em PBS-Tween 0,01%, *overnight*, a 4^{o}C e sob agitação. No dia seguinte, a membrana foi lavada com PBS-Tween 0,01% por 5 minutos e incubada com o anticorpo secundário conjugado à peroxidase e diluído em 5% de leite em PBS-Tween 0,01%, por 2 horas, em temperatura ambiente, sob agitação. Em seguida, foi lavada por 5 vezes de 10 minutos com PBS-Tween 0,01%, seguindo da revelação. Os seguintes anticorpos foram utilizados: NNMT (OTI3D8; *Abcam, UK*) e β-actin (A5316; *Sigma*, USA). Os resultados também foram apresentados em razão da proteína alvo vs. proteína controle, obtidos com o software *Image J* (*NIH, USA*). O teste-t foi utilizado para análise estatística dos dados.

III.2.4.2 Imunohistoquímica

A análise por imunohistoquímica foi realizada apenas para a proteína *NNMT*, em colaboração com a Dra. Isabela Werneck da Cunha, em período no qual trabalhava no Departamento de Anatomia Patológica do AC Camargo Cancer Center. O protocolo foi automatizado utilizando a plataforma Discovery XT (*Ventana Medical Systems, USA*). Em resumo, foram realizadas as seguintes etapas: desparafinização das lâminas em solução de EZ PREP, seguido da recuperação antigênica por calor em tampão Citrato- pH 6.0 a 95°C por 1h. Posteriormente foi realizado o bloqueio da peroxidade endógena utilizando o reagente de bloqueio de peroxidade (*ultraView Universal DAB Inhibitor* (3% H202)) e passagens em soluções de lavagem. Foi então realizada a incubação com anticorpo primário NNMT *Monoclonal Antibody* (OTI3D8; *Abcam, UK*) *overnight* a 4 °C – seguido de passagens em soluções de lavagem. Foi realizado a incubação com polímero HRP (HRP Multimer) com subseqüentes lavagens em tampão, seguido da incubação com o cromógeno Diaminobenzidine (DAB - *Sigma, USA*), passagem em tampão de lavagem e contracorados com Hematoxilina (*Sigma, USA*). Posteriormente, foi realizada a lavagem em água com detergente para retirada do LCS (solução de lamínula), seguida de lavagem em água corrente. Por fim, realizada a última etapa, que é constituída da desidratação e montagem da lâmina.

III.2.5 Metabolômica

A análise por espectroscopia de alta resolução por ressonância magnética nuclear de giro por ângulo mágico (HR-MAS NMR) foi realizada pela Dra. Ljubica Tasic (Departamento de Químico Orgânica, Instituto de Química, UNICAMP). Os espectros de ^{1}H NMR foram adquiridos usando um espectrômetro Bruker Avance (*Bruker BioSpin, Alemanha*) operando a 400 MHz e equipado com a sonda de núcleo duplo de 4 mm para HR-MAS. Os espectros de ^{1}H NMR suprimidos em água e unidimensionais foram gravados com a sequência de pulsos NOESY e 128 repetições, e os espectros editados em T2 foram gravados usando a sequência de pulsos CPMG (*Carr-Purcell-Melboom-Gill*) com 128 repetições. A análise quimiométrica foi realizada usando o *MetaboAnalyst* (www.metaboanalyst.ca). Detalhes para processamento de espectros de RMN, pré-processamento de dados para atribuições de quimiometria e metabólitos foram descritos anteriormente[112,113].

III.2.6 Sequenciamento de transcriptoma por RNA-seq

O sequenciamento de RNA-seq e obtenção de dados primários foram realizados em colaboração com o grupo da Dra. Dirce Carraro e do Dr. Israel Tojal, do Centro de Pesquisas do AC Camargo Cancer Center. O RNA extraído, após tratamento com DNAse I e purificação, foi avaliado por meio de eletroforese capilar (*Bioanalyzer, Agilent Technologies, USA*); amostras que apresentaram integridade adequada (*RNA Integrity Number* ≥ 7) foram encaminhadas para a construção das bibliotecas de cDNA, com o uso do kit *TruSeq®Stranded Total RNA LT- (with Ribo-Zero TM Gold)* (*Illumina, USA*). O primeiro passo do kit envolve a remoção do RNA ribossômico (rRNA) usando oligos biotinilados, específicos para o alvo, combinados com *beads* de remoção de rRNA Ribo-Zero, procedimento eficiente para depletar rRNA citoplasmático e mitocondrial. Após a purificação, o RNA é fragmentado usando cátions divalentes sob temperatura elevada; então, os fragmentos de RNA clivados são copiados em cDNA de fita única utilizando 12 iniciadores de transcriptase reversa e *primers* aleatórios, seguido da síntese da segunda fita de cDNA usando DNA Polimerase I e RNase H. Estes fragmentos de cDNA têm então a adição de uma única base 'A' e subsequente ligação do adaptador. Os produtos são purificados e enriquecidos por PCR para criar a biblioteca final de cDNA. As bibliotecas de cDNA assim construídas foram avaliadas por eletroforese capilar (*Bioanalyzer, Agilent Technologies, USA*) para verificação de tamanho (~500-700 nt) e quantificadas utilizando um kit de quantificação absoluta por PCR (Kapa Biosystems; *Sigma, USA*). Foram montados "*pools*" de bibliotecas, posteriormente encaminhados para sequenciamento na plataforma *NextSeq 500 Mid Output v2 Kit (150 cycles)* (*Illumina, USA*).

A obtenção dos transcriptomas ocorreu em duas levas distintas: 13 amostras (8 hepatoblastomas e 5 amostras de fígado não tumoral) foram sequenciadas no AC Camargo Cancer Center e, posteriormente, 6 amostras adicionais de hepatoblastomas foram sequenciadas como serviço pela empresa *Admera Health* (*USA*).

As análises de bioinformática dos dados de transcriptoma foram realizadas em colaboração com as equipes do Dr. Israel Tojal (AC Camargo Cancer Center) e do Dr. Thiago Mendes (Universidade Federal de Viçosa).

Capítulo IV. Padrão de expressão de genes da via de metilação de DNA no hepatoblastoma

Hepatoblastoma é o tipo tumoral descrito com menor taxa mutacional dentre os tumores sólidos, no qual mutação no gene *CTNNB1* é a alteração genética mais recorrente, levando à ativação aberrante da via de sinalização *Wnt*[43,58]. Deste modo, a epigenética apresenta-se como um mecanismo potencialmente relevante para a gênese e progressão deste tumor.

Trabalho anterior do nosso grupo evidenciou hipometilação global em amostras de hepatoblastomas, em comparação a tecido não-tumoral de fígado adulto e fetal, assim como, hipermetilação em regiões específicas[101]. A DNAm é uma marca epigenética associada à regulação da expressão gênica e sua alteração no câncer é bastante reconhecida[63,64]. Em geral, células cancerosas apresentam hipometilação em sequências repetitivas, resultando na ativação de sítios de transcrição espúrios e retrotransposons; também exibem hipermetilação em regiões promotoras de genes supressores tumorais, resultando em sua inativação[83]. A DNAm ocorre pela ação de proteínas como DNA Metiltransferases (DNMTs) e UHRF1 [82,114] e alterações na atividade e função destas proteínas no câncer estão associadas a modificações na regulação gênica.

A desmetilação do DNA pode ocorrer de forma passiva ou ativa; na primeira, a marca da 5mC é perdida ao longo de consecutivas replicações do DNA sem a atuação da maquinaria de metilação; já na segunda, ocorre ação de proteínas que modificam a citosina metilada[84]. A hidroximetilação do DNA é uma via de desmetilação ativa em que há a oxidação da 5mC e produção de 5hmC [85], sendo que a alteração do nível de 5hmC tem sido bastante descrita em câncer[93]. As proteínas da família TET (do inglês – *Ten-eleven translocation*) são as principais responsáveis pela hidroximetilação do DNA, podendo atuar conjuntamente com DNMT3A para realização deste processo[95,115]. Apesar da hipometilação descrita previamente, a técnica utilizada em nosso trabalho prévio não analisou níveis de hidroximetilação[101], originando dúvida acerca da origem da hipometilação detectada: a alteração de DNAm evidenciada em hepatoblastoma seria derivada de um processo passivo de desmetilação, no qual a 5mC é perdida ao longo das replicações celulares, ou ativo, com perda de 5mC e havendo produção de 5hmC?

Para aprofundar o conhecimento acerca dos mecanismos epigenéticos atuantes no hepatoblastoma, em especial sobre a hipometilação global observada, estudamos o padrão de expressão de genes da maquinaria de metilação (*DNMT1*, *DNMT3A*, *DNMT3B*, *DNMT3L*, *UHRF1*, *TET1*, *TET2* e *TET3*) nestes tumores, assim como, investigamos o nível global de 5hmC nos tumores.

Como resultado, observamos em hepatoblastomas elevada expressão dos genes da maquinaria epigenética, principalmente *UHRF1*, *TET1* e *TET2*, em associação com enriquecimento do conteúdo de 5hmC. Estes resultados indicam a ocorrência em hepatoblastomas de desmetilação ativa do DNA, mediada por *TETs*, provavelmente durante os estágios iniciais do desenvolvimento do fígado. Em combinação com o aumento de expressão de *UHRF1*, culminaria em hipometilação de DNA e concomitante aumento de 5hmC. Além disso, nossos dados sugerem que níveis mais baixos de 5hmC em hepatoblastomas podem estar associados à redução da sobrevida global, destacando o conteúdo de 5hmC como um potencial biomarcador de risco para estes tumores.

Os dados acima descritos foram publicados no artigo *TET upregulation leads to 5-hydroxymethylation enrichment in hepatoblastoma (DOI: 10.3389/fgene.2019.00553)*, na revista *Frontiers in Genetics*, sendo apresentado a seguir.

frontiers in Genetics
doi: 10.3389/fgene.2019.00553

BRIEF RESEARCH REPORT
published: 12 June 2019

TET Upregulation Leads to 5-Hydroxymethylation Enrichment in Hepatoblastoma

Maria Prates Rivas[1], Talita Ferreira Marques Aguiar[1,2], Gustavo Ribeiro Fernandes[3], Luiz Carlos Caires-Júnior[1], Ernesto Goulart[1], Kayque Alves Telles-Silva[1], Monica Cypriano[4], Silvia Regina Caminada de Toledo[4], Carla Rosenberg[1], Dirce Maria Carraro[2], Cecilia Maria Lima da Costa[5], Isabela Werneck da Cunha[6] and Ana Cristina Victorino Krepischi[1]*

[1]Human Genome and Stem-Cell Research Center, Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, São Paulo, Brazil, [2] International Center of Research, A. C. Camargo Cancer Center, São Paulo, Brazil, [3] Department of Biochemistry, Institute of Chemistry, University of São Paulo, São Paulo, Brazil, [4]Department of Pediatrics, Support Group for Children and Adolescents With Cancer (GRAACC), Federal University of São Paulo, São Paulo, Brazil, [5] Department of Pediatric Oncology, A. C. Camargo Cancer Center, São Paulo, Brazil, [6] Department of Pathology, Rede D'Or São Luiz, São Paulo, Brazil

OPEN ACCESS

Edited by:
Carolina Vicente-Dueñas, Instituto de Investigación Biomédica de Salamanca (IBSAL), Spain

Reviewed by:
Gokul Gopinath, Texas A&M University Baylor College of Dentistry, United States
Steven G. Gray, St. James's Hospital, Ireland

***Correspondence:**
Ana Cristina Victorino Krepischi
ana.krepischi@ib.usp.br

Specialty section:
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Genetics

Received: 14 March 2019
Accepted: 24 May 2019
Published: 12 June 2019

Citation:
Rivas MP, Aguiar TFM, Fernandes GR, Caires-Júnior LC, Goulart E, Telles-Silva KA, Cypriano M, de Toledo SRC, Rosenberg C, Carraro DM, da Costa CML, da Cunha IW and Krepischi ACV (2019) TET Upregulation Leads to 5-Hydroxymethylation Enrichment in Hepatoblastoma. Front. Genet. 10:553. doi: 10.3389/fgene.2019.00553

Hepatoblastoma is an embryonal liver tumor carrying few genetic alterations. We previously disclosed in hepatoblastomas a genome-wide methylation dysfunction, characterized by hypermethylation at specific CpG islands, in addition to a low-level hypomethylation pattern in non-repetitive intergenic sequences, in comparison to non-tumoral liver tissues, shedding light into a crucial role for epigenetic dysregulation in this type of cancer. To explore the underlying mechanisms possibly related to aberrant epigenetic modifications, we evaluated the expression profile of a set of genes engaged in the epigenetic machinery related to DNA methylation (DNMT1, DNMT3A, DNMT3B, DNMT3L, UHRF1, TET1, TET2, and TET3), as well as the 5-hydroxymethylcytosine (5hmC) global level. We observed in hepatoblastomas a general disrupted expression of these genes from the epigenetic machinery, mainly UHRF1, TET1, and TET2 upregulation, in association with an enrichment of 5hmC content. Our findings support a model of active demethylation by TETs in hepatoblastoma, probably during early stages of liver development, which in combination with UHRF1 overexpression would lead to DNA hypomethylation and an increase in overall 5hmC content. Furthermore, our data suggest that decreased 5hmC content might be associated with poor survival rate, highlighting a pivotal role of epigenetics in hepatoblastoma development and progression.

Keywords: hepatoblastoma, epigenetics, hydroxymethylation, active demethylation, DNA hypomethylation

INTRODUCTION

Pediatric tumors are inherently different from tumors that develop in adults (Maris and Denny, 2002) since they are supposedly derived from defects of cell differentiation and organogenesis (Anderson, 2006). Embryonal tumors are a special class of childhood cancers characterized by a very early-onset exhibiting histological features that resemble developmental stages of the origin organ (Maris and Denny, 2002).

Cancer development involves the accumulation of genetic mutations over an extended period of time (Hanahan and Weinberg, 2011); however, this model is not suitable for pediatric cancer, which develops in a short temporal window, usually carrying a low number of somatic mutations (Vogelstein et al., 2013; Gröbner et al., 2018). A more comprehensive background takes into account the dysregulation of epigenetic mechanisms (Sandoval and Esteller, 2012), considering the interplay between genetic and epigenetic factors as a key element for tumor development (Soo You and Jones, 2012). Among the epigenetic regulators, DNA methylation, which occurs through the addition of a methyl group to a cytosine (5mC) by DNA methyltransferase enzymes (DNMTs) (Jones, 1999), has been extensively investigated. In tumor cells, well-known DNA methylation disturbances include repression of tumor suppressor genes by promoter hypermethylation, and hypomethylation at repetitive sequences, leading to genomic instability (Łuczak and Jagodzinski, 2006; Rodrigues-Paredez and Esteller, 2011).

Hepatoblastoma is a rare liver embryonal tumor, and the most common form of liver tumor in children (Wiederkehr et al., 2013; López-Terrada et al., 2014), with an incidence of 1-2.4 per million children each year (Paran and La Quaglia, 2009). These tumors carry few genetic alterations, mainly gains of chromosomes 2, 8, 20, and chromosome 18 deletion (Tomlinson and Kappler, 2012; Ozen et al., 2013; Rodrigues et al., 2014), and recurrent *CTNNB1* activating mutations; low frequency mutations of other *Wnt* pathway genes, such as *AXIN1, AXIN2, AXIN3* and *APC*, have also been reported (Tomlinson and Kappler, 2012; Eichenmüller et al., 2014; Jia et al., 2014). In a recent work from our group (Maschietto et al., 2017), the evaluation of hepatoblastoma methylomes disclosed a genome-wide methylation dysfunction, characterized by hypermethylation at specific CpG islands, in addition to a low-level hypomethylation pattern in non-repetitive intergenic sequences, the later finding corroborated by others (Cui et al., 2016).

DNA hydroxymethylation is a kind of DNA modification (Münzel et al., 2011) originated by 5mC conversion to 5-hydroxymethylcytosine (5hmC), in a process catalyzed by proteins encoded by *TET* genes. The 5hmC is an intermediate residue of the active demethylation process and seems to be related to pluripotency maintenance in stem cells and transcriptional activation (Wu and Zhang, 2011; Ko et al., 2011).

To elucidate the underlying mechanisms resulting in DNA methylation loss in hepatoblastomas, we evaluated the expression profiles of the *DNMT* and *TET* genes as well as the global level of 5hmC in a hepatoblastomas cohort.

MATERIALS AND METHODS

Samples

Nineteen hepatoblastoma samples (HBs) and eight paired non-tumoral liver tissue samples were obtained from hepatoblastoma patients submitted to surgical resection in two pediatric cancer institutions (A. C. Camargo Cancer Center and GRAACC, São Paulo, Brazil). This study was carried out in accordance with the recommendations of Research Ethics Committee - A.C. Camargo Cancer Center, Brazil (registration number 1987/14), with written informed consent from all parents/guardian of participants. Clinical features of this HB cohort are described in **Supplementary Table 1**. All patients received pre-surgery chemotherapy according to the SIOPEL protocol[1] and COG protocol[2].

Methods

Gene Expression Analysis

Microfluidics-based electrophoresis (Bioanalyzer, Agilent Technologies; CA, United States) was performed to verify quality, and only RNA samples with RNA Integrity Number (RIN) > 7.0 were used for gene expression analysis. The cDNA was synthesized with the High Capacity RNA-to-cDNA kit (Applied Biosystems, EUA) according to standard procedures. The experiment was designed and interpreted

1http://www.siopel.org/
2https://www.childrensoncologygroup.org/

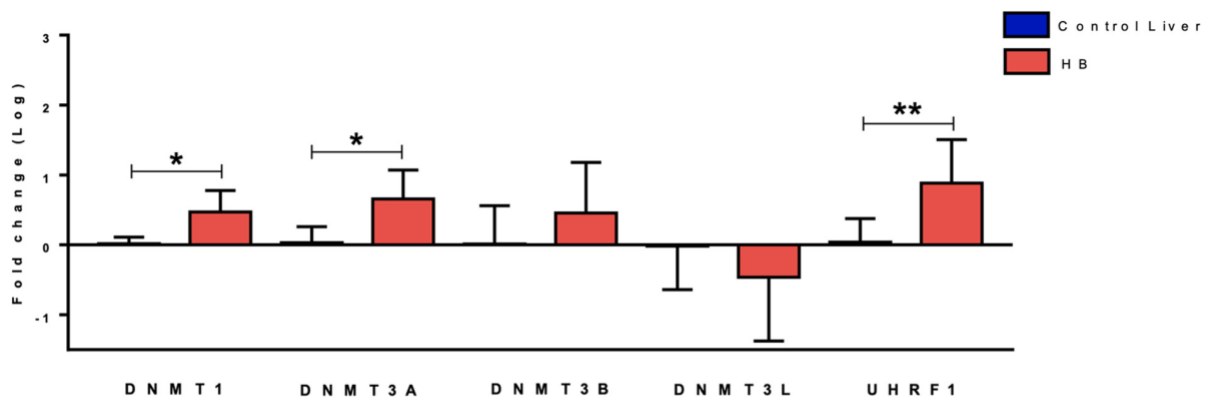


FIGURE 1 | Expression analysis of DNA methylation genes in hepatoblastomas. DNMT1, DNMT3A and UHRF1 upregulation in 19 tumor samples compared with eight non-tumoral liver samples as evidenced by RT-qPCR. Control Liver: non-tumoral liver tissue samples; HB: hepatoblastoma samples. Endogenous gene: 18S rRNA. The Kruskal Wallis post hoc Dunn test with Bonferroni correction was used for statistical analysis. p-value < 0.05, p-value < 0.01.

according to recommendations contained in International Guideline MIQE (The Minimum Information for Publication of Quantitative Real-Time PCR Experiments). The suitable housekeeping gene was chosen using the geNorm algorithm (Vandesompele et al., 2002) after expression analysis of *ACTB*, *GAPDH*, *B2M* genes and 18S ribosomal RNA (18S rRNA). Data were normalized using the expression values of the housekeeping gene 18S rRNA, and all reactions were performed with three technical replicates. The delta-delta C_t ($11C_t$) method was used for data analysis (Livak and Schmittgen, 2001). The Kruskal Wallis *post hoc* Dunn test with Bonferroni correction was used for statistical analysis in the GraphPad Prism 7 software.

5-Hydroxymethylcytosine (5hmC) Quantification

Nanodrop (Thermo Fisher Scientific, United States) was used to access quantity and integrity of DNA samples. The 5hmC global level quantification of 18 HBs and eigth paired non-tumoral liver tissue samples was carried out using the MethylFlash Hydroxymethylated DNA 5-hmC Quantification Kit (Epigentek, United States). The values of each sample are given as percentages relative to internal controls of the kit, according to the manufacturer's recommendation. The *t*-test was used for the statistical analysis.

RESULTS

Gene expression analysis of five DNA methylation genes (*DNMT1, DNMT3A, DNMT3B, DNMT3L*, and *UHRF1*) revealed up-regulation of *DNMT3A* (fold-change mean: 7.0; *p*-value: 0.024) and *DNMT1* (fold-change mean: 3.6; *p*-value: 0.012) expression in the group of 19 HBs as compared to eight paired non-tumoral liver tissues, while *DNMT3B* and *DNMT3L* expression values were not significantly altered (**Figure 1**). Data also showed a very significant increase in *UHRF1* expression in tumor samples (fold-change mean: 20.4; *p*-value: 0.015).

The hydroxymethylation process was evaluated by assessing the expression level of the three *TET* genes, and all of them were found to be significantly upregulated in hepatoblastomas as compared to non-tumoral liver tissue (**Figure 2A**): *TET1* (fold-change mean: 21.5; *p*-value: 0.009), *TET2* (fold-change mean: 7.15; *p*-value: 0.009), and *TET3* (fold-change mean: 4.3; *p*-value: 0.024). Given the observed *TET* genes upregulation in HBs, the global content of 5hmC was evaluated in tumors and control hepatic tissues. Regardless of the variability in total 5hmC content seen across HBs, an overall enrichment of approximately 2.5 (*p*-value 4.2e-05; *t*-test) was observed in tumors (mean: 0.33%; SD: 0.15) relative to non-tumoral livers (mean: 0.13%; SD: 0.07) (**Figure 2B**). Statistical analysis based on the clinical data showed that patients with overall survival 5 years exhibited a 5hmC content lower than the mean of the tumor group (*p*-value 0.02; Fisher's test) (**Figure 3**),

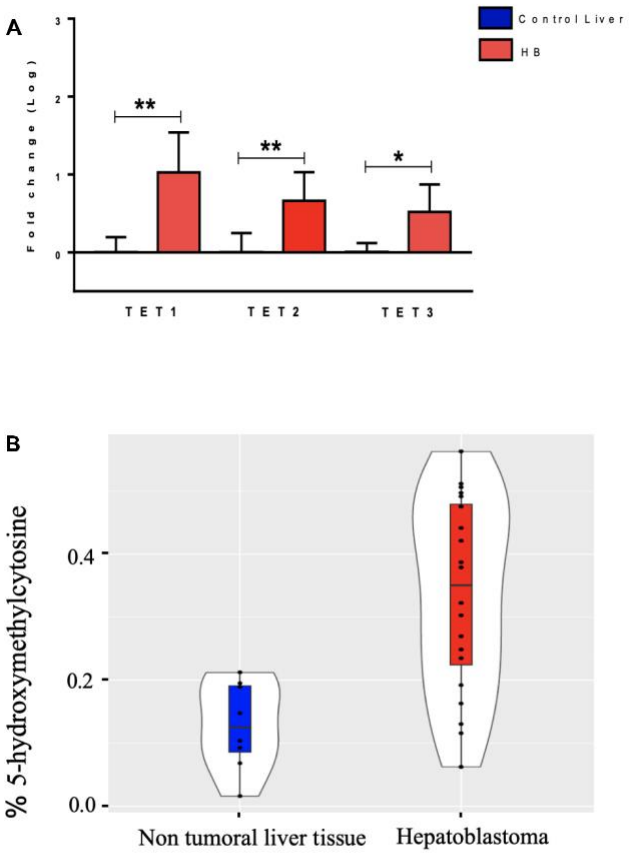


FIGURE 2 | Analysis of DNA hydroxymethylation in hepatoblastomas. **(A)** Expression analysis of DNA hydroxymethylation genes in hepatoblastomas. Upregulation of three genes of the TET family in 19 tumor samples compared with eight non-tumoral liver samples as evidenced by RT-qPCR. Control Liver: non-tumoral liver tissue samples; HB: hepatoblastoma samples. Endogenous gene: 18S rRNA. The Kruskal Wallis post hoc Dunn test with Bonferroni correction was used for statistical analysis. p-value < 0.05; p-value < 0.01. **(B)** 5hmC global level enrichment in hepatoblastoma samples. Violin plots showing the overall level of 5hmC in percentage in hepatoblastomas samples (red) and non-tumoral liver tissues (blue). Hepatoblastoma samples mean: 0.33%, SD: 0.15; non-tumoral liver tissue samples mean: 0.13%; SD: 0.07. The t-test was used for the statistical analysis. p-value: 4.2e-05.

closer to the level that was detected in the group of non-tumoral liver tissues.

DISCUSSION

We previously reported changes in the methylation profile of hepatoblastomas, with global hypomethylation in intergenic non-repetitive sequences, as well as hypermethylation at CpG islands, affecting genes associated with tumor suppression, lipid metabolism, and liver differentiation (Maschietto et al., 2017). Except for the *UHRF1* gene (Beck et al., 2018), here we disclosed for the first time dysregulated expression of genes from the DNA methylation machinery that could be related

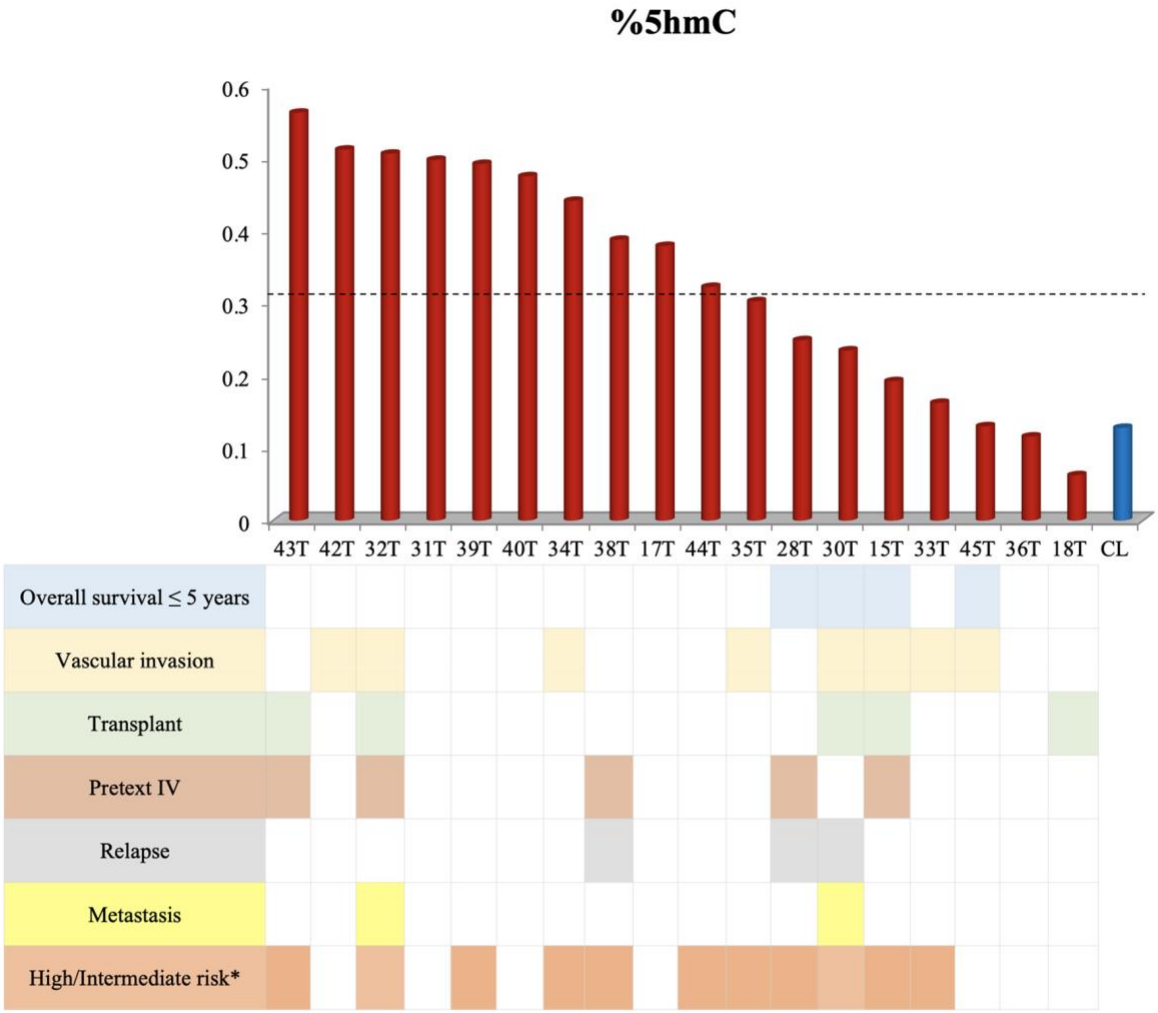


FIGURE 3 | Decreased overall survival is associated with lower levels of 5hmC. This figure shows the 5hmC global content of 18 tumor samples in the columns and their respective clinical data below. Tumor samples with overall survival 5 years present reduced 5hmC content, below the mean of the tumor group (0.33% - represented by the dashed line) and close to the level of non-tumoral hepatic tissues (CL). The Fisher's test was used for the statistical analysis. p-value 0.02.
CL: non-tumoral liver tissue samples; T: hepatoblastoma tumor samples. According to the CHIC criteria (Meyers et al., 2017).

to the detected epigenetic disturbance, such as upregulation of *DNMT3A, DNMT1, UHRF1,* and the *TET* family of genes.

The interaction between UHRF1/DNMT1 plays a pivotal role in the inheritance of DNA methylation patterns, even though the details of this complex network remain under analysis. Thus, *UHRF1* is considered an essential epigenetic regulator for DNA methylation maintenance (Bostick et al., 2007). It is known that UHRF1/DNMT1 are involved in the maintenance of the hypermethylation of tumor suppressor gene promoters in cancer (Alhosin et al., 2016; Ashraf et al., 2017); in particular, overexpression of *UHRF1* was recently reported to silence specific tumor suppressor genes in hepatoblastomas (Beck et al., 2018). Indeed, enhanced *UHRF1* expression was identified in all cancers so far examined (Ashraf et al., 2017). Therefore, it is possible that *DNMT1* and *DNMT3A* upregulation observed in the present study contributes to the hypermethylation at specific CpG islands we previously reported in hepatoblastomas.

Nevertheless, it has also been suggested that impaired interaction between UHRF1 and DNMT1 could be the origin of the global hypomethylation in cancer cells (Hervouet et al., 2010; Pacaud et al., 2014). Moreover, a different line of evidence revealed that *UHRF1* overexpression also acts as a negative regulator by degradation of the *de novo* DNA methylation protein, DNMT3A, as an additional mechanism leading to widespread DNA hypomethylation in cancer cells (Jia et al., 2016). Complementary to these findings, Mudbhary et al. (2014) demonstrated in both zebrafish hepatocytes and hepatocellular carcinoma that *UHRF1* overexpression promotes DNMT1 destabilization and de-localization, driving DNA hypomethylation; further, *UHRF1* increased expression was already associated with poor prognosis in hepatocellular carcinoma, contributing to cell proliferation and metastasis (Liu et al., 2017). These data further support the hypothesis that disruption of normal cooperation of DNA methylases and *UHRF1* can play a dual role in the

patterns of DNA methylation in cancer cells, which mechanisms remain partially elusive. In combination with our results, these evidences suggest that, at least in liver cancer, including hepatoblastomas, the disclosed *UHRF1* overexpression may result in inefficient or impaired DNA methylation process, irrespective of the *DNMTs* expression level, contributing to the global low level hypomethylation pattern in non-repetitive intergenic sequences reported in hepatoblastomas (Cui et al., 2016; Maschietto et al., 2017) and hepatocellular carcinoma (Shen et al., 2013).

The conversion of methylated cytosine to 5hmC, mediated by *TET* enzymes, is the first step in the DNA demethylation active pathway (Wu and Zhang, 2011; Ko et al., 2011; Rasmussen and Helin, 2016). In adult cancer, a decrease in *TET* expression was reported to reduce the 5hmC level in solid and hematological tumors (Lian et al., 2012; Liu et al., 2013; Rawłuszko-Wieczorek et al., 2015; Rasmussen and Helin, 2016). However, the activity of these enzymes in pediatric tumors is poorly explored. Despite the limited size of our cohort, here we showed an increased expression of *TET* genes in hepatoblastomas, suggesting the occurrence of an active demethylation mechanism. Studies with animal models described *TET* overexpression in embryonic development stages repressing differentiation (Ito et al., 2010; Rawłuszko-Wieczorek et al., 2015; Rasmussen and Helin, 2016), as well as in cellular reprogramming of IPS cells (Bagci and Fisher, 2013). Thus, we can speculate that *TET* upregulation in hepatoblastomas indicates the occurrence of a blockage of liver differentiation, which corroborates the long-standing hypothesis of repression of the differentiation pathway of the origin organ leading to embryonal tumors genesis (Maris and Denny, 2002). A functional role for the observed *TET* overexpression was substantiated by the enrichment of the 5hmC content in these tumors, then partially explaining the DNA hypomethylation described in hepatoblastoma, in addition to the possible impaired *de novo* methylation caused by the observed *UHRF1* overexpression. An additional analysis regarding the underlying nature of the DNA sequences with 5hmC enrichment in hepatoblastomas (if intergenic non-repetitive or CpG islands) would add valuable information, and should be made in the future. Remarkably, in hepatocellular carcinoma, a reduced 5hmC content was directly associated with lower overall survival and poor prognosis (Liu et al., 2013). Our findings also suggested that lower levels of 5hmC in hepatoblastomas could be associated with decreased overall survival, highlighting the 5hmC content as a potential biomarker of risk for these tumors. However, these preliminary results require further validation in a larger cohort of hepatoblastomas.

This study addressed possible mechanisms leading to the hypomethylation pattern displayed by hepatoblastomas. Our data revealed disturbance in the expression of genes related to the DNA methylation machinery and enrichment of the 5hmC content in hepatoblastomas. As hepatoblastomas exhibit the lowest mutation rate reported so far for pediatric solid tumors (Gröbner et al., 2018), this analysis of genes of the epigenetic machinery provides evidence that epigenetic disruption is probably key for the onset of this liver embryonal tumor, and highlights the 5hmC content as a potential biomarker for poor overall survival.

DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the **Supplementary Files**.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Research Ethics Committee – A.C. Camargo Cancer Center, Brazil, registration number 1987/14. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved and all the samples were collected and extracted in the A.C.Camargo Cancer Center Bank of Macromolecules, following the technical and ethical procedures of the institution (Campos et al., 2012; Olivieri et al., 2014).

AUTHOR CONTRIBUTIONS

MR and AK conceived the study and participated in its design. MR, LC-J, EG, and KT-S performed the experiments and analyzed the data. GF performed statistical analysis. TA, CdC, IdC, MC, SdT, and DC evaluated patients, collected biological samples, and revised clinical data. MR and AK wrote the manuscript. CR and AK critically revised the manuscript. All authors have read and approved the final version of the manuscript.

FUNDING

This project was funded by the Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP (Project CEPID – Human Genome and Stem Cell Research Center – 2013/08028-1). MR and TA were supported by scholarships FAPESP 2016/23462-8 and 2016/04785-0, respectively. EG and LC-J were also supported by FAPESP (2015/14821-1 and 2017/16283-2, respectively).

ACKNOWLEDGMENTS

We thank the patients and their families for their precious collaboration.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00553/full#supplementary-material

REFERENCES

Alhosin, M., Omran, Z., Zamzami, M., Al-Malki, A. L., Choudhry, H., Mousli, M., et al. (2016). Signalling pathways in UHRF1-dependent regulation of tumor suppressor genes in cancer. *J. Exp. Clin. Cancer Res.* 35:174. doi: 10.1186/ s13046-016-0453-455

Anderson, L. M. (2006). Environmental genotoxicants/carcinogens and childhood cancer: bridgeable gaps in scientific knowledge. *Mut. Res.* 608, 136–156. doi: 10.1016/j.mrgentox.2006.02.016

Ashraf, W., Ibrahim, A., Alhosin, M., Zaayter, L., Ouararhni, K., Papin, C., et al. (2017). The epigenetic integrator UHRF1: on the road to become a universal biomarker for cancer. *Oncotarget* 8, 51946–51962. doi: 10.18632/oncotarget. 17393

Bagci, H., and Fisher, A. G. (2013). DNA demethylation in pluripotency and reprogramming: the role of tet proteins and cell division. *Cell Stem Cell* 13, 265–269. doi: 10.1016/j.stem.2013.08.005

Beck, A., Trippel, F., Wagner, A., Joppien, S., Felle, M., Vokuhl, C., et al. (2018). Overexpression of UHRF1 promotes silencing of tumor suppressor genes and predicts outcome in hepatoblastoma. *Clin. Epigenet.* 10:27. doi: 10.1186/s13148-018-0462-7

Bostick, M., Kim, J. K., Estève, P. O., Clark, A., Pradhan, S., and Jacobsen, S. E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317, 1760–1764. doi: 10.1126/science. 1147939

Campos, A. H. J. F. M., Silva, A. A., De Carvalho Mota, L. D., Olivieri, E. R., Prescinoti, V. C., Patrão, D., et al. (2012). The value of a tumor bank in the development of cancer research in Brazil: 13 years of experience at the A C Camargo Hospital. *Biopreserv. Biobank.* 10, 168–173. doi: 10.1089/bio.2011. 0032.

Cui, X., Liu, B., Zheng, S., Dong, K., and Dong, R. (2016). Genome-wide analysis of DNA methylation in hepatoblastoma tissues. *Oncol. Lett.* 12, 1529–1534. doi: 10.3892/ol.2016.4789

Eichenmüller, M., Trippel, F., Kreuder, M., Beck, A., Schwarzmayr, T., Häberle, B., et al. (2014). The genomic landscape of hepatoblastoma and their progenies with HCC-like features. *J. Hepatol.* 61, 1312–1320. doi: 10.1016/j.jhep.2014. 08.009

Gröbner, S. N., Worst, B. C., Weischenfeldt, J., Buchhalter, I., Kleinheinz, K., Rudneva, V. A., et al. (2018). The landscape of genomic alterations across childhood cancers. *Nature* 555, 321–327. doi: 10.1038/ nature25480

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell Press* 144, 646–674. doi: 10.1016/j.cell.2011.02.013

Hervouet, E., Lalier, L., Debien, E., Cheray, M., Geairon, A., Rogniaux, H., et al. (2010). Disruption of Dnmt1/PCNA/UHRF1 interactions promotes tumorigenesis from human and mice glial cells. *PLoS One* 5:e11333. doi: 10. 1371/journal.pone.0011333

Ito, S., D'Alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466, 1129–1133. doi: 10.1038/ nature09303

Jia, D., Dong, R., Jing, Y., Xu, D., Wang, Q., Chen, L., et al. (2014). Exome sequencing of hepatoblastoma reveals novel mutations and cancer genes in the Wnt pathway and ubiquitin ligase complex. *Hepatology* 60, 1686–1696. doi: 10.1002/hep.27243

Jia, Y., Li, P., Fang, L., Zhu, H., Xu, L., Cheng, H., et al. (2016). Negative regulation of DNMT3A de novo DNA methylation by frequently overexpressed UHRF family proteins as a mechanism for widespread DNA hypomethylation in cancer. *Cell Discov.* 2:16007. doi: 10.1038/celldisc. 2016.7

Jones, P. A. (1999). The DNA methylation paradox. *Trends Genet.* 15, 34–37. doi: 10.1016/S0168-9525(98)01636-9

Ko, M., Bandukwala, H. S., An, J., Lamperti, E. D., Thompson, E. C., Hastie, R., et al. (2011). Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc. Natl. Acad. Sci. U.S.A.* 108, 14566–14571. doi: 10.1073/pnas.1112317108

Lian, C. G., Xu, Y., Ceol, C., Wu, F., Larson, A., Dresser, K., et al. (2012). Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell* 150, 1135–1146. doi: 10.1016/j.cell.2012.07.033

Liu, C., Liu, L., Chen, X., Shen, J., Shan, J., Xu, Y., et al. (2013). Decrease of 5-hydroxymethylcytosine is associated with progression of hepatocellular carcinoma through downregulation of TET1. *PLoS One* 8:e62828. doi: 10.1371/ journal.pone.0062828

Liu, X., Ou, H., Xiang, L., Li, X., Huang, Y., and Yang, D. (2017). Elevated UHRF1 expression contributes to poor prognosis by promoting cell proliferation and metastasis in hepatocellular carcinoma. *Oncotarget* 8, 10510–10522. doi: 10. 18632/oncotarget.14446

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

López-Terrada, D., Alaggio, R., de Dávila, M. T., Czauderna, P., Hiyama, E., Katzenstein, H., et al. (2014). Towards an international pediatric liver tumor consensus classification: proceedings of the Los Angeles COG liver tumors symposium. *Mod. Pathol.* 27, 472–491. doi: 10.1038/modpathol. 2013.8

Łuczak, M. W., and Jagodzinski, P. P. (2006). The role of DNA methylation in cancer development. *Folia Histochem. Cytobiol.* 44, 143–154.

Maris, J. M., and Denny, C. T. (2002). Focus on embryonal malignancies. *Cancer Cell* 2, 447–450. doi: 10.1016/S1535-6108(02)00206-4

Maschietto, M., Rodrigues, T. C., Kashiwabara, A. Y., de Araujo, ÉS. S., Marques Aguiar, T. F., da Costa, C. M. L., et al. (2017). DNA methylation landscape of hepatoblastomas reveals arrest at early stages of liver differentiation and cancer-related alterations. *Oncotarget* 8, 97871–97889. doi: 10.18632/oncotarget. 14208

Meyers, R. L., Maibach, R., Hiyama, E., Häberle, B., Krailo, M., Rangaswami, A., et al. (2017). Risk-stratified staging in paediatric hepatoblastoma: a unified analysis from the Children's hepatic tumors international collaboration. *Lancet Oncol.* 18, 122–131. doi: 10.1016/S1470-2045(16)30598-8

Mudbhary, R., Hoshida, Y., Chernyavskaya, Y., Jacob, V., Villanueva, A., Fiel, M. I., et al. (2014). UHRF1 overexpression drives DNA hypomethylation and hepatocellular carcinoma. *Cancer Cell* 25, 196–209. doi: 10.1016/j.ccr.2014. 01.003

Münzel, M., Globisch, D., and Carell, T. (2011). 5-Hydroxymethylcytosine, the sixth base of the genome. *Angew. Chem. Int. Ed. Engl.* 50, 6460–6468. doi: 10.1002/anie.201101547

Olivieri, E. H. R., Franco, L. de A., Pereira, R. G., Mota, L. D. C., Campos, A. H. J. F. M., and Carraro, D. M. (2014). Biobanking practice: RNA storage at low concentration affects integrity. *Biopreserv. Biobank.* 12, 46–52. doi: 10. 1089/bio.2013.0056

Ozen, C., Yildiz, G., Dagcan, A. T., Cevik, D., Ors, A., Keles, U., et al. (2013). Genetics and epigenetics of liver cancer. *New Biotechnol.* 30, 381–384. doi: 10.1016/j.nbt.2013.01.007

Pacaud, R., Brocard, E., Lalier, L., Hervouet, E., Vallette, F. M., and Cartron, P. F. (2014). The Dnmt1/PCNA/UHRF1 disruption induces tumorigenesis characterized by similar genetic and epigenetic signatures. *Sci. Rep.* 18:4230. doi: 10.1038/srep04230

Paran, T. S., and La Quaglia, M. P. (2009). "Hepatic Tumors in Childhood," in *Pediatric Surgery*, eds P. Puri and M. Höllwarth (Heidelberg: Springer), 727–735. doi: 10.1007/978-3-540-69560-8_75

Rasmussen, K. D., and Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.* 30, 733–750. doi: 10.1101/gad. 276568.115

Rawłuszko-Wieczorek, A. A., Siera, A., and Jagodzinski,´ P. P. (2015). TET proteins in cancer: current 'state of the art'. *Crit. Rev. Oncol. Hematol.* 96, 425–436. doi: 10.1016/j.critrevonc.2015.07.008

Rodrigues, T. C., Fidalgo, F., da Costa, C. M., Ferreira, E. N., da Cunha, I. W., Carraro, D. M., et al. (2014). Upregulated genes at 2q24 gains as candidate oncogenes in hepatoblastomas. *Future Oncol.* 10, 2449–2457. doi: 10.2217/fon. 14.149

Rodrigues-Paredez, M., and Esteller, M. (2011). Cancer epigenetics reaches mainstream oncology. *Nat. Med.* 17, 330–339. doi: 10.1038/nm.2305

Sandoval, J., and Esteller, M. (2012). Cancer epigenomics: beyond genomics. *Curr. Opin. Genet. Dev.* 22, 50–55. doi: 10.1016/j.gde.2012.02.008

Shen, J., Wang, S., Zhang, Y. J., Wu, H. C., Kibriya, M. G., Jasmine, F., et al. (2013). Exploring genome-wide DNA methylation profiles altered in hepatocellular carcinoma using infinium human methylation 450 beadchips. *Epigenetics* 8, 34–43. doi: 10.4161/epi.23062

Soo You, J., and Jones, P. A. (2012). Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell* 22, 9–20. doi: 10.1016/j.ccr.2012.06.008

Tomlinson, G. E., and Kappler, R. (2012). Genetics and epigenetics of hepatoblastoma. *Pediatr. Blood Cancer* 59, 785–792. doi: 10.1002/pbc.24213

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., et al. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3:research0034.1–0034.11.

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A. Jr., and Kinzler, K. W. (2013). Cancer genome landscapes. *Science* 339, 1546–1558. doi: 10.1126/science.1235122

Wiederkehr, J. C., Coelho, I. M., Avilla, S. G., Wiederkehr, B. A., and Wiederkehr, H. A. (2013). Liver tumors in infancy. *Hepatic Surg.* 18, 423–460. doi: 10.5772/ 51764

Wu, H., and Zhang, Y. (2011). Mechanism and functions of TET proteins-mediated 5-methylcytosine oxidation. *Genes Dev.* 25, 2436–2452. doi: 10.1101/gad. 179184.111

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary table. Table 1: Description of the clinical features of 19 hepatoblastoma cases.

ID/gender/age at diagnosis	Histology	AFP ng/ml	Risk stratification*	PRETEXT	Chemoterapy Protocol	Transplant	Metastasis	Relapse	Overall Survival	Premature (low birth weight)	Other features
HB15, F, 18m	Epithelial Embryonal	5668000	Intermediate	4	NA	Yes	No	No	1 year	No	.
HB17, F, 36m	Epithelial Fetal	>400000	Low	1	SIOPEL3	No	No	No	>5 years	No	.
HB18, M, 9m	Epithelial and Mesenchymal mixed	>200000	Low	3	SIOPEL3	Yes	No	No	>5 years	No	.
HB28, M, 17y	Epithelial and Mesenchymal mixed	NA	High	4	SIOPEL4	No	No	Yes	4 years	No	Hepatomegaly at birth
HB30, M, 54m	HB with HCC features - Epithelial Fetal	>1000000	High	2	SIOPEL4	Yes	Lung	Yes	5 years	No	.
HB31, M, 30m	Epithelial Fetal	742000	Low	3	NA	No	No	No	>5 years	No	Non-functional kidney
HB32, F, 36m	Epithelial and Mesenchymal mixed	9328000	High	4	SIOPEL4	Yes	Lung	No	>5 years	No	.
HB33, F, 1m	Epithelial Embryonal and Fetal	28312000	Intermediate	2	SIOPEL3	No	No	No	>5 years	No	Congenital HB and renal agenesis
HB34, F, 19m	Epithelial Fetal	416430	Intermediate risk	3	SIOPEL3	No	No	No	>5 years	No	Mother with Hepatitis C in pregnancy (no treatment)
HB35, M, 26m	Epithelial Fetal	54800	Intermediate risk	3	SIOPEL3	No	No	No	>5 years	No	.
HB36, M, 31m	Epithelial Embryonal and Fetal	76348	Low risk	3	AHEP 0731 - COG	No	No	No	>3 years	Yes	.
HB37, F, 13m	Epithelial Embryonal	1870000	Intermediate risk	2	SIOPEL3	No	No	No	>9 years	No	.
HB38, F, 147m	Epithelial Fetal	643,4	High risk	4	SIOPEL3 and AHEP 0731 - COG	No	No	Yes	>2 years	Yes	Ischemic anoxic neuropathy due to extreme prematurity
HB39, M, 84m	Epithelial with Macrotrabecular pattern	300000	High risk	2	SIOPEL2	No	No	No	6 months	No	Pilocytic astrocytoma (posterior fossa) after HB
HB40, M, 22m	Epithelial Embryonal and Fetal	1842	Low risk	1	SIOPEL3	No	No	No	>5 years	No	.
HB42, M, 45m	Epithelial Fetal	1267	Low risk	1	SIOPEL2	No	No	No	>5 years	No	.
HB43, M, 20m	Epithelial Embryonal	183476	Intermediate risk	4	SIOPEL3	Yes	No	No	>5 years	No	.
HB44, M, 5m	Epithelial and Mesenchymal mixed	300000	Intermediate risk	2	SIOPEL2	No	No	No	>5 years	No	.
HB45, F, 5m	Epithelial Fetal	445611	Low risk	2	SIOPEL3	No	No	No	1 year and 1 month	Yes	.

F: female; M: male; NA: data not available; AFP: Alphafeto protein

*According to the CHIC criteria (MEYERS et al., 2017)

Capítulo V. Papel de marcadores de diferenciação hepática e padrão de metilação de DNA para estratificação de hepatoblastomas

A heterogeneidade intertumoral é um grande desafio para o tratamento do hepatoblastoma e, dado sua raridade, a limitação no número de amostras é um fator que dificulta seu estudo [116]. Além disso, por apresentar poucas mutações *driver* recorrentes, tanto a estratificação de risco como a adequação do tratamento ainda são discussões recorrentes neste tumor. Para oferecer alternativas a estas questões, alguns grupos têm proposto subclassificações moleculares, com o uso de assinaturas gênicas[117,118]. Em 2008, Cairo et al. propôs a classificação de agressividade tumoral baseado no padrão de expressão de 16 genes, denominado assinatura C1 e C2; em resumo, foi evidenciado que a β-catenina está elevada em ambas as assinaturas, sendo que a diferença principal ocorre na expressão de marcadores de hepatócitos (*CYP2E1*, *HPD* e *AQP9*), ciclo celular (*E2F5*, *BUB1*, *DLG7*), células hepáticas imaturas (*AFP*, *NLE1* e *DUSP9*), metabolismo celular (*APOC4*, *APCS*, *ALDH2* e *GHR*), bem como genes associados a via de complemento (*C1S*), imunoglobulina (*IGSF1*) e proteína ribossomal (*RPL10A*) [117]. Posteriormente, Sumazin et al. propôs três subtipos moleculares para estratificação de risco do hepatoblastoma; em resumo, estes subtipos apresentaram diferenças de expressão de marcadores de células hepáticas imaturas (*LIN28B*, *SALL4*, *AFP*), e genes associados a vias amplamente alteradas no câncer (*TERT e TP53*), metabolismo (*NFE2L2*) e sistema hepatobiliar (*NOCHT1* e *HNF*) [118]. Por outro lado, Hook et al. sugere uma assinatura composta por quatro genes (*HSD17B6*, *ITGA6*, *VIM* e *TOP2A*) para classificar o hepatoblastoma em três subgrupos: C1, C2A e C2B, sendo o gene *TOP2A* proposto como um biomarcador de hepatoblastoma altamente proliferativo [55].

Apesar destas assinaturas já propostas, os determinantes genéticos da heterogeneidade clínica deste câncer ainda não estão claros e, vale ressaltar, os biomarcadores prognósticos sugeridos nestes estudos anteriores não apenas apresentaram limitações na reprodutibilidade de sua correlação com os resultados do tratamento, mas também não são claramente compreendidas as suas implicações biológicas[58].

A partir da evidência de alteração da metilação em hepatoblastoma, o nível dessa marca epigenética tem sido recentemente usada em novos modelos para classificação tumoral. Uma análise do grau de metilação indicou que hepatoblastomas apresentam duas assinaturas epigenéticas (Epi-CA e Epi-CB) [57]. Além disso, a amplificação do *locus*

oncogênico 14q32 *DLK1-DIO3* em parte dos hepatoblastomas permitiu delinear assinaturas de expressão moderada ou forte de genes mapeados nesta região. Usando ambos os achados, uma estratificação de risco molecular de três categorias foi proposta (MRS-HB)[57]. Usando uma análise multiômica integrada de hepatoblastoma, Sekiguchi et al. propôs ainda outra classificação de hepatoblastomas em três grupos de acordo com o nível de metilação: F, E1 e E2, realizando associação dos subgrupos aos estágios de diferenciação de hepatócitos[58].

Considerando a importância da epigenética no contexto de tumores pediátricos, uma vez que apresentam número de mutações somáticas reduzidas comparada a tumores em adultos[4], nosso grupo avaliou a utilização do padrão de expressão de genes da maquinaria epigenética como elementos-chave para a estratificação de hepatoblastomas. Com base na teoria do bloqueio da diferenciação como origem dos tumores embrionários[9], também utilizamos para esta análise marcadores das fases de diferenciação hepática (célula-tronco, endoderme definitiva, hepatoblasto e hepatócito), a fim de determinar em qual delas teria ocorrido o bloqueio no desenvolvimento e subsequente gênese tumoral.

Desta forma, este estudo teve por objetivo classificar hepatoblastomas de acordo com 16 marcadores de diferenciação hepática (*AFP*, *ALB*, *CK18*, *CK19*, *CXCR4*, *CYP3A4*, *FOXA2*, *G6PC*, *HNF1A*, *HNF4A*, *NANOG*, *POU5F1*, *SOX17*, *TDO2*, *UGTA1* e *SOX2*) e oito genes associados à maquinaria de DNAm (*TET1*, *TET2*, *TET3*, *DNMT1*, *DNMT3A*, *DNMT3B*, *DNMT3L* e *UHRF1*). Neste estudo, foi possível a estratificação das amostras tumorais em três grupos, destacando a importância de 13 genes para o estabelecimento desta classificação. Os resultados permitiram correlacionar os grupos identificados com distintos níveis de metilação global de DNA, que estão de acordo com a fase de diferenciação hepática em que as células tumorais se encontrariam, o que corrobora a hipótese de bloqueio de diferenciação do órgão de origem. Deste modo, o trabalho evidencia novamente a importância da maquinaria epigenética no desenvolvimento destes tumores e também como ferramenta de estratificação do hepatoblastoma, assim como propõe assinaturas tumorais associadas a diferentes fases de diferenciação hepática.

Os dados acima descritos foram publicados no artigo *DNA methylation as a key epigenetic player for hepatoblastoma characterization (DOI:* 10.1016/j.clinre.2021.101684*)*, na revista *Clinics and Research in Hepatology and Gastroenterology,* sendo apresentado a seguir.

Clinics and Research in Hepatology and Gastroenterology 45 (2021) 101684

Available online at ScienceDirect www.sciencedirect.com

Elsevier Masson France EM|consulte www.em-consulte.com/en

ORIGINAL ARTICLE

DNA methylation as a key epigenetic player for hepatoblastoma characterization

Maria Rivas [a,1], Talita Aguiar [a,b,1], Gustavo Fernandes [a], Renan Lemes [a], Luiz Caires-Júnior [a], Ernesto Goulart [a], Kayque Telles-Silva [a], Mariana Maschietto [c], Monica Cypriano [d], Silvia de Toledo [d], Dirce Carraro [e], Isabela da Cunha [f], Cecilia da Costa [g], Carla Rosenberg [a], Ana Krepischi [a,*]

[a] *Human Genome and Stem Cell Research Center, Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, São Paulo, SP, Brazil*
[b] *Department of Urology - NYU Grossman School of Medicine, New York City, NY, USA* [c] *Research Center, Boldrini Children's Hospital, Campinas, Brazil*
[d] *Department of Pediatrics, Adolescent and Child with Cancer Support Group (GRAACC), Federal University of São Paulo, SP, Brazil*
[e] *International Center for Research, A. C. Camargo Cancer Center, SP, Brazil* [f] *Pathology Department, Rede D'OR-São Luiz, SP, Brazil*
[g] *Department of Pediatric Oncology, A. C. Camargo Cancer Center, SP, Brazil*

Available online 20 April 2021

KEYWORDS
Hepatoblastoma;
Stratification;
Epigenetics;
Methylation;
Hepatocyte differentiation

Abstract
Background: Hepatoblastoma (HB) is a rare embryonal liver tumor of children. Although intrinsic biological differences between tumors can affect prognosis, few groups have studied these differences. Given the recent increased attention to epigenetic mechanisms in the genesis and progression of these tumors, we aimed to classify HB samples according to the stages of liver development and DNA methylation machinery.
Basic procedures: We evaluated the expression of 24 genes associated with DNA methylation and stages of hepatocyte differentiation and global DNA methylation. Using bioinformatics tools and expression data, we propose a stratification model for HB.
Main findings: Tumors clustered into three groups that presented specific gene expression profiles of the panel of DNA methylation enzymes and hepatocyte differentiation markers. In addition to reinforcing these embryonal tumors' molecular heterogeneity, we propose that a

Abbreviations: HB, hepatoblastoma; SIOPEL, International Childhood Liver Tumours Strategy Group; COG, Children Oncology Group; GRAACC, Support Group for Adolescents and Children with Cancer; iPSC, induced pluripotent stem cell; CHIC, Children's Hepatic Tumor International Collaboration; HCC, hepatocellular carcinoma; Epi-CA/CB, epigenetic signatures of cluster A and B; MRS-HB, molecular risk stratification for hepatoblastoma.

* Corresponding author.
E-mail address: ana.krepischi@ib.usp.br (A. Krepischi).
[1] These two authors contributed equally.

https://doi.org/10.1016/j.clinre.2021.101684

panel of 13 genes can stratify HBs (*TET1*, *TET2*, *TET3*, *DNMT1*, *DNMT3A*, *UHRF1*, *ALB*, *CYP3A4*, *TDO2*, *UGT1A1*, *AFP*, *HNF4A*, and *FOXA2*). DNA methylation machinery participates in the characterization of HBs, directly reflected in diverse DNA methylation content. The data suggested that a subset of HBs were similar to differentiated livers, with upregulation of mature hepatocyte markers, decreased expression of DNA methylation enzymes, and higher global methylation levels; these findings might predict worse outcomes.

Conclusions: HBs are heterogeneous tumors. Despite using a small cohort of 21 HB samples, our findings reinforce that DNA methylation is a robust biomarker for this tumor type.

Introduction

Hepatoblastoma (HB) is a rare liver tumor occurring in childhood [1]. Hypothetically, it arises from hepatocyte precursors [2]. There is a global trend toward increased incidence of HB in recent years [3]. The increase in premature birth may be responsible, as prematurity correlates with low birth weight, a known risk factor for HB [4]. The overall 5-year survival rate of children with HB is 70% [5,6]; however, patients who do not respond to standard treatment have very poor survival [7—10]. HBs are heterogeneous, with several histologic types [2] that frequently display combinations of epithelial, stromal, mesenchymal, and undifferentiated cells [11]. Despite this clinical heterogeneity, a few recur-rent genetic lesions are reported in HBs that present the lowest burden of somatic mutations [12]. Epigenetic mechanisms appear to be critical factors for the genesis and progression of this tumor. In our previous work [13], an evaluation of HB methylomes revealed a lower level of DNA methylation in non-repetitive sequences than non-tumoral livers, a finding corroborated by others [14]. Hypermethylation in CpG islands are mapped to genes associated with tumor suppression, lipid metabolism, and liver differentiation. We also identified changes in gene expression of the epigenetic machinery in HBs that support a model of an active demethylation process in this tumor, probably during the early stages of liver development [15].

HB is an embryonal tumor, and the central hypothesis of origin proposes that cancer arises from a failure of cell differentiation, leading to the disruption of the normal process of organ development [16]. Therefore, some childhood tumors might mimic embryonic cellular activity; as an example, in most HBs, there are elevated levels of a-fetoprotein that the fetal liver produces under normal circumstances [17].

Based on the impact of molecular differences on tumor prognosis, several models of HB classification have been proposed. Evaluating the expression pattern of 16 genes related to hepatocyte markers, cell cycle regulation, hepatic stem/progenitor markers, and metabolism, Cairo et al. classified HBs in two groups, called C1 and C2 signatures [18]. Sumazin et al. showed that this model did not predict outcomes when applied in other cohorts; they presented a HB risk stratification based on the differential activation of hepatic stem/progenitor markers, genes related to cancer pathways, metabolism, and the hepatobiliary system [19]. Another study proposed an alternative signature comprising four genes, classifying HB in three subgroups (C1, C2A, and C2B) [20]. More recent comprehensive studies added methylation data to HB classification models. Using this epigenetic mark, investigators stratified HBs into two signatures (Epi-CA and Epi-CB) [21]. The evaluation of the 14q32 *DLK1-DIO3* locus delineated signatures of moderate or high expression of genes mapped to this region; using both findings, researchers created a molecular risk stratification with three categories (MRS-HB). Finally, applying an integrated multi-omics analysis, Sekiguchi et al. proposed an HB classification in three groups according to methylation level (F, E1, and E2), associating these subgroups with hepatocyte differentiation stages [22].

Here, we focus on DNA methylation and a model of embryonal tumors' origin to classify HB samples according to the different stages of liver development. We based this model on gene expression patterns of markers of hepatocyte differentiation combined with enzymes of the DNA methylation machinery and global methylation levels. Using these data, we provide a framework to stratify HBs, correlating the molecular biomarkers with clinical features.

Material and methods

Patients

From the biobanks from A.C. Camargo Cancer Center and GRAACC, we recovered fresh-frozen primary HBs (n = 21) and matched non-tumoral liver tissues (n = 9) derived from tumor resection surgeries. All patients (n = 21) received pre-surgery chemotherapy accord ing to either SIOPEL (http://www.siopel.org/) or COG (https://www.childrensoncologygroup.org/) protocols, and clinicians followed them with clinical examination, imaging tests, and measurements of alpha-fetoprotein levels. The Ethics Committees of the A.C. Camargo Cancer Center and GRAACC approved the study. We collected biological samples after obtaining written informed consent from parents.

Cell cultures

We obtained induced pluripotent stem cell lines (iPSC) derived from three control individuals' human peripheral blood following the procedure reported by Okita et al. [23]. Hepatic differentiation was performed in these iPSC lines as described [24] to collect cells at various differentiation

stages (definitive endoderm, hepatoblasts, and precursor hepatocyte cells).

RNA isolation

We obtained RNA samples from the biobank of the A.C. Camargo Cancer Center and GRAACC; samples were extracted from tissues using the QIASymphony DNA Mini kit (*QIAGEN*) and the RNeasy Mini Kit (*QIAGEN*), according to the manufacturer's recommendations. The quality of RNA samples was verified using microfluidics-based electrophoresis (Bioanalyzer, Agilent Technologies; CA, USA), and only RNA samples with RNA Integrity Number (RIN) > 7.0 were used.

Gene expression analysis

Expression of 24 genes was evaluated by qRT-PCR (AFP, ALB, CK18, CK19, CXCR4, CYP3A4, FOXA2, G6PC, HNF1A, HNF4A, NANOG, POU5F1, SOX2, SOX17, TDO2, UGTA1, TET2, TET3, DNMT1, DNMT3A, DNMT3L, SOX2, TET1, and UHRF1 - primers upon request) using SYBR Green system. We performed all runs in the QuantStudio 5 (Applied Biosystems, USA). The cDNA was synthesized using the High-Capacity RNA-to-cDNA kit (Applied Biosystems, USA). Experiments were based on the International Guideline MIQE (The Minimum Information for Publication of Quantitative Real-Time PCR Experiments). The housekeeping genes were chosen using the geNorm algorithm [25] after expression analysis of ACTB, GAPDH, B2M genes, and 18S ribosomal RNA (18S rRNA). Data were normalized using the housekeeping genes' expression values 18S rRNA and B2M, and all reactions were performed with three technical replicates. The delta-delta Ct (Ct) method was used for data analysis [26].

Statistical analysis

Statistical analysis of gene expression data was applied to $\log_2$ -transformed values [27]. A gene was considered differentially expressed between tumors and controls if the p-value was lower than 0.05 (FDR-corrected ANOVA). Sample clustering was performed using the k-means algorithm with 1000 random permutations and four different centroids. Centroid numbers were estimated using the elbow method and UPGMA hierarchical clustering; input data were scaled, centered and PCA transformed input data before the clustering procedure. Differentially expressed genes were identified by analysis of variance (ANOVA) [28] between sample sets 1, 2 and 3. We performed all analyses using R (version 3.4.4), with the following packages: STRINGR 1.4.0, GGPLOT2 2.3, GRID 3.4.3, gridExtra 2.3, and GGBIPLOT 0.55.

Pairwise gene expression correlation was calculated using Pearson's method and the data was displayed using the heatmap.2 function (packages: STRINGR 1.4.0, GGPLOT2 2.3, GRID 3.4.3, gridExtra 2.3, and GGBIPLOT 0.55) with the same setting as the previously described UPGMA.

Unsupervised hierarquical clustering was performed using gene expression data from previous studies [18—20] with the heatmap function from the R Statistical Package v4.0.3 using default parameters such as Euclidean distance and the complete linkage method.

Methylation analysis and global methylation level

We recovered DNA methylation data from our previous study with the same cohort of HB samples using the HM450K methylation array (Illumina) [13]. DNA samples were extracted using phenol-chloroform and bisulfite conversion was performed using the EZ DNA Methylation kit (Zymo Research). The global mean methylation level was calculated as the average of all normalized beta-values of the probes in the HM450K platform for each sample.

Spearman correlation analysis of gene expression ($\log_2$ relative expression) and DNA methylation (M-value) was applied for each gene using the average of beta-values derived from all CpG sites mapped in the promoter (located at TSS150, TSS200, promoter, first exon regions). To evaluate the 14q32 *DLK1-DIO3* locus, we calculated the average of beta-values derived from all CpG sites mapped in this region and applied a *t*-test to compare the methylation level of each set of HB with the non-tumoral livers.

In silico analysis

To explore the biological pathways related to the studied genes, we used the Reactome Pathway Database (https://reactome.org/ — Version 72 Released), with the Homo sapiens species (GRCh37/hg19) as reference. A binomial test was used to calculate the probability for each result and the *p*-values were corrected for multiple testing (Benjamini—Hochberg procedure) [29,30] with a significant adjusted *p*-value < 0.05. We subjected 13 genes (*TET1*, *TET2*, *TET3*, *DNMT1*, *DNMT3A*, *UHRF1*, *ALB*, *CYP3A4*, *TDO2*, *UGT1A1*, *AFP*, *HNF4A*, and *FOXA2*) to the Reactome Pathways analysis (with and without interactors), searching for pathways related to biological mechanisms as well as to cancer.

Results

We analyzed 21 HB samples from different patients (13 males and eight females). All tumors were primary lesions surgically removed after neoadjuvant chemotherapy. The average age at diagnosis was 40 months, and four of these patients exhibited pulmonary metastases at diagnosis. Table 1 presents the details of clinical characteristics and previously detected somatic driver mutations (data derived from Aguiar et al., 2020 [31]). *CTNNB1* alterations were detected in a 45% frequency in this HB cohort; because this evaluation derived from exome data and Sanger sequencing of DNA samples, we could not exclude the possibility that very low-frequency *CTNNB1* deletions/mutations are missing.

Gene expression of differentiation markers was evaluated for specific stages: stem cells (*NANOG*, *POU5F1*, *SOX2*), definitive endoderm (*CXCR4*, *FOXA2*, *SOX17*), hepatoblasts (*AFP*, *CK19*, *HNF4A)*, and hepatocytes (*ALB*, *CYP3A4*, *CK18*, *HNF1A*, *G6PC*, *TDO2*, *UGT1A1*). In addition, the gene expression data of genes associated with the DNA methylation machinery (*DNMT1*, *DNMT3A*, *DNMT3B*, *DNMT3L*, *TET1*, *TET2*, *TET3*, *UHRF1*) were recovered from our previous study [15]. Pairwise correlation analysis of the 24 genes expression pointed to three groups with related expression patterns

Table 1 Clinical features of 21 hepatoblastoma cases.

Set	ID	Age at diagnosis	Gender	Histopathology	PRETEXT	Metas-tasis	Risk stratification[a]	Transplant	Relapse	Deceased	*TERT* pro-moter mutations[b]	*CTNNB1* mutations[b]
SET1	HB33	1 month	F	Epithelial embryonal and fetal	2	No	Intermediate	No	No	No	Negative	G34V
SET1	HB35	26 months	M	Epithelial fetal	3	No	Intermediate	No	No	No	Negative	S29F/D32A
SET1	HB43	20 months	M	Epithelial embryonal	4	No	Intermediate	Yes	No	No	Negative	A5 Y142del
SET2	HB18	9 months	M	Epithelial and mesenchymal mixed	3	No	Low	Yes	No	No	Negative	T41A
SET2	HB28	17 years	M	Epithelial and mesenchymal mixed	4	No	High	No	Yes	Yes	C250T	Negative
SET2	HB32	36 months	F	Epithelial and mesenchymal mixed	4	Lung	High	Yes	No	No	Negative	Negative
SET2	HB37	13 months	F	Epithelial embryonal	2	No	Intermediate	No	No	No	DNA not avaible	DNA not avaible
SET2	HB38	12 years	F	Epithelial fetal	4	No	High	No	Yes	No	Negative	Negative
SET2	HB40	22 months	M	Epithelial embryonal and fetal	1	No	Low	No	No	No	Negative	A21 Ser33del
SET2	HB42	45 months	M	Epithelial fetal	1	No	Low	No	No	No	Negative	Negative
SET2	HB44	5 months	M	Epithelial and mesenchymal mixed	2	No	Intermediate	No	No	No	Negative	Negative
SET2	HB45	5 months	F	Epithelial fetal	2	No	Low	No	No	Yes	Negative	A5 A80del
SET3	HB15	18 months	F	Epithelial embryonal	4	No	Intermediate	Yes	No	Yes	Negative	A5 A80del

Table 1 (*Continued*)

Set	ID	Age at diagnosis	Gender	Histopathology	PRETEXT	Metastasis	Risk stratification[a]	Transplant	Relapse	Deceased	*TERT* promoter mutations[b]	*CTNNB1* mutations[b]
SET3	HB17	36 months	F	Epithelial fetal	1	No	Low	No	No	No	Negative	Negative
SET3	HB30	54 months	M	HB/HCC features	2	Lung	High	Yes	Yes	Yes	C250T	Negative
SET3	HB31	30 months	M	Epithelial fetal	3	No	Low	No	No	No	Negative	A5 Y142del
SET3	HB34	19 months	F	Epithelial fetal	3	No	Intermediate	No	No	No	Negative	Negative
SET3	HB36	31 months	M	Epithelial fetal	3	No	Low	No	No	No	Negative	Negative
SET3	HB39	7 years	M	Epithelial with macrotrabecular focus	2	No	High	No	No	Yes	Negative	Negative
SET3	HB41	21 months	M	Epithelial fetal	3	Lung	High	No	No	No	Negative	Negative
SET3	HB46	28 months	M	Epithelial and mesenchymal mixed	4	Lung	High	No	No	No	Negative	G34E

F: female; M: male.

[a] According to the CHIC criteria [47,48].

[b] The characterization of CTNNB1 and TERT mutations was recovered from our previous work (Aguiar et al., 2020 [31]), based on whole-exome data and/or Sanger sequencing of exons 3 and 4.

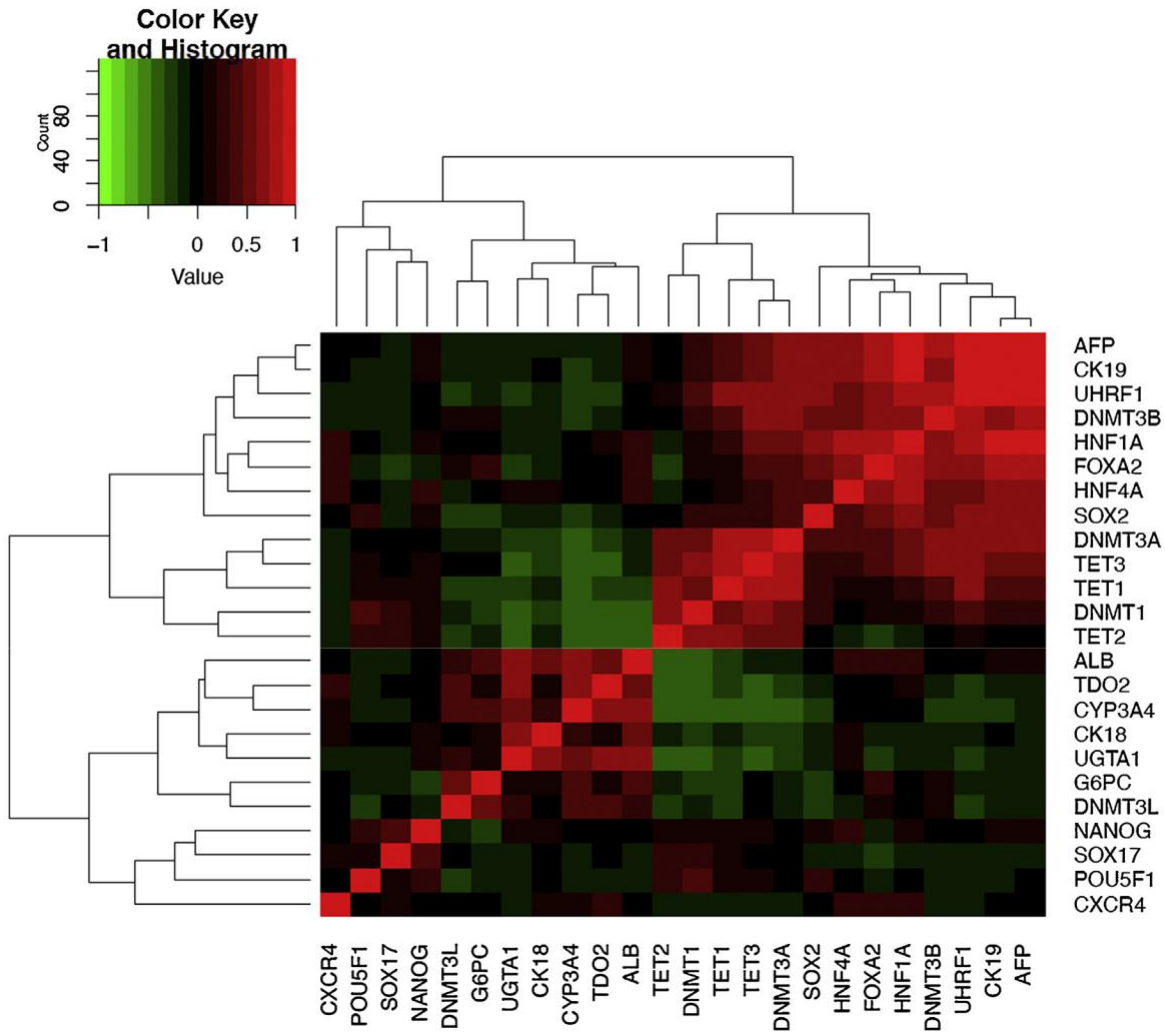


Figure 1 Pairwise gene expression correlation. For each gene (24), based on the average of samples in each compendium (full and partial), we measure for pairwise gene association using Pearson's correlation, and the data was displayed using the heatmap.2 function. The analysis evidenced three groups of genes with similar expression patterns: 1. *AFP*, *CK19*, *UHRF1*, *DNMT3B*, *HNF1A*, *FOXA2*, *HNF4A*; 2. *SOX2*; *DNMT3A*, *TET3*, *TET1*, *DNMT1*, and 3. *TET2*; *ALB*, *TDO2*, *CYP3A4*, *CK18*, and *UGT1A1*. In the heatmap, the green color represents a reduction in gene expression, while the red color indicates an increase in expression.

(Fig. 1): (1) *AFP*, *CK19*, *UHRF1*, *DNMT3B*, *HNF1A*, *FOXA2*, *HNF4A*, *SOX2*; (2) *DNMT3A*, *TET3*, *TET1*, *DNMT1*, *TET2*; and (3) *ALB*, *TDO2*, *CYP3A4*, *CK18*, and *UGT1A1*. Interestingly, gene expression of mature hepatocyte markers (1) presented an inverse correlation to genes' pattern from the epigenetic machinery (2). The next step was to determine whether these differences in the gene expression pattern would occur between the samples, enabling their clustering.

Based on the similarity of the gene expression profiles, samples were clustered into four groups (Fig. 2), and the ideal number of groups was evaluated using UPGMA hierarchical clustering (Supporting Information Figure S1A and S1B) and K-means (Supporting Information Figure S2). The Set-1 cluster (in black) consisted of hepatoblasts and hepatocyte-like cells and only three intermediate-risk HBs (33T, 35T, and 43T), suggesting a gene expression profile similar to transitional phases of hepatocyte differentiation. A second group (Set-2, in red) contained nine HB samples (18T, 28T 32T, 37T, 38T, 40T, 42T, 44T, and 45T), a group of clinically heterogeneous tumors, with two cases of late diagnosis, one of them carrying a somatic mutation in the *TERT* promoter. The larger group (Set-3, in green) clustered all nine non-tumoral liver samples plus nine HBs (15T, 17T, 30T, 31T, 34T, 36T, 39T, 41T, and 46T), suggesting that these tumors exhibited gene expression profiles similar to those of differentiated livers. Six of nine patients in the Set-3 were classified as intermediate/high-risk: three died from the disease, three developed pulmonary metastases, and one had HB/HCC features (carrying a *TERT* promoter mutation). The Set-2 cluster presented a partial overlap to Set-1 and Set-3, pointing to mixed characteristics of gene expression in a more advanced hepatocyte differentiation phase than the Set-1 group. Finally, we detected an isolated group consisting of only the iPSC and the definitive endoderm cells (Set-4, in blue), exhibiting the most distinctive gene expression pattern when compared to the other groups, as expected.

We evaluated the clinical parameters of risk stratification regarding tumor distribution in the three sets; there were no

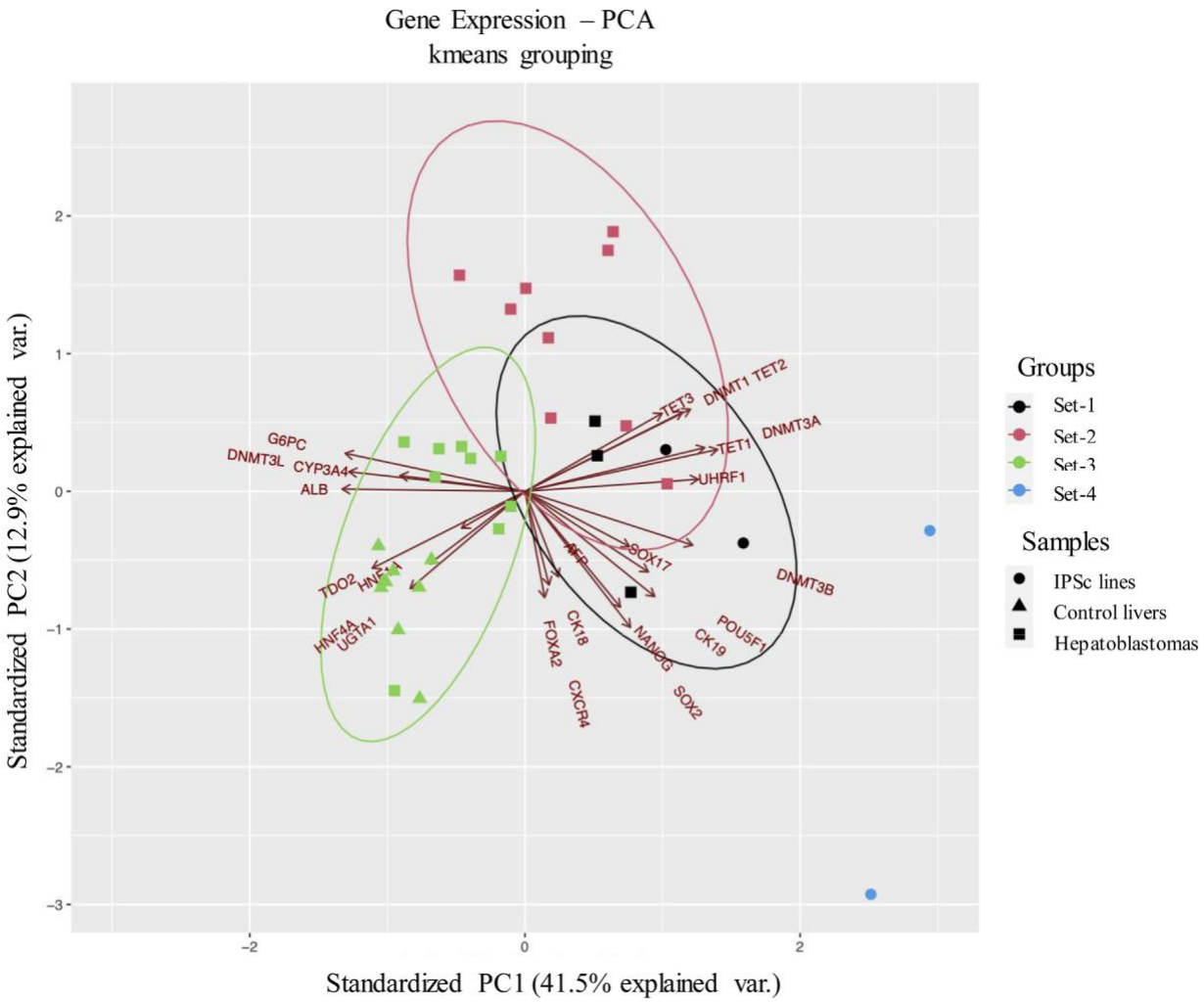


Figure 2 Analysis of PCA K-means of hepatoblastomas and control liver samples based on their expression profile of 24 genes. The analysis was performed grouping HBs and non-tumoral liver samples with similar gene expression levels, which resulted in four clusters, indicated by color legends. The triangles indicate samples of non-tumoral tissues; the squares indicate HB samples; circles indicate samples derived from cell culture (iPSC lines). The samples grouped in Set-1 are represented in black; samples grouped in Set-2 are in red; in green, samples grouped in Set-3; in blue, samples grouped in Set-4.

significant associations. Set-1 presented a predominance of embryonal epithelial tumors, Set-2 presented mixed epithelial and mesenchymal, and Set-3 presented fetal epithelial samples. Although Set-3 appeared to consist of tumors with the worst parameters (death by disease, metastasis, and *TERT* mutations), the Kaplan–Meier survival curve did not show a significant association (Supporting Information Figure S3).

The expression analysis indicated that only 13 out the 24 genes presented significant differences between Set-1, Set-2, and Set-3 groups (Supporting Information Figure S4): two hepatoblasts markers (*AFP* and *HNF4*A – Fig. 3A and B), a definitive endoderm marker (*FOXA2* – Fig. 3C), four hepatocyte markers (*ALB, CYP3A4, TDO2*, and *UGT1A1* – Fig. 3D–G), and six genes of the DNA methylation machinery (*TET1, TET2, TET3, DNMT1, DNMT3A,* and *UHRF1* – Fig. 3H–M). The gene expression analysis suggested that Set-1 tumors had upregulation of hepatoblast and definitive endoderm markers and high expression of DNA methylation genes, primarily *UHRF1* and *TET1*. The Set-2 tumors exhibited upregulation of the DNA methylation genes and low expression of markers of both the intermediary stages of differentiation and mature hepatocyte genes. Finally, in Set-3, samples showed downregulation of the DNA methylation genes, associated with high expression of mature hepatocyte gene markers.

The comparison of the global DNA methylation between the three groups (Fig. 4A) showed that Set-3 presented over-all methylation levels similar to non-tumoral liver samples, in agreement with expression data. The samples from the Set-1 that presented expression profiles similar to embryonic stages of the hepatocyte differentiation had lower global DNA methylation levels than the other two groups. The Set-2 samples containing HBs with transitional gene expression profiles compared to the others also had intermediate mean levels of DNA methylation.

We performed a correlation analysis of gene expression and DNA methylation levels in promoter regions

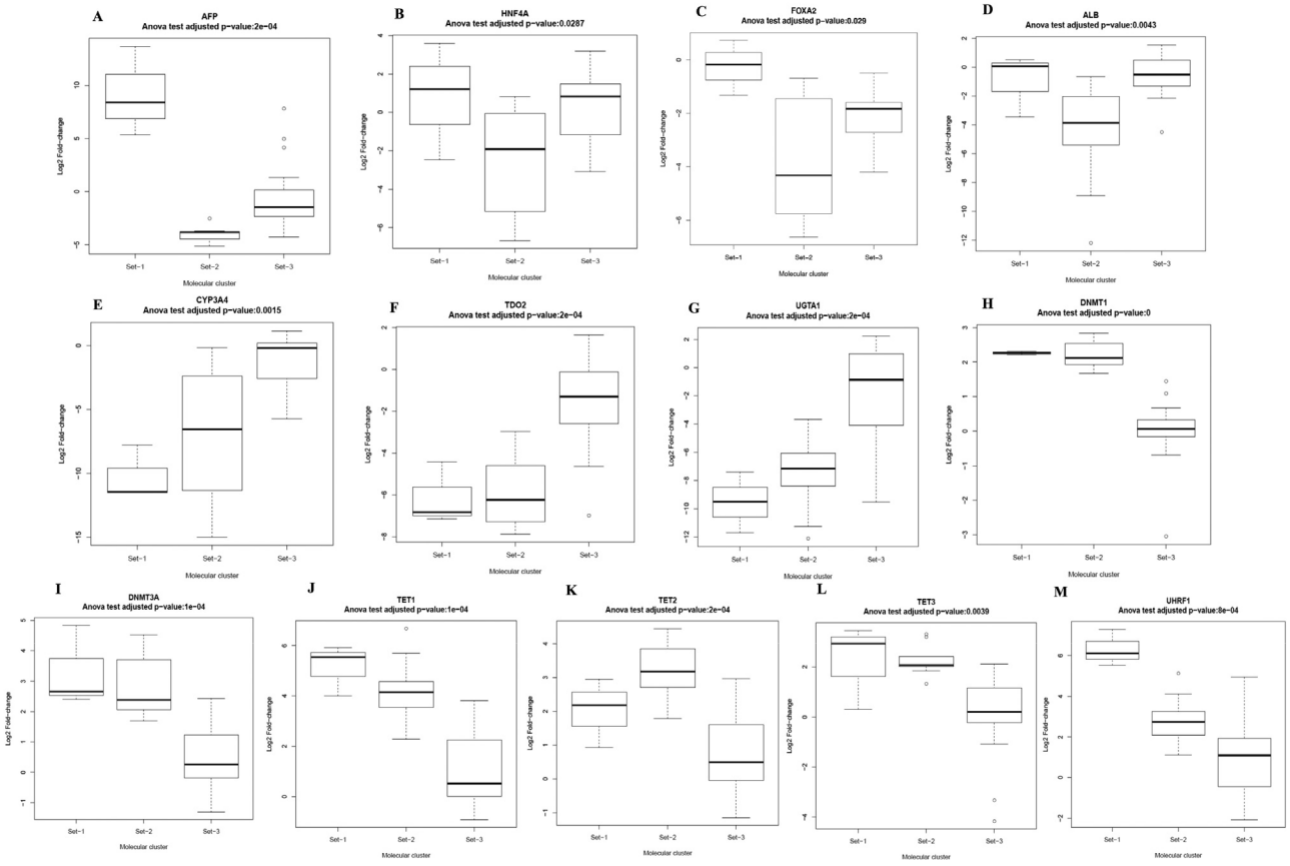


Figure 3 Analysis of gene expression of 13 genes in hepatoblastoma groups according to their clustering. Two genes were associated hepatoblast markers (*AFP* and *HNF4A* — Fig. 3A and B), a definitive endoderm marker (*FOXA2* — Fig. 3C), four hepatocyte markers (*ALB*, *CYP3A4*, *TDO2*, and *UGT1A1* — Fig. 3D—G) and six with epigenetic machinery (*TET1*, *TET2*, *TET3*, *DNMT1*, *DNMT3A*, and *UHRF1* — Fig. 3H—M). The gene expression analysis evidenced that Set-1 tumors have upregulation of hepatoblast and definitive endoderm markers and high expression of the DNA methylation genes, mainly *UHRF1* and *TET1*. The Set-2 tumors exhibit upregulation of the DNA methylation genes and low expression of markers of both the intermediary stages of differentiation and mature hepatocyte genes. Set-3, samples show down-regulation of the DNA methylation genes, associated with high expression of mature hepatocyte marker genes.

for the 13 genes that presented differential expression between the three Sets. Although not highly correlated, two genes exhibited significant results (Supporting Information Table S1). *FOXA2* expression negatively correlated with the DNA methylation level of the promoter region (Spear-man Rank Correlation: −0.632; $p < 0.008$), while *HNF4A* positively correlated (Spearman Rank Correlation: 0.6421; $p < 0.007$).

Using the list of 13 genes, the Reactome Pathways analysis (without interactors) detected nine pathways (Supporting Information Table S2), seven of which involved epigenetic processes and mechanisms. Expanding the Reactome Pathways analysis with the 13 genes' interactors, 158 pathways related to biological mechanisms were identified (Supporting Information Table S3). The top five had associations with mRNA processing/splicing, apoptosis, and protein localization. Considering cancer pathways, the top five out of 33 significant pathways (Supporting Information Table S4) were endosomal/vacuolar pathway, *TP53* regulation of transcription of DNA repair genes, signaling by *BRAF* and *RAF* fusions, constitutive signaling by aberrant *PI3K* in cancer, and oncogenic MAPK signaling. We also evaluated the genes' intersection in these top pathways (with and without interactors) with the list of genes mapped to the previously detected differentially-methylated regions in the same HB group [13]. Only five genes exhibited altered DNA methylation in HBs; two were hypomethylated (*RXRA* and *NUAK1*), and three were hypermethylated (*BMP4*, *AHRR*, and *MAP3K8*).

Additionally, we explored the 14q32 *DLK1-DIO3* locus's methylation data, previously reported to present predominant DNA hypomethylation in HBs [21]. Evaluating the 32 CpG sites at 14q32 that Carrillo-Reixach et al. identified as hypomethylated in HBs, we found 13 CpGs significantly hypomethylated in tumors in our data with three of them (cg13551098; cg08051604; cg18910298) exhibiting methylation changes >20% (Supporting Information Figure S5). We also calculated the average of the beta-values of methylation for each set of tumors, and performed a *t*-test comparing each set with control livers [Set-1 ($p < 3.46E{-}09$), Set-2 ($p < 6.35E{-}07$) and Set-3 ($p < 0.002362$)]. This analysis showed that Set-3 HBs presented the smallest difference in the DNA methylation level at the 14q32 *DLK1-DIO3* region compared to control livers (Fig. 4B).

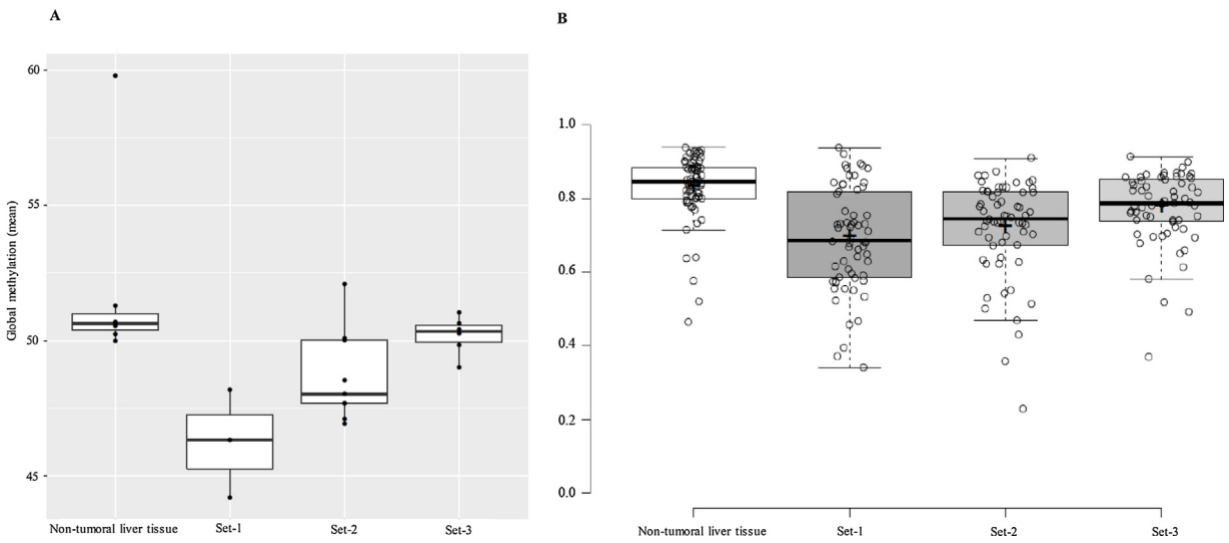


Figure 4 Methylation analysis of the three different clusters of hepatoblastomas and liver control samples. (A) Global methylation level. In the y-axis, the percentage of methylation is indicated for each sample (dots). The boxplots referring to the overall methylation level show that the average content of this epigenetic mark differs between different HB clusters and non-tumoral liver tissues. (B) DNA methylation evaluation of the 14q32 DLK1-DIO3 locus. The different HB sets and respective differences in comparison to the non-tumoral liver tissues are depicted: Set-1 (*p*-value < 3.46E−09), Set-2 (*p*-value < 6.35E-07), and Set-3 (*p*-value < 0.002362); the *y*-axis corresponds to the beta-values of DNA methylation ranging from 0 to 1.

The proposed gene signature was evaluated in other HB cohorts using previously published expression data from Cairo et al. [18], Sumazin et al. [19], and Hooks et al. [20] (Supporting Information Figure S6). Using RNA-Seq data from Hooks et al. [20], the panel of 13 genes grouped samples in three clusters, one exclusively composed of control livers and two others with HBs; in one of these HB clusters, tumors clustered with a large number of control samples (Supporting Information Figure S6A), a finding that is similar to the Set-3 we described. The analyses using Cairo et al.
[18] and Sumazin et al. [19] data were incomplete because they did not cover four genes; despite this limitation, in both datasets, a subgroup of HBs also clustered with control liver samples, showing a possible resemblance with the Set-3 (Supporting Information Figure S6B and S6C).
Fig. 5 summarizes our findings and proposes the stratification of this cohort of HBs.

Discussion

The scenario of few recurrent driver mutations in HBs represents a challenge for risk stratification and adjustment of the therapeutic regimen, and for this reason, there have been proposals for molecular sub-classifications, including gene signatures [2,18,19]. Here, we focused on epigenetic markers (expression of DNMTs and TETs as well as global DNA methylation) and genes related to the stages of liver differentiation to perform HB stratification. Hepatocytes are the major cell type of the liver, accounting for ~70% of the adult organ's mass. The anterior portion of the hepatic diverticulum gives rise to the liver from endoderm cells that differentiate in the bi-potential cells known as hepatoblasts [32,33]. Liver growth and hepatocyte maturation are processes regulated by genes acting intrinsically in these cells [34]. HB might develop from hepatoblasts, the pre-cursor cells of hepatocytes [2]; however, microscopically, HBs are heterogeneous, rarely composed of only one cell type, often exhibiting combinations of epithelial, stromal, mesenchymal, and undifferentiated cells components [11]. Genes associated with DNA methylation in HBs exhibited a directly correlated expression pattern, which is inverse to the expression pattern of mature hepatocyte markers, highlighting the central role of this specific epigenetic pathway in liver differentiation and HB stratification (Fig. 1). Our data also disclosed a direct correlation of *UHRF1* and *AFP* expression in tumors previously described for HCC [35], possibly representing a new finding to be explored in HBs. Additionally, there was a similar expression pattern of *HFN4A* and *FOXA2*, which are nuclear factors that present tissue co-expression and cooperate for hepatic pathway cell commitment [36], emphasizing the critical role of the hepatic differentiation blockage in HB development.
In addition to reinforcing these embryonal tumors' molecular heterogeneity, we proposed a panel of 13 genes for this HB stratification (*TET1*, *TET2*, *TET3*, *DNMT1*, *DNMT3A*, *UHRF1*, *ALB*, *CYP3A4*, *TDO2*, *UGT1A1*, *AFP*, *HNF4A*, and *FOXA2*). Tumors could be clustered in three groups (Fig. 2), which presented specific gene expression profiles of the panel of DNA methylation enzymes and hepatocyte differentiation markers. Remarkably, we found that the DNA methylation machinery exerts a key role in the characterization of HBs, directly reflected in diverse DNA methylation content. Six of the eight genes associated with DNA methylation were determinants for HB stratification, demonstrating the importance of the epigenetic machinery for this embryonal liver cancer's biology (Fig. 3H—M).
Considering this panel of 13 genes, we observed in HBs that only *FOXA2* expression inversely correlated with the DNA methylation level in the promoter region, as expected if promoter methylation levels controlled gene expression. In

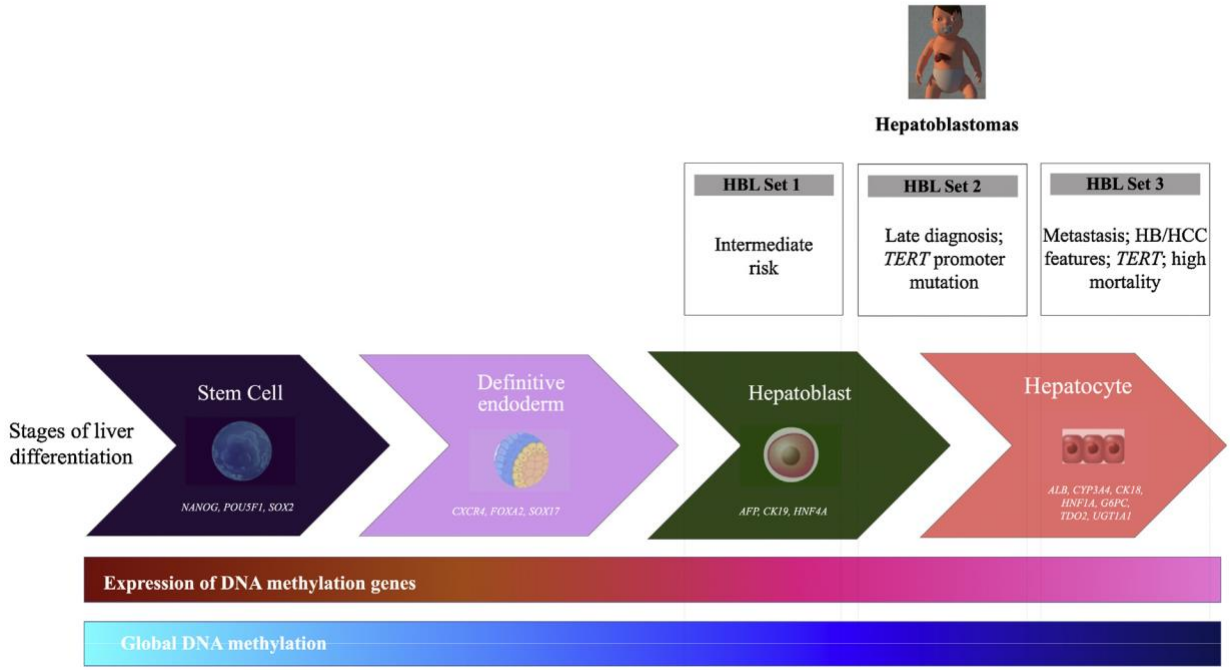


Figure 5 Hepatoblastoma stratification according to global content of DNA methylation and expression profile of DNA methylation enzymes and markers of hepatocyte differentiation. The middle depicted differentiation stages from stem cells to hepatocytes, with the respective genes used in the present study as markers of each stage. The blue bar indicates the global DNA methylation level, with the darkest color corresponding to the highest level of the epigenetic mark; similarly, the red gradient bar indicates the level of expression of DNA methylation genes (*DNMT1*, *DNMT3A*, *DNMT3B*, *DNMT3L*, *TET1*, *TET2*, *TET3*, and *UHRF1*), with the darkest one corresponding to the highest level. Clusters of HBs (Set-1, 2, and 3) are indicated on the top.

the intersection of the pathways related to the 13 genes with the previously detected differentially-methylated regions in the same HB group [13], five genes exhibited altered DNA methylation in HBs, four of which (*BMP4*, *AHRR*, *MAP3K8*, and *NUAK1*) already associated with HB or liver metabolism in general [37—40], including our previous work [31], in which we observed somatic mutations in *BMP4* and *AHRR*. In particular, the hypermethylated *RXRA* is a well-established drug target for pharmacologic interventions and therapeutic applications in cancer [41—44].

At the beginning of embryonic development, the epigenetic machinery acts by decreasing the overall level of DNA methylation, allowing the expression of genes associated with pluripotency [45]. During the organ's maturation, there is an increase in DNA methylation, a process that culminates in the selective expression of tissue specific genes associated with cell differentiation promotion. Accordingly, the iPSC and definitive endoderm cell lines cluster separately (Set-4), reinforcing the notion that HBs consist of cells already compromised with hepatocyte differentiation (Fig. 2). Sekiguchi et al. recently emphasized the relationship with the stages of hepatic differentiation. The authors showed three clusters (F, E1, and E2), with E1 and E2 presenting a gene expression and methylation pro-file similar to liver progenitor cells at a more immature state. In contrast, the group F likely arose from hepatoblasts at a relatively mature stage [22], with expression profile and DNA methylation with high similarity to the adult liver. In our analysis, the Set-3 presented downregulation of the DNA methylation genes and high expression of mature hepatocyte markers, similar to control livers; accordingly, we hypothesized that they likely derived from cells compromised with more advanced stages of hepatocyte differentiation. We also observed a global DNA methylation content in Set-3 tumors similar to that of non-tumoral livers, corroborating the hypothesis of origin in the late stages of hepatocyte differentiation (Fig. 4A). Considering that Set-3 consisted of HBs from patients who are deceased and developed metastases, as well as one sample characterized as HB/HCC features, we believe that the significant hepatocyte markers (*UGT1A1*, *TDO2*, and *CYP3A4*), upregulated in this group, may deserve further investigation as biomarkers of prognosis. Using expression data from previous studies [18—20], we showed that the proposed 13 gene panel allows tumor stratification, although with limitations possibly related to the use of different techniques to obtain gene expression data (RT-PCR, microarrays, and RNA-Seq), and the fact that our panel of genes was not completely covered in two of the cohorts. More importantly, this *in silico* analysis reinforces the hypothesis of a sub-set of tumors with more differentiated biological profiles that would correspond to Set-3 in the stratification we proposed. However, we observed a difference in the prognosis associated with these Set-3 HBs supposedly derived from later hepatoblast stages, comparing our findings with the literature [22], which showed the opposite; this discrepancy needs clarification. Several confounding factors may explain the discrepancy, including the use of surgical samples after neoadjuvant chemotherapy versus biopsy samples and different ethnic backgrounds.

There are recent proposals of the use of DNA methylation for HB stratification [21,22]. Sekiguchi's E1/E2 clusters presented a high AFP level, HNF4a/CEBPA-binding regions hypermethylation, and fetal liver like expression profile, albeit distinguishable by different patient age distributions [22]. In our case, a small group of intermediate-risk tumors exhibited high expression of the epigenetic machinery genes and hepatoblast markers (Set-1), a profile suggestive of a precursor stage of hepatocyte differentiation. This HB group presents the highest expression of *TET1*, *TET3*, and *UHRF1* among all tumors (Fig. 3); in a recent work, we proposed an active demethylation process in HBs, significantly associated with upregulation of *TETs* and *UHRF1*
[15] and, accordingly, Set-1 tumors exhibited marked global DNA hypomethylation (Fig. 4A). Evidence of the association between high expression of *TET1* and *TET2* genes and strong global hypomethylation in HBs was also later demonstrated by Carillo et al. [21] in the Epi-CB signature, reinforcing the role of active demethylation.
The global methylation of Epi-CB signature was similar to that of early embryonal/fetal phases of liver development [21], our hypothesis for the HB Set-1 samples based on our findings. The Set-2 contains clinically heterogeneous HBs exhibiting an intermediate expression profile between the Set-1 and Set-3 clusters, with high expression of the DNA methylation genes and downregulation of hepatoblasts and mature hepatocyte markers (Fig. 3). The Set-2 also has a global DNA methylation level intermediary between Set-1 and Set-3 (Fig. 4A), which is in accordance with the hypothesis that these tumor samples have a molecular profile transitional between hepatoblasts and mature hepatocytes. The Epi-CA signature also described this finding of milder hypomethylation in HB samples [21]. The literature also highlighted the 14q32 *DLK1-DIO3* locus, associating hypomethylation and local gene expression changes [21,46]. Accordingly, we also detected 14q32 hypomethylation at *DLK1-DIO3* locus, and our analysis revealed that Set-3 HBs present a methylation level in this region similar to those of control samples, supporting the hypothesis of biological similarity to differentiated livers (Fig. 4B).
Taken together, our findings deepen the understanding of the underlying mechanism associated with DNA methylation changes in HBs, pinpointing specific DNA methylation enzymes as crucial players because their expression pattern allowed the clustering of tumors in different sets.

Conclusion

We showed that stratification of HBs in different groups can be performed using the expression pattern of a panel of genes of the epigenetic machinery and markers of liver differentiation stages. These groups have DNA methylation levels according to their presumptive stage of cell differentiation, suggesting that the clinical heterogeneity described for this tumor type also occurs at the epigenetic level. Therefore, despite the use of a small cohort of 21 HB samples, our data reinforce the notion that DNA methylation is a robust biomarker for HBs.

Ethical approval

Samples were recovered from patients enrolled in two Brazilian cancer institutions: A.C. Camargo Cancer Center and GRAACC. The Research Ethics Committee of the respective institutions approved this research using these biological samples, and all samples were collected after informed signed consent was obtained from parents or legal guardians.

Funding

The present study was supported by grants from FAPESP (CEPID — Human Genome and Stem Cell Research Center 2013/08028-1; 2018/21047-9; fellowships 2015/06281-7, 2016/04785-0, 2016/23462-8) and CNPq (141625/2016-3). The funders had no roles in the study design, data collection, analysis, decision to publish, or manuscript preparation.

Authors contribution

MR, TA, and AK conceived the study and were in charge of overall direction and planning. MR and TA conceived and planned the experiments. RL, GF, and MM helped with the computational framework and analyzed the data as bioinformatics. LCJ, EG, and KTS contributed to sample preparation, particularly samples from cell lines derived from IPSc. MC and CC are the Oncology Pediatric physicians responsible for collecting the clinical information. The first one from GRAACC and the second from A.C. Camargo Cancer Center. ST, DC, CR, and AK are specialists in genetics and molecular data, helping to conceive each result's biological hypotheses. IC is the pathologist responsible for reviewing each case to confirm the diagnosis based on histological samples. MR, TA, and AK wrote the manuscript with input from all authors.

Conflicts of interest

We declare that we have no conflicts of interest.

Acknowledgments

We thank patients and their families for participating in the study. This manuscript has been released as a Pre-Print at Authorea. June 01, 2020. DOI: 10.22541/au.159103640.00224438.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.clinre.2021.101684.

References

[1] Heck JE, Meyers TJ, Lombardi C, Park AS, Cockburn M, Reynolds P, et al. Case-control study of birth characteristics and the risk of hepatoblastoma. Cancer Epidemiol 2013;37:390—5, http://dx.doi.org/10.1016/j.canep.2013.03.004.

[2] López-Terrada D, Alaggio R, De Dávila MT, Czauderna P, Hiyama E, Katzenstein H, et al. Towards an international pediatric liver tumor consensus classification: proceedings of the Los Ange-les COG liver tumors symposium. Mod Pathol 2014;27:472—91, http://dx.doi.org/10.1038/modpathol.2013.80.

[3] Howlader N, Noone A, Krapcho M, Miller D, Bishop K, Altekruse S, et al. SEER Cancer Statistics Review 1975—2013. Natl Cancer Inst 2013.

[4] Turcotte LM, Georgieff MK, Ross JA, Feusner JH, Tomlinson GE, Malogolowkin MH, et al. Neonatal medical exposures and characteristics of low birth weight hepatoblastoma cases: a report from the Children's Oncology Group. Pediatr Blood Can-cer 2014;61:2018—23, http://dx.doi.org/10.1002/pbc.25128.

[5] Perilongo G, Malogolowkin M, Feusner J. Hepatoblastoma clin-ical research: lessons learned and future challenges. Pediatr Blood Cancer 2012, http://dx.doi.org/10.1002/pbc.24217.

[6] Trobaugh-Lotrario AD, Katzenstein HM. Chemotherapeutic approaches for newly diagnosed hepatoblastoma: past, present, and future strategies. Pediatr Blood Cancer 2012, http://dx.doi.org/10.1002/pbc.24219.

[7] Pritchard J, Brown J, Shafford E, Perilongo G, Brock P, Dicks-Mireaux C, et al. Cisplatin, doxorubicin, and delayed surgery for childhood hepatoblastoma: a successful approach — results of the first prospective study of the International Society of Pediatric Oncology. J Clin Oncol 2000;18:3819—28, http://dx.doi.org/10.1200/JCO.2000.18.22.3819.

[8] Douglass EG, Reynolds M, Finegold M, Cantor AB, Glicksman A. Cisplatin, vincristine, and fluorouracil therapy for hepato-blastoma: a pediatric oncology group study. J Clin Oncol 1993, http://dx.doi.org/10.1200/JCO.1993.11.1.96.

[9] Ortega JA, Douglass EC, Feusner JH, Reynolds M, Quinn JJ, Finegold MJ, et al. Randomized comparison of cis-platin/vincristine/fluorouracil and cisplatin/continuous infu-sion doxorubicin for treatment of pediatric hepatoblastoma: a report from the Children's Cancer Group and the Pediatric Oncology Group. J Clin Oncol 2000;18:2665—75, http://dx.doi.org/10.1200/JCO.2000.18.14.2665.

[10] Meyers RL, Czauderna P, Otte JB. Surgical treat-ment of hepatoblastoma. Pediatr Blood Cancer 2012, http://dx.doi.org/10.1002/pbc.24220.

[11] Weinberg AG, Finegold MJ. Primary hepatic tumors of child-hood. Hum Pathol 1983;14:512—37.

[12] Ma X, Liu Y, Liu Y, Alexandrov LB, Edmonson MN, Gawad C, et al. Pan-cancer genome and transcriptome analyses of 1,699 pae-diatric leukaemias and solid tumours. Nature 2018;555:371—6, http://dx.doi.org/10.1038/nature25795.

[13] Maschietto M, Rodrigues TC, Kashiwabara AY, de Araujo ÉSS, Aguiar TFM, da Costa CML, et al. DNA methylation landscape of hepatoblastomas reveals arrest at early stages of liver dif-ferentiation and cancer-related alterations. Oncotarget 2016, http://dx.doi.org/10.18632/oncotarget.14208.

[14] Yamada N, Yasui K, DohiO, Gen Y, Tomie A, Kitaichi T, et al. Genome-wide DNA methylation analysis in hepatocellular carcinoma. Oncol Rep 2016;35:2228—36, http://dx.doi.org/10.3892/or.2016.4619.

[15] Rivas MP, Aguiar TFM, Fernandes GR, Caires-Júnior LC, Goulart E, Telles-Silva KA, et al. TET upreg-ulation leads to 5-hydroxymethylation enrichment in hepatoblastoma.Front Genet 2019;10:1—7, http://dx.doi.org/10.3389/fgene.2019.00553.

[16] Maris JM, DennyCT. Focus on embryonal malignancies. Cancer Cell 2002;2:447—50, http://dx.doi.org/10.1016/S1535-6108(02)00206-4.

[17] Kim EH, Koh KN, Park M, Kim BE, Im HJ, Seo JJ. Clinical features of infantile hepatic hemangioendothelioma. Korean J Pediatr 2011, http://dx.doi.org/10.3345/kjp.2011.54.6.260.

[18] Cairo S, Armengol C, De Reynies A, Wei Y, Thomas E, Renard C-A, et al. Hepatic stem-like phenotype and interplay of Wnt/beta-catenin and Myc signaling in aggres-sive childhood liver cancer. Cancer Cell 2008;14:471—84, http://dx.doi.org/10.1016/j.ccr.2008.11.002.

[19] Sumazin P, Chen Y, Trevino˜ LR, Sarabia SF, Hampton OA, Patel K, et al. Genomic analysis of hepatoblastoma identi-fies distinct molecular and prognostic subgroups. Hepatology 2017;65:104—21, http://dx.doi.org/10.1002/hep.28888.

[20] Hooks KB, Audoux J, Fazli H, Lesjean S, Ernault T, Dugot-Senant N, et al. New insights into diagnosis and therapeutic options for proliferative hepatoblastoma. Hepatology 2018, http://dx.doi.org/10.1002/hep.29672.

[21] Carrillo-Reixach J, Torrens L, Simon-Coma M, Royo L, Domingo-Sabat M, Abril-Fornaguera J, et al. Epigenetic footprint enables molecular risk stratification of hepa-toblastoma with clinical implications. J Hepatol 2020, http://dx.doi.org/10.1016/j.jhep.2020.03.025.

[22] Sekiguchi M, Seki M, Kawai T, Yoshida K, Yoshida M, Isobe T, et al. Integrated multiomics analysis of hepatoblastoma unravels its heterogeneity and pro-vides novel druggable targets. Npj Precis Oncol 2020, http://dx.doi.org/10.1038/s41698-020-0125-y.

[23] Okita K, Yamakawa T, Matsumura Y, Sato Y, Amano N, Watanabe A, et al. An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells. Stem Cells 2013;31:458—66, http://dx.doi.org/10.1002/stem.1293.

[24] Hay DC, Zhao D, Fletcher J, Hewitt ZA, McLean D, Urruticoechea-Uriguen A, et al. Efficient differentiation of hep-atocytes from human embryonic stem cells exhibiting markers recapitulating liver development in vivo. Stem Cells 2008, http://dx.doi.org/10.1634/stemcells.2007-0718.

[25] Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantita-tive RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002.

[26] Livak KJ, Schmittgen TD. Analysis of relative gene expres-sion data using real-time quantitative PCR and the 2(−Delta Delta C(T)) method. Methods 2001;25:402—8, http://dx.doi.org/10.1006/meth.2001.1262.

[27] Edmunds RC, McIntyre JK, Luckenbach JA, Baldwin DH, Incar-dona JP. Toward enhanced MIQE compliance: reference residual normalization of qPCR gene expression data. J Biomol Techn 2014;25:54—60, http://dx.doi.org/10.7171/jbt.14-2502-003.

[28] Hardeo S, Mohammed AI. The analysis of variance: fixed, random and mixed models. Birkhauser; 2000, http://dx.doi.org/10.1007/978-1-4612-1344-4.

[29] Fabregat A, Sidiropoulos K, Viteri G, Forner O, Marin-Garcia P, Arnau V, et al. Reactome pathway analysis: a high-performance in-memory approach. BMC Bioinform 2017;18:142, http://dx.doi.org/10.1186/s12859-017-1559-2.

[30] Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, et al. The reactome pathway knowledgebase. Nucleic Acids Res 2016;44:D481—7, http://dx.doi.org/10.1093/nar/gkv1351.

[31] Aguiar TFM, Rivas MP, Costa S, Maschietto M, Rodrigues T, Sobral de Barros J, et al. Insights into the somatic mutation bur-den of hepatoblastomas from Brazilian patients. Front Oncol 2020;10:556, http://dx.doi.org/10.3389/fonc.2020.00556.

[32] Houssaint E. Differentiation of the mouse hepatic primordium. I. An analysis of tissue interactions in hepatocyte differentiation. Cell Differ 1980, http://dx.doi.org/10.1016/0045-6039(80)90026-3.

[33] Medlock ES, Haar JL. The liver hemopoietic environment: I. Developing hepatocytes and their role in fetal hemopoiesis. Anat Rec 1983, http://dx.doi.org/10.1002/ar.1092070105.

[34] Zorn. Liver development. StemBook; 2008. https://doi.org/10.3824/stembook.1.25.1.

[35] Mudbhary R, Hoshida Y, Chernyavskaya Y, Jacob V, Vil-lanueva A, Fiel MI, et al. UHRF1 overexpression drives DNA hypomethylation and hepatocellular carcinoma. Cancer Cell 2014, http://dx.doi.org/10.1016/j.ccr.2014.01.003.

[36] Lau HH, Ng NHJ, Loo LSW, Jasmen JB, Teo AKK. The molecular functions of hepatocyte nuclear factors — in and beyond the liver. J Hepatol 2018;68:1033—48, http://dx.doi.org/10.1016/j.jhep.2017.11.026.

[37] Bach DH, Park HJ, Lee SK. The dual role of bone mor-phogenetic proteins in cancer. Mol Ther Oncolyt 2018, http://dx.doi.org/10.1016/j.omto.2017.10.002.

[38] Vogel CFA, Ishihara Y, Campbell CE, Kado SY, Nguyen-Chi A, Sweeney C, et al. A protective role of aryl hydrocarbon recep-tor repressor in inflammation and tumor growth. Cancers (Basel) 2019, http://dx.doi.org/10.3390/cancers11050589.

[39] Vogel CFA, Haarmann-Stemmann T. The aryl hydro-carbon receptor repressor — more than a simple feedback inhibitor of AhR signaling: clues for its role in inflammation and cancer. Curr Opin Toxicol 2017, http://dx.doi.org/10.1016/j.cotox.2017.02.004.

[40] Sun X, Gao L, Chien HY, Li WC, Zhao J. The regulation and function of the NUAK family. J Mol Endocrinol 2013, http://dx.doi.org/10.1530/JME-13-0063.

[41] Evans RM, Mangelsdorf DJ. Leading edge review nuclear recep-tors, RXR, and the Big Bang. Cell 2014.

[42] Szanto A, Narkar V, Shen Q, Uray IP, Davies PJA, Nagy L. Review retinoid x receptors: X-ploring their (patho)physiological functions. Cell Death Differ 2004, http://dx.doi.org/10.1038/sj.cdd.4401533.

[43] Altucci L, Leibowitz MD, Ogilvie KM, de Lera AR, Gronemeyer H. RAR and RXR modulation in cancer and metabolic disease. Nat Rev Drug Discov 2007, http://dx.doi.org/10.1038/nrd2397.

[44] H.K., W.W. Peroxisome proliferator-activated receptors: a link between endocrinology and nutrition? Trends Endocrinol Metab 1993.

[45] Combes AN, Whitelaw E. Epigenetic reprogramming: enforcer or enabler of developmental fate? Dev Growth Differ 2010, http://dx.doi.org/10.1111/j.1440-169X.2010.01185.x.

[46] Honda S, Chatterjee A, Leichter AL, Miyagi H, Minato M, Fujiyoshi S, et al. A microRNA cluster in the DLK1-DIO3 imprinted region on chromosome 14q32.2 is dysreg-ulated in metastatic hepatoblastomas. Front Oncol 2020, http://dx.doi.org/10.3389/fonc.2020.513601.

[47] Czauderna P, Haeberle B, Hiyama E, Rangaswami A, Krailo M, Maibach R, et al. The Children's Hepatic tumors Inter-national Collaboration (CHIC): novel global rare tumor database yields new prognostic factors in hepatoblastoma and becomes a research model. Eur J Cancer 2016;52:92—101, http://dx.doi.org/10.1016/j.ejca.2015.09.023.

[48] Meyers RL, Maibach R, Hiyama E, Haberle B, Krailo M, Ran-gaswami A, et al. Risk-stratified staging in paediatric hepato-blastoma: a unified analysis from the Children's Hepatic tumors International Collaboration. Lancet Oncol 2017;18:122—31, http://dx.doi.org/10.1016/S1470-2045(16)30598-8.

SUPPORTING INFORMATION

DNA METHYLATION AS A KEY EPIGENETIC PLAYER FOR HEPATOBLASTOMA CHARACTERIZATION

Maria Rivas*[a], Talita Aguiar*[a,b], Gustavo Fernandes[a], Renan Lemes[a], Luiz Caires-Júnior[a], Ernesto Goulart[a], Kayque Telles-Silva[a], Mariana Maschietto[c], Monica Cypriano[d], Silvia de Toledo[d,] Dirce Carraro[e], Isabela da Cunha[f], Cecilia da Costa[g], Carla Rosenberg[a], Ana Krepischi**[a]

a. Human Genome and Stem Cell Research Center, Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, São Paulo, SP, Brazil.

b. Department of Urology - NYU Grossman School of Medicine, New York City, NY, USA.

c. Research Center, Boldrini Children's Hospital, Campinas, BR

d. Department of Pediatrics, Adolescent and Child with Cancer Support Group (GRAACC), Federal University of São Paulo, SP, Brazil.

e. International Center for Research, A. C. Camargo Cancer Center, SP, Brazil.

f. Pathology Department, Rede D'OR-São Luiz, SP, Brazil

g. Department of Pediatric Oncology, A. C. Camargo Cancer Center, SP, Brazil

* These two authors contributed equally

** Corresponding author:

Dr. Ana Cristina Victorino Krepischi

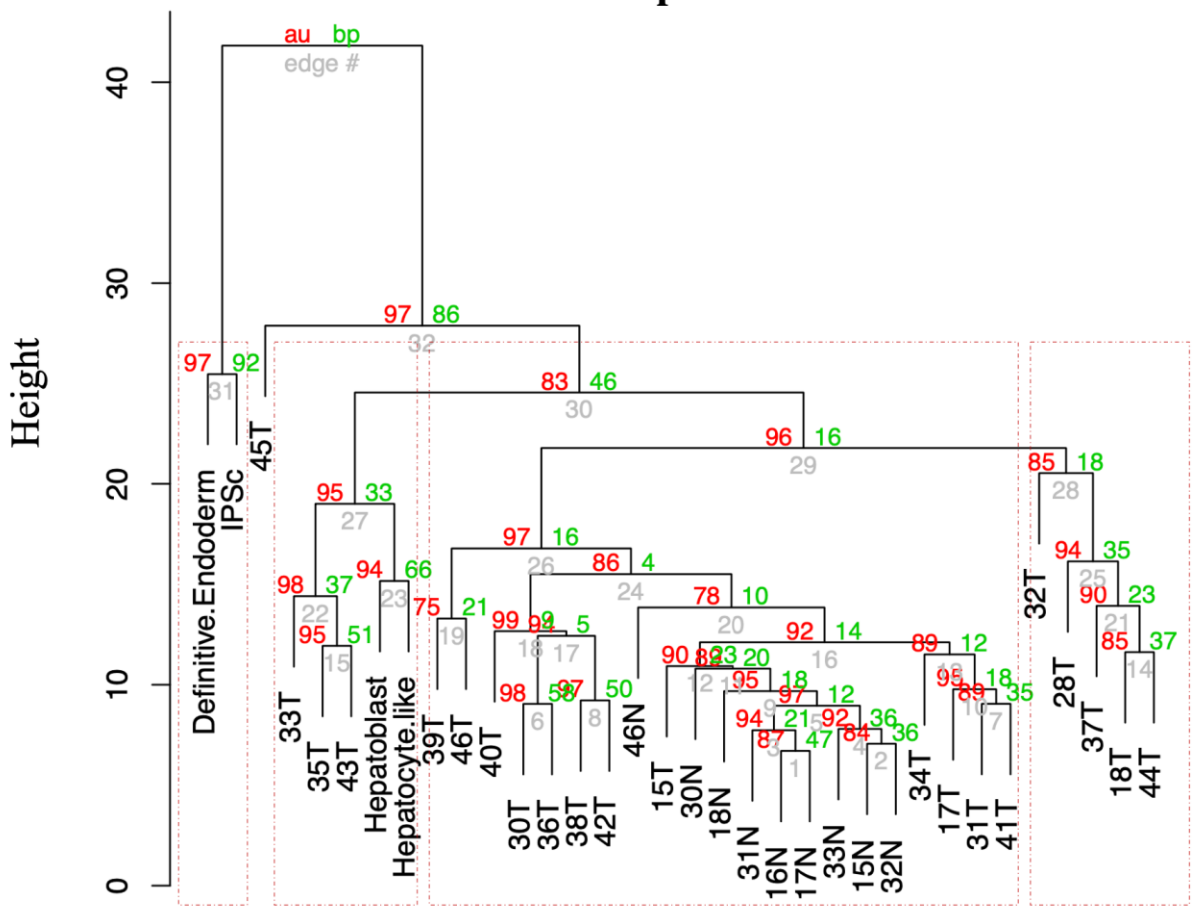
Human Genome and Stem-Cell Research Center, Department of Genetics and Evolutionary Biology - Institute of Biosciences, University of São Paulo, São Paulo, Brazil

Phone: 55 11 3091 7573

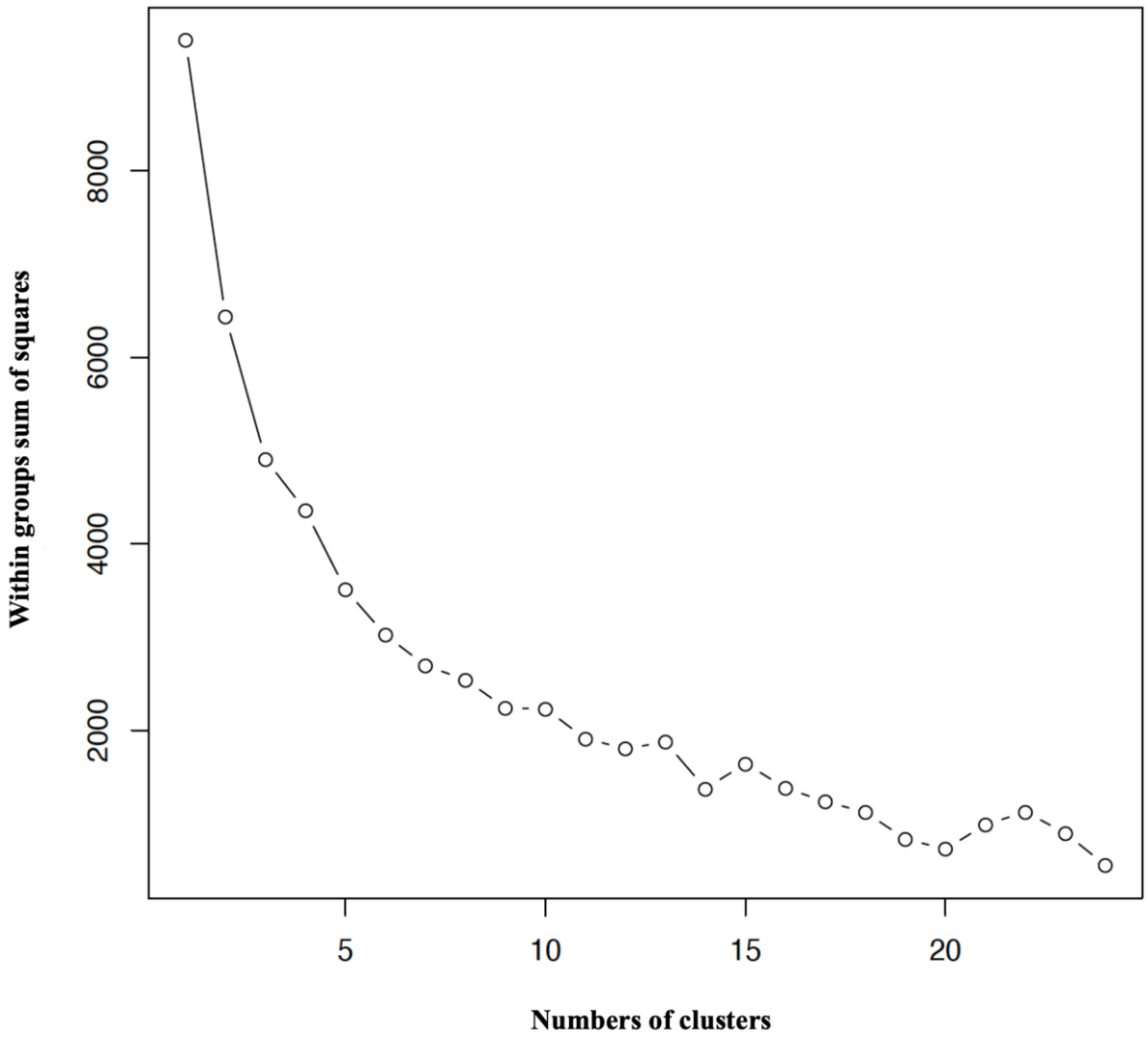
email: ana.krepischi@ib.usp.br

Supporting Information Figure S1. Samples hierarchical clustering by UPGMA and Elbow method. A. Dendogram showing the distance between samples and similarities to form the clusters. The squares in the figure highlight the four clusters formed. AU (printed in red color) is the “approximately unbiased” p-value for nonselective inference. BP value (printed in green color) is the "bootstrap probability" value. One can consider that clusters (edges) with high AU values (e.g. 95%) are strongly supported by data. **B.** The elbow method is used to determine the number of principal components or clusters on dataset. This method analyzes the percentage of variation explained as a function of the number of clusters. To choose a number of clusters that offers a better modeling of the data, the percentage of variation explained by the clusters must be plotted against the number of clusters, with the first clusters adding a lot of information, but at some point the marginal gain will fall, giving an angle on the graph. The number of clusters is chosen at this point, hence the "elbow criterion". This "elbow" cannot always be identified unambiguously. For the case of the clusters observed in this study, the graph shows that the elbow starts from cluster 4.

A. **Hierarchical Clustering based on log2 fold-change and 10000 replications**

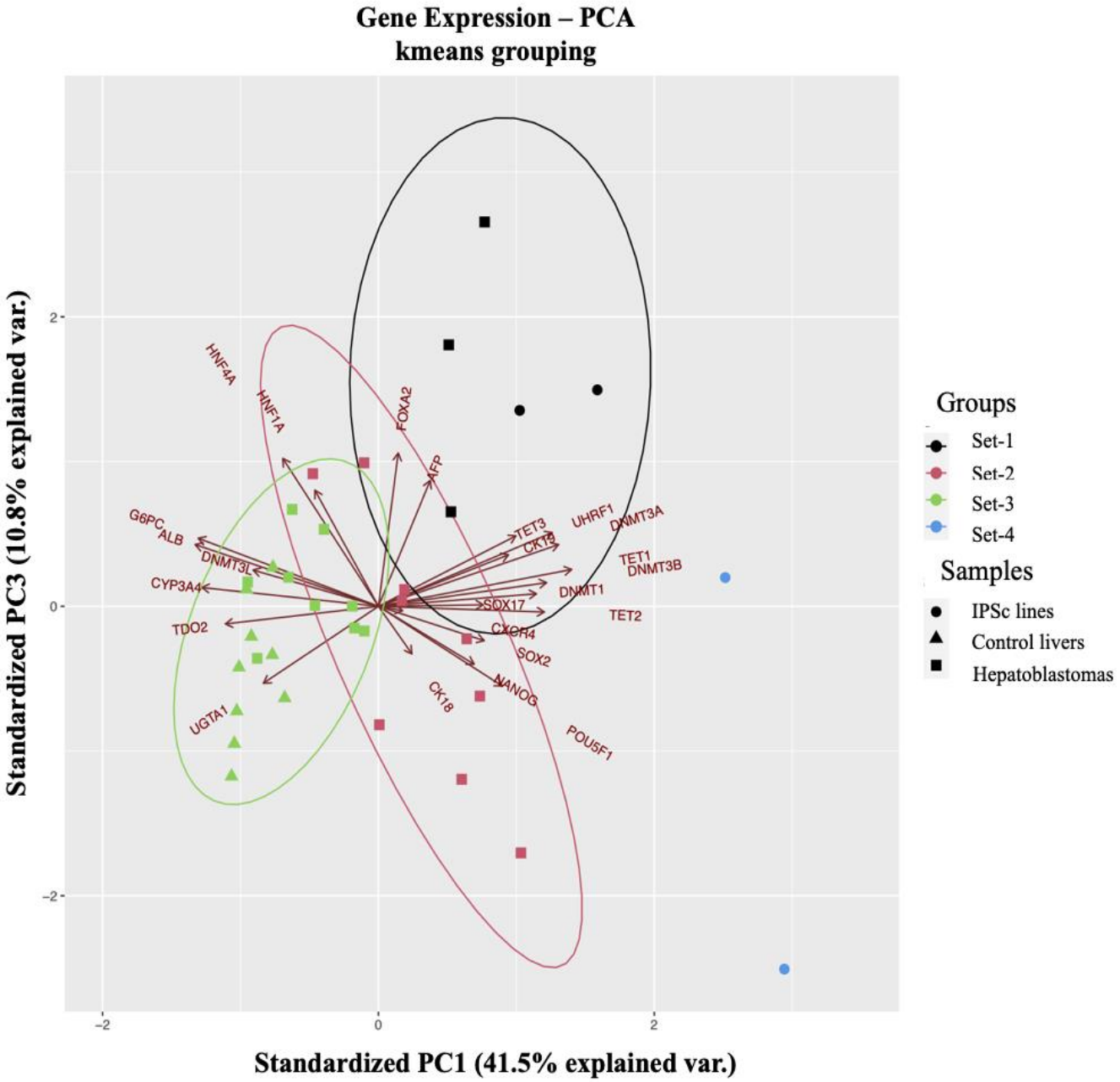


B.

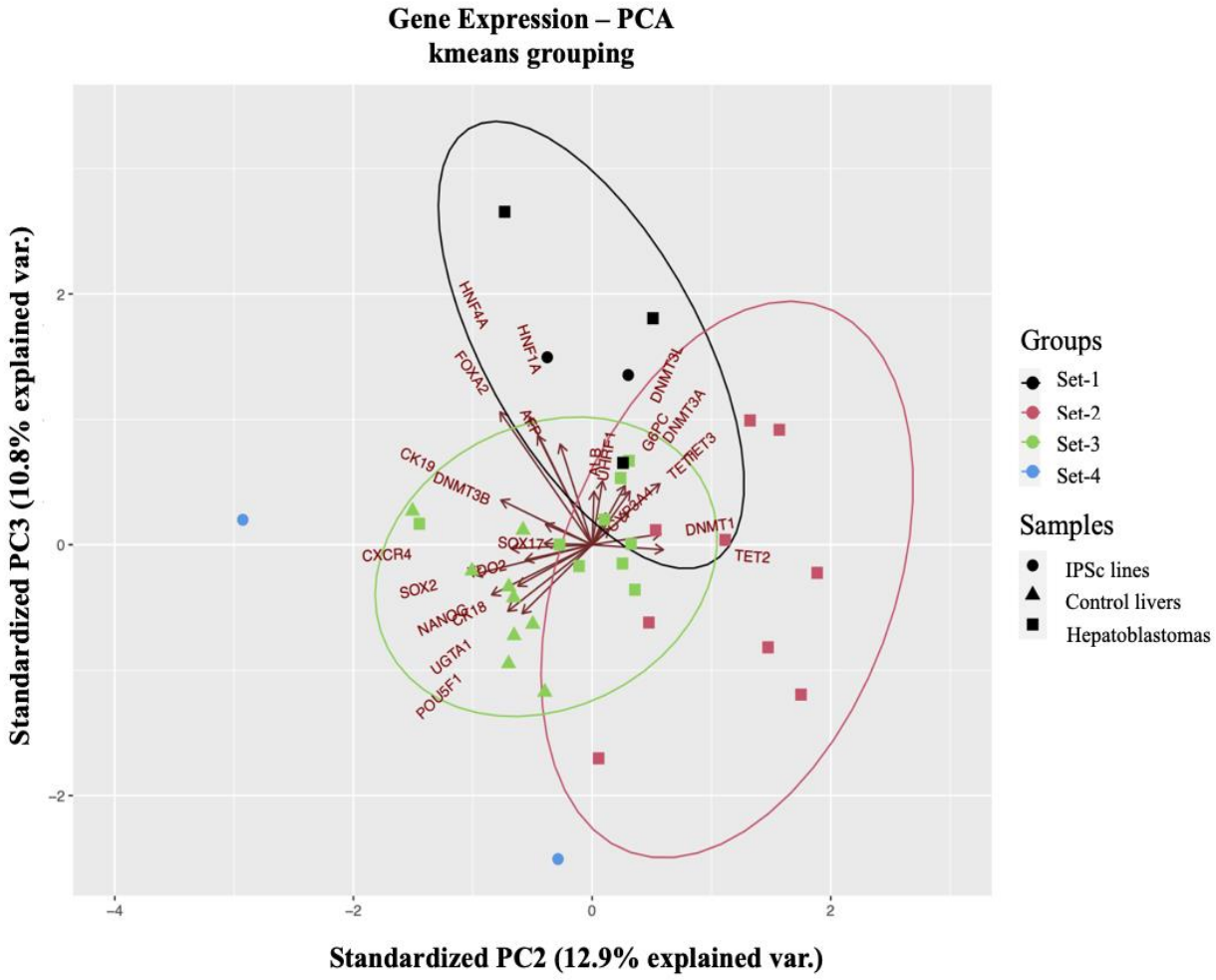

Within groups sum of squares
2000
4000
6000
8000
5
10
15
20
Numbers of clusters

Supporting Information Figure S2. Analysis of PCA K-means of hepatoblastomas and control liver samples based on their expression profile of 24 genes. The analysis was performed grouping HBs and non-tumoral liver samples with similar levels of gene expression which resulted in four clusters, indicated by color legends. In the figure, the triangles indicate samples of non-tumoral tissues; the squares indicate HB samples; circles indicate samples derived from cell culture (iPSC lines). The samples grouped in Set-1 are represented in black; the samples grouped in Set-2 are red; in green, samples grouped in Set-3; in blue, samples grouped in Set-4. **A.** represents the PC1 and PC3 and **B.** PC2 and PC3.

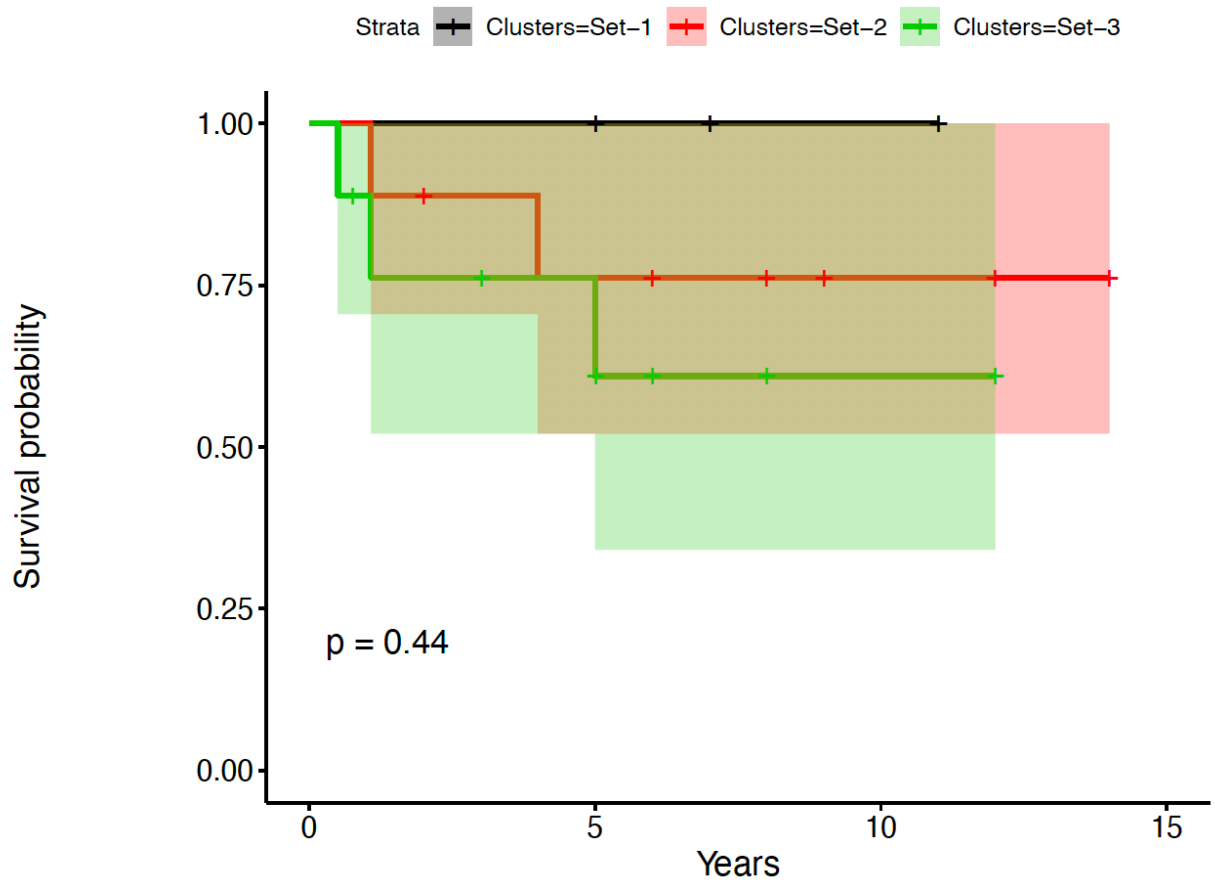
A.


Gene Expression – PCA
kmeans grouping
Standardized PC3 (10.8% explained var.)
Standardized PC1 (41.5% explained var.)
2
0
−2
−2
0
2
HNF4A
HNF1A
FOXA2
AFP
G6PC
ALB
DNMT3L
CYP3A4
TDO2
UGTA1
TET3
CK19
UHRF1
DNMT3A
TET1
DNMT3B
DNMT1
SOX17
TET2
CXCR4
SOX2
NANOG
CK18
POU5F1
Groups
Set-1
Set-2
Set-3
Set-4
Samples
IPSc lines
Control livers
Hepatoblastomas

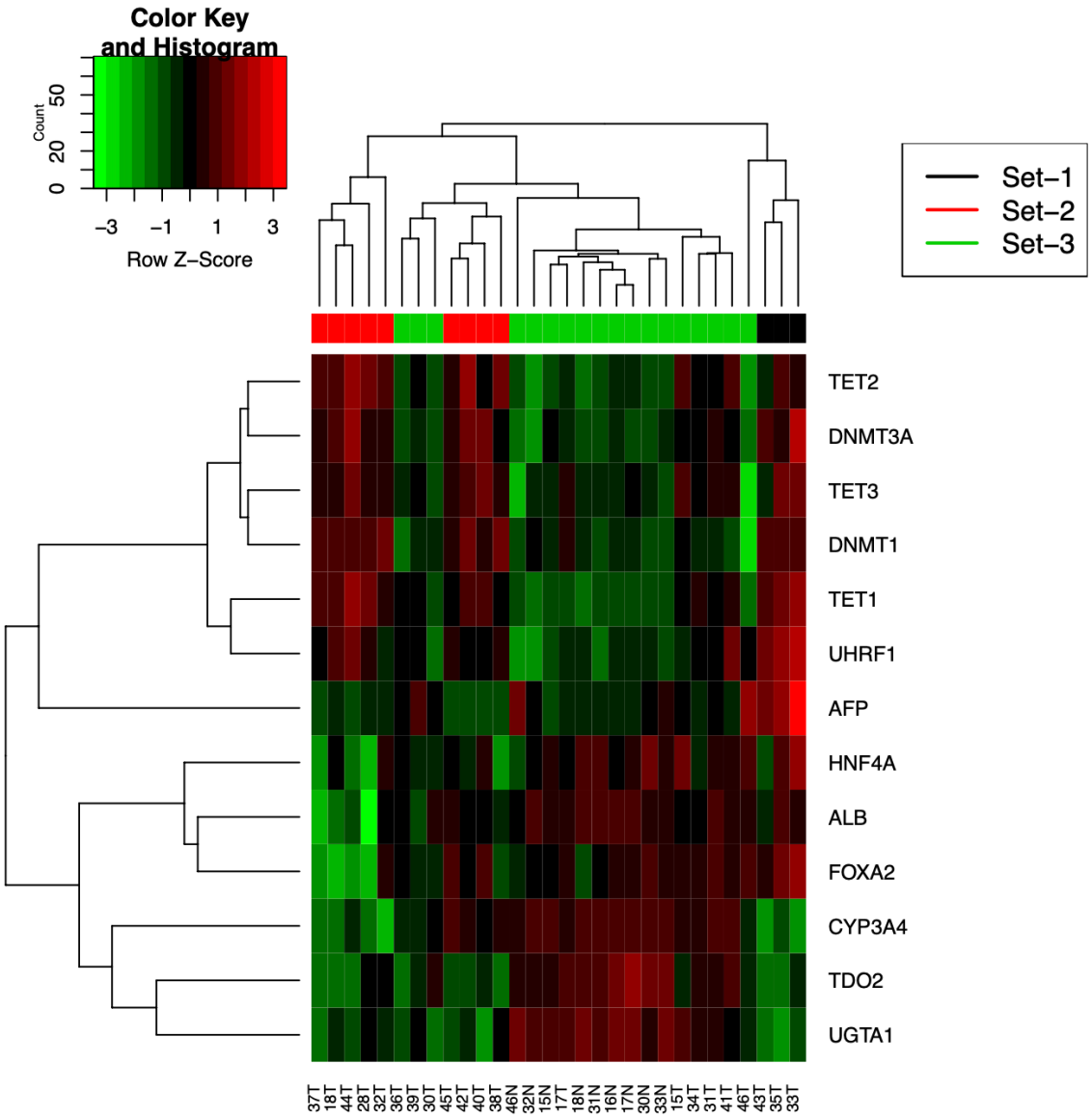
B.


Gene Expression – PCA
kmeans grouping
Standardized PC3 (10.8% explained var.)
Standardized PC2 (12.9% explained var.)
HNF4A
HNF1A
FOXA2
AFP
DNMT3L
DNMT3A
G6PC
TET3
CK19
DNMT3B
CXCR4
SOX17
SOX2
NANOG
CK18
UGTA1
POU5F1
DNMT1
TET2
Groups
Set-1
Set-2
Set-3
Set-4
Samples
IPSc lines
Control livers
Hepatoblastomas

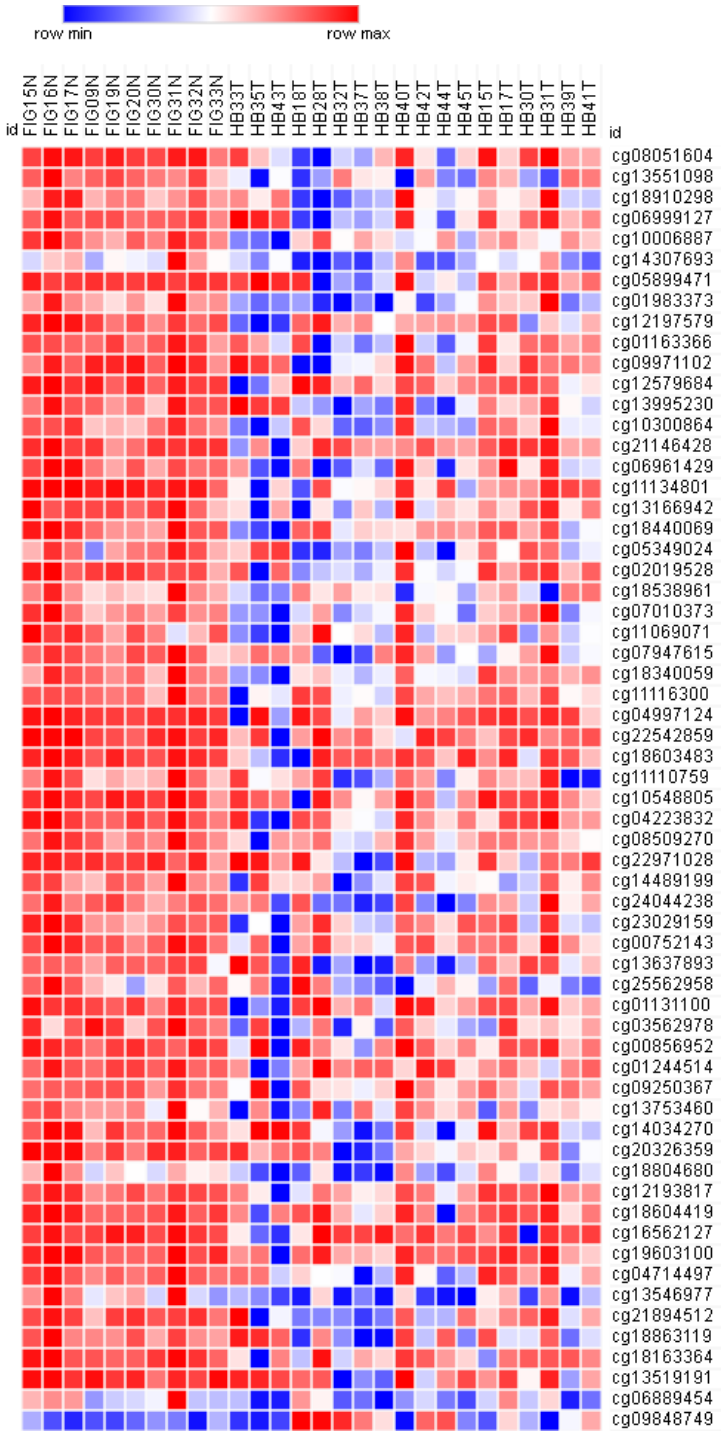
Supporting Information Figure S3. Kaplan-Meier survival curve. The graph below shows the analysis of the survival curve of the studied HB patients. The gray color corresponds to the Set-1, red corresponds to the Set-2 and green to Set-3. Despite the non-significant statistical test (p> 0.05), it is possible to observe that Set-3 has the highest number of deaths in 10 years.



Supporting Information Figure S4. Heatmap analysis highlighting the grouping of the HB samples according to the thirteen genes with the highest level of expression differences. In black, samples of Set-1; in red, samples grouped in Set-2; in green, samples of Set-3. Each tumor is represented in one column and genes are depicted in lines. In the heatmap, the green color represents a reduction in gene expression, while the red color indicates an increase in expression. T: Tumors samples; N: Non-tumor liver samples.



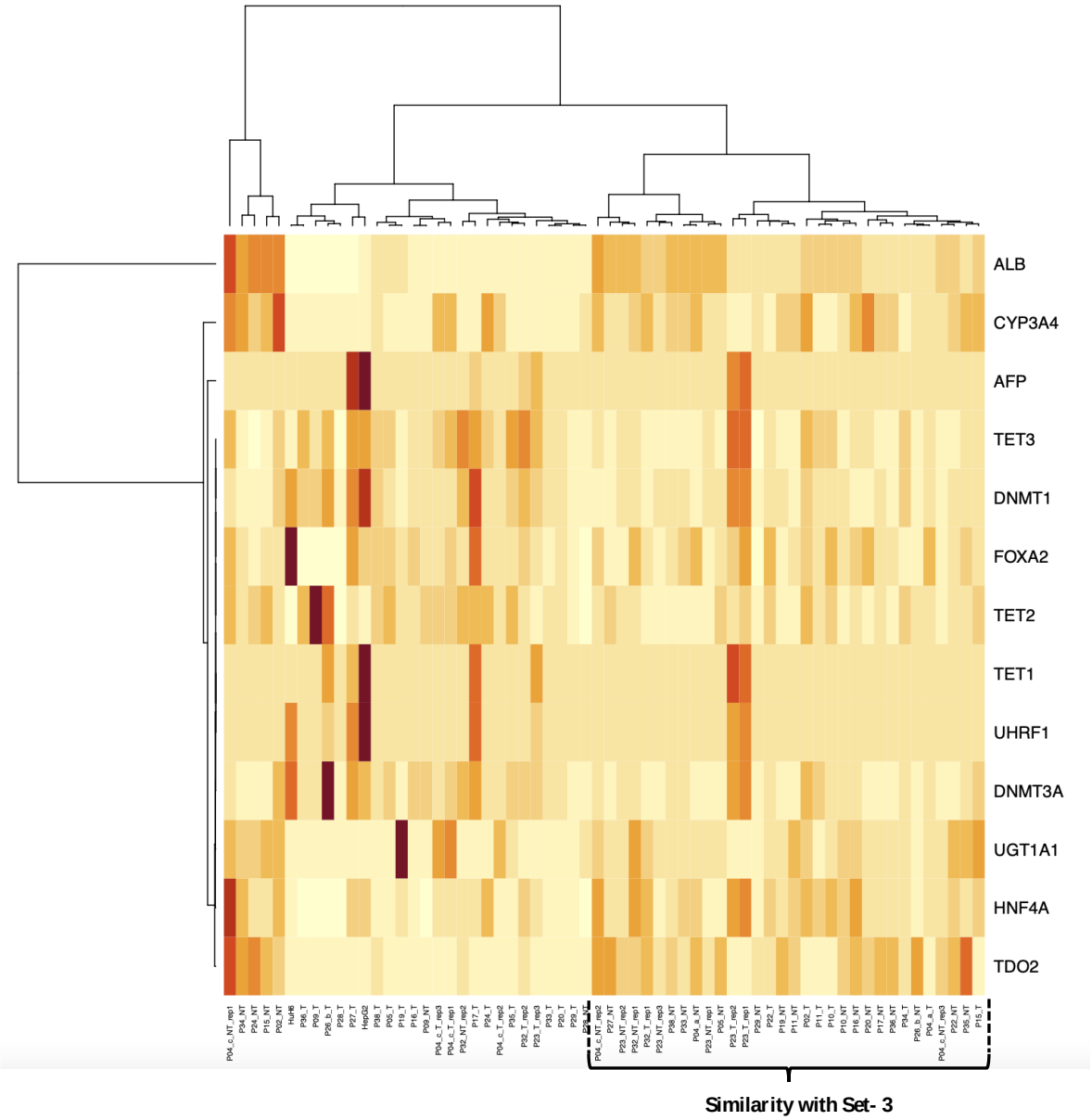
Supporting Information Figure S5. Evaluation of the DNA methylation profile of the 14q32 *DLK1-DIO3* locus. Evaluating the 32 CpG sites at 14q32 that Carrillo-Reixach et al. [21] identified as hypomethylated in HBs, we found 13 sites significantly hypomethylated in tumors, with three CpGs (cg13551098; cg08051604; cg18910298) exhibiting methylation changes >20%.



Supporting Information Figure S6. Evaluation of the proposed HB signature of 13 genes using previously published gene expression data. Unsupervised hierarchical clustering was performed using gene expression data from previous studies by means of heatmap function from R Statistical Package v4.0.3 using default parameters, such as Euclidean distance and complete linkage method. Each tumor is represented in one column and genes are depicted in lines; in the heatmap, darker color represents higher gene expression. **A**. Hooks et al. [20] - analysis was performed with the set of 13 genes; **B**. Sumazin et al. [19] - analysis was performed using the nine available genes; **C**. Cairo et al. [18] - performed using the nine available genes.

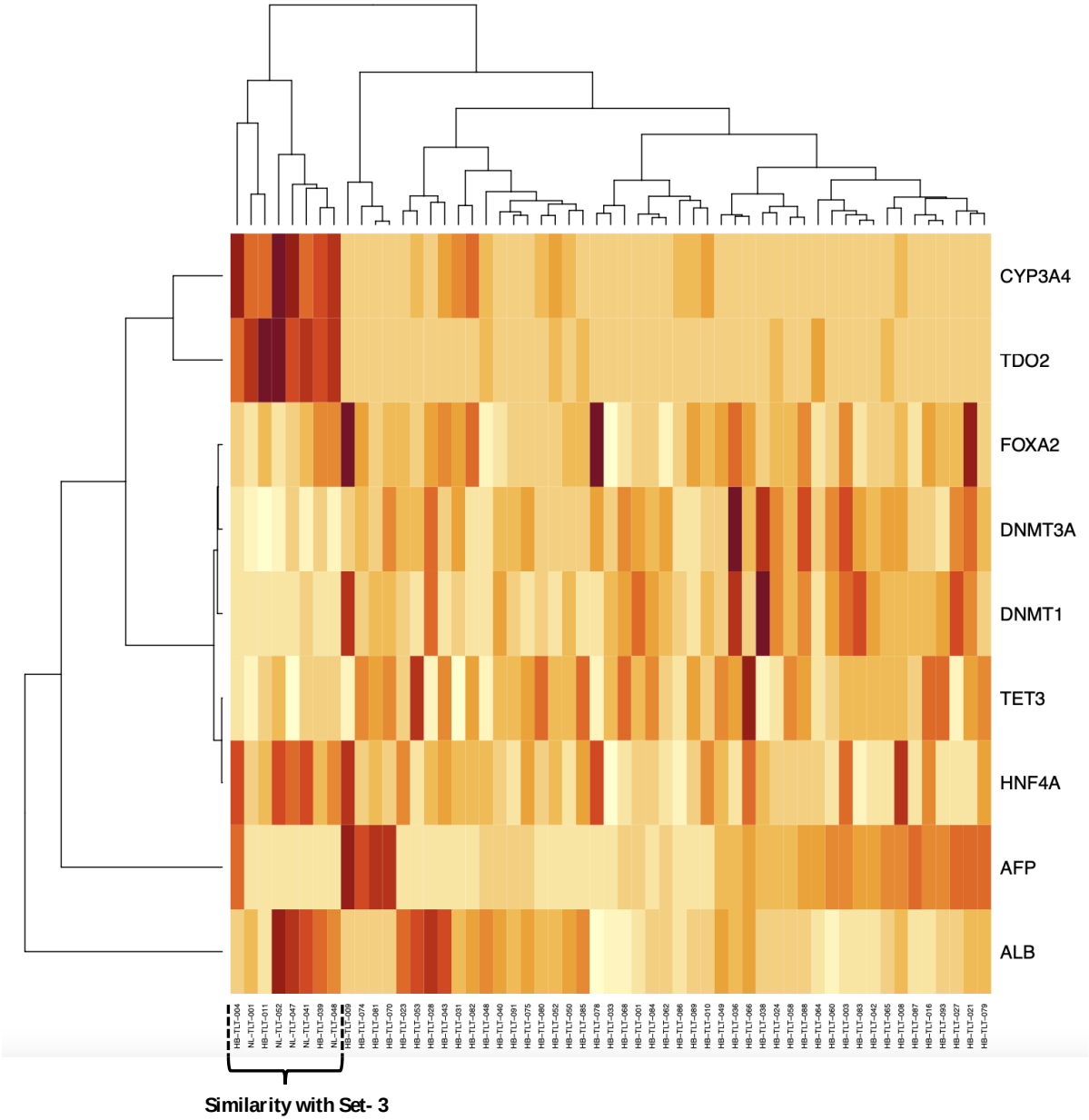
A.

Hooks. et al. 2018 (n=62)



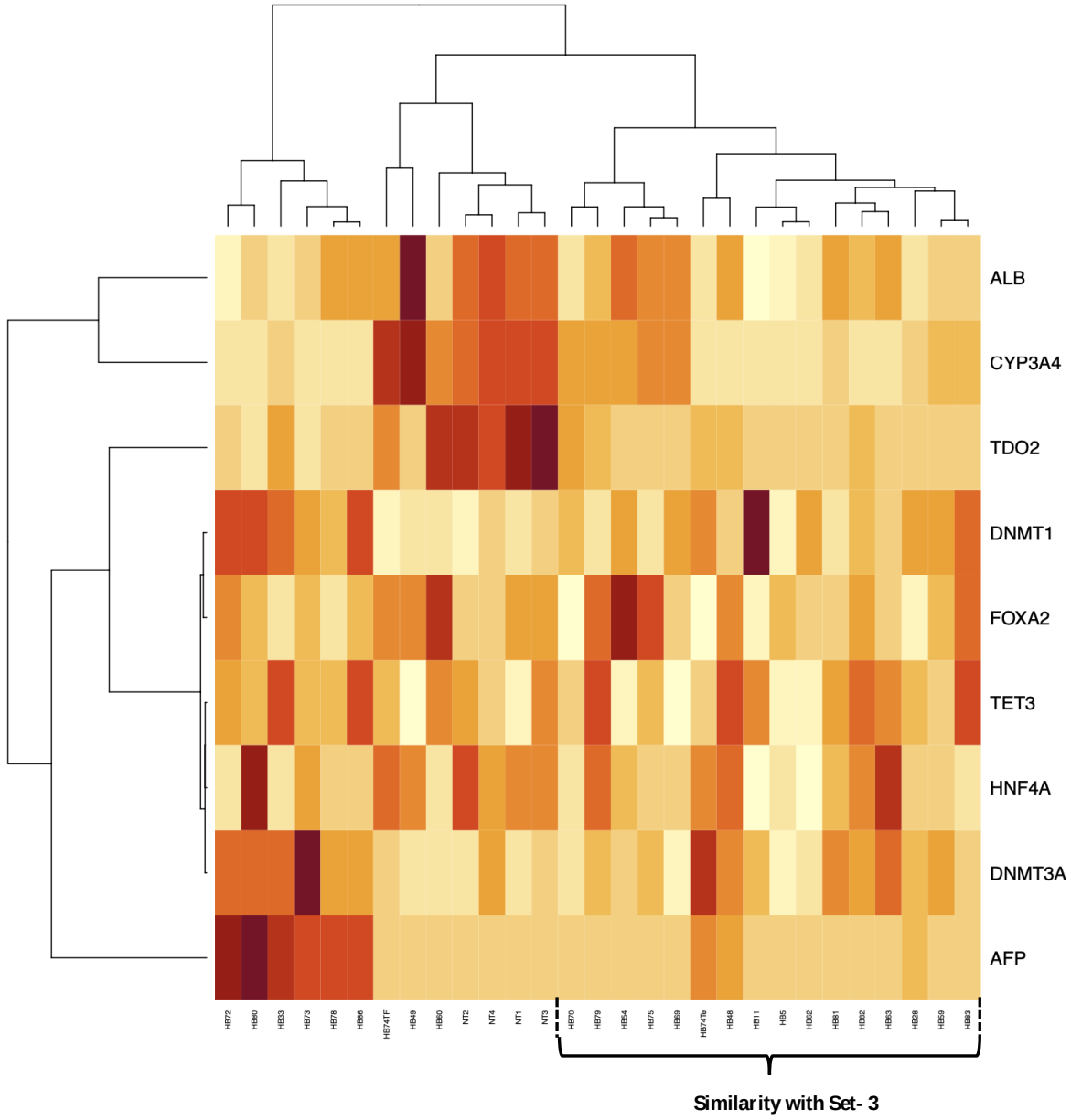
B.

Sumazin. et al. 2017 (n=55)



C.

Cairo. et al. 2008 (n=29)



Capítulo VI. Modulação da expressão gênica de *NNMT* por hipermetilação de promotor e efeito metabólicos em hepatoblastomas

Em trabalho prévio do nosso grupo, além de demonstrar alteração no padrão global de metilação entre amostras de hepatoblastoma e fígado não-tumoral, foram apresentados genes diferencialmente metilados nos tumores. Os genes que apresentam hipometilação em hepatoblastoma estão relacionados a vias de metabolismo energético, interação celular, regulação negativa do ciclo celular, além de proto-oncogenes. Já os genes que apresentam hipermetilação são relacionados ao metabolismo de Acetil-CoA, desenvolvimento do fígado, senescência, bem como transporte de ácidos graxos[101]. A maioria dos sítios hipermetilados mapeados estão localizados em ilhas CpG e, devido as suas funções, o gene *NNMT* (nicotina N-metiltransferase) obteve destaque nas análises seguintes.

Originalmente, *NNMT* foi identificada como enzima responsável pelo processo de N-metilação da nicotinamida, produzindo 1-metilnicotinamida (1MNA), que é precursor de NAD^+, um cofator doador de elétrons para o complexo mitocondrial e diversas reações oxidorredutases[119]. Estudos mais recentes expandiram o papel de NNMT, sendo atualmente associada à regulação de vias metabólicas em hepatócitos e adipócitos pelo uso de grupos doadores de metil na geração de metabólitos ativos[120–124].

A superexpressão de *NNMT* foi descrita em câncer de mama, glioblastoma, tireóide papilar, entre outros[125–127]. Ulanovskaya e colaboradores observaram altos níveis de *NNMT* e um dos seus produtos 1MNA em linhagens agressivas de câncer, exercendo impacto sobre o potencial de metilação destas células. Os autores propuseram que o produto 1MNA atuaria como armazenador de grupos metil, que as células cancerosas têm capacidade limitada de recuperar; desta forma, NNMT seria capaz de regular o estado de metilação dos tumores[128].

De maneira contrastante, em carcinoma hepatocelular foi descrita a redução da expressão de *NNMT* quando comparado ao tecido hepático não-tumoral[121]. Além disso, os autores observaram que a expressão de *NNMT* nos tumores é significativamente correlacionada com o estágio tumoral, de modo que quanto maior a expressão de *NNMT*, menor a sobrevida global e menor o tempo de vida livre de doença. Apesar do mecanismo originando a redução de NNMT e sua atuação em câncer de fígado ainda não terem sido esclarecidos, recentemente Shin et al. revelou que a redução de NNMT é capaz de promover autofagia e contribuir para sobrevivência e crescimento de células tumorais do fígado, indicando um possível caminho de ação deste gene para o câncer de fígado em adultos[129].

Contudo, o papel e padrão de expressão do *NNMT* em câncer de fígado pediátrico ainda é pouco explorado.

Dado que a hipermetilação na região promotora de um gene pode ser relacionada ao silenciamento da expressão do mesmo, impedindo o acesso da maquinaria de transcrição[78], direcionamos o estudo para investigar em hepatoblastomas se a hipermetilação do promotor de *NNMT* observada poderia levar à desregulação de sua expressão gênica e proteica. O impacto potencial das alterações em NNMT foi avaliado pela análise do perfil metabólico de hepatoblastomas por espectroscopia de ressonância magnética.

Como resultado, evidenciamos a redução significativa de *NNMT* em hepatoblastomas, com nível de expressão duas ordens de magnitude menor nas amostras de tumor e linhagens celulares de hepatoblastoma, comparadas a linhagens celulares de carcinoma hepatocelular. Além disto, análise de *Pearson* demonstrou correlação inversa entre o nível de metilação do promotor de *NNMT* e a expressão deste gene. Ao nível proteico, a redução significativa de NNMT foi observada nos tumores, com forte correlação entre a expressão do gene e da proteína. É importante ressaltar que a maior expressão de *NNMT* foi estatisticamente associada ao diagnóstico tardio em hepatoblastoma, uma reconhecida variável clínica de pior prognóstico. Por fim, a análise metabolômica não direcionada detectou metabolismo lipídico aberrante em hepatoblastomas, sugerindo que a diminuição de *NNMT* pode estar associada à redução detectada do conteúdo lipídico em amostras de hepatoblastoma, possivelmente relacionada a concentrações alteradas de metabólicos ligados a função do NNMT ou precursores de NAD^+ [123]. Os dados apresentados aqui mostraram pela primeira vez que a redução de NNMT ocorre em hepatoblastomas, fornecendo suporte adicional à teoria de que a regulação negativa desta proteína é um fenômeno geral em câncer de fígado.

Os resultados foram publicados no artigo *Hepatoblastomas exhibit marked NNMT downregulation driven by promoter DNA hypermethylation*, na revista *Tumor Biology (DOI:10.1177/1010428320977124)*, sendo apresentado a seguir.

Original Article

Tumor Biology

Tumor Biology
December 2020: 1–14

DOI: 10.1177/1010428320977124
journals.sagepub.com/home/tub
SAGE

Hepatoblastomas exhibit marked NNMT downregulation driven by promoter DNA hypermethylation

Maria Prates Rivas[1*], Talita Ferreira Marques Aguiar[1*], Mariana Maschietto[2], Renan B Lemes[1], Luiz Carlos Caires-Junior[1], Ernesto Goulart[1], Kayque Alves Telles-Silva[1], Estela Novak[3,4], Lilian Maria Cristofani[3], Vicente Odone[3], Monica Cypriano[5], Silvia Regina Caminada de Toledo[5], Dirce Maria Carraro[6], Melissa Quintero Escobar[7], Hana Lee[8], Michael Johnston[8], Cecilia Maria Lima da Costa[9], Isabela Werneck da Cunha[10,11], Ljubica Tasic[7], Peter L Pearson[1], Carla Rosenberg[1], Nikolai Timchenko[8] and Ana Cristina Victorino Krepischi[1]

Abstract

Hepatoblastomas exhibit the lowest mutational burden among pediatric tumors. We previously showed that epigenetic disruption is crucial for hepatoblastoma carcinogenesis. Our data revealed hypermethylation of nicotinamide N-methyltransferase, a highly expressed gene in adipocytes and hepatocytes. The expression pattern and the role of nicotinamide N-methyltransferase in pediatric liver tumors have not yet been explored, and this study aimed to evaluate the effect of nicotinamide N-methyltransferase hypermethylation in hepatoblastomas. We evaluated 45 hepatoblastomas and 26 non-tumoral liver samples. We examined in hepatoblastomas if the observed nicotinamide N-methyltransferase promoter hypermethylation could lead to dysregulation of expression by measuring mRNA and protein levels by real-time quantitative polymerase chain reaction, immunohistochemistry, and Western blot assays. The potential impact of nicotinamide N-methyltransferase changes was evaluated on the metabolic profile by high-resolution magic angle spinning nuclear magnetic resonance spectroscopy. Significant nicotinamide N-methyltransferase downregulation was revealed in hepatoblastomas, with two orders of magnitude lower nicotinamide N-methyltransferase expression in tumor samples and hepatoblastoma cell lines than in hepatocellular carcinoma cell lines. A specific TSS1500 CpG site (cg02094283) of

[1]Human Genome and Stem Cell Research Center, Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, São Paulo, Brazil
[2]Research Center, Boldrini Children's Hospital, Campinas, Brazil
[3]Pediatric Cancer Institute (ITACI) at the Pediatric Department, São Paulo University Medical School, São Paulo, Brazil
[4]Molecular Genetics—São Paulo's Blood Center, São Paulo, Brazil
[5]Department of Pediatric, Adolescent and Child with Cancer Support Group (GRAACC), Federal University of São Paulo, São Paulo, Brazil
[6]International Center for Research, A.C. Camargo Cancer Center, São Paulo, Brazil
[7]Department of Organic Chemistry, Institute of Chemistry, University of Campinas, Campinas, Brazil
[8]Department of Surgery, Division of General and Thoracic Surgery, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

[10]Department of Pathology, Rede D'OR São Luiz, São Paulo, Brazil
[11]Department of Pathology, A.C. Camargo Cancer Center, São Paulo, Brazil
*These authors contributed equally to this work.

Corresponding author:
Ana Cristina Victorino Krepischi, Human Genome and Stem Cell Research Center, Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, São Paulo 05508-090, Brazil.
Email: ana.krepischi@ib.usp.br

nicotinamide N-methyltransferase was hypermethylated in tumors, with an inverse correlation between its methylation level and nicotinamide N-methyltransferase expression. A marked global reduction of the nicotinamide N-methyltransferase protein was validated in tumors, with strong correlation between gene and protein expression. Of note, higher nicotinamide N-methyltransferase expression was statistically associated with late hepatoblastoma diagnosis, a known clinical variable of worse prognosis. In addition, untargeted metabolomics analysis detected aberrant lipid metabolism in hepatoblastomas. Data presented here showed the first evidence that nicotinamide N-methyltransferase reduction occurs in hepatoblastomas, providing further support that the nicotinamide N-methyltransferase downregulation is a wide phenomenon in liver cancer. Furthermore, this study unraveled the role of DNA methylation in the regulation of nicotinamide N-methyltransferase expression in hepatoblastomas, in addition to evaluate the potential effect of nicotinamide N-methyltransferase reduction in the metabolism of these tumors. These preliminary findings also suggested that nicotinamide N-methyltransferase level may be a potential prognostic biomarker for hepatoblastoma.

Keywords
Hepatoblastoma, nicotinamide N-methyltransferase, hypermethylation, epigenetics, low lipids, metabolomics

Date received: 17 April 2020; accepted: 29 October 2020

Introduction

Hepatoblastoma (HBL) is the most common primary liver tumor in children, accounting for ;1% of pediatric cancers.[1] However, it is a very rare disease with an incidence of 2.6 cases per 1 million children aged 18 years and younger.[2] HBL presents a relatively normal genomic background with the lowest mutational burden reported among pediatric solid tumors.[3] These embryonal hepatic tumors usually carry cytogenetic alterations, mostly aneuploidies involving gains of chromosomes 2, 8, and 20.[4–6] A few somatic mutations have already been recognized as drivers of HBL tumorigenesis, mainly as activators of the WNT pathway, with recurrent mutations in CTNNB1.[7–10]

This relative paucity of molecular biomarkers poses a challenge to proper risk stratification, and gene expression signatures have been reported in recent years, providing clues about specific HBL subtypes.[7,11,12] A 16-gene signature discriminated two HBL subgroups that resemble early and late phases of liver development.[13] This signature stratified tumors into one group presenting fairly well-differentiated histology and favorable prognosis, and another group with a poorly differentiated histology and worse prognosis. Another model that considered differential activation of hepatic progenitor cell markers and metabolic pathways stratified HBLs in three risk groups.[7] Recently, a new approach adding epigenomics to genomic/transcriptomic data resulted in an HBL risk stratification model composed of three subgroups, based on the degree of hypomethylation, as well as the expression pattern of genes located at the 14q32 locus.[14]

In a previous work, we explored the role of epigenetic mechanisms in HBL by analyzing changes in DNA methylation (DNAm) in comparison to control embryonic and differentiated liver samples;[15] a widespread and non-stochastic pattern of global low-level hypomethylation was disclosed in tumors, with enrichments at intergenic CpG sites. Loss of DNAm in HBL was also reported by Cui et al.[16] Furthermore, aberrant DNAm in specific loci has been described in HBL samples, suggesting that epigenetic alterations are an important mechanism associated with their development.[17,18]

Our previous work evidenced that most of the detected hypermethylated sites were mapped to CpG islands,[15] and a specific gene has drawn our attention due to its role in liver metabolism. The promoter region of NNMT (nicotinamide N-methyltransferase), a highly expressed gene in adipocytes and hepatocytes, was found to be hypermethylated in HBL. A previous study had shown that NNMT methylation detected in the fetal liver is lost in the differentiated liver and inversely correlated with gene expression.[19] This cytosolic enzyme was initially related to the N-methylation of nicotinamide, purines, and other structural analogues[20] using S-adenosyl methionine (SAM) as a methyl group donor, and the best known function of NNMT is associated with the biotransformation of drugs and xenobiotic compounds. SAM is the universal methyl donor for DNA, histones, non-histone proteins, lipids, and other metabolites, and the transfers of a reactive methyl group by NNMT to nicotinamide generate S-adenosyl homocysteine and the metabolic product 1-methylnicotinamide (1MNA).[21] This activity generates a methyl sink of 1MNA, which leads to the depletion of SAM and reduces the global methylation potential of the cell.[22]

To examine whether NNMT expression in HBL could be controlled by promoter hypermethylation, we determined the methylation pattern of CpG sites mapped at the promoter region and assessed the mRNA and protein NNMT expression level in tumors. Furthermore, we explored the potential impact of

NNMT changes on the metabolic profile of HBL in relation to non-tumoral control liver (NTCL) samples.

Patients and methods

Samples

The samples included in this study were HBLs surgically removed between 2016 and 2019 from patients of cancer hospitals and were selected from the biobanks for presenting at least .80% of tumoral cellularity. Thirty HBL (23 fresh-frozen and 15 paraffin block—for some samples, we had material available in both conditions) samples and 11 NTCL samples were recovered from patients enrolled in three Brazilian cancer institutions: A.C. Camargo Cancer Center, GRAACC, and ITACI (São Paulo, SP, Brazil); clinical data are summarized in Table 1. The validation cohort consisted of 15 fresh-frozen HBL samples and 15 NTCL samples from the Cincinnati Children's Hospital Medical Center in Cincinnati (USA); clinical data of this group are summarized in Table 2. Supplementary Table 1S indicates which sample was used for each experimental procedure.

The Research Ethics Committee of the respective institutions approved this research and the use of biological samples, and all samples were collected after informed signed consents from parents or children legal guardians.

Human induced pluripotent stem cell (iPSC) lines derived from the peripheral blood from two healthy individuals were obtained according to Okita et al.[23] Using iPSCs, hepatic differentiation was performed as described by Hay et al.[24] to obtain definitive endoderm, hepatoblasts, and precursor hepatocyte (hepatocyte-like) cells. The generation of iPSC lines derived from healthy individuals[25] was approved by the Ethics Committee of the Institute of Biosciences, University of São Paulo, Brazil (Protocol Number 1.294.118). The liver cancer cell lines, SNU-387 (ATCC CRL-2237), SNU-423 (ATCC CRL-2238), SNU-449 (ATCC CRL-2234), and SNU-475 (ATCC CRL-2236), were acquired from ATCC (USA). Following the recommended protocol, the cell lines were grown in Roswell Park Memorial Institute Medium (RPMI 1640; Thermo Fisher, USA) supplemented with 10% Fetal Bovine Serum (FBS; ThermoFisher), in conditions at 37LC and 5% CO_2. The HBL cell lines HEPG2 (ATCC HB-8065) and C3A (ATCC CRL-10741) were grown in Minimum Essential Medium (MEM; ThermoFisher) supplemented with 10% FBS (ThermoFisher) and maintained at the same conditions.

Gene expression

Total RNA was isolated from HBL and NTCL samples using the RNeasy Mini Kit (QIAGEN, Germany) according to the manufacturer's recommendations. Microfluidics-based electrophoresis (Bioanalyzer, Agilent Technologies, USA) was performed to verify the quality, and only RNA samples with RNA Integrity Number (RIN) 7.0 were used. Complementary DNA (cDNA) was synthesized with the High Capacity RNA-to-cDNA Kit (Applied Biosystems, USA) according to the standard procedures. NNMT expression was evaluated by reverse transcription-quantitative polymerase chain reaction (RT-qPCR) in 35 HBL samples and 24 NTCL samples using the TaqMan System (Applied Biosystems), and the data were normalized to the expression level of the housekeeping gene 18S ribosomal RNA (18S rRNA). All reactions were performed in tri-plicate in a total volume of 10 mL containing 5 mL Master Mix, 1.5 mL water, 50 ng/mL cDNA template, and 0.5 mL of the gene-specific TaqMan Assay Probe Mixture. The delta–delta Ct (DDCt) method was used for data analysis,[26] and the Kruskal–Wallis test and the post hoc Dunn test with Bonferroni correction were used for statistical analyses using GraphPad Prism 7 software.

Protein analysis

Qualitative protein analysis was performed in 14 available HBL samples from the Brazilian cohort by immunohistochemistry using the antibody anti-NNMT antibody (OTI3D8; Abcam, UK); reactions were automated in the BenchMark Ultra-VENTANA equipment. Images were obtained from Aperio Digital Pathology Slide Scanners - AT2 (Leica Biosystems, USA). HBL samples were classified as either positive or negative for NNMT protein expression.

Western blot assays were conducted using samples from the USA cohort. Protein extracts were isolated from available 10 HBL and 5 NTCL samples as previously described.[27,28] This assay was performed in duplicate, where proteins (50 mg) were loaded onto a 4% to 20% gradient gel (BioRad, USA) and transferred to a nitrocellulose membrane (BioRad). The membranes were probed with anti-NNMT (OTI3D8; Abcam) and anti-beta-actin (ab8227; Abcam) antibodies. The results of Western blotting are also presented as ratios of the protein to the loading control, which were obtained by using ImageJ software (NIH, USA).[29] The t-test was used in the data analysis.

Metabolomics by nuclear magnetic resonance

Ten Brazilian HBL samples and eight NTCL samples were available for analysis by high-resolution magic angle spinning nuclear magnetic resonance (HR-MAS NMR) spectroscopy. ^{1}H NMR spectra were acquired using a Bruker Avance spectrometer (Bruker BioSpin, Germany) operating at 400 MHz and equipped with

Table 1. Clinical features of the 30 Brazilian HBL cases.

ID/gender/ age at diagnosis	Histology	AFP (ng/mL)	Risk stratification	PRETEXT	Chemotherapy protocol	Transplant	Metastasis	Relapse	Deceased	Other features
HB02, M, 1 m	Epithelial and mesenchymal mixed	NA	Low risk	2	SIOPEL3	No	No	No	No	
HB05, M, 21 m	Epithelial and mesenchymal mixed	NA	High risk	4	NA	Yes	No	No	No	
HB15, F, 18 m	Epithelial embryonal	5,668,000	Intermediate risk	4	NA	Yes	No	No	Yes	
HB16, M, 9 m	Epithelial fetal	824	Intermediate risk	4	SIOPEL3	No	No	No	No	
HB17, F, 36 m	Epithelial fetal	.400,000	Low risk	1	SIOPEL3	No	No	No	No	
HB18, M, 9 m	Epithelial and mesenchymal mixed	.200,000	Low risk	3	SIOPEL3	Yes	No	No	No	
HB28, M, 17 y	Epithelial and mesenchymal mixed	NA	High risk	4	SIOPEL4	No	No	Yes	Yes	Hepatomegaly at birth
HB30, M, 54 m	HB with HCC features	.1,000,000	High risk	2	SIOPEL4	Yes	Lung	Yes	Yes	
HB31, M, 30 m	Epithelial fetal	742,000	Low risk	3	NA	No	No	No	No	Non-functional kidney
HB32, F, 36 m	Epithelial and mesenchymal mixed	9328000	High risk	4	SIOPEL4	Yes	Lung	No	No	
HB33, F, 1 m	Epithelial embryonal and fetal	28,312,000	Intermediate risk	2	SIOPEL3	No	No	No	No	Congenital HB and unilateral renal agenesis
HB34, F, 19 m	Epithelial fetal	416,430	Intermediate risk	3	SIOPEL3	No	No	No	No	Mother with Hepatitis C in pregnancy (no treatment)
HB35, M, 26 m	Epithelial fetal	54,800	Intermediate risk	3	SIOPEL3	No	No	No	No	
HB36, M, 31 m	Epithelial embryonal and fetal	76,348	Low risk	3	AHEP 0731—COG	No	No	No	No	
HB37, F, 13 m	Epithelial embryonal	1,870,000	Intermediate risk	2	SIOPEL3	No	No	No	No	
HB38, F, 147 m	Epithelial fetal	643.4	High risk	4	SIOPEL3/AHEP 0731—COG	No	No	Yes	No	Ischemic anoxic neuropathy due to extreme prematurity
HB39, M, 84 m	Epithelial with macrotrabecular pattern	300,000	High risk	2	SIOPEL2	No	No	No	Yes	Pilocytic astrocytoma (posterior fossa) after HB
HB40, M, 22 m	Epithelial embryonal and fetal	1842	Low risk	1	SIOPEL3	No	No	No	No	
HB42, M, 45 m	Epithelial fetal	1267	Intermediate risk	1	SIOPEL2	No	No	No	No	
HB43, M, 20 m	Epithelial embryonal	183,476	Intermediate risk	4	SIOPEL3	Yes	No	No	No	
HB44, M, 5 m	Epithelial and mesenchymal mixed	300,000	Intermediate risk	2	SIOPEL2	No	No	No	No	
HB45, F, 5 m	Epithelial fetal	445611	Intermediate risk	2	SIOPEL3	No	No	No	Yes	
HB46, M, 28 m	Epithelial and mesenchymal mixed	.200,000	High risk	4	SIOPEL6	No	Lung	No	No	Syndromic patient (craniosynostosis and developmental delay)

(continued)

Table 1. Continued

ID/gender/ age at diagnosis	Histology	AFP (ng/mL)	Risk stratification	PRETEXT	Chemotherapy protocol	Transplant	Metastasis	Relapse	Deceased	Other features
HB47, M, 7 m	Epithelial fetal	653,190	NA	NA	SIOPEL3	No	No		No	
HB48, M, 24 m	Epithelial embryonal and fetal	60,500	Low risk	2	SIOPEL6	No	No		No	
HB49, M, 7 m	Epithelial and mesenchymal mixed	65,000	High risk	4	SIOPEL3	Yes	Lung		No	
HB50, M, 1 m	Epithelial and mesenchymal mixed	NA	Intermediate risk	1	AHEP 0731—COG	No	No		Yes	Congenital HB—extreme prematurity
HB72, M, 5 m	Epithelial fetal	2,565,530	Intermediate risk	4	AHEP 0731—COG	No	No	No	No	Congenital HB
HB79, M, 9 m	Epithelial fetal	.50,000	High risk	4	SIOPEL4	Yes	No	No	No	
HB81, M, 20 m	Epithelial and mesenchymal mixed	.100,000	High risk	4	SIOPEL4	No	Lung	No	No	

AFP: alphafeto protein; F: female; m: months; y: years; M: male; NA: not available; HB: hepatoblastoma; HCC: hepatocellular carcinoma.

Table 2. Clinical Features of the 15 CCHMC HBL cases.

ID/gender/age at diagnosis	Histology	AFP (ng/mL)	Risk stratification	PRETEXT	Chemotherapy protocol	Transplant	Metastasis	Relapse	Deceased	Other features
9T, F, 4 y	Epithelial embryonal	304.2	High risk	4	C5VD, ICE	No	Yes	Yes	No	
13T, F, 2 y	Epithelial embryonal and fetal	2526	High risk	4	AHEP0731, ICE, pazopanib	Yes	Yes	Yes	No	
14T, M, 2 y	Epithelial fetal	.300,000	Intermediate risk	4	C5VD	Yes	No	No	No	
16T, M, 1 y	Epithelial with macrotrabecular pattern	358	Low risk	1	AHEP0731	No	No	No	No	
18T, F, 2 y	Epithelial and mesenchymal mixed	88,285	Intermediate risk	3	SIOPEL6	No	No	No	No	Trisomy 18
21T, M, 3 y	HB with HCC features	845,140	High risk	4	C5VD	Yes	No	No	No	
23T, F, 1 y	Epithelial and mesenchymal mixed	38,932	High risk	4	AHEP0731	Yes	No	No	No	
24T, F, 3 y	HB with HCC features	.200,000	High risk	4	AHEP1531	Yes	Yes	No	No	Ruptured tumor
25T, M, 2 y	Epithelial fetal	.1,000,000	High risk	4	SIOPEL4	Yes	Yes	No	No	
26T, F, 2 y	Epithelial embryonal and fetal	19,202	Intermediate risk	4	AHEP1531	Yes	No	No	No	Extreme prematurity
27T, F, 2 y	Epithelial and mesenchymal mixed	2234	Intermediate risk	2	AHEP1531	No	Yes	No	No	
28T, F, 2 y	Epithelial and mesenchymal mixed	1179	Intermediate risk	3	AHEP0731	No	Yes	No	No	
29T, M, 2 y	Epithelial embryonal	.300,000	High risk	4	AHEP0731, C5VD	No	Yes	Yes	No	
31T, M, 2 y	Epithelial embryonal and fetal	7052	High risk	2	AHEP1531	No	Yes	No	No	
33T, F, 8 y	Epithelial embryonal and fetal	.300,000	High risk	2	AHEP1531	No	No	No	No	Trisomy 18

CCHMC: Cincinnati Children's Hospital Medical Center; AFP: alphafeto protein; F: female; y: years; M: male; NA: not available; HB: hepatoblastoma; HCC: hepatocellular carcinoma.

the double nuclei 4 mm probe for HR-MAS. One-dimensional water-suppressed ^{1}H NMR spectra were recorded with the nuclear overhauser effect spectroscopy (NOESY) pulse sequence and 128 repeats, and the T_2-edited spectra were recorded using the CPMG (Carr–Purcell–Melboom–Gill) pulse sequence with 128 repetitions. All spectra were recorded at a magic angle spinning frequency of 3.5 kHz and at 293 K. Chemometrics analysis was performed using MetaboAnalyst (www.metaboanalyst.ca). Details for NMR spectra processing, data pre-processing for chemometrics, and metabolites' assignments were previously described.[30,31]

Statistical analysis

To evaluate overall survival rates, Kaplan–Meier method was applied by means of the survfit function from R statistical package and the p-value from log-rank test was reported. Nonparametric Mann– Whitney–Wilcoxon rank-sum test was applied to investigate differences in NNMT gene expression levels (RT-qPCR) between pairs of groups. The Pearson correlation coefficient (r) between RT-qPCR relative quantification (RQ) values and metabolite variable importance in projection (VIP) scores (1HR NMR HR-MAS) per metabolite was estimated applying the Spearman non-parametric test.

Results

The DNAm values of HBL and NTCL samples were recovered from our previous work.[15] These data show that non-tumoral liver samples present a homogeneous methylation pattern, and HBL samples present a tendency of hypermethylation, although with heterogeneity. The HM450K platform contains 11 CpG sites located at the NNMT sequence, from which 10 reached all quality control parameters (Supplementary Table 2; Figure 1(a)). In Figure 1(a), boxplots of each NNMT CpG site are depicted showing the level of DNAm of HBL and NTCL samples. These boxplots and the heat-map (Figure 1(c)) showed that the CpG sites located at the 5#of the NNMT are more methylated in HBL samples than in NTCL samples. Three out of these 10 CpG sites are located at TSS1500, but only one (cg02094283, in green) was considered differentially methylated for both paired and unpaired analyses.

Significant downregulation of NNMT expression was observed in the group of 20 HBL samples com-pared to 9 NTCL samples ($p < 0.001$; Figure 2(a)). Only four HBL samples exhibited expression levels equivalent to NTCL samples, and two of them were derived from HBL patients diagnosed older than 12 years. The remaining 16 HBL samples presented NNMT expression levels similar to hepatoblasts. The analysis in iPSC samples showed a progressive increase in NNMT expression during in vitro hepatocyte differentiation; as expected, NNMT expression was very low in stem cells and derived definitive endoderm cells, while hepatoblasts exhibited a modest increase in expression, with hepatocyte-like cells presenting NNMT expression levels similar to NTCL samples. To confirm the reduction in the NNMT expression, we used a validation cohort of paired 15 HBL and NTCL samples, and a decrease in the expression of NNMT was also detected in tumor samples ($p < 0.001$; Figure 2(a)). Moreover, we found that NNMT expression was two orders of magnitude lower in the HBL cell lines HEPG2 and C3A than in four hepatocellular carcinoma (HCC) cell lines. The TSS1500 CpG (cg02094283) was used for the correlation analysis of DNAm (M values) versus NNMT expression (logged relative expression values), which is showed in Figure 2(b); in tumors, there is an inverse correlation between the methylation level of the TSS1500 cg02094283 and NNMT expression ($R = –0.8271$; $p < 0.0001$). In addition, using data from a published study in HBLs,[7] which evaluated 50 tumor samples, we could observe that NNMT was also more expressed in the NTCL samples, corroborating our expression findings (Supplementary Figure 1S).

Qualitative NNMT analysis by immunohistochemistry was performed in 15 Brazilian HBL. In hepatocytes of NTCL samples, NNMT protein labeling was observed in the membrane, cytoplasm, and nucleus. A decrease in the level of NNMT protein was validated in the majority of the HBL (Figure 3), which was detected in several cellular components, although the loss of protein labeling was heterogeneous in some tumors (Supplementary Table 3S); one sample (HB50) did not exhibit change regarding NNMT protein expression.

A Western blot assay was used in the validation cohort as another approach to verify the expression levels of the NNMT protein (Figure 4(a)). As already revealed by NNMT immunohistochemistry, the

NNMT protein presented a wide reduction in HBL samples compared to NTCL samples ($p < 0.001$; Figure 4(b)), although its decrease was not homogeneous. In addition, the Pearson analysis showed a strong correlation of the gene and protein NNMT expression in tumor samples ($R = 0.93$; $p = 9.2e–05$; Figure 4(c)).

Statistically significant differences on NNMT gene expression depending on HBL histological subtypes were not detected (Supplementary Figure 2S). Using CHIC (Children's Hepatic tumors International Collaboration) parameters,[32,33] the HBL cases were stratified according to their risk (Table 1). To evaluating the overall survival rates, Kaplan–Meier method was applied considering the level of NNMT expression (Supplementary Figure 3S), considering the first 36 months after diagnosis and the level of NNMT expression; the mean RQ value of the NNMT

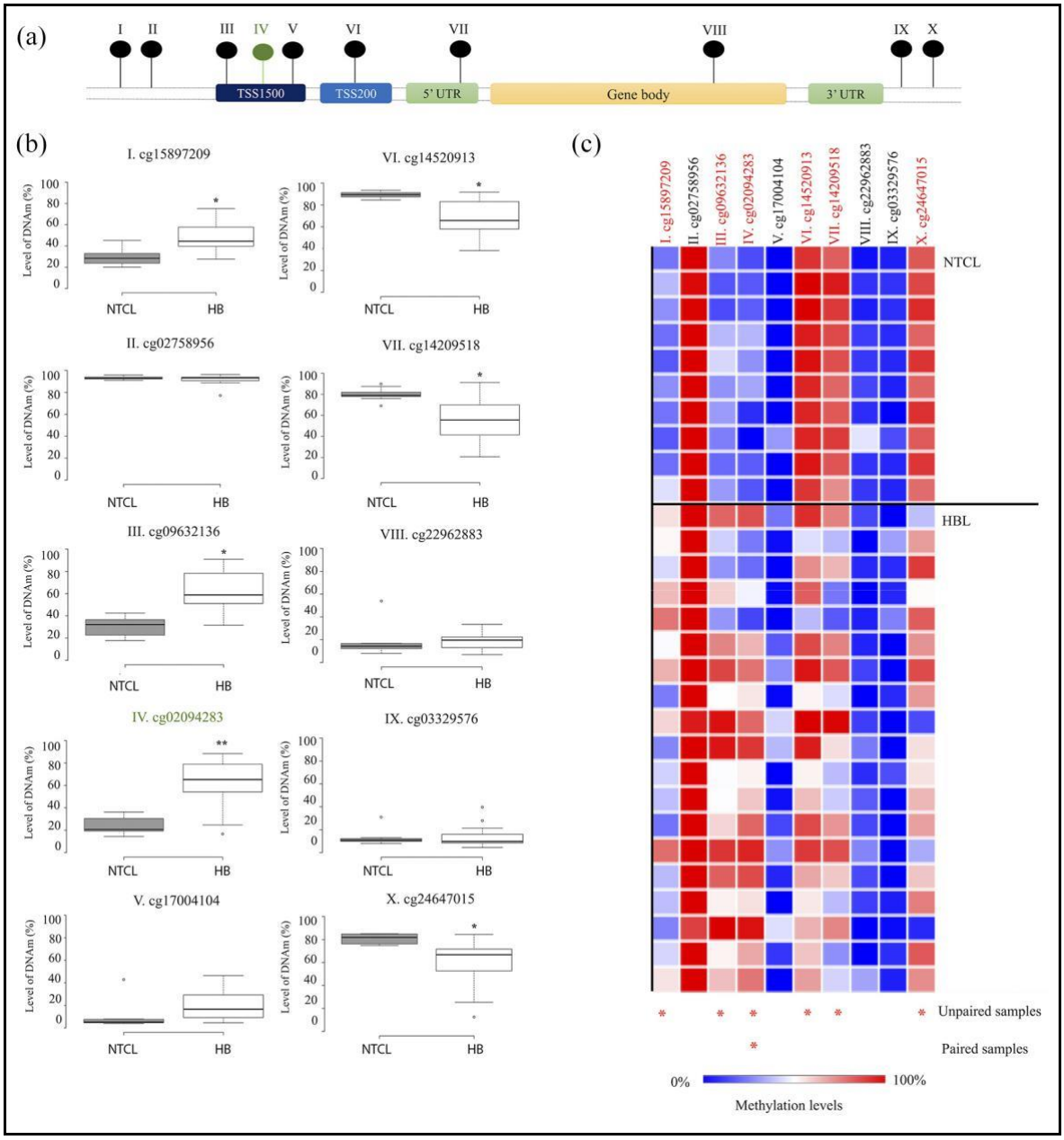


Figure 1. DNA methylation (DNAm) pattern of NNMT CpG sites in hepatoblastomas (HBLs) and non-tumoral control liver (NTCL) samples. The HM450K platform contains 11 CpG sites located at the NNMT sequence, from which 10 reached all quality control parameters; DNAm beta values (ranging from 0% to 100%) of the 10 CpG sites mapped at the NNMT gene sequence were retrieved from the Infinium Illumina 450K BeadChip arrays from previous study[15] in 19 HBL and 10 NTCL samples. Beta values were transformed into M values before performing comparison between groups, employing an empirical Bayesian framework linear model from limma. (a) The regions of the NNMT (TSS1500, TSS200, 5# UTR, gene body, and 3# UTR) are shown with the associated CpG sites depicted as numbered lollypops, according to their genomic coordinates; in green, the TSS1500 cg02094283 was differentially methylated for both paired and unpaired analyses.(b) Boxplots of each CpG site are presented showing the level of DNAm (%) of HBL and NTCL samples (median values indicated by lines); the asterisks (*) indicate the CpGs with significant differentially methylation (p\0.03) between unpaired samples, and one of them (cg02094283) presents significant differentially methylation (p\0.04) in both paired and unpaired analyses (**). (c) Heatmap showing the level of DNAm of the 10 NNMT CpG sites (the identification of each CpG site is given above in the columns) in group of tumors and NTCL; above is indicated the CpG site and below a DNAm level scale, in which blue and red correspond to lower and higher methylation content, respectively. The red asterisks marked the six differentially methylated CpG sites; 3 out of 10 CpG sites are located at TSS1500, and only one (cg02094283) was considered differentially methylated in both paired and unpaired analyses (data generated based on the results from Maschietto et al. 2016[15]).

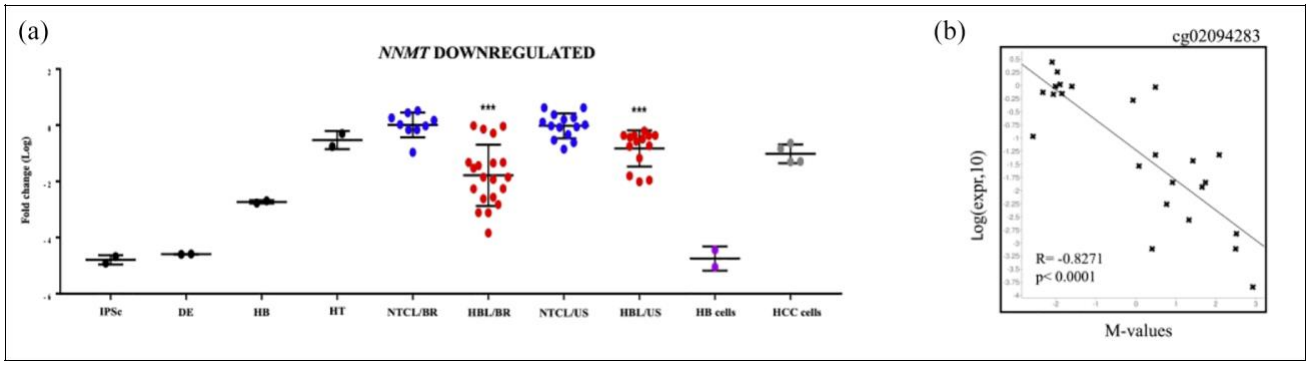


Figure 2. Downregulation of NNMT expression in hepatoblastomas (HBLs): (a) NNMT expression analyzed by reverse transcription-quantitative polymerase chain reaction (RT-qPCR). In black dots, iPSC: induced pluripotent stem cells (iPSCs) derived from healthy individuals; DE: definitive endoderm cells; HB: hepatoblast cells; HT: hepatocyte-like cells. In blue, non-tumoral control liver (NTCL) samples from Brazilian (BR) and American (USA) cohorts. In red, tumors: HBL samples from Brazilian (BR) and American (USA) cohorts; liver cancer cell lines: in purple, HBL cell lines (HEPG2 and C3A); in gray, hepatocellular carcinoma cell lines (SNU-387, SNU-423, SNU-449, and SNU-475). For iPSC, DE, HB, and HT showing two samples for each one, corresponding to two different patients. (b) TSS1500 CpG (cg02094283) was used for the correlation analysis of DNA methylation (M values) versus NNMT expression (logged relative expression values); in tumors, there is an inverse correlation between the methylation level of the TSS1500 cg02094283 and NNMT expression ($R = –0.8271$; $p < 0.0001$).

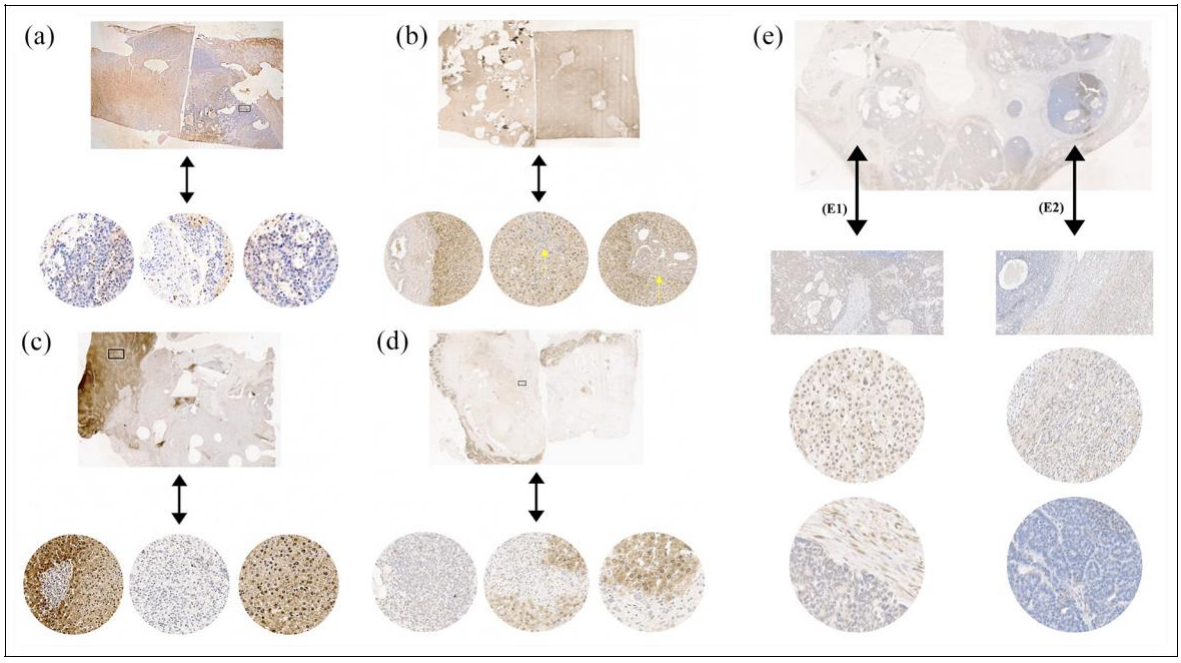


Figure 3. Reduction in the expression of the NNMT protein evaluated in HBLs by immunohistochemistry. In hepatocytes, the detection of the NNMT protein in the cytoplasm and nucleus is expected (brown labeling). Panels A to E represent five HBL cases, with images obtained from Aperio Digital Pathology Slide Scanners with increases of 500, 300, and 200 mm. (a) HB31 tumor showing loss of NNMT expression in some regions. (b) HB18 exhibiting loss of expression in some regions, with complete nuclear NNMT absence in specific regions, indicated by the yellow arrow. (c) HB15 and (d) HB32 exhibiting the loss of expression in some regions. (e) HB33, a congenital tumor presenting variable NNMT expression in the epithelial-fetal region (E1), and a total absence of the NNMT labeling in the epithelial-embryonal component (E2).

expression of all HBL samples was obtained (mean RQ = 0.15) and then HBL was classified as Group A: high expression level (above 0.15; n = 15 tumors) or Group B: low expression level (below 0.15; n = 20 tumors). No statistical difference was observed in HBL exhibiting higher or lower NNMT expression level than

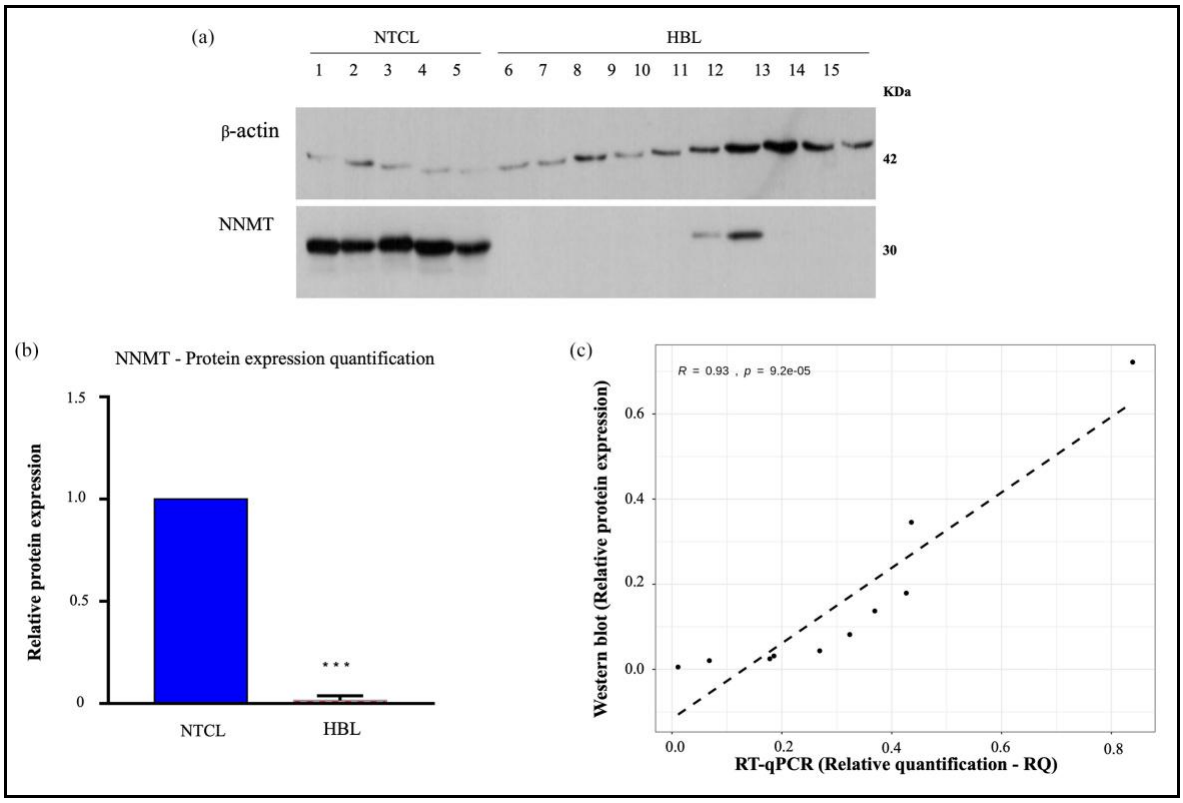


Figure 4. NNMT protein expression is decreased in HBLs in comparison to non-tumoral control liver (NTCL) samples by Western blot assay: (a) Immunoblotting analysis in NTCL (1–5) and HBL (6–15) samples showed a decrease in the NNMT protein in HBL. (b) Bar graph showing the NNMT protein level as a ratio to b-actin (***p-value < 0.001). (c) Correlation analysis between relative protein expression and RT-qPCR (reverse transcription-quantitative polymerase chain reaction) relative quantification (RQ) values per sample; the Pearson correlation coefficient and corresponding p-value are written in the top-left part of the graph. The dashed line represents the linear regression line.

the mean of the group. Following, NNMT expression level was compared between pairs of groups according to the clinical variables associated with HBL prognosis parameters (risk stratification, age at diagnosis, alpha-feto protein (AFP) dosage level in the range of 101–1000 ng/mL, PRETEXT IV, vascular invasion, metastasis, transplantation and relapse; Supplementary Figure 4Sa). It was detected a statistically significant difference related to the age at diagnosis ($p < 0.015$; Supplementary Figure 4Sb); patients diagnosed with more than 8 years old presented significantly higher NNMT expression. In addition, although not significant, the groups of patients with AFP levels 101 to 1000 ng/mL and high risk also appeared to be associated with higher NNMT expression.

We also investigated 10 Brazilian HBL samples and 8 NTCL samples using metabolomics by 1HR NMR HR-MAS, aiming to identify changes in metabolites that could be related to the reduction in the NNMT protein levels. HBL showed to be more heterogeneous than NTCL samples, as could be seen in the Partial least squares-discriminant analysis (PLS-DA) results (Figure 5(a)). The highest loadings and VIP values were observed for lipids, aromatic amino acids, and other metabolites that could be linked to NNMT lower activity in HBL (Figure 5(b) and (c)). In summary, 15 metabolites were detected with different concentrations between tumors and NTCL samples (boxplots in Figure 5(c)), 10 of them exhibiting increase in HBLs, and 5 with reduction. The five metabolites detected with reduced levels in tumors compared to NTCL samples were the amino acid tryptophan and four peaks of the triglyceride lipid class, used as an energy reserve source, both saturated and unsaturated fatty acids (–CH3 (C18) cholesterol; –CH3; –CH– glyceryl; –CH=CH). Among the 10 metabolites detected with increased concentrations, there are formate, some amino acids (tyro-sine, alanine, phenylalanine), and several peaks of structural lipids, such as phospholipids, which are part of the composition of cell membranes. Typical ^{1}H NMR HR-MAS spectra of HBL and NTCL samples are illustrated in Figure 5(d), where the most important metabolites (peaks 1–16) are marked.

Six HBLs and three NTCL samples were investigated using both methods ^{1}HR NMR HR-MAS and RT-qPCR. Considering all studied samples, the Pearson correlation coefficients (r) between NNMT RT-qPCR RQ values and metabolite VIP scores (^{1}HR

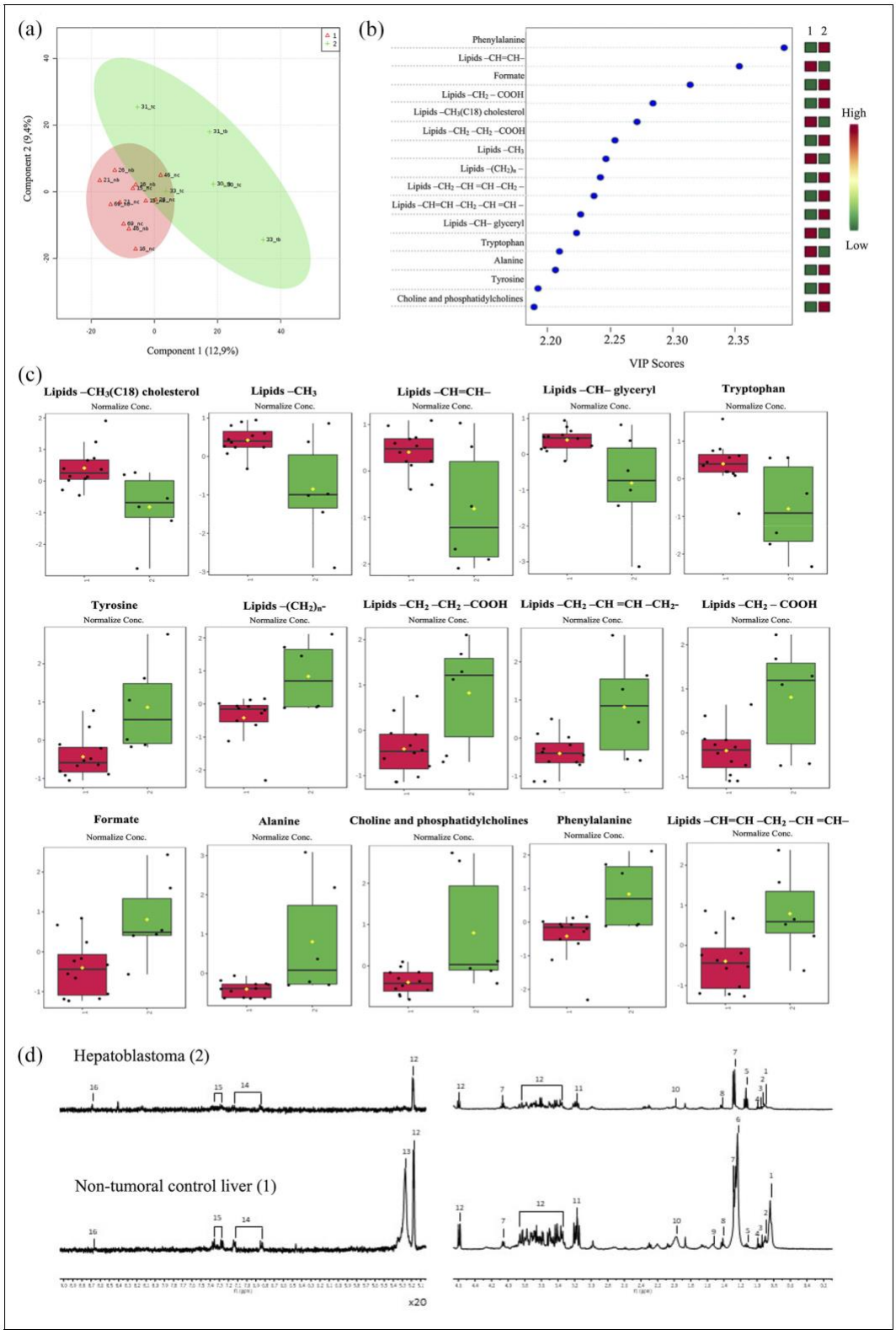


Figure 5. Metabolomics data obtained from HBL and non-tumoral control liver (NTCL) samples: (a) Scores of the PLS-DA model obtained using T^2-edited ^{1}H NMR (CMPG) HR-MAS spectral data, constructed from data of NTCL samples (red triangles, and indicated as Group 1) and tumor (green plus sign, and indicated as Group 2). (b) The VIP values are indicated, where the colored boxes in the right show the relative concentrations of the corresponding metabolite in each group under study. (c) The boxplots show the HBL and NTCL samples levels of the top-15 altered metabolites (Student's t-test, $p < 0.05$); black dots represent the concentrations of each sample, and the yellow diamond the average values of the group (HBL illustrated in green and NTCL samples in red). (d) ^{1}H NMR HR-MAS spectrum of HBL (upper) and NTCL samples (lower); two regions are presented between 0.20 and 4.60 ppm and amplified 20 times in the regions between 5.10 and 9.00 ppm.

NMR HR-MAS) showed strong positive correlation ($p > 0.697$) in all five metabolites which are reduced in HBLs (Supplementary Table 4S and Supplementary Figure 5S), one of them statistically significant (the lipid class –CH3; $p < 0.05$). In addition, all 10 metabolites with increased levels in HBLs exhibited a negative correlation with NNMT expression, although not significant.

Discussion

NNMT was originally identified as the enzyme responsible for the methylation of nicotinamide (NAM), producing 1MNA, which is a precursor for NAD^+, a cofactor known for donating electrons to the mitochondrial complex I and to multiple oxidoreductases.[34] Recent studies have expanded the role of NNMT, which has been increasingly associated with the regulation of multiple metabolic pathways in adipose and liver tissues through the consumption of methyl donor groups and the generation of active metabolites.[35–39]

NNMT overexpression has been reported in breast cancer, glioblastoma, and papillary thyroid cancer, among others.[20,40,41] Ulanovskaya et al.[42] previously revealed that NNMT overexpression leads to an increase in 1MNA in various aggressive cancer cell lines; the authors proposed that the accumulation of 1MNA caused by NNMT overexpression would sequester methyl groups in cancer cells, leading to diminished methylation with changes in the epigenetic profiles.

However, it has been shown that NNMT is reduced in HCC; Kim et al.[39] analyzed 120 patients with HCC and demonstrated that NNMT expression was reduced in the majority of the examined samples and correlated with poor prognosis. Recently, these findings were corroborated in HCC samples, with the detection of significant NNMT downregulation;[43] in addition, NNMT expression was found to be heterogeneous, and tumors exhibiting high NNMT protein levels presented unfavorable prognostic features, such as vascular invasion and distant metastasis. Although the precise mechanisms by which the reduction in NNMT could contribute to liver cancer are not known, Shin et al.[44] recently revealed that NNMT depletion enhances autophagy and contributes to liver cancer cell survival and tumor growth, thus providing new insights into the mechanisms of liver cancer in adult patients. Nevertheless, very little is known about the expression pattern and role of NNMT in pediatric liver cancer.

Previous work from our group identified hypermethylation at the promoter region of NNMT in HBL.[15] As DNAm is a widely recognized epigenetic mechanism for the regulation of gene expression,[45] we speculated that NNMT dysregulation by epigenetic changes might be a plausible factor contributing to HBL tumorigenesis, deserving further investigation.

Accordingly, we examined whether DNAm alterations in HBL were correlated with NNMT expression. Significant NNMT downregulation was observed in these embryonal liver tumors compared to NTCL tissues in two different HBL cohorts. Furthermore, a specific TSS1500 CpG site (cg02094283) of NNMT was hypermethylated in HBL in both paired and unpaired analyses, with an inverse correlation between its methylation level and NNMT expression, with other neighbor CpG sites following a similar methylation pattern.

In addition, regarding the reduction in NNMT observed by Shin et al.[44] in liver cancer cells, we found two orders of magnitude lower NNMT expression in HBL cell lines compared to HCC cell lines. In general, although the level of the NNMT protein was heterogeneous among HBL, remarkable reductions in NNMT mRNA and protein levels were observed using RT-qPCR, immunohistochemistry, and Western blot assays. We found a strong correlation between gene and protein expression of NNMT in tumor samples, indicating that the decrease in mRNA directly regulates the protein level in HBL.

In the HBL samples with available data, NNMT expression did not significantly impact the 36-month survival, probably because of the small sample size. However, higher NNMT expression was associated with known clinical variables of HBL prognosis. HBL from patients diagnosed with more than 8 years old exhibited higher NNMT level ($p < 0.015$), maybe because these tumors arose from more differentiated hepatocytes. It is also worthy to highlight that tumors from patients with AFP levels, 101 to 1000 ng/mL, and tumors from patients classified as high risk also appeared to be associated with higher NNMT expression.

Cancer cells have the ability to generate energy in a nutrient-deficient environment, and aberrant metabolism has become a characteristic hallmark of malignancies,[46–48] including shifts in the metabolism of lipids.[47–49] Using untargeted metabolomics by NMR, variation in many metabolites was detected in HBLs, and our data clearly showed changes in the lipid content, with an increase of the structural class of lipids, commonly detected in cancer cells because they are used for the synthesis of membranes (cytoplasmic and organelles).[50,51] However, lipids used as an alternative source of energy were found to be reduced in HBLs with correlation with NNMT expression, suggesting that these tumors are using this class of molecules for their own energy demand for tumor progression.[46–49,51]

Structural changes in membranes, cell signaling, gene expression, protein distribution, and disruption of energy homeostasis are caused by lipid changes, and these changes may have consequences in autophagy, necrosis, apoptosis, proliferation, differentiation, growth, and chemotherapy resistance.[52–54]

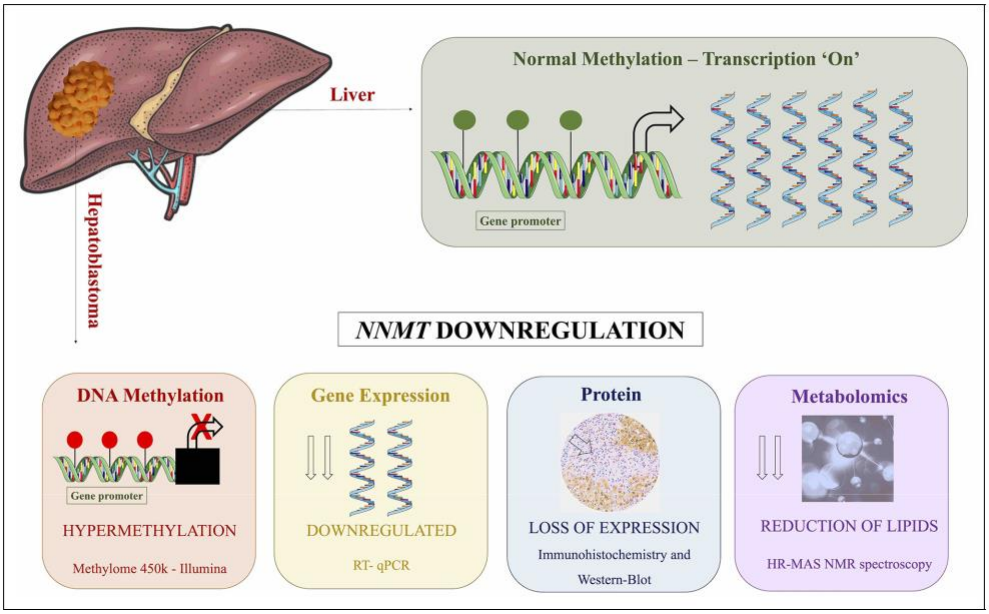


Figure 6. Main findings of this study showing effects of NNMT depletion in hepatoblastomas modulated by DNA methylation.

Hepatocytes present the highest levels of 1MNA,[21] evidencing a role for NNMT in the liver. Functionally, 1MNA increases SIRT1,[55] leading to the suppression of fatty acid and cholesterol synthesis in normal hepatocytes and resulting in decreased triglyceride and cholesterol contents as well as liver inflammation.[35] Therefore, it is expected that NNMT downregulation would impact the SIRT1 function; nevertheless, the relation between NNMT and SIRT1 is complex and modulated by several factors. Adaptations of these normal conditions were reported in cancer cells, with abnormal feedback for increased or decreased NNMT content, as well as changes in other pathways of lipid metabolism, as a source of energy.[34–37,56] Recent studies[36,52,53] have shown that NNMT activity consumes SAM, which decreases the methylation of selected epigenetic marks, with cell-specific effects such as fat accumulation in adipocytes. Furthermore, similar to our findings, a decrease in the lipid content in liver cells, especially affecting the polyunsaturated fatty acids, was previously reported following NNMT downregulation,[57] providing an indirect evidence that supports the hypothesis that NNMT downregulation could diminish the lipid content in HBLs. Even though this hypothesis is appealing, the functional impact of NNMT downregulation should be evaluated in hepatocytes and their precursors to be causally associated with the detected lower lipid content in HBL.

This study provides further evidence that the reduction of NNMT in liver cancer, previously observed by others, is a wide phenomenon, now also validated in the embryonal tumor HBL. Furthermore, we have shown for the first time that HBL presented reduction in the level of specific classes of lipids. Taken together, our data highlight the role of DNAm in the regulation of NNMT expression in HBL and suggest a possible effect of NNMT depletion on the lipid metabolism.

The main findings of this study are summarized in Figure 6.

Acknowledgements

The authors thank the patients and their families for participating in the study.

Author contributions

M.P.R., T.F.M.A., M.M., and A.C.V.K. conceived the study and participated in its design. M.P.R., T.F.M.A., M.M., L.C.C-.J., E.G., K.A.T-.S., R.B.L., E.N., L.M.C., V.O., H.L., M.J., L.T., N.T., and A.C.V.K. performed the collection and assembly of data. M.P.R., T.F.M.A., M.C., R.B.L., S.R.C.d.T., D.M.C., C.M.L.d.C., I.W.d.C., M.Q.E., L.T., P.L.P., C.R., N.T., and A.C.V.K. realized the data analysis and interpretation. M.P.R., T.F.M.A., M.M., L.T., C.R., P.L.P., and A.C.V.K. wrote the manuscript. All authors have read and approved the final version of the manuscript.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial sup-port for the research, authorship, and/or publication of this article: The present study was supported by grants from

FAPESP (CEPID—Human Genome and Stem Cell Research Center 2013/08028-1; 2018/21047-9; 2018/06510-4; fellow-ships 2015/06281-7, 2016/04785-0, 2016/23462-8) and CNPq (141625/2016-3). The funders had no roles in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Ethical approval

Samples were recovered from patients enrolled in three Brazilian cancer institutions: A.C. Camargo Cancer Center, GRAACC and ITACI (São Paulo, Brazil), and Cincinnati Children's Medical Hospital Center in Cincinnati (USA). The Research Ethics Committee of the respective institutions approved the research for the used biological samples, and all samples were collected after the informed signed consents were obtained from parents or legal guardians. For induced pluripotent stem cells (iPSCs), the Ethics Committee of the Instituto de Biociências at Universidade de São Paulo, Brazil (Protocol Number 1.294.118), approved the present study, including experimental procedures involving samples from human subjects.

Guarantor

M.P.R., T.F.M.A., and A.C.V.K. are listed as guarantors of the paper.

ORCID iDs

Mariana Maschietto https://orcid.org/0000-0003-2589-3186

Estela Novak https://orcid.org/0000-0002-5732-4845 Peter L Pearson https://orcid.org/0000-0002-5654-4264 Ana Cristina Victorino Krepischi https://orcid.org/0000-0003-2931-8605

Supplemental material

Supplemental material for this article is available online.

References

1. Czauderna P and Garnier H. Hepatoblastoma: current understanding, recent advances, and controversies. F1000res 2018; 7: 53.
2. Fajardo-Gutie´rrez A, Jua´rez-Ocan˜a S, Gonza´lez-Mir-anda G, et al. Incidence of cancer in children residing in ten jurisdictions of the Mexican Republic: importance of the cancer registry (a population-based study). BMC Cancer 2007; 7: 68.
3. Grobner SN, Worst BC, Weischenfeldt J, et al. The land-scape of genomic alterations across childhood cancers. Nature 2018; 555: 321–327.
4. Stejskalova E, Malis J, Snajdauf J, et al. Cytogenetic and array comparative genomic hybridization analysis of a series of hepatoblastomas. Cancer Genet Cytogenet 2009; 194: 82–87.
5. Rodrigues TC, Fidalgo F, da Costa CM, et al. Upregu-lated genes at 2q24 gains as candidate oncogenes in hepa-toblastomas. Future Oncol 2014; 10(15): 2449–2457.
6. Rushton J and Lopez-Terrada D. Molecular and genetic basis of childhood cancer. Cancer Biomark 2010; 9(1–6): 211–234.
7. Sumazin P, Chen Y, Trevin˜o LR, et al. Genomic analysis of hepatoblastoma identifies distinct molecular and prog-nostic subgroups. Hepatology 2017; 65(1): 104–121.
8. Curia MC, Zuckermann M, De Lellis L, et al. Sporadic childhood hepatoblastomas show activation of beta-cate-nin, mismatch repair defects and p53 mutations. Mod Pathol 2008; 21(1): 7–14.
9. Koch A, Denkhaus D, Albrecht S, et al. Childhood hepatoblastomas frequently carry a mutated degradation targeting box of the beta-catenin gene. Cancer Res 1999; 59: 269–273.
10. Udatsu Y, Kusafuka T, Kuroda S, et al. High frequency of beta-catenin mutations in hepatoblastoma. Pediatr Surg Int 2001; 17(7): 508–512.
11. Maris JM and Denny CT. Focus on embryonal malig-nancies. Cancer Cell 2002; 2(6): 447–450.
12. Eichenmu¨ller M, Trippel F, Kreuder M, et al. The geno-mic landscape of hepatoblastoma and their progenies with HCC-like features. J Hepatol 2014; 61(6): 1312–1320.
13. Cairo S, Armengol C, De Reynies A, et al. Hepatic stem-like phenotype and interplay of Wnt/beta-catenin and Myc signaling in aggressive childhood liver cancer. Can-cer Cell 2008; 14: 471–484.
14. Carrillo-Reixach J, Torrens L, Simon-Coma M, et al. Epigenetic footprint enables molecular risk stratification of hepatoblastoma with clinical implications. J Hepatol 2020; 73(2): 328–341.
15. Maschietto M, Rodrigues TC, Kashiwabara AY, et al. DNA methylation landscape of hepatoblastomas reveals arrest at early stages of liver differentiation and cancer-related alterations. Oncotarget 2016; 8(58): 97871–97889.
16. Cui X, Liu B, Zheng S, et al. Genome-wide analysis of DNA methylation in hepatoblastoma tissues. Oncol Lett 2016; 12(2): 1529–1534.
17. Rumbajan JM, Maeda T, Souzaki R, et al. Comprehen-sive analyses of imprinted differentially methylated regions reveal epigenetic and genetic characteristics in hepatoblastoma. BMC Cancer 2013; 13: 608.
18. Honda S, Minato M, Suzuki H, et al. Clinical prognostic value of DNA methylation in hepatoblastoma: four novel tumor suppressor candidates. Cancer Sci 2016; 107(6): 812–819.
19. Bonder M, Kasela S, Kals M, et al. Genetic and epige-netic regulation of gene expression in fetal and adult human livers. BMC Genomics 2014; 15: 860.
20. Li S, Qiao L, Yang Z, et al. Prognostic value of nicotina-mide N-methyltransferase expression in patients with solid tumors: a systematic review and meta-analysis. Front Physiol 2018; 9: 1407.
21. Eckert MA, Coscia F, Chryplewicz A, et al. Proteomics reveals NNMT as a master metabolic regulator of cancer-associated fibroblasts. Nature 2019; 569(7758): 723–728.
22. Aksoy S, Szumlanski CL and Weinshilboum RM. Human liver nicotinamide N-methyltransferase. cDNA cloning, expression, and biochemical characterization. J Biol Chem 1994; 269: 14835–14840.
23. Okita K, Yamakawa T, Matsumura Y, et al. An efficient nonviral method to generate integration-free human-

induced pluripotent stem cells from cord blood and peripheral blood cells. Stem Cells 2013; 31(3): 458–466.

24. Hay DC, Zhao D, Fletcher J, et al. Efficient differentiation of hepatocytes from human embryonic stem cells exhibiting markers recapitulating liver development in vivo. Stem Cells 2008; 26(4): 894–902.

25. Goulart E, De Caires-Junior LC, Telles-Silva KA, et al. Adult and iPS-derived non-parenchymal cells regulate liver organoid development through differential modulation of Wnt and TGF-b. Stem Cell Res Ther 2019; 10(1): 258.

26. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002; 3(7): RESEARCH0034.

27. Valanejad L, Cast A, Wright M, et al. PARP1 activation increases expression of modified tumor suppressors and pathways underlying development of aggressive hepatoblastoma. Commun Biol 2018; 1: 67.

28. Cast A, Valanejad L, Wright M, et al. C/EBPalpha-dependent preneoplastic tumor foci are the origin of hepatocellular carcinoma and aggressive pediatric liver cancer. Hepatology 2018; 67: 1857–1871.

29. Schneider CA, Rasband WS and Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Meth-ods 2012; 9(7): 671–675.

30. Costa TBBC, Lacerda ALT, Mas CD, et al. Insights into the effects of crack abuse on the human metabolome using a NMR approach. J Proteome Res 2019; 18: 341–348.

31. Tian Y, Xu T, Huang J, et al. Tissue metabonomic phenotyping for diagnosis and prognosis of human colorec-tal cancer. Sci Rep 2016; 6: 20790.

32. Meyers RL, Maibach R, Hiyama E, et al. Risk-stratified staging in paediatric hepatoblastoma: a unified analysis from the Children's Hepatic tumors International Colla-boration. Lancet Oncol 2017; 18(1): 122–131.

33. Czauderna P, Haeberle B, Hiyama E, et al. The Children's Hepatic tumors International Collaboration (CHIC): novel global rare tumor database yields new prognostic factors in hepatoblastoma and becomes a research model. Eur J Cancer 2016; 52: 92–101.

34. Bogan KL and Brenner C. Nicotinic acid, nicotinamide, and nicotinamide riboside: a molecular evaluation of NAD + precursor vitamins in human nutrition. Annu Rev Nutr 2008; 28: 115–130.

35. Pissios P. Nicotinamide N-methyltransferase: more than a vitamin B3 clearance enzyme. Trends Endocrinol Metab 2017; 28(5): 340–353.

36. Hong S, Moreno-Navarrete JM, Wei X, et al. Nicotinamide N-methyltransferase regulates hepatic nutrient metabolism through Sirt1 protein stabilization. Nat Med 2015; 21(8): 887–894.

37. Xu J, Capezzone M, Xu X, et al. Activation of nicotinamide N-methyltransferase gene promoter by hepatocyte nuclear factor-1beta in human papillary thyroid cancer cells. Mol Endocrinol 2005; 19(2): 527–539.

38. Shabalin K, Nerinovski K, Yakimov A, et al. NAD metabolome analysis in human cells using 1 H NMR spectro-scopy. Int J Mol Sci 2018; 19(12): 3906.

39. Kim J, Hong SJ, Lim EK, et al. Expression of nicotinamide N-methyltransferase in hepatocellular carcinoma is associated with poor prognosis. J Exp Clin Cancer Res 2009; 28(1): 20.

40. Kratz CP, Franke L, Peters H, et al. Cancer spectrum and frequency among children with Noonan, Costello, and cardiofacio-cutaneous syndromes. Br J Cancer 2015; 112: 1392–1397.

41. You Z, Liu Y and Liu X. Nicotinamide N-methyltransferase enhances the progression of prostate cancer by stabilizing sirtuin 1. Oncol Lett 2018; 15(6): 9195–9201.

42. Ulanovskaya OA, Zuhl AM and Cravatt BF. NNMT promotes epigenetic remodeling in cancer by creating a metabolic methylation sink. Nat Chem Biol 2013; 9(5): 300–306.

43. Li J, You S, Zhang S, et al. Elevated N-methyltransferase expression induced by hepatic stellate cells contributes to the metastasis of hepatocellular carcinoma via regulation of the CD44v3 isoform. Mol Oncol 2019; 13(9): 1993–2009.

44. Shin JH, Park CW, Yoon G, et al. NNMT depletion contributes to liver cancer cell survival by enhancing autop-hagy under nutrient starvation. Oncogenesis 2018; 7: 58.

45. Jones PA. The DNA methylation paradox. Trends Genet 1999; 15: 34–37.

46. Baenke F, Peck B, Miess H, et al. Hooked on fat: the role of lipid synthesis in cancer metabolism and tumour development. Dis Model Mech 2013; 6(6): 1353–1363.

47. Butler L, Perone Y, Dehairs J, et al. Lipids and cancer: emerging roles in pathogenesis, diagnosis and therapeutic intervention. Adv Drug Deliv Rev. Epub ahead of print 14 July 2020. DOI: 10.1016/j.addr.2020.07.013.

48. Warburg O, Wind F and Negelein E. The metabolism of tumors in the body. J Gen Physiol 1927; 8: 519–530.

49. Luo X, Cheng C, Tan Z, et al. Emerging roles of lipid metabolism in cancer metastasis. Mol Cancer 2017; 16: 76.

50. Currie E, Schulze A, Zechner R, et al. Cell metabolism perspective cellular fatty acid metabolism and cancer. Cell Metab 2013; 18: 153–161.

51. Stephenson DJ, Hoeferlin LA and Chalfant CE. Lipidomics in translational research and the clinical significance of lipid-based biomarkers. Transl Res 2017; 189: 13–29.

52. Huang C and Freter C. Lipid metabolism, apoptosis and cancer therapy. Int J Mol Sci 2015; 16: 924–949.

53. Zaytseva YY, Harris JW, Mitov MI, et al. Increased expression of fatty acid synthase provides a survival advantage to colorectal cancer cells via upregulation of cellular respiration. Oncotarget 2015; 6: 18891–18904.

54. Beloribi-Djefaflia S, Vasseur S and Guillaumond F. Lipid metabolic reprogramming in cancer cells. Oncogenesis 2016; 5: e189.

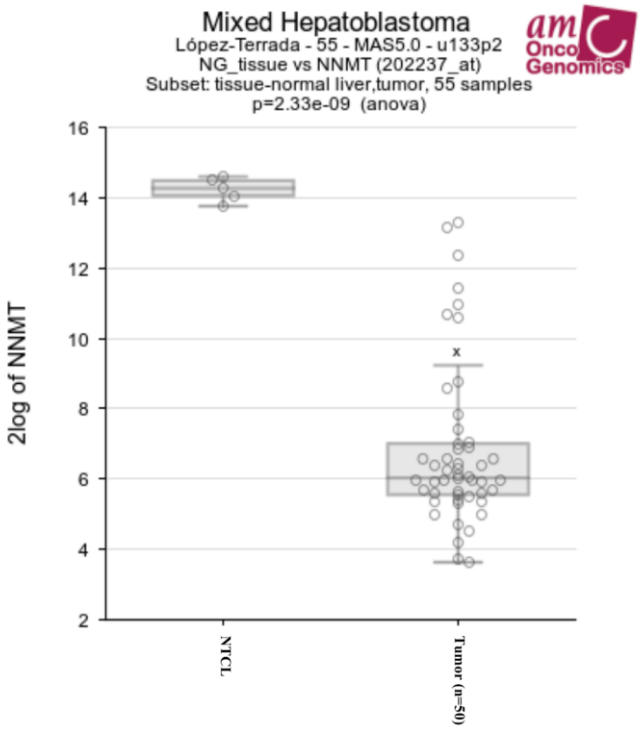
55. Yamamoto H, Schoonjans K and Auwerx J. Sirtuin functions in health and disease. Mol Endocrinol 2007; 21(8): 1745–1755.

56. Rodgers JT, Lerin C, Haas W, et al. Nutrient control of glucose homeostasis through a complex of PGC-1a and SIRT1. Nature 2005; 434: 113–118.

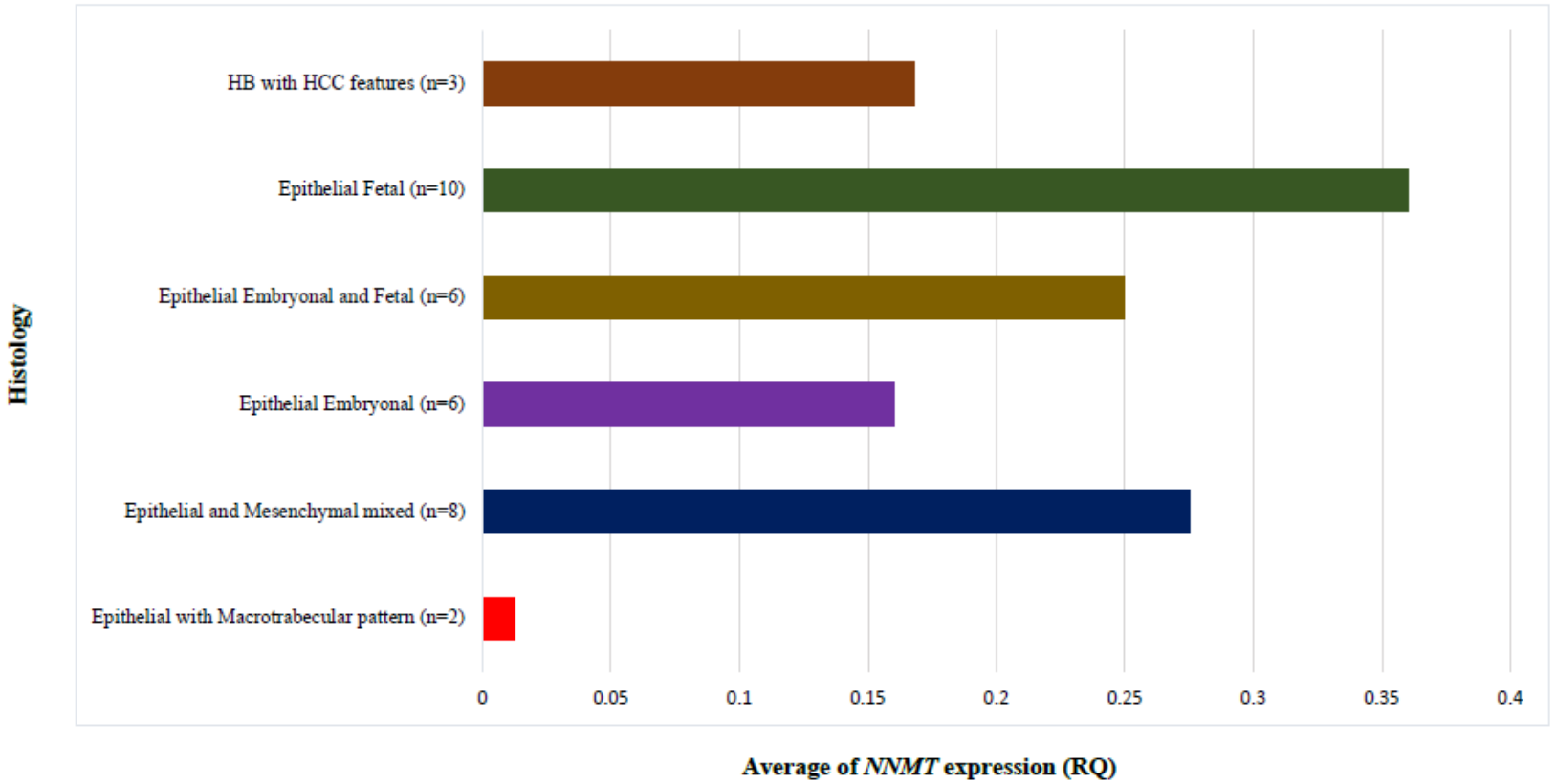
57. Komatsu M, Kanda T, Urai H, et al. NNMT activation can contribute to the development of fatty liver disease by modulating the NAD + metabolism. Sci Rep 2018; 8(1): 8637.

One Way Analysis of variance (ANOVA):

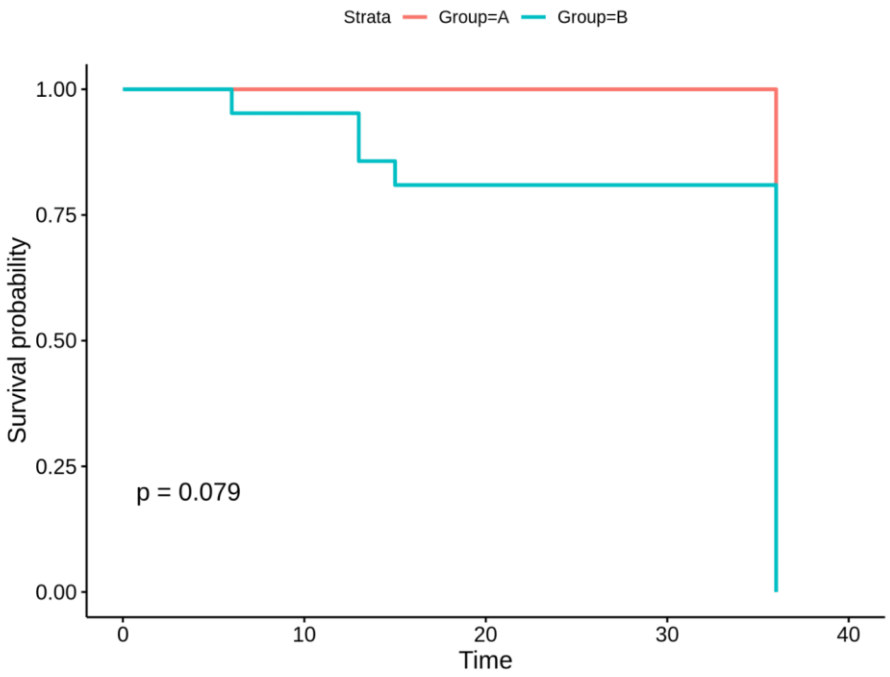
ANOVA	sum_square	df	mean_square	F	p-value
Between	251.178	1	251.178	51.521	2.33e-09
Within	258.389	53	4.875	-	-



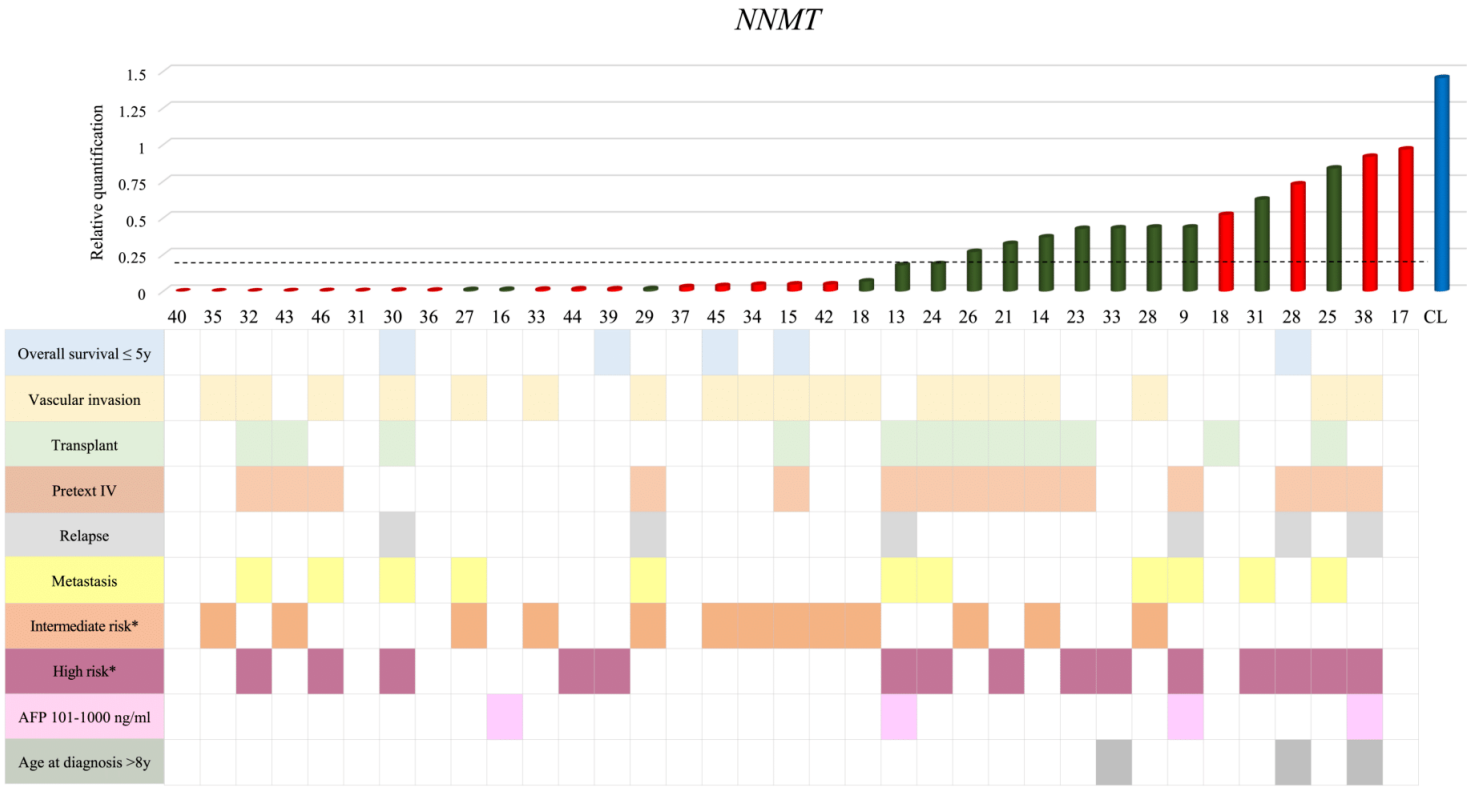
Supplementary Figure 1S. *NNMT* expression in hepatoblastomas and control liver recovered from the literature. Using data from a published study in hepatoblastomas (Sumazin et al. 2017), wich evaluated 50 tumor samples, it is possible to observe that *NNMT* was also more expressed in the NTCL samples, corroborating our findings.



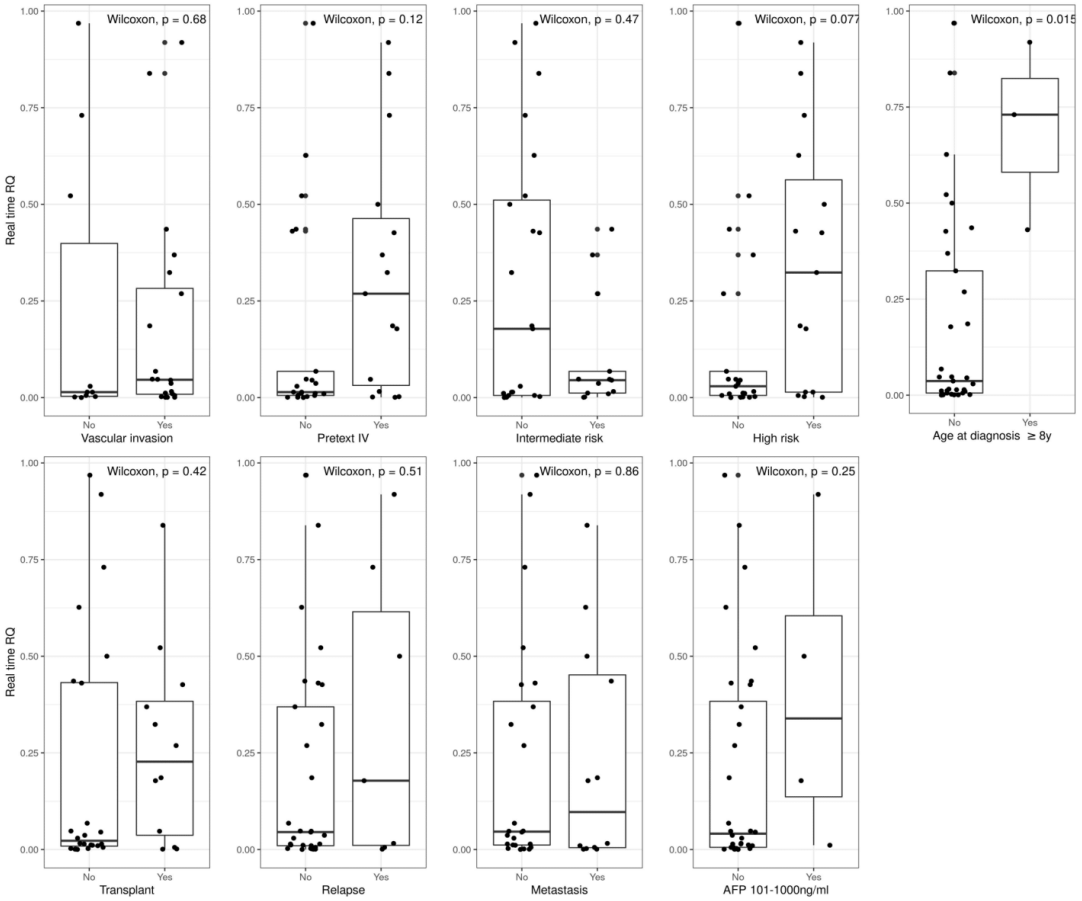
Supplementary Figure 2S: *NNMT* average expression level (by RT-qPCR) of tumors according to the HBL histology. Each horizontal bar shows the average of the relative quantitative (RQ) values of *NNMT* expression for the samples of each histological subtype.



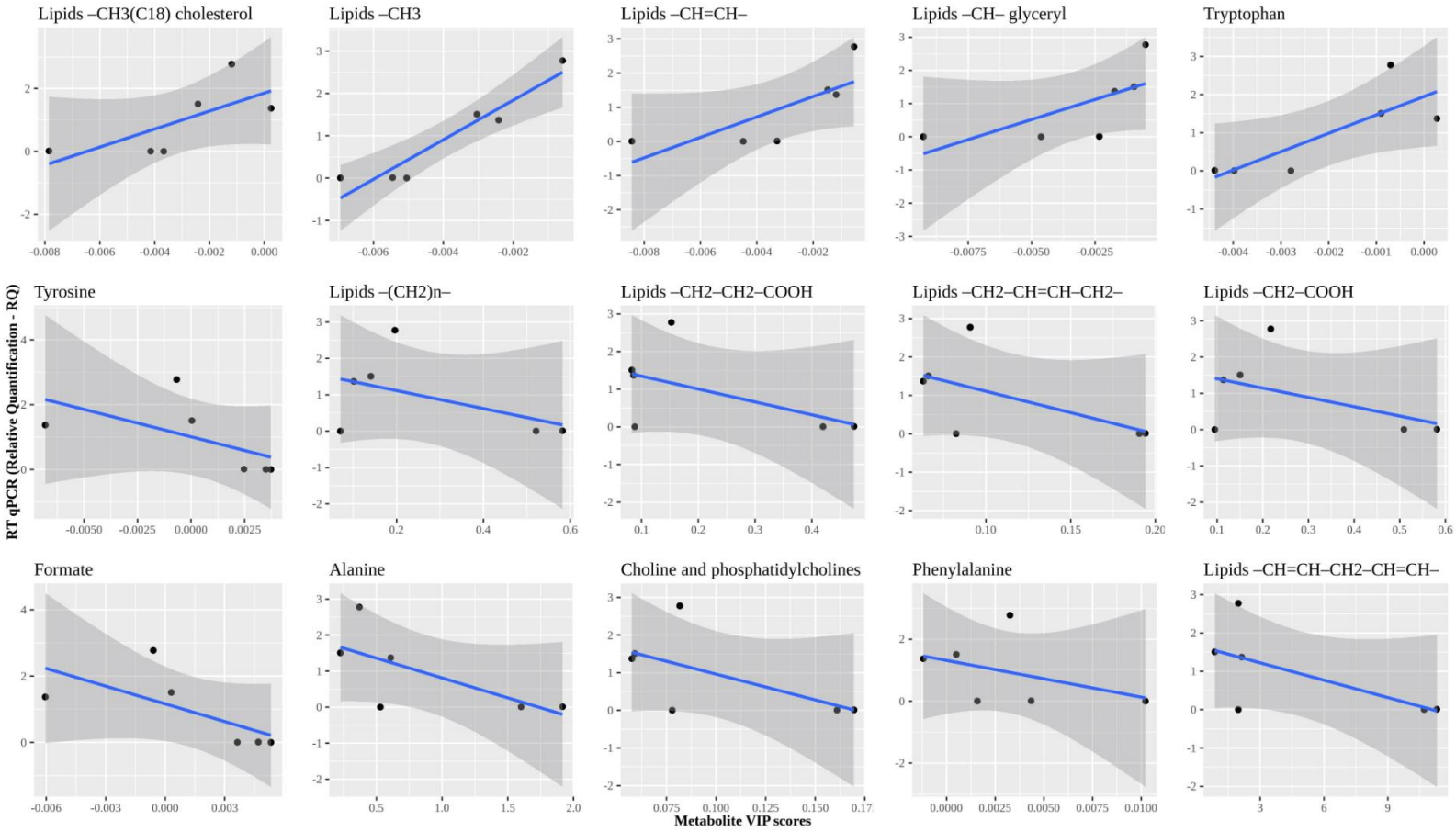
Supplementary Figure 3S: Survival analysis using Kaplan-Meier analysis based on the level of *NNMT* expression as evaluated by RT-qPCR. Survival analysis was performed considering the survival probability in the first 36 months after diagnosis of each patient and the level of expression of *NNMT*; the mean RQ value of the *NNMT* expression of all HBL samples was obtained (Mean RQ=0.15) and then HBLs were classified as Group A: high expression level (above 0.15, orange line; n=15 tumors) or Group B: low expression level (below 0.15, blue line; n=20 tumors). No statistical difference was observed in the survival of patients carrying HBL with high or low *NNMT* expression.



Supplementary Figure 4SA. Level of *NNMT* expression and clinical characteristics of patients with hepatoblastomas. The upper panel shows the *NNMT* expression level of 35 hepatoblastoma (HBL) samples (relative quantification - RQ values); Brazilian HBL samples are represented in red bars, American HBL samples are indicated in green bars, and control samples are blue. The dashed line represents the mean *NNMT* expression value of the group of HBLs. The clinical characteristics of each HBL sample are given below (Intermediate and High risk* according to CHIC stratification).



Supplementary Figure 4SB: Statistical analysis of clinical variables of HBL prognosis and *NNMT* expression levels. It was applied the nonparametric Mann-Whitney-Wilcoxon rank-sum test. *NNMT* expression was compared between pairs of groups according to the clinical variables associated with HBL worse prognosis (risk stratification, age at diagnosis >8 years, AFP dosage level 101-1000 ng/mL, PRETEXT IV, vascular invasion, metastasis, transplantation and relapse). For each variable, the x-axis exhibits the presence or absence of the characteristic in each sample, and the y-axis is the *NNMT* RQ (relative quantification) value obtained by RT-qPCR of samples. Statistical difference was observed only related to the age at diagnosis ($p<0.015$); patients diagnosed with more than 8 years-old presented significantly higher *NNMT* expression.



Supplementary Figure 5S: Pearson correlation analysis applying the Spearman non-parametric test between NNMT RT-qPCR relative quantification (RQ) values and 15 metabolites VIP scores (1HR NMR HR-MAS). The regression lines of the data using all samples (six HBL and three NTCL samples) are represented in blue. The Pearson correlation coefficients (ρ) showed strong positive correlation ($\rho \geq 0.697$) in five out of 15 metabolites when all samples are considered (graphs in first row).

Capítulo VII. Análise de transcriptoma de hepatoblastomas

Poucos grupos exploraram o transcriptoma de hepatoblatomas utilizando RNA-Seq, em geral com foco na classificação/estratificação tumoral [55,57,58]. No entanto, ainda não há consentimento de assinatura molecular para estratificar esse tumor. A análise do transcriptoma tumoral disponibiliza não só informações de genes codificadores de proteínas (mRNA), bem como informa o perfil de expressão de genes não codificadores (ncRNA), os quais têm papel crescentemente reconhecido em câncer. Cerca de 85% do genoma humano é transcrito em ncRNA [130,131], categoria de genes composta por duas classes principais, RNAs pequenos não codificantes (sncRNAs) e RNAs longos não codificantes (lncRNAs) [132]. Em hepatoblastoma, existem exemplos isolados da relevância potencial de ncRNAs. *MEG3/MEG8* e pequenos RNAs nucleolares da família C/D box (*SNORD113* e *SNORD114*) possuem associação com graus de agressividade tumoral e assinatura de pior prognóstico [57,133]. No entanto, a relevância das alterações de ncRNA em hepatoblastoma como reguladores epigenéticos potenciais merece uma investigação mais aprofundada.

No presente projeto a construção das bibliotecas de cDNA foi realizada com o uso do kit *TruSeq®Stranded Total RNA LT-* (*with Ribo-Zero TM Gold*) (Illumina). Ocorre remoção de RNA ribossômico (rRNA) usando oligos biotinilados, específicos para o alvo, combinados com *beads* de remoção de rRNA Ribo-Zero, procedimento em geral eficiente para depletar rRNA citoplasmático e mitocondrial. O sequenciamento realizado foi *paired-end*, ou seja, ambas as extremidades das sequências são obtidas, o que aumenta a acurácia na determinação das sequências em geral. Neste estudo, mesmo considerando a abordagem para depleção ribossomal, ao final grande parte do sequenciamento de algumas amostras resultou nesse tipo de sequência; assim, utilizamos como critério para análises posteriores de expressão gênica a obtenção de no mínimo 5 milhões de *reads* por amostra. Deste modo, ao utilizar este critério, 19 amostras (14 hepatoblastomas e 5 amostras de fígado não tumoral) foram selecionadas para as análises apresentadas a seguir.

Com o objetivo de compreender melhor os processos biológicos alterados, realizamos análises de bioinformática a partir dos dados de RNASeq obtidos. Foi identificado um grupo de 1492 genes com expressão diferencial entre tumores e amostras controle (differentially expressed genes - DEGs), 1.031 deles regulados positivamente e 461 com diminuição de expressão. A lista de DEGs é composta por 920 genes codificadores de proteínas (558 com aumento de expressão e 362 com diminuição de expressão). Além deste grupo de genes

codificadores de proteínas, observamos uma forte desregulação de ncRNAs, principalmente regulados positivamente, possivelmente associados à regulação direta de genes codificadores de proteínas. Oitenta e cinco dos genes com expressão diferencial são expressos exclusivamente em amostras tumorais, sendo aproximadamente 55% deles ncRNAs. Dentre os 50 principais genes, aqui definidos como aqueles com a maior alteração de expressão, 25 genes com aumento de expressão em hepatoblastomas e 25 com diminuição, sendo que 42% correspondem a ncRNAs. Dessa lista top 50, três (*MAGED4, SNO144-16 e RP11-431J24.2*) tinham expressão detectável exclusivamente em tumores; adicionalmente, 16 destes genes com maior diferença de expressão entre tumores e fígado normal já haviam sido identificados anteriormente como alterados neste grupo tumoral (*GCK, HMGA2, INS-IGF2, DKK1, CPA6, TTC36, LIX1, PART1, DKK4, CDCA7, GNG4, THRSP, CNDP1, CYP2B7P, ASPG e HAO2*), embora quase nenhum tenha sido investigado em profundidade.

Considerando a lista total de DEGs, os processos biológicos foram enriquecidos no sentido de regulação negativa do metabolismo, principalmente de amina, nicotinamida e metabolismo de lipídios / ácidos graxos; também há alteração de mecanismos epigenéticos, como metilação do DNA, ncRNAs e silenciamento de genes por miRNA; citotoxicidade e vias de sinalização, incluindo via de sinalização mediada por citocinas. Uma rede de interação miRNA-mRNA foi construída para explorar uma possível regulação de genes codificadores de proteínas por miRNAs. A análise identificou quatro genes (miR-186, miR-214, miR-377 e miR-494) conectados a mais de 10 mRNAs. Os principais processos biológicos relacionados a esta rede miRNA-mRNA foram associados também a metabolismo (10 processos) e reações de oxidação (5 processos) de lipídios e carboidratos. Em conclusão, a análise RNASeq evidencia uma grande alteração de vias metabólicas em hepatoblastomas, incluindo lipídios, aminas e nicotinamidas, o que retoma resultados anteriores de nosso grupo. Também revelou grande modificação na rede de expressão de ncRNAs neste tipo de tumores.

Os resultados são apresentados a seguir em caráter preliminar, posto que análises adicionais visando avaliar interação lncRNAs e mRNAs, assim como correlação da expressão gênica com dados de metilação do DNA serão ainda realizadas. Optamos pela apresentaçaõ em formato de manuscrito em preparação (*Transcriptome dysregulation in hepatoblastomas: the pivotal role of noncoding RNAs and detection of altered lipid metabolism).*

Transcriptome dysregulation in hepatoblastomas: the pivotal role of noncoding RNAs and detection of altered lipid metabolism

Maria Prates Rivas[1], Edson Mario de Andrade[2], Talita Ferreira Marques Aguiar[1], Sara Ferreira Pires[1], Alexandre Defelicibus[3], Bruna Barros[3], Estela Novak[4], Lilian Maria Cristofani[4], Vicente Odone[4], Monica Cypriano[6], Silvia Regina Caminada de Toledo[5], Isabela Werneck da Cunha[6,7], Cecilia Maria Lima da Costa[8], Dirce Maria Carraro[3], Israel Tojal[3], Tiago Antônio de Oliveira Mendes[2], Ana Cristina Victorino Krepischi[1*]

1. Human Genome and Stem-Cell Research Center, Department of Genetics and Evolutionary Biology,
Institute of Biosciences, University of São Paulo, São Paulo, Brazil.
2. Departament of Biochemistry and Molecular Biology, Federal University of Viçosa, Minas Gerais, Brazil
3. International Center for Research, A. C. Camargo Cancer Center, São Paulo, Brazil.
4. Pediatric Cancer Institute (ITACI) at the Pediatric Department, São Paulo University Medical School, São Paulo, Brazil.
5. Department of Pediatric, Adolescent and Child with Cancer Support Group (GRAACC), Federal University of São Paulo, São Paulo, Brazil.
6. Department of Pathology, Rede D'OR-São Luiz, São Paulo, Brazil.
7. Department of Pathology, A. C. Camargo Cancer Center, São Paulo, Brazil.
8. Department of Pediatric Oncology, A. C. Camargo Cancer Center, São Paulo, Brazil.

* Corresponding author:
Ana Cristina Victorino Krepischi
Human Genome and Stem-Cell Research Center, Department of Genetics and Evolutionary Biology - Institute of Biosciences, University of São Paulo, São Paulo, Brazil
Phone: 55 11 3091 7573
e-mail: ana.krepischi@ib.usp.br

Keywords: Hepatoblastoma, RNA-Seq, transcriptome, noncoding RNA, lipid, metabolism

Abstract

Hepatoblastoma is the most common primary liver tumor in children, accounting for ~1% of pediatric cancers. Over the past decade, sequencing studies revealed that hepatoblastomas exhibited the lowest mutational burden among solid human cancers. RNA sequencing has significantly evolved, becoming a paramount approach for transcriptome profiling of tumors. Few groups have explored HBL transcriptomes by RNA sequencing, and related studies have focused mainly on risk stratification; however, still there is no consensual molecular signature to stratify this tumor type.

Aiming to understanding the altered processes and relevant biological pathways in hepatoblastomas, we performed RNA sequencing data of 14 hepatoblastomas, comparing with control liver samples. The analysis detected 1492 differentially expressed genes (DEGs) in HBLs, 1031 of them upregulated and 461 downregulated. The list of DEGs comprised 920 protein coding genes (558 upregulated and 362 downregulated), corresponding to 62% of the number of genes with changes in expression. In addition to differentially expressed protein-coding genes, we observed a strong dysregulation of non-coding RNAs, mostly upregulated, possibly associated with direct regulation of protein-coding genes. 85 genes were exclusively expressed in HBL samples, and approximately 55% of them are ncRNAs. Among the top 50 genes defined as those with the highest expression change, 25 upregulated and 25 downregulated genes, three of them (*MAGED4, SNO144-16, and RP11-431J24.2*) were exclusively expressed in HBLs and 16 were previously reported in HBLs, such as *GCK, HMGA2, INS-IGF2, DKK1, CPA6, TTC36, LIX1, PART1, DKK4, CDCA7, GNG4, THRSP, CNDP1, CYP2B7P, ASPG,* and *HAO2*); of note, 42% of this top 50 DEGs correspond to ncRNAs.

Considering the total list of DEGs, enriched biological processes were negative regulation of oxidation and metabolism inhibition, such as amine, nicotinamide, and lipid metabolism. Conversely, upregulated biological processes in HBLs were linked with the regulation of cell communication and signaling pathways, as well as cell differentiation, and embryonic and developmental processes. The clusters revealed by the PPI network

highlighted one hub protein (F13A1), nineteen bottleneck nodes (ACADS, TH, HSD17B10, ECHS1, GSTZ1, SHMT1, GLS2, RBP2, MTR, NAT8L, ELOVL4, GGPS1, HES1, LIMD1, GCK, DCP2, MSH2, CAMK2B, and CNOT6,) and eleven hub-bottleneck (CDA, CCND1, HDAC2, CCNB1, AGO1, AKT1, IRF7, PSMB8, CD36, GNAL, and GPLD1). These proteins are associated with clusters were also involved with biological processes of negative regulation of metabolism, mostly amine, nicotinamide and lipid/fatty acid metabolism; epigenetic mechanisms, such as DNA methylation, ncRNAs and gene silencing by miRNA; cytotoxicity, and signaling pathways, including cytokine-mediated signaling pathway.

Considering the amount of ncRNA dysregulation we detected in the HBLs transcriptomes, a miRNA-mRNA interaction network was built to explore a possible regulation of protein coding genes in HBLs by miRNAs, and 18 out of 20 upregulated miRNAs in HBLs was used to obtain a network with 30 processed miRNAs and 82 downregulated mRNAs. The miR-186, miR-214, miR-377, and miR-494 are connected with more than 10 mRNAs, suggesting an important role in the regulation of protein coding genes in HBLs. The major biological processes related to this miRNA-mRNA network were associated with metabolism (10 processes) and oxidation reactions (5 processes) of lipids and carbohydrates.

In conclusion, the RNASeq analysis reinforced the relevance of disruption of metabolic pathways in HBLs, including lipids, amines, and nicotinamides, and disclosed a strong ncRNA transcriptome alteration in this type of tumors, highlighting a new epigenetic player in the control of gene expression, along with DNA methylation. Finally, our findings presented four miRNAs as key players in HBLs (miR-186, miR-214, miR-377, and miR494), and two mRNAs (*THRSP* and *GCK*) possibly related to crucial dysregulation of biological processes.

Introduction

Cancer is a complex condition driven by the acquisition of an increasing number of somatic genomic changes at diverse levels of complexity [4]. Hepatoblastomas (HBL), the most common childhood liver cancer, stand out as the solid tumor with the lowest rate of somatic mutations, presenting 1-7 single nucleotide variant per tumor genome [43,49], indicating that HBL tumorigenesis mostly reliant on larger chromosomal gains and losses [44,134–137] and epigenetic changes [57,101,138]. However, currently there is no widely used HB classification, given the clinical and genetic heterogeneity, and the prognostic biomarkers already suggested with clinical correlation presented reproducibility limitations [55,57,58,117,118]. Thus, the search for key biological pathways involved in HBL development and progression is a challenge to the field yet.

Ultimately, all genomic changes in cancer, such as SNV and chromosomal aberrations as well as epigenetic modifications, will result in crucial deviations of the expression of the genome in a particular cell/tissue type, and the result can be better captured by evaluating the pattern of downstream molecules, including RNA, protein, and metabolites [139,140]. RNA sequencing (RNA-seq) is a widespread next-generation sequencing technique, providing an unbiased global picture of the tumoral heterogeneity, being successfully applied in the disclosure of the molecular bases of many oncological diseases [56,141]. Only a few groups have explored the HBL transcriptome using RNA-seq, the first of them proposing a HBL stratification model based on the expression pattern of four genes (hydroxysteroid 17-β dehydrogenase 6, integrin α 6, topoisomerase 2-α, and vimentin) [55]. Recently, transcriptome data were integrated with methylome and genomic analyses, revealing hypomethylation and overexpression of the 14q32 *DLK1/DIO3* locus as a possible marker of tumor aggressiveness [57]. Sekiguchi et al. also integrated multiomic analyses allowing the clusterization of HBLs in groups according to the stage of tumor cell differentiation [58].

Beyond changes in protein-coding genes mainly associated with WNT pathway and hepatocyte differentiation markers, the role of noncoding RNAs (ncRNA) is poorly explored in HBL transcriptomes. About 85% of the human genome is transcribed in ncRNA [130,131], which is a gene category composed of two major classes , the small non-coding RNAs (sncRNAs) and long non-coding RNAs (lncRNAs) [132]. In HBLs, there are isolated and remarkably examples of the potential relevance of ncRNAs. *MEG3/MEG8* (2 maternally expressed non-coding genes), and small nucleolar RNAs of the C/D box family (namely *SNORD113* and S*NORD114*) were suggested for classification of tumor aggressiveness [57], and isolated sncRNAs have been associated with tumor suppression or oncogenic activity in HBL, with prognostic value [133]. However, the relevance of ncRNA changes in HBLs as potential epigenetic regulators deserves further investigation.

Metabolic dysfunction driven by genomic changes are also largely unknown in HBLs. Recently, we demonstrated NNMT downregulation [142], which is an enzyme responsible for the nicotinamide N-methylation process and associated with the regulation of metabolic pathways in hepatocytes [121]. Metabolomics analysis detected a reduction in the lipid content of tumors, possibly associated with the NNMT downregulation and suggesting a connection with worse prognosis [142].

In the present HBL transcriptome study, on the analysis revealed a central role for downregulation of metabolic pathways, mostly associated with lipids, and a global ncRNA upregulation.

Materials and Methods

Tumor specimens and RNA extraction

Fourteen fresh-frozen primary HBL and five control liver samples (non-tumoral matched liver tissues) were obtained surgically between 2016 and 2019 from patients enrolled in three Brazilian cancer institutions: A.C. Camargo Cancer Center, GRAACC, and ITACI (São Paulo, SP, Brazil). Clinical data are summarized in **Table 1**. This study was approved by the Research Ethics Committee of the respective institutions, and samples were collected after signed informed consent from patients's parents or legal guardians.

Total RNA samples were isolated using the RNeasy Mini Kit (Qiagen, Hilden, Germany), according to the manufacturer's protocol. Microfluidics-based electrophoresis (Bioanalyzer, Agilent Technologies, USA) was performed to verify the RNA quality and only samples with RIN (RNA Integrity Number) $\geq$ 7.0 were used.

Table 1. Description of the clinical features of the 14 hepatoblastoma cases.

ID/gender/age at diagnosis	Histology	AFP ng/ml	Risk stratification*	PRETEXT	Chemoterapy Protocol	Metastasis	Relapse	Decesead
HB15, F, 18m	Epithelial Embryonal	5668000	Intermediate risk	4	NA			Yes
HB28, M, 17y	Epithelial and Mesenchymal mixed	NA	High risk	4	SIOPEL4, SIOPEL5 (Recidive)		Yes	Yes
HB30, M, 54m	HB with HCC features	>1000000	High risk	2	SIOPEL4	Lung	Yes	Yes
HB31, M, 30m	Epithelial Fetal	742000	Low risk	3	NA			
HB35, M, 26m	Epithelial Fetal	54800	Intermediate risk	3	SIOPEL3			
HB38, F, 147m	Epithelial Fetal	643,4	High risk	4	SIOPEL3 and AHEP 0731 - COG		Yes	
HB39, M, 84m	Epithelial with Macrotrabecular pattern	300000	High risk	2	SIOPEL2			Yes
HB40, M, 22m	Epithelial Embryonal and Fetal	1842	Low risk	1	SIOPEL3			
HB41, M, 19m	Epithelial Fetal	201733	High risk	3	AHEP 0731 - COG	Lung		
HB42, M, 45m	Epithelial Fetal	1267	Low risk	1	SIOPEL2			
HB46, M, 28m	Epithelial and Mesenchymal mixed	>200000	High risk	4	SIOPEL6	Lung		
HB47, M, 10m	Epithelial Embryonal	409596	High risk	2	SIOPEL6	Lung	Yes	Yes
HB48, M, 7m	Epithelial and Mesenchymal mixed	122131	Low risk	1	SIOPEL6			
HB79, M, 19m	Epithelial Fetal	>50000	High risk	4	SIOPEL4			

F: female; M: male; NA: not available; AFP: Alphafeto protein; PRETEXT: PRE-Treatment EXTent of tumor

Risk stratification* according CHIC (Czauderna et al., 2016 and Meyers et al., 2017)

RNA sequencing and alignment

The total RNA extracted from the 19 liver samples were used to build the cDNA libraries using the TruSeq®Stranded Total RNA LT- kit (with Ribo-Zero TM Gold) (Illumina, USA). Sequencing was performed on the NextSeq 500 platform Mid Output v2 Kit (150 cycles) (Illumina, USA).

The FASTQ files were aligned against the ribosomal reference sequence (NCBI, 12/2017) using the BWA software [143] version 0.7.17-r1188, in MEM mode, with the standard parameters, except for the -t 4 parameters. Reads not aligned to ribosomal sequences went to the alignment step against the reference sequence of the human genome (version GRCh37 - hg19) using the STAR software [144], version 2.6.1a_08-27. All samples presented a minimum of 5 million reads, and the reading count per gene was calculated using the HTSeq software [145] version 0.11.2. The annotation database (GTF file) used was the Ensembl file in version 87 in the same build as the human genome reference (GRCh37). The FPKM metrics was used to normalize the count with a Python 3 script developed by the Bioinformatics Laboratory of the A. C. Camargo Cancer Center.

RNASeq analysis

Principal Component Analysis (PCA)

A PCA analysis was performed using the FPKM values, on a $\log_{10}$ scale, with the function *prcomp* [146]. The ideal number of clusters was determined by the function *fviz_nbclust*, and the PCA representation, with the definition of the clusters, was generated with the function *fviz_cluster* of the *factoextra* package [147]. In addition, the summary of contributions for the first 10 components was obtained using the function f*viz_contrib*, from the *factoextra* package.

Differential Gene Expression Analysis

Differential gene expression analysis of HBLs versus control liver samples was performed using DEseq2 software [148] version 1.18.1, available in the Bioconductor package for R version 3.4.1, following the good practices listed in the Standard Workflow available in the literature [149]. The parameters used in this analysis were: *test_method Wald*, *fit_method parametric*, *padj BH*, and *normalization vst*. For the calculation of saturation and expression, minimum coverage of 1 (one) read was considered. Cutoff values of 0.05 for the adjusted p-value (FDR) and $\log_2$ fold-change < -1 or > 1 were used for selection of a robust list of differentially expressed genes (DEGs).

Volcano plot and Vein diagram

The result of the differential expression analysis comparing HBLs and control liver samples was summarized in a volcano plot. The values of log2 fold-change and FDR generated by DESeq2 were used for plotting the graph in an R environment with the functions plot, point, and textxy [146,150]. A Venn diagram was elaborate to depict the number of shared or exclusive expressed genes in HBLs and control liver samples among the DEGs.

Heatmap

A heatmap was used to analyze the expression profile and grouping of the DEGs between HBL and control liver samples. The gene expression levels between HBL and control liver samples were normalized by FPKM and converted to a z-score scale before using the *Heatmap* function, from the *ComplexHeatmap* package, in an R environment [151], with the parameter km=3. An analysis of functional enrichment was performed to identify the enriched biological processes in the groups of genes in the generated clusters.

Manhattan plot

The Manhattan plot was used to represent the genomic mapping of DEGs and their respective $\log_2$ fold-change and FDR values. This analysis was implemented with command

line routines in bash/awk and in an R environment, with tools from the *karyoploteR* package [152].

Gene annotation enrichment analysis

An analysis of functional enrichment was performed with the set of the up and downregulated DEGs, in the comparison between HBLs and control liver samples. This analysis was carried out with the *enrichGO* function of the *clusterProfiler* package, with the parameters: *de, OrgDb = "org.Hs.eg.db", ont = "all"* [153]. Furthermore, the *dotplot* function was used to plot the main enriched biological processes, cellular components, and molecular functions considering this set of genes.

To identify the tendency of a given biological process to be increased (mainly up-regulated genes) or reduced (mainly down-regulated genes), an analysis with the *GOplot* package (using the *GOCircle* and *GOChord* functions) was performed. In addition, a functional enrichment analysis was performed using the Cytoscape's BiNGO plugin, with the list of DEGs as input (the BiNGO default settings were maintained). The output of BiNGO was formatted, using Python script and bash/awk routine, according to the standard required by GOplot [154]. Thus, the circular diagram (GOCircle) was obtained to represent ten relevant biological processes. Also, a string diagram (GOChord) was obtained to represent the genes in the processes shown in the GOCircle which have the biggest fold-changes. This information was used to generate a summarized network in which were represented only the interactions of proteins encoded by the DEGs, their centralities, clusters, and enriched biological processes.

Protein-protein Interaction (PPI)

A study based on protein-protein interaction networks was conducted with the aim of determining centrality and identifying the functional modules with proteins encoded by DEGs. The network of *H. sapiens* was downloaded from the STRING database and the interactions were filtered (confidence greater than 900) for subsequent analyses. The central

proteins for the system (Hub, Bottleneck, and Hub-Bottleneck) were identified using an analysis for the betweenness centrality and degree, with the functions *betweenness* and *degree* of the igraph package [155]. A subnet for first-degree interactions of proteins encoded by DEGs was used in the cluster analysis in an R environment, applying the *fastgreedy.community* function [155]. The identification and selection of biological enriched functional processes of protein clusters was performed using the BiNGO plugin from the Cytoscape program [156], followed by protein selection associated with relevant biological processes.

miRNA-mRNA interaction network

Based on the miRNAs and mRNAs with differential expression between HBLs and control liver samples, a miRNA-mRNA interaction network was built. Data from the miRTarBase [157] and TargetScan databases were used [158]. This analysis presents the limitation of not being able to identify the processed miRNA (3p or 5p) and, therefore, the network was built considering both forms of mature/processed miRNA. The miRNA-mRNA interactions of these databases were joined with the Merge Network tool of the Cytoscape program, and the interactions between differentially expressed miRNAs and mRNAs were selected using routine commands in bash/awk. The main enriched biological processes were represented in a dot plot, obtained with the *dot plot* function of the *clusterProfiler* package.

Results

The clinical features of the cohort of the 14 patients with HBL are described in Table 1. The mean age at diagnosis was 24 months, excluding one patient who was diagnosed at 17 years (HB28), one at 12 years (HB38) and one at 7 years (HB39). Four patients were born premature (<37 weeks) and low birth weight (<900g). According to CHIC classification, eight patients were diagnosed with high-risk HBL [33,159] and four presented with pulmonary metastasis at diagnosis. All received neoadjuvant chemotherapy protocols followed by tumor resection or transplantation. Four patients died due to HBL and one patient (HB39) died after a second primary tumor (fossa pilocytic astrocytoma).

PCA and cluster analyses showed that two groups were the optimal number to clustering the liver samples (**Supplementary Material S1**). The differential expression analysis detected 1492 DEGs (FDR <0.05 and log2 fold-change <-1 or >1) in HBLs compared to control liver samples. The heatmap analysis based on the DEGs supported the existence of two major clusters of samples, one of them exclusively composed of HBLs, and a second one containing all control liver samples plus one HBL (**Figure 1A**). To delineate the expression profile of DEGs according to their genomic mapping, we used the Manhattan plot (**Figure 1B**), which exhibited an evenly distribution of the DEGs along all chromosomes; however, one cluster of upregulated genes could be highlighted at the cytoband 14q32.31.

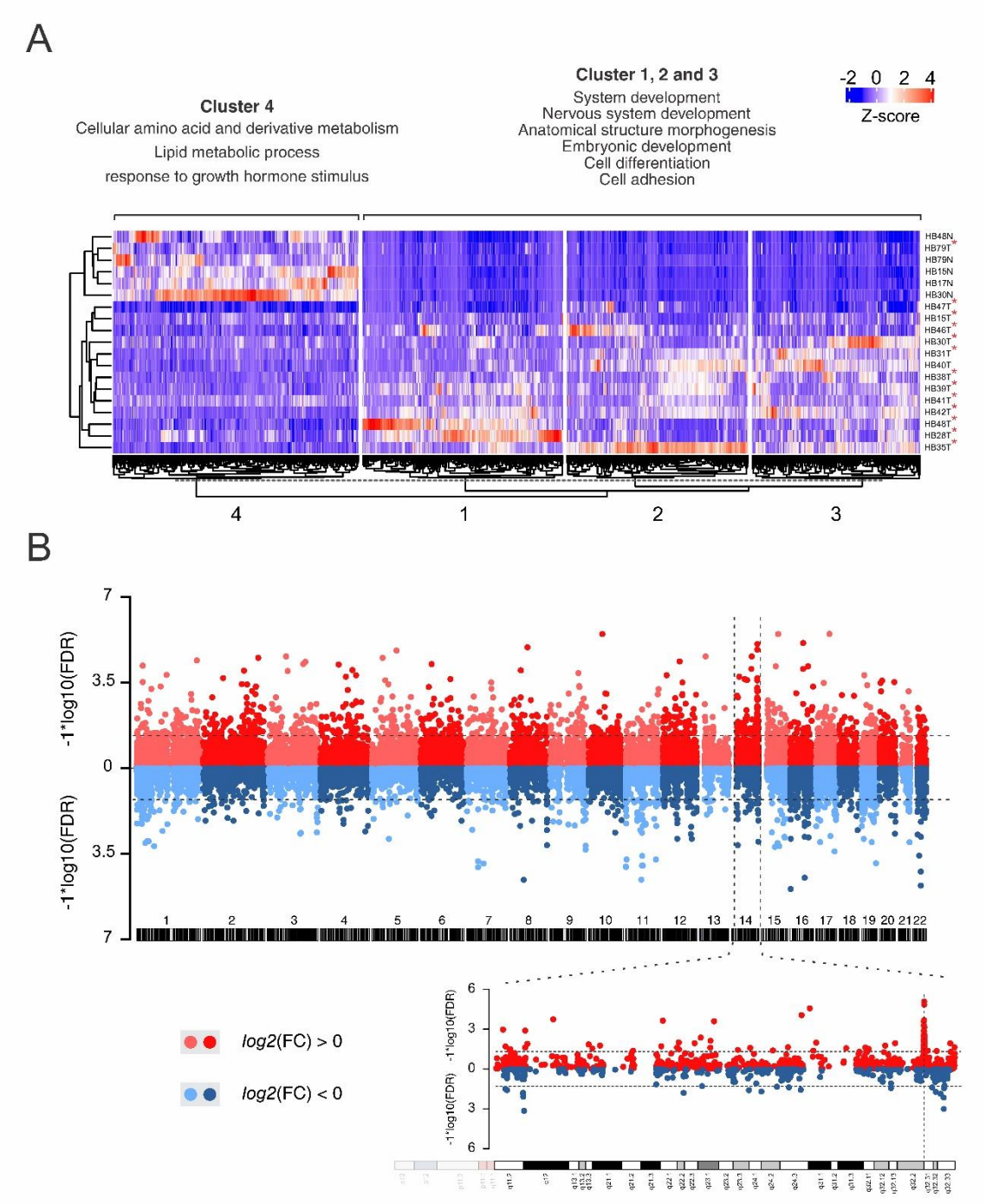


Figure 1. RNASeq gene expression analysis in HBLs. (A) A heatmap representation of gene expression (normalized by Z score, legend of colors in the figure, blue indicates downregulated genes and red indicated upregulated) for HBLs (T; samples with a red asterisk) and control liver samples (N). Four major clusters were separated using the k-means method; the main biological processes enriched in Cluster 4 and Clusters 1, 2, and 3 are indicated at right. (B) Representation of the mapping of the differentially expressed genes along the chromosomes, from 1 to 22; dots in red and blue represent genes with log_2 fold-change above (upregulated) or below (downregulated) 0, respectively. One peaks of upregulated genes can be visualized in chromosomes 14, at 14q32 (detailed below the plot).

Among the detected 1492 DEGs, 1031 were upregulated and 461 were downregulated (**Figure 2A**). The list of DEGs (**Table 2**) comprised 920 protein coding genes (558 upregulated and 362 downregulated), corresponding to 62% of the number of genes with changes in expression. Different classes of non-coding RNAs (ncRNAs) accounts for 22% of the total list of DGEs, mostly upregulated, such as 258 long ncRNAs (lcnRNAs) (199 upregulated and 59 downregulated), and 73 small ncRNAs (68 upregulated sncRNAs, of which 20 were miRNAs, and 5 downregulated).

The Venn Diagram presents the number of DEGs that have shared or exclusive expression in the HBL and control liver groups (**Figure 2B**). The list of the 85 genes exclusively expressed in HBL samples can be found in the **Supplementary Table S1** and approximately 55% of these genes correspond to ncRNAs.

lncRNA	Upregulated	AC105760.3, RP11-16K12.1, RBMS3-AS2, RP11-401P9.6, AC007277.3, AC008277.1, AC009948.5, AC093818.1, AC007362.3, ZNF32-AS1, RP1-167F1.2, AC015923.1, AF178030.2, AC105760.2, RP11-492E3.2, SNAP25-AS1, RP3-508I15.9, RP13-476E20.1, RP3-399L15.3, RP5-981O7.2, RP11-175O19.4, AC005592.2, DLX6-AS1, AC078883.3, RNASEH2B-AS1, AC098617.1, RP5-1073O3.7, TSPEAR-AS1, RBMS3-AS3, INTS6-AS1, NIFK-AS1, RP4-668G5.1, RP11-431J24.2, RP4-584D14.6, PVRL3-AS1, RP11-115C21.2, RP13-578N3.3, RP11-296I10.3, STX18-AS1, RP3-510L9.1, CTB-174D11.1, MEF2C-AS1, RP11-798M19.3, RP11-436H11.3, MAGI2-AS1, RP11-724M22.1, HOXA-AS2, RP4-791M13.3, CACNA1C-AS4, RP11-434H14.1, RP11-611E13.2, RP4-755D9.1, RP11-16B13.1, RP11-242P2.1, RP11-401P9.1, RP11-1299A16.3, RNF157-AS1, PLCE1-AS1, KCNQ1OT1, CTC-203F4.2, RP11-380L11.4, TMCC1-AS1, RP11-435O5.4, RP11-396C23.4, RP11-422P24.10, PART1, LINC00670, RP11-474D1.3, RASAL2-AS1, RP11-361F15.2, AL163953.2, AL132709.5, AC026150.5, RP1-69D17.4, AC006037.2, ZNF883, LINC00341, RP11-95P13.1, AL132709.8, RP11-187C18.2, LINC00327, RP4-655J12.4, RP11-154D17.1, USP27X-AS1, RP11-423O2.5, RP11-86H7.7, RP11-89N17.4, RP11-32F11.2, BCYRN1, RP11-1008C21.2, LINC00312, ZBTB20-AS2, RN7SL832P, RP5-839B4.8, RAB30-AS1, CTB-73N10.1, CTC-428G20.3, RP11-184M15.1, LINC00616, CTD-2015A6.2, LINC01091, RP11-77P16.4, LINC01021, AC141928.1, RP11-427M20.1, RP11-150O12.1, RP11-242J7.1, RP11-396O20.1, LINC01059, RP11-890B15.3, PWAR6, RP11-350F4.2, RP1-97G4.1, RP11-114F10.2, RP11-18J9.3, MEG8, CTD-2341M24.1, CTD-2292M16.8, LINC00639, LINC00924, RP11-323I15.5, CASC7, RP11-327J17.1, RP11-49C24.1, RP11-490M8.1, RP11-42I10.1, RP11-4F5.2, CASC14, CTD-2196E14.9, RP11-252A24.7, RP13-192B19.2, RP11-566E18.3, RP11-142G1.2, RP11-111K18.2, GS1-279B7.1, SNORD116-20, RP11-93I21.3, RP1-80N2.3, CTA-14H9.5, CTC-462L7.1, RP11-401P9.4, RP11-589P10.5, RP11-599B13.3, RP11-720L2.4, RP11-674P19.2, RP11-466P24.7, RP11-242D8.1, CTD-2528L19.6, RP11-397A16.2, LINC00906, RP11-147L13.8, CTC-471F3.6, RP11-343N15.5, RP11-218C14.8, RP11-415J8.7, AC006129.4, RP5-1024G6.8, RP11-546J1.1, RP11-154H23.3, CTC-523E23.11, RP11-1109F11.3, RP11-909M7.3, RP5-1112D6.7, RP11-242F4.2, RP11-7F18.2, AC005618.6, CASC15, RP11-22L13.1, XXbac-BPGBPG55C20.2, RP11-344P13.6, RP11-399K21.14, RP1-74M1.3, RP11-523G9.3, RP11-314B1.2, RP11-672L10.6, ALMS1-IT1, RP11-231L11.3, RP11-150O12.6, RP11-705C15.3, RP11-395N3.1, RP11-815I9.4, RP11-715F3.2, CTC-429P9.5, CTD-2132N18.4, SNORD109A, SNORD64, RP11-701H24.7, RP3-368A4.5, RP3-368A4.6, XXbac-B444P24.13, AC074286.1, RP11-156E6.1, RP11-4O1.2, RP11-119F7.5, RP11-57H14.4, RP1-228H13.5, RP11-244O19.1, CTD-3064M3.3, RP11-340F14.5
	Downregulated	FAM99B, AF124730.4, CTA-85E5.10, LINC00659, RP11-157P1.5, AC016999.2, AC113189.5, LINC01137, AC061961.2, RP5-1103B4.3, ST3GAL6-AS1, PRKAG2-AS1, USP30-AS1, RP11-396F22.1, RP11-1069G10.1, RP11-394B2.6, RP11-303E16.3, RP11-669E14.4, RP11-936I5.1, RAB11B-AS1, LINC00311, LINC00313,LINC00319, SPATA41, LRCOL1, RP4-710M16.2, LINC00365, RP11-431K24.3, RP5-1119A7.14, RP5-888M10.2, AP001043.1, RP11-135A1.3, RP11-172E9.2, RP11-88H9.2, RP1-27K12.2, AC016768.1, AP001044.2, RP11-503C24.4, RP11-466L17.1, RP4-763G1.2, AC079922.3, RP11-279F6.1, RP11-273G15.2, RP11-328K4.1, LINC01093, RP11-434D9.1, LINC00605, RP11-320M2.1, AL161668.5, LINC00543, RP11-883G14.1, RP11-285E9.6, RP11-794M8.2, RP11-84C13.1, RP11-1002K11.1, RP3-395M20.12, RP11-47A8.5, RP11-467I20.6, RP11-110H1.4
sncRNA	Upregulated	MIR494, MIR370, MIR377, MIR381, MIR411, MIR544A, MIR655, MIR186, MIR299, MIR329-2, MIR654, MIR214, MIR323B, MIR770, MIR758, AL023913.1, MIR300, MIR889, MIR541, AL117190.1, MIR4712, AL132709.3, RN7SKP255, RN7SKP124, RN7SKP80, SNORD114-7, SNORD114-5, SNORD114-16, SNORD114-19, SNORD113-6, SNORD114-10, SNORD113-8, SNORD114-23, SNORD114-28, SNORD114-25, SNORD114-2, SNORD114-4, SNORD113, SNORD114-13, SNORD114-17, SNORD113-4, SNORD113-3, SNORD113, SNORD114-24, SNORD114-20, SNORD114-18, SNORD113-1, SNORD114-12, SNORD116-14, SNORD116-18, SNORD116-19, SNORD113-2, SNORD113, snoU13, SNORD112, SNORD116-26, SNORD116-27, SNORD112, SNORD112, SNORD112, SNORD112, SNORD114-21, SNORD113-5, RNU5D-1, RNU6-960P, RNU6-652P, RNU7-84P, RNU2-10P
	Downregulated	AL592188.7, RN7SL25P, RN7SL406P, SNORD88A, RNU6-723P
protein coding	Upregulated	TMEM98, DLX6, ETV1, CDKL5, ADAM22, EXTL3, SLC7A14, CLDN11, DPEP1, SLC38A5, ATP1A2, NRXN3, KITLG, ADAMTS6, DKK3, LAMA3, RASGRF1, TARBP1, BCAT1, CDH3, CDON, CTNNA2, MPPED2, ATP1B3, IKBKAP, ATP2B1, MGAT4A, TRIB2, RHOBTB1, ADD2, PAX2, NAALAD2, FGF20, CLUL1, STX7, RFX3, RBL1, AFP, ARG2, C1QTNF3, DLG3, SERTAD4, TRPM3 OXCT1, AGBL5, MECOM, SCAMP1, AQP6, NOX4, MTMR2, MT3, DOCK3, CPXM1, CHGB, LYZ, NRCAM, ITGA6, CD200, CMA1, AGO1, CBX5, SUCO, MSH2, PCSK5, ADRBK2, SEZ6L, TTLL1, ISM2, SIX4, NDRG3, PFDN4, NKAIN4, FERMT1, MID1, PAGE4, CAB39L, FGF9, HSDL1, SLC6A2, IQCH, MYEF2, ZDHHC2, DKK4, BNIP3L, CDK6, HOXA3, FKTN, TMEM245, TGFBR1, LHX6, PIP5K1B, RAB11FIP2, BMPR1A, DKK1, KRT23, SLC6A4, MMD, RAB34, USP46, ODAM, NEIL3, TBC1D19, CCDC34, KIAA1377, VWF, ASIC1, GOLT1B, ST8SIA1, COL9A1, MDFI, PTK7, CLIC5, MEP1A, TBX18, PCDHB2, KIF20A, PCDHB3, THBS4, CNOT6, RASGRF2, PDE4D, STC2, RBP2, SLC25A36, FGF12, CSPG5, NEK11, PLCH1, ADAM23, REG1A, EPHA4, PARD3B, DNAJC6, MTR, MFAP2, CCDC181, MROH9, TNFSF4, LPPR4, ELOVL4, CCND2, PTBP3, PROX2, EPCAM, MSX2, CCDC170, PCDHB8, PCDHB10, PCDHB14, KCNJ5, TEX11, TAS2R10, POPDC2, KIF18A, GJB6, KIAA1549, NECAB1, COL10A1, RAB9B, EDN3, LYPD3, F13A1, COL21A1, RUNX2, SSUH2, RAP2A, BMP4, GRPR, ARMCX1, TNFRSF19, CASD1, RASL11B, GAD1, FOXJ1, SGOL1, GDPD2, PLVAP, SULT4A1, SLC7A10, SH3BGRL, TOP2A, LRRC9, ABHD12B, ATP8A2, GPALPP1, BEX1, EPHB2, DUSP26, UNC79, ADAMDEC1, CCNB1, LRP4, MST4, RNFT2, CD36, KIAA1009, GDF11, SERPINE2, CKAP2, LECT1, ZFP37, CTSV, DSCC1, NOV, LRRC1, GCM1, RAB30, TRPC6, FXYD2, UACA, STRA6, CYP19A1, KIF20B, MSTN, CDK15, MDH1B, SENP7, BBS7, PDE5A, PRDM5, FRAS1, LEF1, PAPSS1, ERP27, ALG10, PIANP, COL2A1, LGR5, NEDD1, ZNF740, C14orf37, UBE2Q2, FANCI, TICRR, MFGE8, NKD1, CDH13, RASL10B, RHBDL3, GNAL, RNF157, FKBP10, ZFP14, HUNK, PRDM16, FHAD1, SLC44A3, GJA5, CRABP2, XPR1, VASH2, ATP8B2, KCNN3, NPHP1, GALNT13, SCN1A, TMEFF2, CDCA7, DLX1, UNC80, ZNF660, CORIN, SSBP2, LIX1, MMS22L, SDK1, ZNF92, CXorf36, ZNF711, GPR119, GABRQ, SNTG1, TACC1, ATP6V0D2, RSPO2, SHC3, MS4A2, HMGA2, SLC7A11, MIPOL1, ANKRD50, ANO4, TEX9, RNF144A, ZNF385D, BEND6, PTPN14, PFKM, GRIA4, SPARCL1, GJA1, GGPS1, CETN3, CCDC148, RBMS1, MSI2, GDPD1, THY1, ANGPT1, CCSAP, MAGED4, APCDD1, PIEZO2, HSPA13, GRIA1, ALS2CR11, ACPL2, AFF2, WIF1, MAP3K7CL, SFR1, CACNA2D3, DGKI, SHROOM4, TMSB15B, PPP1R9A, DYNC1I1, SV2A, LRRC36, CD300LG, BEND5, PRKAA2, ZSWIM5, RAVER2, VANGL2, OLFML2B, SPATA17, FAM84A, DUSP19, MSX1, GABRB1, IHH, TRAT1, SERPINI1, CLASP2, EIF5A2, CPA3, INTU, TMEM144, ITGA2, EDIL3, HTR4, GRPEL2, GPX8, FOXQ1, STXBP5, ZNF704, C8orf48, CTHRC1, CPA6, TMEM246, MELK, ARMC3, PHYHIPL, TMEM52B, KIAA1462, VWA2, SALL2, HIF1AN, TMEM130, PRTG, C18orf54, MAP1A, ZNF610, CASKIN1, NAALADL1, SLC22A11, PBK, GNG4, NT5DC2, BMI1, ADAM9, AXIN2, UGT3A2, NPNT, TCTN2, TSPAN5, STXBP6, B3GALNT1, IL1RAPL1, SLC38A11, LINGO1, ROBO1, AHSP, SP7, CYTL1, PAQR8, PDGFD, HS6ST2, PYGO1, ZNF439, ZBTB26, MROH2B, ZIK1, RHNO1, PRND, REG3A, GAP43, FRMD3, MYOZ2, ZNF738, DCP2, NBEA, GXYLT2, CCDC96, DEC1, DMRT2, ZNF449, OLR1, THAP2, NUDT4, PIFO, ZCCHC12, IGDCC3, TMEM167A, ISX, MCTP1, ETV4, UNC119B, DLEU1, ZNF404, FOXL1, CDK5R1, ZBTB34, RIMKLA, ZBTB38, IRX3, TBL1XR1, ZNF713, FAM101A, SLITRK4, HTR1D, SHISA2, MAP3K15, ZNF572, BBS12, FANCB, EDDM3A, ANKRD62, PLAG1, RFX7, RTKN2, PNMAL1, HIST2H4B, FIGN, CACNB4, NXPH3, HS3ST4, IGIP, GPIHBP1, GJC1, PCP4, HIST2H2AA3, KREMEN1, PNMA3, FAM3B, C1orf64, HIST2H4A, KCNQ3, ADRA2C, PRKD1, MAP7D2, ZBTB7C, ZFP90, RFX6, WBP5, PARPBP, C1orf227, NAT8L, DLEU7, ZBTB6, EDARADD, ZDHHC17, NAP1L1, TSPYL4, MAGED4B, ZNF286A, GABRD, LIN28B, NHS, CERKL, FSIP2, SHISA6, ZNF292, SP6, COLCA1, SRGAP3, FAM217B, ZNF107, ZNF781, MME, HDAC2, KRT39, KPNA5, SLC39A10, ZMYM1, PGAP1, GYPE, PCDHB11, KLHL14, HIST1H2BF, SLC22A12, DNM3, ZNF248, AC135178.1, ZNF607, ZNF334, PEG3, ZNF280B, ATL1, ZNF28, EPS8L3, FAM169A, PAX9, GRM3, DZIP3, MAGEE1, ARMCX6, AL365202.1, HIST2H3C, HIST2H2AA4, RBM20, IFIT1B, LAYN, C4orf46, ATP6AP1L, SERF1B, C21orf49, DOK6, ATP10A, TRIM71, DIO2, KLHL23, UGT2B11, UBD, COLCA2, EML6, KCNU1, PPT2, SLC26A6, SAPCD1, C21orf37, FAM27E3, AC109135.1, TDGF1, ZNF512, HS3ST5, GYPB, PABPC4L, EID1, AC005477.1, RP11-80A15.1, CTD-2054N24.2, RP11-89K11.1, FRRS1L, TMEM178B, AC117395.1, AC007382.1
	Downregulated	C19orf60, CCL3, AASS, IYD, ETV7, SYT7, IGF1, TYMP, C6, PNKP, PQLC2, RTN4R, PARP3, CAMK2B, DGAT2, ISOC2, EPN1, NTN1, PHKA1, CECR5, RNF126, SREBF1, HSD17B10, SELO, MGLL, SCARF1, RAP1GAP, CXCL2, SLC27A5, CD82, HSD17B14, CETP, GNAO1, SIGLEC1, ANKRD24, HSD3B7, ATP5D, IGFALS, RASSF7, SEC14L2, SEC14L3, PRODH, CYP2D6, SUN2, PLEK2, GSTZ1, SLC10A1, TELO2, PCK2, DCAF11, NTSR1, LPIN2, SLC25A15, MGRN1, ELMO3, FAM173A, METRN, FAH, CA2, DECR1, PYCRL, NFKBIB, PEX11G, TJP3, BBC3, TMEM205, GCDH, MOGAT3, GCK, APBA1, SHB, CPEB3, FBXL15, ENO3, HGFAC, FAM149A, B3GAT1, SIAE, ATG2A, CCND1, HPX, SLC22A18, TCIRG1, TRPV4, GPLD1, GHR, NNT, HES1, FAHD2A, STEAP3, PCSK4, NDUFS7, MLPH, HAO2, TSPAN1, FAAH, MUC5B, GDA, ZC2HC1C, SOCS2, PEX16, PQLC1, ACADS, C12orf44, SARDH, DBH, IL13RA2, PEPD, NQO2, GNMT, MT1G, MT2A, HECTD3, ECHS1, TPST2, IVD, SHBG, NRL, INS-IGF2, APOE, PPAN, ANGPTL6, C19orf66, GCH1, PEMT, MACROD1, SEC14L4, MRO, CMPK2, IL15RA, GGACT, GLS2, PPP1R1A, MIP, EGLN1, FPGS, SLC37A4, FDX1, BCL2L10, SLC3A1, AGPAT9, TAPBPL, INHBE, RITA1, SDSL, RDH16, FITM1, CYP11A1, FURIN, GFOD2, RNF165, DPP9, IFITM3, AKT1, ZNF787, CTU1, FCN3, CREG1, NAT8, LIMD1, SRD5A1, CRHBP, NAPRT1, HMCN2, IDI2, HABP2, SERPING1, NUDT22, CLEC1A, CNDP1, KCTD14, THRSP, G6PC2, ZNF837, TCTEX1D1, UBALD1, LURAP1L, ANKH, NAT2, SLC28A1, ANKRD9, UBE2L6, COX6A2, NRG1, CPT2, DHRS1, SLC38A10, XDH, SPATA2L, CDA, PINK1, MPZ, FBXW5, PTMS, C1R, UROC1, GNE, FNDC5, CBS, CRYAA, ADAMTS13, LY6E, PGLYRP2, ALDH16A1, SIGLEC11, SLC25A45, AKR7A3, ALPL, OXER1, HAAO, FCAMR, TDRD10, IL17RC, SLC51A, NDST3, HPGD, TMUB1, FBP1, CLDN3, SMCO4, ABTB2, ASPG, PLEKHF1, MOGAT2, C16orf46, C16orf45, LDHD, MESP1, TMEM92, SLC27A4, ENDOG, VKORC1, GPT, RILP, RAB26, MFSD2A, HFE2, ATF5, AGPAT2, MT1E, DCXR, GUSB, GPRC5C, CD14, TRABD, NRTN, TRIM8, KCND3, TCEA2, CYP4F22, SUCLG2, TTC36, TMEM134, RAB43, MYD88, ESRRA, RHOD, SYT12, SPTBN2, ZDHHC24, ZG16, SLC22A1, PPM1E, APOF, POLD4, MSRA, BAIAP2, DMRTA1, PLA2G16, HORMAD2, PXMP2, C8G, SHMT1, CASKIN2, SLC25A22, JUN, MAMDC4, ODF3B, STAP2, GRINA, OPLAH, CITED4, TH, TIGD2, CLRN3, ACBD4, GPR88, SLC35C1, IDH2, TMEM259, ENPP7, FBXL6, C1S, CLEC4G, TSKU, SLC25A18, DGCR6, NPBWR1, NPW, OAF, CIB1, ADSSL1, SIGIRR, NRBP2, RAB11B, KPNA7, C17orf70, IRF7, C11orf35, PRR5, VSIG10L, KANK3, CTD-2600O9.1, CHP1, IDO2, NOXA1, SH2D5, C9orf163, SULT1A1, PLXNB2, ARRDC1, SULT1A2, IL27, OGDHL, GSTK1, SLC2A10, BCO2, CFD, NRAP, FAM109A, PIM3, INF2, PSMB8, ASPDH, CTD-2210P24.4, MT1H, DPF3, COL6A6, LTC4S, LCAT, GBP7, C8orf82, CEACAM16, ANG, ATAD3C, ADRA2B, GLTPD1, C2CD4D, C19orf24, TMEM238, AC104667.3, GET4, PDXP, RP11-181C3.1, RP11-422N16.3, CHCHD10, SLC22A18AS, PDF, RP11-162A12.2, FXYD1, AC093677.1, RP11-986E7.7

Table 2: Differently expressed protein coding, long non-coding RNA, and small non-coding RNA genes. The cutoff values used were 0.05 for the adjusted p-value (FDR) and fold-change 2.

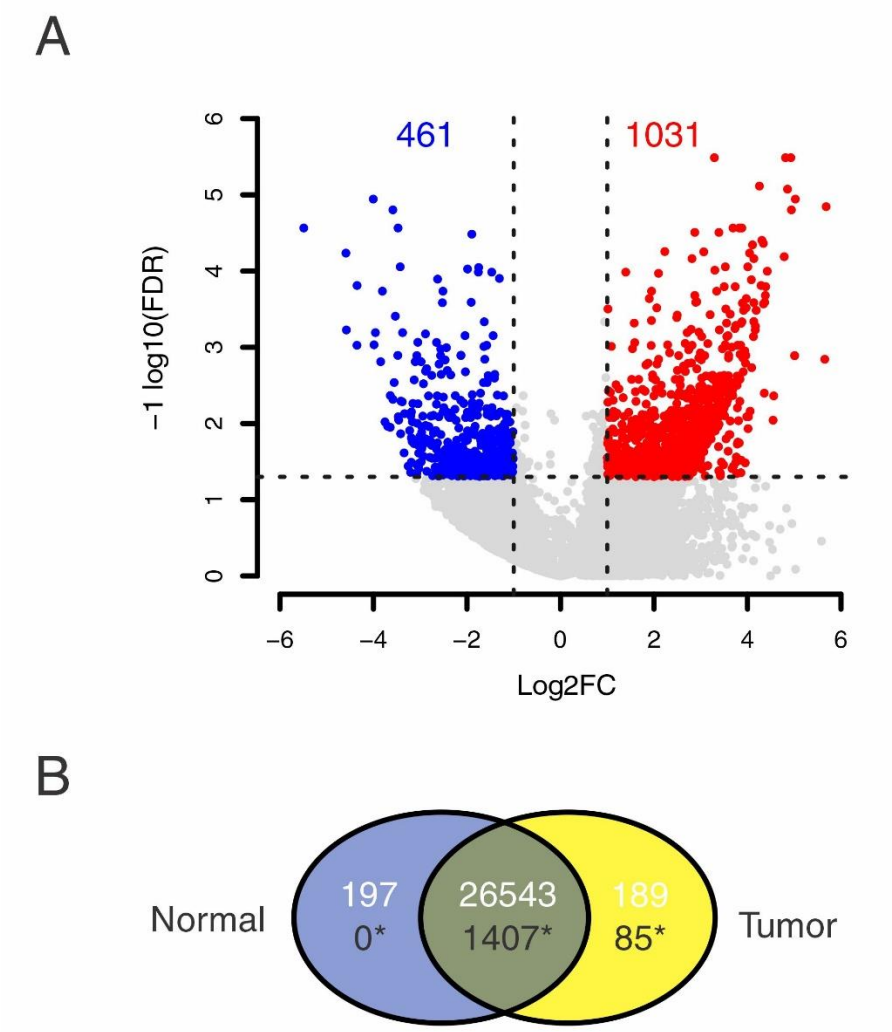


Figure 2. Differentially expressed genes in HBL. (A) A volcano plot showing all expressed genes in HBL samples. Genes differentially expressed (DEG) in HBLs compared to control livers with statistical significance and $log_2FC < -1$ or > 1, are represented as blue (down-regulated genes) or red (up-regulated genes) dots; the gray-colored dots represent genes with no statistical differences. (B) A Venn diagram showing the groups of shared and exclusive expressed genes of HBLs and controls. In black with "*", the number of genes presenting statistical difference (FDR <0.05 and $log_2FC > -1$ or <1) in the transcriptome analysis; numbers in white correspond to genes without statistical difference.

Table 3 exhibit the top 50 genes defined as those with the highest expression change, 25 upregulated and 25 downregulated genes in HBLs compared to control livers, and associated information. Considering the protein coding genes, 16 were directly associated with HBL according to literature: *GCK* [160], *HMGA2* [161], *INS-IGF2* [162], and *DKK1* [163], in addition to a large group recently described in a bioinformatics (*CPA6, TTC36, LIX1, PART1, DKK4, CDCA7, GNG4, THRSP, CNDP1, CYP2B7P, ASPG,* and *HAO2*) [164]. Of note, 42% of this top 50 DEGs correspond to ncRNAs. Three genes among the top 50 list (*MAGED4*, *SNO144-16*, and *RP11-431J24.2*) were exclusively expressed in HBLs, and none of them have been already reported.

Considering the large portion of deregulated ncRNAs, we built a miRNA-mRNA interaction network to explore a possible regulation of protein coding genes in HBLs by miRNAs. In this analysis, the network links the functions of protein coding mRNA with miRNAs. Using 18 out of 20 upregulated miRNAs in HBLs was possible to obtain a network with 30 processed miRNAs (blue nodes) and 82 downregulated mRNAs (red nodes) were involved (**Figure 3A**). Among the miRNAs, miR-186, miR-214, miR-377, and miR-494 are connected with more than 10 mRNAs, suggesting an important role of these miRNAs in the regulation of protein coding genes in tumors. The top 20 biological processes with the highest level of significance (padj) associated with the network's mRNAs can be found in **Figure 3B;** the major biological processes related to this miRNA-mRNA network are associated with metabolism (10 processes) and oxidation reactions (5 processes) of lipids and carbohydrates.

Table 3. Top 50 differentially expressed genes in HBLs (higher fold-changes).

	Gene	Gene summary	log2 fold-change	padj	Gene Info
upregulated	*SNORD114-21*	small nuclear RNA associated with methylation	5,682091848	1,43E-05	Imprinted snoRNA 14q(II-21; cluster II) expressed from the maternal allele; higher expression levels in the brain and the endometrium (https://doi.org/10.1093/hmg/11.13.1527) | GO:0006396: RNA processing
	RP11-492E3.2	long non coding RNA	5,652763776	0,001436457	Antisense RNA; ENSG00000227619 (Ensembl)
	CPA6	encodes a member of the peptidase M14 family of metallocarboxypeptidases	5,023753631	1,14E-05	Metallocarboxypeptidase (GeneCards) | GO:0006508: Proteolysis | Mutations or promoter methylation associated with seizures and epilepsy (https://doi.org/10.1016/j.eplepsyres.2013.10.007; https://doi.org/10.1371/journal.pone.0123180) | Increased expression in hepatocellular carcinoma; promotes proliferation and metastasis via activation of the AKT-mediated signaling pathway (https://doi.org/10.1007/s11596-019-2098-z) | Upregulated in the early stages of oral squamous cell carcinoma (https://doi.org/10.1016/j.oraloncology.2007.10.011)
	SNORD113-4	small nuclear RNA associated with methylation	5,009450797	0,001287522	Imprinted snoRNA 14q(I-4; cluster I) expressed from the maternal allele; higher expression levels in the brain and the endometrium (https://doi.org/10.1093/hmg/11.13.1527) | GO:0006396: RNA processing
	AC007277.3	long non coding RNA	5,006882432	0,001287522	Antisense RNA; ENSG00000213981 (Ensembl) | Antisense To MYO3B (GeneCards)
	LIX1	limb and CNS expressed 1	4,939276878	1,57E-05	GO:0097352: Autophagosome maturation | Upregulated in gastrointestinal stromal tumours (https://doi.org/10.1111/jcmm.15569) | Regulates the pro-proliferative factor YAP1 (https://doi.org/10.1186/s12915-016-0257-2)
	IGDCC3	immunoglobulin superfamily DCC subclass member 3	4,927716587	3,26E-06	GO:0050885: Neuromuscular process controlling balance | Enhancer-gene pair 1944-IGDCC3 methylated in epigenetically-affected colorectal cancer (https://doi.org/10.18632/oncotarget.8473)
	SNORD114-20	small nuclear RNA associated with methylation	4,856381927	8,43E-06	Imprinted snoRNA 14q(II-20; cluster II) expressed from the maternal allele; higher expression levels in the brain and the endometrium (https://doi.org/10.1093/hmg/11.13.1527) | GO:0006396: RNA processing
	DKK1	encodes a member of the dickkopf family of proteins	4,813007254	3,26E-06	Tumor suppressor (TSGene) | Member of the Dickkopf family; Inhibitor of the Wnt signaling pathway (GeneCards; https://doi.org/10.1074/jbc.274.27.19465) | GO:0008083: Growth factor activity; GO:0007275: Multicellular organism development | Upregulated in hepatoblastoma (https://doi.org/10.1097/01.lab.0000059926.66359.bd) | Increased expression has also been observed in several cancers such as lung cancer, hepatocellular carcinoma, esophageal squamous cell carcinoma, gastric cancer, breast cancer, among others (https://doi.org/10.2147/cmar.s275172)
	RP1-74M1.3	long non coding RNA	4,785288061	6,48E-05	lincRNA; ENSG00000272796 (Ensembl)

	CDH3	encodes a classical cadherin of the cadherin superfamily	4,561553198	0,004343233	Member of the cadherin superfamily of cell adhesion proteins (GeneCards) Z Reactome R-HSA-421270.5: Cell-cell junction organization | Upregulated in breast cancer (https://doi.org/10.1007/s10549-019-05477-5), gallbladder carcinoma (https://doi.org/10.1016/j.prp.2014.01.014), thyroid cancer (https://pubmed.ncbi.nlm.nih.gov/32655830/), colorectal cancer (https://doi.org/10.18632/oncoscience.370), among others | Downregulated in hepatocellular carcinoma (https://pubmed.ncbi.nlm.nih.gov/25337260/)
	PART1	prostate androgen-regulated transcript 1	4,547262613	0,009047934	lncRNA (GeneCards) | Increased expression in breast cancer (https://doi.org/10.4149/gpb_2020008), pancreatic cancer (https://doi.org/10.1186/s12957-021-02232-3), non-small-cell lung cancer (https://doi.org/10.1002/jcb.29714), hepatocellular carcinoma (https://doi.org/10.2147/ott.s259962), among others
	DKK4	encodes a member of the dickkopf family of proteins	4,42392665	0,000100471	Member of the Dickkopf family; Inhibitor of the Wnt signaling pathway (GeneCards; https://doi.org/10.1111/jcmm.16372) | GO:0007275: Multicellular organism development | Upregulated in renal carcinoma (https://doi.org/10.1002/cncr.25666), colorectal cancer (https://doi.org/10.1111/j.1349-7006.2009.01272.x) and pancreatic cancer (https://doi.org/10.1007/s13277-015-4379-x) | Downregulated in hepatocellular carcinoma (https://doi.org/10.1038/onc.2011.580)
	SLC7A11-AS1	long non coding RNA	4,389239117	0,000160135	Antisense RNA; Antisense to SLC7A11 (GeneCards) | Upregulated in lung cancer (https://doi.org/10.2147/ott.s253082) | Downregulated in gastric cancer (https://doi.org/10.18632/oncotarget.22486)
	SNORD114-13	small nuclear RNA associated with methylation	4,381968536	0,000208928	Imprinted snoRNA 14q(II-13; cluster II) expressed from the maternal allele; higher expression levels in the brain and the endometrium (https://doi.org/10.1093/hmg/11.13.1527) | GO:0006396: RNA processing
	RP11-431J24.2	long non coding RNA	4,368201534	0,000256996	Antisense RNA; ENSG00000238178 (Ensembl) | Antisense To MAGEB17 (GeneCards)
	CDCA7	cell division cycle associated 7	4,358962994	0,003999118	Member of the cell division cycle associated family; c-Myc target gene (GeneCards) | GO:0006355: Regulation of transcription, DNA-templated; GO:0006915: Apoptotic process; GO:0042127: Regulation of cell proliferation | Overexpression in a variety of cancers, such as renal cell carcinoma (https://doi.org/10.1186/s12935-021-01834-x), breast cancer (https://doi.org/10.1002/ijc.31766), lung adenocarcinoma (https://doi.org/10.1016/j.prp.2019.152559) and colorectal cancer (https://doi.org/10.3892/mmr.2020.11089)
	MAGED4	MAGE family member D4	4,341087017	0,000270014	Member of the melanoma-associated antigen family (GeneCards)| Elevated expression levels in breast cancer (https://dx.doi.org/10.3892%2Fol.2019.10722), esophageal squamous cell carcinoma (https://doi.org/10.21873/anticanres.13807), hepatocellular carcinoma (https://doi.org/10.1002/jso.23440), glioma (https://doi.org/10.7314/apjcp.2014.15.8.3495), among others

	HMGA2	encodes a protein that belongs to the non-histone chromosomal high mobility group (HMG) protein family	4,339765949	4,34E-05	Oncogene (ONGene) | Transcriptional regulator; regulation of cell cycle (GeneCards) | GO:0001837: Epithelial to mesenchymal transition; GO:0003131: Mesodermal-endodermal cell signaling; GO:0006325: Chromatin organization; GO:0007095: Mitotic G2 DNA damage checkpoint | Reactome R-HSA-2559586.3: DNA Damage/Telomere Stress Induced Senescence | KEGG hsa05202: Transcriptional misregulation in cancer | Affects cell proliferation, metastasis, EMT, telomere length maintenance and evasion of apoptosis (https://doi.org/10.3892/ijo.2019.4856) | Upregulated in lung adenocarcinoma (https://doi.org/10.1111/1759-7714.12476), breast cancer (https://doi.org/10.1002/jcp.28397), bladder cancer (https://doi.org/10.1111/cpr.12096), among others
	GNG4	G protein subunit gamma 4	4,307933033	3,95E-05	GTPase involved in transmembrane signaling (GeneCards) | GO:0007165: Signal transduction; GO:0007186: G protein-coupled receptor signaling pathway; GO:0030308: Negative regulation of cell growth | KEGG hsa04062: Chemokine signaling pathway | Upregulated in gastric cancer and related liver metastasis (https://doi.org/10.1038/s41416-021-01366-1), colon cancer (https://doi.org/10.2147/cmar.s173240) and non-small-cell lung cancer (https://doi.org/10.1155/2021/6696198)
	RP4-655J12.4	long non coding RNA	4,289677231	0,0001547	lincRNA; ENSG00000233154 (Ensembl) | LINC01762 (GeneCards)
	RP11-401P9.4	long non coding RNA	4,254929541	7,68E-06	Antisense RNA; ENSG00000260249 (Ensembl) | Antisense to SNX20 (GeneCards)
	PRND	prion like protein doppel	4,18681406	0,000328201	Membrane glycoprotein (GeneCards) | Reactome R-HSA-392499.7: Metabolism of proteins | Role in promoting angiogenesis (https://doi.org/10.1242/dev.193094, https://doi.org/10.1172/jci83427) | Polymorphisms implicated in neurological disorders (https://doi.org/10.1007/s004390100591)
	SNORD114-4	small nuclear RNA associated with methylation	4,170782057	0,000531715	Imprinted snoRNA 14q(II-4; cluster II) expressed from the maternal allele; higher expression levels in the brain and the endometrium (https://doi.org/10.1093/hmg/11.13.1527) | GO:0006396: RNA processing
	SNORD114-16	small nuclear RNA associated with methylation	4,167348041	0,000592038	Imprinted snoRNA 14q(II-16; cluster II) expressed from the maternal allele; higher expression levels in the brain and the endometrium (https://doi.org/10.1093/hmg/11.13.1527) | GO:0006396: RNA processing
downregulated	*NPBWR1*	neuropeptides B and W receptor 1	-5,490135811	2,72E-05	Interacts with opioid ligands (GeneCards) | GO:0007165: Signal transduction; GO:0007186: G protein-coupled receptor signaling pathway; GO:0007218: Neuropeptide signaling pathway; GO:0007268: Chemical synaptic transmission; GO:0019222: Regulation of metabolic process | Reactome R-HSA-372790.4: Signaling by GPCR; Reactome R-HSA-162582.11: Signal transduction | KEGG hsa04080: Neuroactive ligand-receptor interaction

	HORMAD2	HORMA Domain containing 2	-4,589310678	5,82E-05	GO:0051177: Meiotic sister chromatid cohesion; GO:0051321: Meiotic cell cycle | Hypermethylation and decreased expression in thyroid cancer; restored expression slowed down cell proliferation and increased apoptosis (https://doi.org/10.1111/jcmm.13680)
	BCL2L10	BCL2 Like 10	-4,579128675	0,000593448	Member of the BCL-2 protein family of apoptosis modulators (GeneCards) | GO:0042981: Regulation of apoptotic process; GO:0008630: Intrinsic apoptotic signaling pathway in response to DNA damage; GO:2001243: Negative regulation of intrinsic apoptotic signaling pathway | Downregulated in hepatoma (https://doi.org/10.18632/aging.101737), hepatocellular carcinoma (https://doi.org/10.1002/mc.22580) and gastric cancer (https://doi.org/10.3892/or_00000814)
	AC022596.2	long non coding RNA	-4,352911344	0,000940723	lncRNA; ENSG00000283033 (GeneCards)
	GCK	glucokinase	-4,350541402	0,0001547	Member of the hexokinase family of hexose 6-phosphotransferases; catalyzes the conversion of glucose into glucose-6-phosphate (GeneCards) | GO:0001678: Cellular glucose homeostasis; GO:0005975: Carbohydrate metabolic process | KEGG hsa00010: Glycolysis / Gluconeogenesis; KEGG hsa05230: Central carbon metabolism in cancer; KEGG hsa01100: Metabolic pathways | Abnormal metabolism and glucose uptake is a hallmark of cancer (https://doi.org/10.1016/j.cmet.2015.12.006)
	IGFALS	insulin like growth factor binding protein acid labile subunit	-4,004112021	1,14E-05	Tumor suppressor (TSGene) | GO:0007155: Cell adhesion; GO:0007165: Signal transduction; GO:0044267: Cellular protein metabolic process | KEGG hsa04935: Growth hormone synthesis, secretion and action | Reactome R-HSA-392499.7: Metabolism of proteins | Downregulated in hepatocarcinomas; affecting IGF signaling pathways (https://doi.org/10.1016/j.jhep.2013.10.014)
	CRYAA	crystallin alpha A	-3,987797835	0,000928453	Member of the heat shock family; HSPB4; acts similarly to chaperones; response to stress (GeneCards) | GO:0007601: Visual perception; GO:0032387: Negative regulation of intracellular transport; GO:0042026: NOT protein refolding; GO:0043066: Negative regulation of apoptotic process; GO:0050821: Protein stabilization | KEGG hsa04141: Protein processing in endoplasmic reticulum | Downregulated in pancreatic cancer (https://doi.org/10.18632/oncotarget.11668)
	RP4-763G1.2	long non coding RNA	-3,95784965	0,000642558	lincRNA; LINC01702; ENSG00000235200 (LNCipedia)
	SEC14L3	SEC14 like lipid binding 3	-3,845861696	0,001545014	Similar to the phosphatidylinositol (phospholipid) transfer protein SEC14 from Saccharomyces cerevisiae; involved in the biogenesis of transport vesicles (GeneCards) | GO:0008289: Lipid binding
	TTC36	Tetratricopeptide repeat domain 36	-3,808668156	0,000183617	Chaperone HBP21; binds to the heat shock protein HSP70 (GeneCards) | Acts as tumor suppressor (downregulated) in hepatocellular carcinoma; capable of inducing apoptosis under stress conditions (https://doi.org/10.1093/carcin/bgv116) | Downregulated in gastric cancer (https://doi.org/10.7150/jca.47292)
	RP11-135A1.3	long non coding RNA	-3,753508148	0,009529464	lincRNA; ENSG00000229964 (Ensembl)

RPS2P6	ribosomal protein S2 pseudogene 6	-3,695821857	0,010890416	RPS2 pseudogene (GeneCards)
INS-IGF2	INS-IGF2 readthrough	-3,640924657	0,01123384	Encodes one protein coding transcript and one nonsense-mediated decay transcript, both imprinted and expressed from the paternal allele (GeneCards) | GO:0005179: Hormone activity; GO:0007165: Signal transduction | AMP-activated protein kinase (AMPK) signaling; Type II diabetes mellitus (BioSystems)
LINC01093	long non coding RNA	-3,640155177	0,004301224	lincRNA; ENSG00000249173 (GeneCards) | Downregulated in liver tumors (https://doi.org/10.1007/s00432-020-03378-5P, https://doi.org/10.3892/etm.2019.8046)
CTA-85E5.10	long non coding RNA	-3,584516563	1,57E-05	Antisense RNA; ENSG00000227117 (Ensembl) | Antisense to HORMAD2 and MTMR3 (GeneCards)
THRSP	thyroid hormone responsive	-3,583675648	0,004826747	Lipogenesis regulator; expressed in the liver, adipocytes and lactating mammary gland; responsive to hormones and dietary substrates (GeneCards) | GO:0046890: Regulation of lipid biosynthetic process; GO:0006629: Lipid metabolic process | Reactome R-HSA-1430728.10: Metabolism; Reactome R-HSA-556833.7: Metabolism of lipids
CNDP1	carnosine dipeptidase 1	-3,556646028	0,00290187	Member of the M20 family of metalloproteases (GeneCards) | GO:0006508: Proteolysis; GO:0032268: Regulation of cellular protein metabolic process | KEGG hsa01100: Metabolic pathways; KEGG hsa00330: Arginine and proline metabolism; KEGG hsa00410: β-alanine metabolism; KEGG hsa00340: Histidine metabolism | Low levels in plasma from patients with gastrointestinal cancer, poor prognosis and cachexia (https://doi.org/10.1371/journal.pone.0123566), glioblastoma (https://doi.org/10.1371/journal.pone.0046153) and aggressive, metastatic prostate cancer (https://doi.org/10.1074/mcp.m110.001560)
AC111186.1	misc_RNA	-3,531879942	0,000391167	WIPI2 processed pseudogene; ENSG00000270993 (GeneCards)
RP11-434D9.1	long non coding RNA	-3,47894726	0,001278478	lincRNA; ENSG00000249364 (Ensembl) | Uncharacterized LOC101928858 (GeneCards)
SYT7	synaptotagmin 7	-3,47429767	2,72E-05	Member of the synaptotagmin family of calcium-dependent membrane proteins, involved in synaptic transmission (GeneCards) | GO:0006887: Exocytosis; GO:0016192: Vesicle-mediated transport | Reactome R-HSA-6794362.4: Protein-protein interactions at synapses; Reactome R-HSA-112316.7: Neuronal system; Reactome R-HSA-6794361.4: Interactions of neurexins and neuroligins at synapses | Upregulated in lung cancer (https://doi.org/10.1016/j.prp.2020.153101, https://doi.org/10.1042/bsr20181298), osteosarcoma (https://doi.org/10.14670/hh-18-174), colorectal cancer (https://doi.org/10.7150/jca.25098) and glioblastoma https://doi.org/10.3892/mmr.2017.7723)
RP11-467I20.6	long non coding RNA	-3,470037624	0,008162224	lncRNA (lncRNome)
CYP2B7P	cytochrome P450 family 2 subfamily B member 7 pseudogene	-3,459388422	0,009013014	CYP2B7 pseudogene (GeneCards)

	CAMK2B	calcium/calmodulin dependent protein kinase II beta	-3,427338291	8,80E-05	Role in calcium signaling and synaptic plasticity (GeneCards) | GO:0004674: Protein serine/threonine kinase activity; GO:0004683: Calmodulin-dependent protein kinase activity | Epigenetically silenced in breast cancer cell lines (https://doi.org/10.5483/bmbrep.2011.44.8.523) | Downregulated in glioblastoma multiforme (https://doi.org/10.3892/or.2019.7022)
	ASPG	Asparaginase	-3,418259938	0,013645656	Lysophospholipase, transacylase, platelet-activating factor acetylhydrolase and asparaginase (asparagine hydrolase) (GeneCards) | Reactome R-HSA-71291.6: Metabolism of amino acids and derivatives; Reactome R-HSA-1430728.10 : Metabolism; Reactome | GO:0006520: Cellular amino acid metabolic process; GO:0006629: Lipid metabolic process; GO:0008652: Cellular amino acid biosynthetic process; GO:0016042: Lipid catabolic process | Cytotoxic potential against leukemia cell lines, via inducing apoptosis (https://doi.org/10.1371/journal.pone.0178174)
	HAO2	Hydroxyacid Oxidase 2	-3,411989453	0,005085007	GO:0016491: Oxidoreductase activity; GO:0003824: Catalytic activity; GO:0006625: Protein targeting to peroxisome; GO:0009062: Fatty acid catabolic process | KEGG hsa01100: Metabolic pathways; KEGG hsa04146: Peroxisome (redox signalling, lipid homeostasis); KEGG hsa01200: Carbon metabolism | Differential expression in liver and kidney (https://doi.org/10.1074/jbc.275.17.12590)

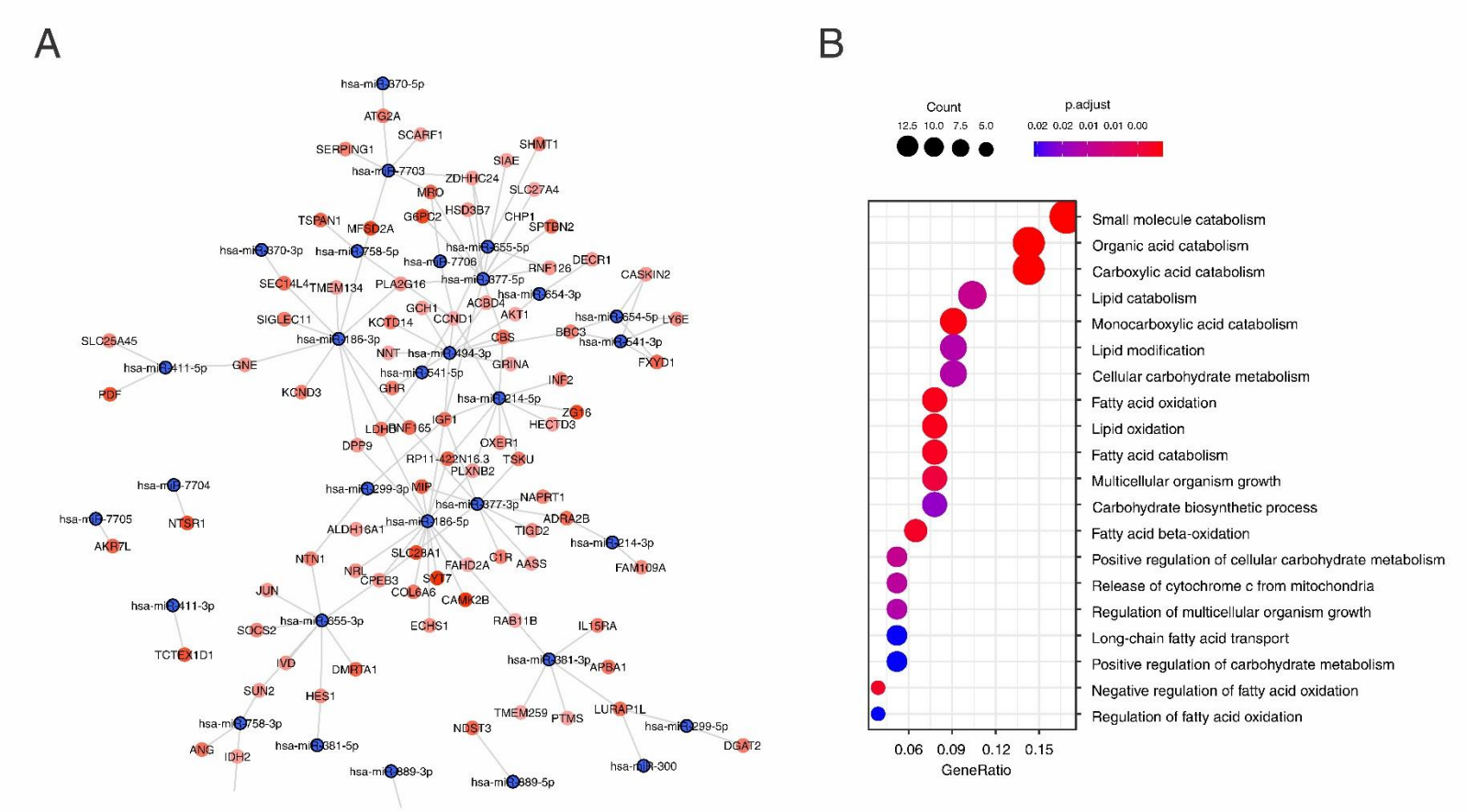


Figure 3. Association between miRNA and protein-coding genes in HBLs. (A) miRNA-mRNA interaction network. Blue dots represent miRNAs upregulated in HBL, red dots represent mRNAs downregulated in HBL. miRNA: micro-RNA; mRNA: messenger RNA. (B) Dotplot of biological processes. The dot size represents the number of mRNAs in the network associated with the process; the dot color is linked to the level of the adjp-value. The gene ratio refers to the value of the number of mRNAs associated with the biological process by the total number of mRNAs in the network.

The **Figure 4A** presents the major enriched up and down-regulated biological process, cellular components and molecular functions according to the list of genes with altered expression in HBLs. The **Figure 4B** depicts ten selected biological processes which were found to be statistically disrupted in the HBL transcriptome; in the GoCircle plot, it is possible to observe that biological processes presenting predominance of negative regulation were associated with reduced oxidation and metabolism inhibition, such as amine, nicotinamide, and lipid metabolism. Conversely, upregulated biological processes in HBLs were linked with the regulation of cell communication and signaling pathways, as well as cell differentiation, and embryonic and developmental processes. In the GoChord graph (**Figure 4C**) are depicted nine of the enriched biological processes highlighted above, and ten of the 50 top genes which are associated with them, five with increased expression and five with reduction; it is important to notice that only coding protein genes could be associated with biological processes.

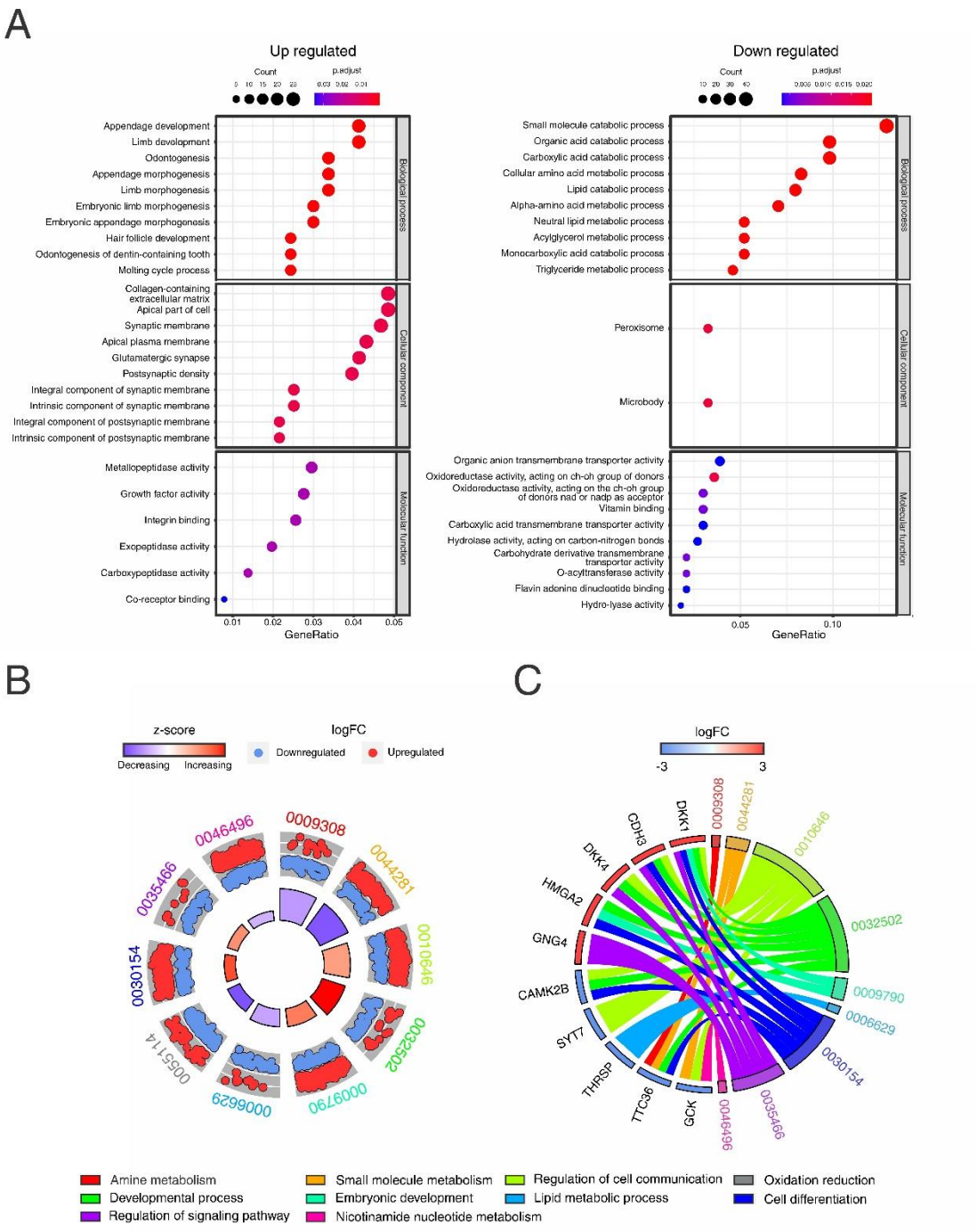


Figure 4. The gene ontology enrichment. (A) Dotplot presenting the top ten functional enrichment analyses of DEGs in biological process, cellular component, and molecular function in HBL samples. The dot size refers to the number of genes in the DEG list associated with the event, and the color is associated with the p-adj value. At left, upregulated genes, at right, downregulated genes. (B) A GOCircle showing the biological processes enriched for the differentially expressed genes of HBLs. The outer circle shows a scatter plot of the log_2FC for each gene under the gene ontology (GO) terms; red dots indicate upregulated genes and blue dots, downregulated genes. The inner circle shows the z-score, a measure of up-or downregulation predominance of

each identified process. GO identification – 0009308: Amine Metabolic process; 0044281: Small molecule metabolic process; 0010646: Regulation of cell communication; 0032502: Developmental process; 0009790: Embryonic development; 0006629: Lipid metabolic process; 0055114: Oxidation reduction; 0030154: Cell differentiation; 0035466: Regulation of signaling pathway; 0046496: Nicotinamide nucleotide metabolic process. (C) The chord diagram shows ten out of 50 genes with major expression changes associated with nine of the ten biological processes. Blue and red bars next to the genes represent the down or upregulated fold-change, respectively. Biological processes are identified by specific colors and, the association with the gene is demonstrated by the connecting chord.

The PPI network displayed 6,431 nodes and 239,279 connectors with degree mean and betweenness mean of 42.0 and 7026.0, respectively. In this network, seven clusters stood out, in relation to the enriched biological processes (**Supplementary Material S2)**. Those proteins were categorized into i) common (C): the degree and betweenness values were below the global average of the network; ii) hub (H): the degree value was above the global average of the network; iii) bottleneck (B): the value for betweenness centrality was above the global average value of the network; and iv) hub-bottleneck: both the degree and betweenness values were higher than the global network average. From this network and analysis, it was possible to generate the subnet shown in **Figure 5**, highlighting the interaction of central proteins coded by DEGs, among others. Centrality betweenness and degree measures showed that this subnet was composed by one hub protein (F13A1), nineteen bottleneck nodes (ACADS, TH, HSD17B10, ECHS1, GSTZ1, SHMT1, GLS2, RBP2, MTR, NAT8L, ELOVL4, GGPS1, HES1, LIMD1, GCK, DCP2, MSH2, CAMK2B, and CNOT6,) and eleven hub-bottleneck (CDA, CCND1, HDAC2, CCNB1, AGO1, AKT1, IRF7, PSMB8, CD36, GNAL, and GPLD1). These proteins are associated with clusters involved with biological processes of negative regulation of metabolism, mostly amine, nicotinamide and lipid/fatty acid metabolism; epigenetic mechanisms, such as DNA methylation, ncRNAs and gene silencing by miRNA; cytotoxicity, and signaling pathways, including cytokine-mediated signaling pathway.

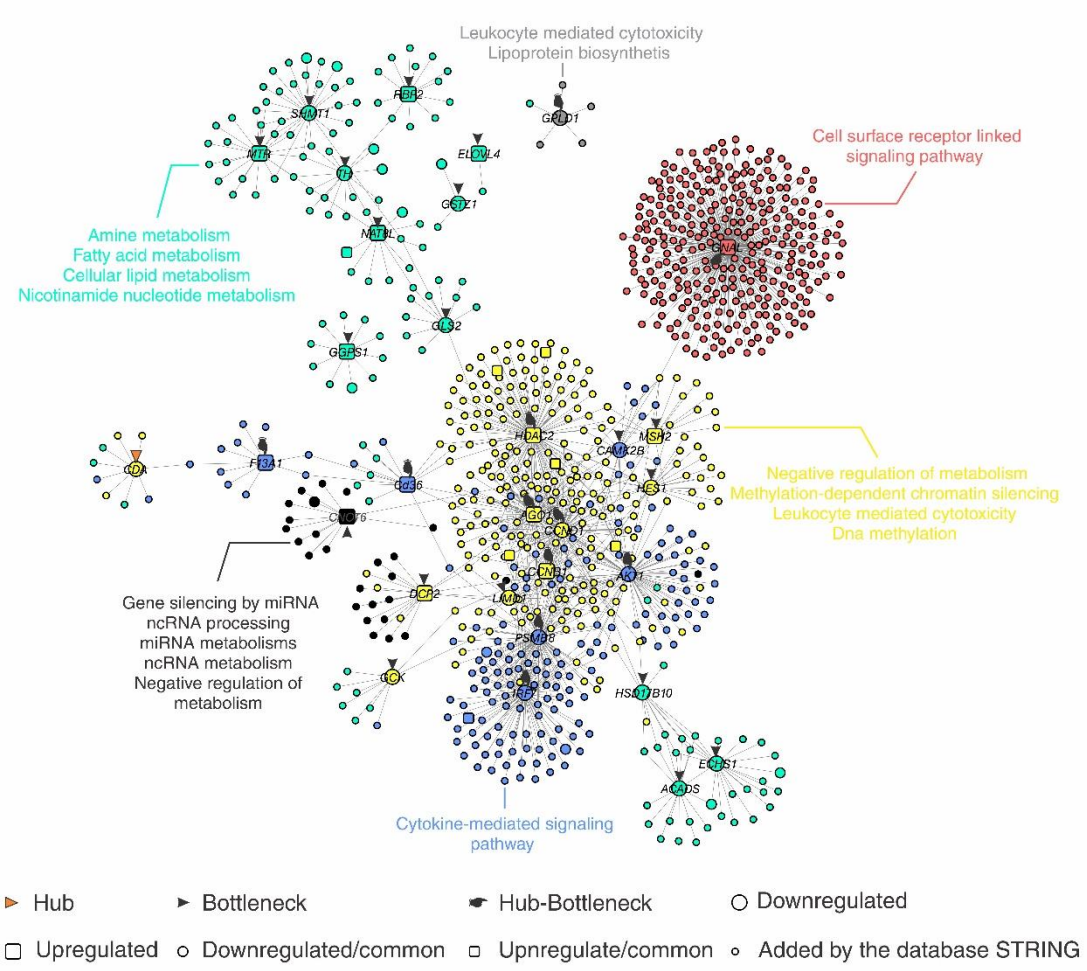


Figure 5. Protein-protein interaction network. PPI highlighting centralities and proteins that interact directly with these nodes. Colors represent different biological processes. Square and round dots indicate nodes up- and down-regulated, respectively. Centrality measures showed that the networks were composed by one hub protein (orange triangle), nineteen bottleneck nodes (black triangle) and eleven hub-bottleneck (black finger).

Discussion

The analysis of PCA and heatmap showed separation of the total samples in two clusters, one exclusively tumoral and the other composed by control liver samples and one tumor (79T). It is noteworthy that this specific tumor did not carry somatic coding mutations or copy number alterations, reinforcing a strong similarity with the group of control liver samples.

The heatmap analysis underscored four sets of genes with similar expression patterns in each cluster (**Figure 1A**). Genes from clusters 1, 2, and 3, which are upregulated in HBLs, are related to nervous system, embryonic development, cell differentiation, and adhesion process, consistent with the main hypothesis of HBL genesis based on disruption of the liver cells differentiation process [22,23,165]. A recent study showed that nervous system act as an essential part of the cancer microenvironment, and cancer stem cells derived from colorectal and gastric carcinomas were able to develop into neurons involved in tumor growth [166]. The identified cluster 4, composed of genes linked to metabolic processes and response to growth hormone stimulus with down regulation revealed a crucial role of metabolism disruption in HBLs. Changes in HBL metabolism are still poorly described and understood [167,168]. We previously showed the occurrence of lipid reduction and nicotinamide metabolism changes in HBLs [142], and this RNASeq analysis reinforces and expands these observations.

Genes with changes in expression in HBLs are widely distributed on chromosomes (**Figure 1B**), but the 14q32 region showed a cluster of upregulated genes. The overexpression of a 300 kb region located in the 14q32 *DLK1/DIO3* locus, previously reported, associated with DNA hypomethylation, indicates a direct role of this epigenetic mechanism in regulating the expression of genes in this region [57]. This locus was also associated with a major increase in the expression of a cluster of miRNAs and snoRNA of the C/D box family (namely SNORD113 and SNORD114), especially in metastatic tumors [102]. Our evidence corroborated these recent findings regarding the 14q32 locus on HBL. Carillo et. al also selected four genes mapped to this locus (*DLK1*, *MEG3*, *SNORD113-3*,

SNORD114-22) to classify tumors according to the level of 14q32 gene expression [57]. Here, three of these genes (*DLK1*, *SNORD113-3*, *SNORD114-22*) were found to be upregulated in the HBL group.

From the 1492 DEGs detected in HBLs, 85 were exclusively expressed in tumors, 47 of them being non-coding RNA from lncRNA or miRNA classes. For decades, transcripts with no potential for protein coding have been considered "noise" or "junk" accumulated across evolution [130,169]. However, high-throughput sequencing technologies revealed RNA-mediated gene regulation, and changes of the normal expression pattern of these RNAs were linked in several instances with pathological conditions [170]. Altered ncRNA expression have been observed in colorectal cancer [171], hepatocellular carcinoma [172], gastric cancer [173], and renal cell carcinoma [174], indicating that aberrant expression of ncRNAs probably contributes to carcinogenesis. A cluster analysis of HBLs showed that the microRNA profile can distinguish the tumor from the control liver tissues, indicating the importance of ncRNA in this pediatric cancer [175].

MiRNAs expression is frequently deregulated in many tumors and their specific target genes determine whether they act as tumor suppressors or oncogenes in human cancers [176]. The miRNA-mRNA analysis found that 18 of the differentially expressed miRNAs here detected in HBLs were connected RNAm regulation of 82 genes. The overexpression of miR-186 was linked with 26 down-regulated mRNA in tumor samples; recently, Cui et al. described the reduction of this gene in HBL and targeted an N6-methyladenosine gene (*MTTL3*), affecting tumor progression [177]. miR-214 was increased in tumor samples and linked with the downregulation of 12 mRNA. Gyugos *et al*. reported downregulation of this gene in HBL fetal samples [178], and this expression pattern was also evidenced by Vastrad *et al.* [164]; the miRNA-mRNA analysis made by Vastrad et al. [164] was based on mRNAs with different expressions in HBL to infer possible regulators from miRNA databases, while the present work was based in upregulated miRNAs to link with mRNA targets from the DEG list. The overexpression of miR-377 was associated with 22 down-regulated mRNA in tumor samples. Recently publications supported downregulation of this

gene in an HBL dataset and cell lines [164,179]. MiR-494 was linked to downregulation of 11 mRNA genes, and this gene has not yet been described in HBL; however, overexpression of miRNA-494 has been reported in HCC, considered an epigenetic regulator associated with tumor invasiveness, proliferation and migration [180,181]. In our work, miRNAs found to be upregulated were associated with oncogenic activities through the regulation of downregulated mRNAs. However, the analysis of the HBL literature showed that three of the highlighted miRNA presented an opposite expression pattern compared to our findings, which will require a future discussion regarding the techniques used and intrinsic differences of the cohorts. On the other hand, the upregulation of *miR-494* can be associated with oncogenic activities. In HCC, this miRNA was linked to TET1 activity and tumor progression [180]. Previous work by our group has shown that HBL has increased expression of the TET family genes (*TET1*, *TET2*, and *TET3*) associated with active demethylation activity in this tumor. The miR-494, a novelty in HBL biology, is a candidate for future analysis, aiming to explored a possible regulation of TET1. Recently, bioinformatics analyses in HBLs have highlighted changes in biological processes associated with small-molecule catabolism, organic acid metabolism, lipid metabolism, and oxidation-reduction reactions linked to the down-regulating genes in this tumor [164].

To analyze biological processes, cellular components, and altered molecular functions in HBL, we carry out gene ontology analysis (**Figure 4A**), observing that positively regulated biological processes were associated with development, morphogenesis, odontogenesis, and molting cycle process. All these increased processes are linked to developmental events, and associated with changes of genes such as *DKK1* [182], *BMP4* [183], *RUNX2* [184], *LEF1* [185], homeobox genes such as *DLX1*, *DLX6*, *MSX1* [186], among others of the DEG list involved in these processes. The HBL origin hypothesis is mostly associated with a developmental block during the liver cell differentiation process [22,23]; our RNAseq findings showed the genetic changes in HBLs interfering with developmental pathways. On the other hand, all the negatively regulated biological processes in HBL were associated with metabolism and oxidation. Through non-targeted metabolomic analysis, our group

recently detected reduced lipid content in HBLs [142]. Wang et al. demonstrated metabolic changes in HBLs associated with negative regulation of SLC10A1 and linked its effect to disturbances in the control of cell proliferation [168], a gene also downregulated in our data.

Membrane components are the majority of cellular components with upregulation which can be associated with the increase of genes such metallopeptidase as *ADAM22* [187], *ADAM23* [188], and *ADAM9* [189], while peroxisome and microbody are the downregulated cellular components in HBL. GO functional annotation analysis made by Vastrad et al. from a HBL microarray dataset showed that of the upregulated genes, the cellular components were mainly involved in DNA packaging complex and chromosomal region, while for the downregulated genes, the cellular components were associated with blood microparticle and endoplasmic reticulum membrane [164]. For the molecular functions, the metallopeptidase activity is increased in HBL, which is linked to the upregulation of the metallopeptidase family genes (*ADAM22*, *ADAM23*, *ADAM9*, *ADAMDEC1*, and *ADAMTS6*). The upregulated carboxypeptidase genes (*CPA3*, *CPA6*, and *CPXM1*) are also associated with this molecular function alteration, and are key players for the increase of the carboxypeptidase activity in HBLs. On the other hand, the oxidoreductase activity found in the list of negatively regulated molecular functions in HBLs was also observed in Vastrad et. al. [164].

Upregulated genes from the Dickkopf family (*DKK1* and *DKK4*) were associated with the development process, regulation of cell communication, and regulation of signaling pathways. The overexpression of *DKK1* was described in HBL as a novel biomarker candidate related to uncontrolled *wingless*/WNT signaling [163]. Analyzes of colorectal cancer, gastric cancer, pancreatic cancer, and renal cancer showed increased expression of *DKK4*, and only in hepatocellular carcinoma was observed a reduction in expression, demonstrating that effects of *DKK4* in tumorigenesis and progression can be different in the various kind of tumors [190,191]. In this way, the upregulation of *DKK4* in HBL is a novelty, also corroborated in a recent study [164].

The *HMGA2* upregulation is associated with the development process, embryonic development and cell differentiation. Polymorphisms of this gene were recently linked to the

advancement of HBL [192], and the increased expression was also shown in microarray dataset analysis [164]. *CDH3* encodes a classical cadherin component of glycoproteins. The upregulation of this gene is associated with cellular communication, developmental process, cell differentiation, and regulation of signaling pathway, which was not described in HBL before. The Guanine nucleotide-binding protein 4 (*GNG4*) increase is linked to regulation of signaling pathway, and also involved with the methylation process [164].

THRSP is downregulated in HBL and associated with the process of lipid metabolism. The expression of this gene is observed in the liver and adipocytes, and our group previously showed hypermethylation in its promoter region in HBL samples [101], indicating control by epigenetic mechanism interfering in the lipid metabolism in the tumor. *TTC36* downregulation evidenced in HBL was associated with alterations in the metabolism of amine and small molecules, development process and cell differentiation. This gene is associated with tumoral suppressor activities, and its inactivation in gastric cancer might promote cell proliferation, at least in part, through activating the Wnt/β-catenin signaling pathway [193]. Glucokinase (*GCK*) contributes to the reprogramming of energy metabolism in cancer cells [194]. The present work showed the negative regulation of this gene linked to changes in the metabolism of small molecules and nicotinamide and regulation of cellular communication. The downregulation of *STY7* was associated with alterations in cell communication; overexpression of this gene was reported in gastric and hepatic cancer [195,196], however, it is the first evidence in HBL. The downregulation of *CAMK2B* was associated with alterations in the regulation of cell communication, development process and cell differentiation; it was linked with promoter regulation in breast cancer [197], but there is no description in liver cancer. It is important to note that both genes, *STY7* and *CAMK2B*, showed a reduction in expression associated with an increase in the miR-186, and they are predicted to be targets of this miRNAs.

Analysis of the protein network showed which proteins encoded by the DEGs are decisive in the pathways and mechanisms highlighted throughout this work (**Figure 5**). Histone deacetylase 2 (HDAC2) is a well-known epigenetic player, crucial for embryonic

development [198]. Recent publications showed that epigenetic silencing by histone deacetylases (HDAC) might be critical for development of pediatric liver cancer [199]. Particularly, elevation of HDAC1 and HDAC2 proteins was found in a large group of patients with HBLs [200]. In addition to corroborating the literature data with evidence of increased HDAC2 expression, we highlight here that this protein presents a hub-bottleneck centrality associated with methylation-dependent chromatin silence, DNA methylation, and negative regulation of metabolism. Also classified as hub-bottleneck centrality protein, Cyclin D1 (CCND1) presented downregulation in tumor samples. CCND1 is a major regulator of the cell cycle transition from G1 phase to S phase [201], which indicates that the reduction of this protein interfere directly in the control of cell proliferation in HBL. Polymorphism of this gene has also been described to contribute to tumor development in children with HBL [201], and here we evidenced that CCND1 protein was associated with methylation-dependent chromatin silence, DNA methylation, negative regulation of metabolism, leukocyte mediated cytotoxicity, and cytokine-mediated signaling pathway. It is important to note that the regulation of *CCND1* by mir-494 was observed in HBL in the miRNA-mRNA analysis. GCK and CAMK2B were classified with this centrality degree presenting downregulation in HBL. The first one was associated with metabolism and cytokine pathways, and the second was associated with epigenetic process as DNA methylation and gene silencing as well as cytokine-mediated signaling. The mRNA encoding these proteins is part of the ten genes highlighted in the analysis of the GOcircle associated with processes of interest, and *CAMK2B* was also predicted to be regulated by miR-186 in the miRNA-mRNA analysis. In this way, *GCK* and *CAMK2B* are candidates for future analyzes. The downregulated CDA was the only hub protein detected in our analysis; CDA was associated with proteins from multiple clusters as metabolism pathways, cytokine-mediated signaling, and epigenetic mechanisms, so far, not reported in HBLs.

Conclusion

Despite the limited cohort, the present work showed key biological pathways in the tumorigenesis of HBL. We showed deregulation in the gene expression profile in HBL with

a predominance of upregulated genes, highlighting 85 genes exclusively expressed in tumors. Different techniques have shown that processes associated with oxidation and metabolism are reduced in tumors. Changes in lipid and nicotinamide metabolism, pathways previously described by our group, reinforced the importance of these processes in HBL tumorigenesis, and *THRSP* and *GCK* gene were highlighted as possible candidates for future analyses. From the total list of DEGs, it was noticed a strong presence of non-coding RNAs. The importance of ncRNA as an epigenetic mechanism involved in tumorigenesis is widely described, however, in HBL few studies have focused on this area. We found four upregulated miRNAs (miR-186, miR-214, miR-377, and miR-494) in HBLs linked to reduced expression of at least ten genes. Processes associated with dysregulation of miRNAs in HBLs were mainly linked to metabolism. Thus, the present work presents changes in metabolic pathways as well as its regulation by epigenetic mechanisms associated with miRNA as a future target for understanding tumorigenesis in HBL.

SUPPLEMENTARY MATERIAL

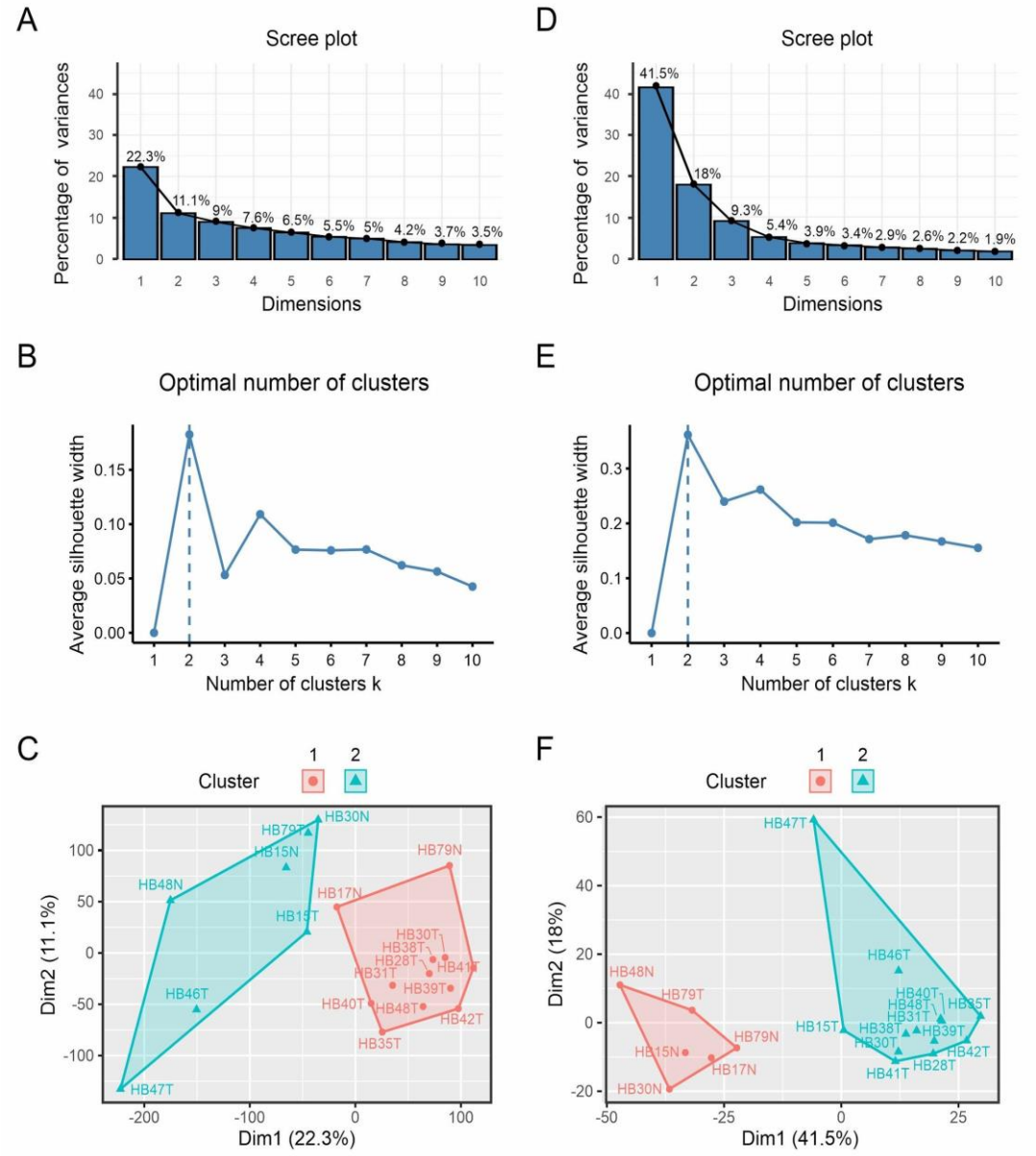
Transcriptome dysregulation in hepatoblastomas: the pivotal role of noncoding RNAs and detection of altered lipid metabolism

Maria Prates Rivas[1], Edson Mario de Andrade[2], Talita Ferreira Marques Aguiar[1], Sara Ferreira Pires[1], Alexandre Defelicibus[3], Bruna Barros[3], Estela Novak[4], Lilian Maria Cristofani[4], Vicente Odone[4], Monica Cypriano[6], Silvia Regina Caminada de Toledo[5], Dirce Maria Carraro[3], Isabela Werneck da Cunha[6,7], Cecilia Maria Lima da Costa[8], Israel Tojal[3], Tiago Antônio de Oliveira Mendes[2], Ana Cristina Victorino Krepischi[1*]

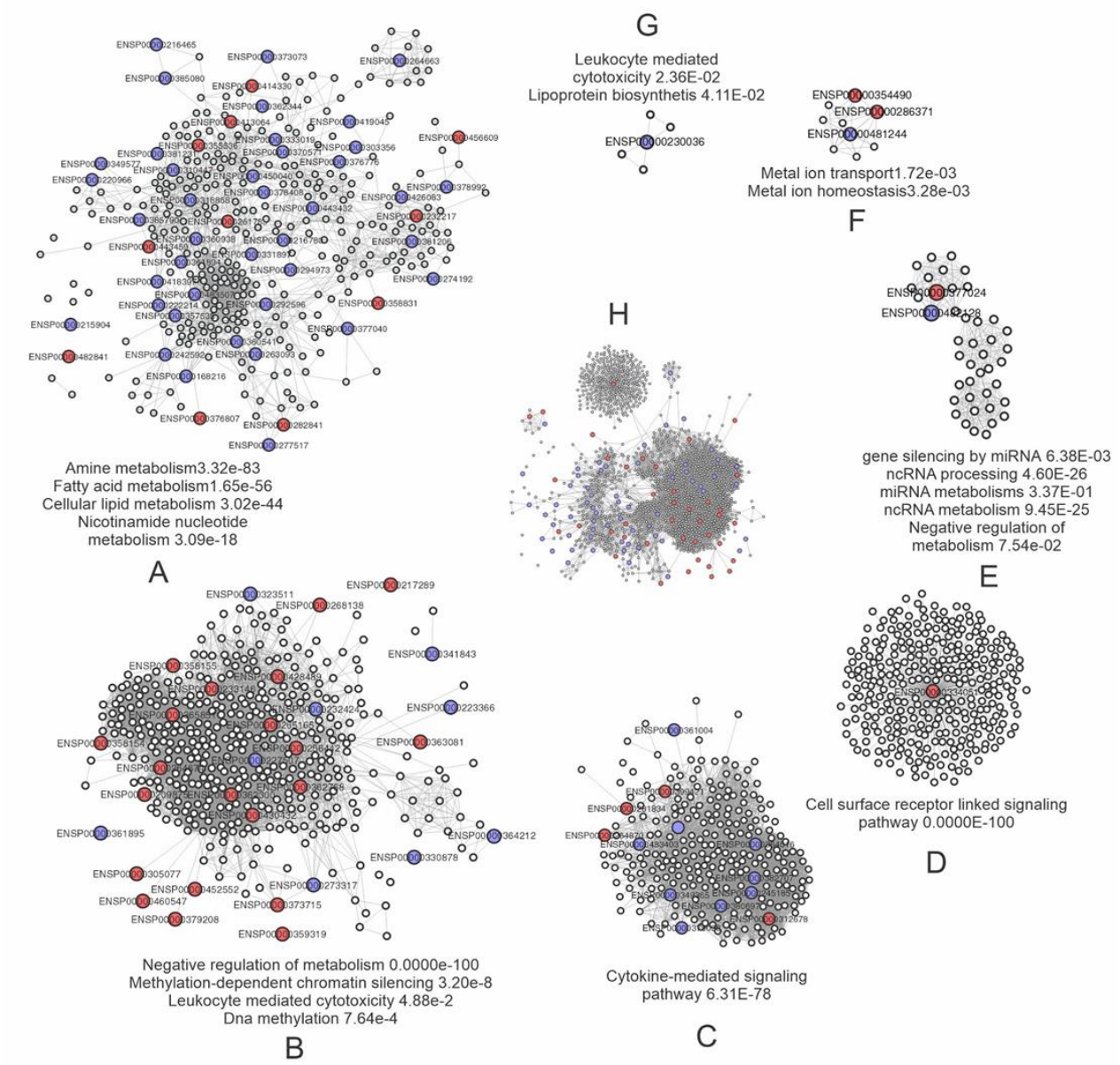
1. Human Genome and Stem-Cell Research Center, Department of Genetics and Evolutionary Biology,
Institute of Biosciences, University of São Paulo, São Paulo, Brazil.
2. Departament of Biochemistry and Molecular Biology, Federal University of Viçosa, Minas Gerais, Brazil
3. International Center for Research, A. C. Camargo Cancer Center, São Paulo, Brazil.
4. Pediatric Cancer Institute (ITACI) at the Pediatric Department, São Paulo University Medical School, São Paulo, Brazil.
5. Department of Pediatric, Adolescent and Child with Cancer Support Group (GRAACC), Federal University of São Paulo, São Paulo, Brazil.
6. Department of Pathology, Rede D'OR-São Luiz, São Paulo, Brazil.
7. Department of Pathology, A. C. Camargo Cancer Center, São Paulo, Brazil.
8. Department of Pediatric Oncology, A. C. Camargo Cancer Center, São Paulo, Brazil.

* Corresponding author:
Ana Cristina Victorino Krepischi
Human Genome and Stem-Cell Research Center, Department of Genetics and Evolutionary Biology - Institute of Biosciences, University of São Paulo, São Paulo, Brazil
Phone: 55 11 3091 7573
e-mail: ana.krepischi@ib.usp.br

Supplementary Figure S1. Principal Component Analysis (PCA) and clustering of HBL and control liver samples. The PCA analysis was applied to the total RNASeq data (A, B, and C), and to the differentially expressed genes (DEGs) (FDR <0.05 and log2 fold-change) (D, E, and F). In A and D, the contribution in total variance of the first 10 dimensions (components) are represented (bars). The graphs in B and E showed the ideal number of clusters (dashed vertical lines) both for the total RNASeq data (B) and the DEGs (E). Finally, the graphs C and F correspond to the PCAs highlighting the two identified clusters and the respective grouped samples; the cluster 1 samples are represented by solid red dots and the cluster 2 samples are represented by blue triangles (HB15T, HB28T, HB30T, HB31T, HB35T, HB38T, HB39T, HB40T, HB41T, HB42T, HB46T, HB47T, HB48T, and HB79T are tumors samples; HB15N, HB17N, HB30N, HB48N, and HB79N are control liver samples).



Supplementary Figure S2. Protein-protein interaction (PPI) network. The network represents the enriched processes in each cluster and all proteins that interact directly (A-G). Proteins with a degree of centrality (hub, bottleneck, or hub-bottleneck) were represented as larger dots. Down or upregulation of genes encoding proteins with centrality are shown in blue and red, respectively. The main network is shown in the center of the figure (H).



Supplementary Table S1. Differentially expressed genes with exclusive expression in HBL samples.

Gene	log2 fold-change	padj	Gene Info
PAX2	2,938164153	0,024433155	Oncogene (ONGene) | Transcription factor expressed during development; acts in proliferation and tissue regeneration in breast, kidney and female reproductive tract stem cells (https://doi.org/10.1158/1940-6207.capr-15-0121-t) | Important for the development of nervous system, eyes, ears, kidney and mammary gland (https://doi.org/10.1016/j.diabres.2019.107792) | Wnt / Hedgehog / Notch pathways; Embryonic and Induced Pluripotent Stem Cell Differentiation Pathways and Lineage-specific Markers (PathCards) | Increased expression observed in prostate cancer (https://doi.org/10.1002/pros.22687; https://doi.org/10.1111/febs.12959), breast cancer (https://doi.org/10.1038/sj.onc.1205172) and chemotherapy-resistant pancreatic ductal adenocarcinoma (https://doi.org/10.1371/journal.pone.0223554)
CMA1	3,965601266	0,001278478	Peptidase (serin protease); expressed in mast cells (GeneCards) | KEGG hsa04614: Renin-angiotensin system; regulation of blood pressure | Reactome R-HSA-1474228.4: Degradation of the extracellular matrix; R-HSA-2980736.2: Peptide hormone metabolism; R-HSA-1592389.2: Activation of Matrix Metalloproteinases | Expression associated with infiltration of immune markers in gastric cancer (https://doi.org/10.1080/08916934.2020.1735371)
PAGE4	2,801849943	0,037108861	Member of the G antigen gene family; expressed in fetal and reprodutive tissues such as prostate, testis, uterus, and in related cancers (GeneCards; http://www.ncbi.nlm.nih.gov/pmc/articles/pmc4219281/; https://doi.org/10.1186/s13046-019-1032-3) | GO:0006979: Response to oxidative stress; GO:0032872: Regulation of stress-activated MAPK cascade; GO:0043066: Negative regulation of apoptotic process; GO:0045893: Positive regulation of transcription, DNA-templated
AC105760.3	3,275014125	0,012278856	Antisense RNA; ENSG00000124835 (Ensembl) | Uncharacterized LOC93463 (GeneCards)
LECT1	2,830544021	0,036593432	Transmembrane protein; growth regulator; positive control of chondrocyte growth and inhibition of angiogenesis (GeneCards) | GO:0001501: Skeletal system development; GO:0001886: Endothelial cell morphogenesis; GO:0006029: Proteoglycan metabolic process | Downregulated in gastric adenocarcinoma (https://doi.org/10.3892/ijo.2015.3081) and chondrosarcoma (https://doi.org/10.1016/s0014-5793(99)01201-6) | Potential of suppressing tumorigenesis in osteosarcoma (https://doi.org/10.3892/mmr.2017.7629) and breast cancer (https://doi.org/10.3892/mmr.2016.5079) | Overexpressed in Ewing sarcoma (https://doi.org/10.1002/1878-0261.12057)
GCM1	2,961328139	0,027948918	Transcription factor (GeneCards) | Human early embryo development (PathCards) | KEGG hsa04928: Parathyroid hormone synthesis, secretion and action | Epigenetic regulator correlated with the histone marks H3K79me2/me3; decreased expression in high-grade serous ovarian cancer (https://doi.org/10.1038/s41467-018-03590-5) | Target of the Wnt/β-catenin signalling pathway (https://doi.org/10.1038/ncomms1551)
DLX1	3,176921225	0,004904049	Transcription factor; regulator of cytokine signaling pathways, such as TGF-β (GeneCards) | DNA damage/telomere stress induced senescence; Regulation of nuclear SMAD2/3 signaling (PathCards) | GO:0000122: Negative regulation of transcription by RNA polymerase II; GO:0007275: Multicellular organism development; GO:0009790: Embryo development | Overexpressed in prostate cancer; related to the Wnt/β-catenin pathway (https://doi.org/10.1186/s12894-021-00810-x) and the TGF-β/Smad4 signaling pathway (https://doi.org/10.1111/jcmm.14402)
ATP6V0D2	3,436284992	0,007365906	Subunit of an ATPase complex (GeneCards) | KEGG hsa04151: PI3K-Akt signaling pathway; KEGG hsa01100: Metabolic pathways; KEGG hsa05165: Human papillomavirus infection | Reactome R-HSA-168256.7: Immune system | Promotes tumorigenesis in esophagus cancer; related to epithelial-mesenchymal transition (https://doi.org/10.1007/s10388-020-00735-8) | Downregulation in tumor-associated macrophages results in tumor progression (https://doi.org/10.1172/jci123027)

MAGED4	4,341087017	0,000270014	Member of the melanoma-associated antigen family (GeneCards)| Elevated expression levels in breast cancer (https://dx.doi.org/10.3892%2Fol.2019.10722), esophageal squamous cell carcinoma (https://doi.org/10.21873/anticanres.13807), hepatocellular carcinoma (https://doi.org/10.1002/jso.23440), glioma (https://doi.org/10.7314/apjcp.2014.15.8.3495), among others
LRRC36	3,452629657	0,00637598	Member of the leucine-rich repeat superfamily | Other members present increased expression levels in cancer, such as LRRC15 in osteosarcoma (https://doi.org/10.1002/jor.24848)
GABRB1	3,775478148	0,028584338	Subunit of a chloride channel (GABA) related to inhubitory synaptic transmission (GeneCards) | Reactome R-HSA-112314.8: Neurotransmitter receptors and postsynaptic signal transmission | KEGG hsa04727: GABAergic synapse; KEGG hsa04080: Neuroactive ligand-receptor interaction | GO:0007165: Signal transduction; GO:0007214: Gamma-aminobutyric acid signaling pathway; GO:0007268: Chemical synaptic transmission | Associated with psychiatric disorders (https://doi.org/10.1016/j.ynstr.2021.100307; https://doi.org/10.1002/ana.25619) |High expression in colon tumor (https://dx.doi.org/10.3892%2Fol.2020.11573)
ARMC3	2,755630634	0,04233934	Beta-catenin-like protein (GeneCards) | GO:0005515: Protein binding
TMEM130	3,586231628	0,023112954	Transmembrane protein (GeneCards) | GO 0005515: Protein binding
MYOZ2	3,144748553	0,017606836	Sarcomeric protein; involved in calcineurin signaling (GeneCards)| GO:0000122: Negative regulation of transcription by RNA polymerase II; GO:0007519: Skeletal muscle tissue development; GO:0045214: Sarcomere organization | High expression levels reported in gastric cancer (https://doi.org/10.26355/eurrev_201809_15921)
ISX	1,83188755	0,029923689	Transcription factor (GeneCards) | Intestine specific; Metabolism of vitamin A (https://doi.org/10.1074/jbc.m707928200; https://doi.org/10.1073/pnas.1714963114) | Increased expression levels in hepatocellular carcinoma (https://doi.org/10.1158/0008-5472.can-12-2795)
LINC00670	3,042323484	0,022953977	lincRNA (GeneCards)
ANKRD62	3,194933485	0,014080593	Member of the group of Ankyrin repeat domain containing proteins (GeneCards) | GO:0005515: Protein binding | Other members are differentially expressed in colorectal cancer (ANKRD12; downregulated; https://doi.org/10.1186/1756-9966-32-35), colon cancer (ANKRD6 ; upregulated; https://doi.org/10.21147/j.issn.1000-9604.2021.01.10) and breast cancer (ANKRD22; upregulated; https://doi.org/10.17305/bjbms.2020.4701)
C1orf227	2,870009759	0,031887107	SPATA45; Spermatogenesis-associated protein (GeneCards) | Other SPATA proteins are involved in pro-inflammatory TNF receptor signalling in ovarian cancer (https://doi.org/10.1111/cas.13955) and endometrial cancer (https://doi.org/10.1093/carcin/bgaa034)
MIR494	3,468901916	0,00647755	Tumor suppressor (TSGene) | KEGG hsa05206: MicroRNAs in cancer | Expression levels upregulated in hepatocellular carcinoma; related to increased cell proliferation, migration and invasion, and decreased apoptosis; increases resistance of cancer cells to the chemotherapeutic agent sorafenib by targeting PTEN; regulates the Pi3K/Akt pathway; (https://doi.org/10.3892/or.2015.4030) | High expression levels previously reported in other cancer types such as retinoblastoma (https://doi.org/10.3892/ol.2020.11749), non-small cell lung cancer (https://doi.org/10.2147/ott.s278233) and gastrointestinal stromal tumors (https://doi.org/10.1002/ijc.31235)
KRT39	2,752089167	0,04185929	Structural protein (keratin); part of the cytoeskeleton of epithelial cells (GeneCards) | Reactome R-HSA-1266738.7: Developmental biology | KEGG hsa04915: Estrogen signaling pathway | Expressed during hair follicle differentiation (https://doi.org/10.1038/sj.jid.5700734)
PAX9	3,189818185	0,013970758	Transcription factor expressed during development (GeneCards) | Wnt/Hedgehog/Notch pathway; Mesodermal commitment pathway (PathCards) | GO:0006357: Regulation of transcription by RNA polymerase II; GO:0007275: Multicellular organism development | Downregulated in alcohol-induced oro-oesophaeal squamous cell carcinoma (https://doi.org/10.1002/path.5041; https://doi.org/10.1002/path.4998) | Coamplified with TTF-1 and NKX2–8 in lung cancer (https://doi.org/10.1073/pnas.0900827106)
MIR411	3,006417055	0,024322725	GO:0032701: Negative regulation of interleukin-18 production; GO:0035195: Gene silencing by miRNA | Upregulated in hepatocellular carcinoma (https://doi.org/10.1016/j.biopha.2015.01.001), lung cancer (https://doi.org/10.1158/1078-0432.ccr-13-3348, https://doi.org/10.1007/s13277-015-4425-8) and osteosarcoma (https://doi.org/10.1007/s11356-018-1331-9) | Downregulated in gastric cancer (https://doi.org/10.26355/eurrev_201904_17697), ovarian cancer

			(https://doi.org/10.3892/etm.2020.8899), renal cell cancer (https://doi.org/10.5301/ijbm.5000261), breast cancer (https://doi.org/10.3892/mmr.2016.5645), among others
SNORD114-5	3,50164421	0,006185698	snoRNA (GeneCards) | GO:0006396: RNA processing
SNORD114-16	4,167348041	0,000592038	snoRNA (GeneCards) | GO:0006396: RNA processing
SNORD113-6	3,676478367	0,003570901	snoRNA (GeneCards) | GO:0006396: RNA processing
RNU6-960P	2,810270135	0,039280518	RNU6-960 pseudogene (GeneCards)
RN7SKP124	2,805909637	0,029653262	RN7SK pseudogene (GeneCards)
RNU6-652P	2,737716043	0,036392312	RNU6-652 pseudogene (GeneCards)
MIR299	2,834184408	0,036838905	GO:0035195: Gene silencing by miRNA; GO:0035278: miRNA mediated inhibition of translation | Downregulated in gastric cancer (https://doi.org/10.1515/biol-2021-0022), nasopharyngeal carcinoma (https://pubmed.ncbi.nlm.nih.gov/31786874/), oral squamous cell (https://doi.org/10.1177/1533033819874803) and breast cancer (https://doi.org/10.3892/or.2020.7486)
MIR329-2	3,115574959	0,018919756	KEGG hsa05010: Alzheimer disease | GO:0035195: Gene silencing by miRNA; GO:0038061: NIK/NF-kappaB signaling; GO:0090051: Negative regulation of cell migration involved in sprouting angiogenesis; GO:1900747: Negative regulation of vascular endothelial growth factor signaling pathway
MIR654	3,448126688	0,007624028	GO:0035195: Gene silencing by miRNA | Downregulated in colorectal cancer; possible correlation with the mTOR signaling pathway (https://doi.org/10.3892/ol.2019.10508) and ovarian cancer (https://doi.org/10.2147/cmar.s229013)
AL023913.1	2,698339196	0,048685684	lncRNA; ENSG00000286936 (GeneCards)
RPS3P6	3,212386593	0,013092012	RPS3 pseudogene (GeneCards)
KB-1683C8.1	2,701263487	0,048940832	Ribosomal Protein S2 (RPS2) Pseudogene; ENST00000492365 (Ensembl)
RP11-53I6.1	2,944130903	0,027272586	lncRNA; ENSG00000263917 (Ensembl)
MIR889	2,887034168	0,032950776	Promotes cell proliferation in colorectal cancer (https://doi.org/10.26355/eurrev_201904_17695), non-small cell lung cancer (https://doi.org/10.1016/j.gene.2019.02.077) and esophageal squamous cell carcinomas (https://doi.org/10.1016/j.febslet.2015.03.027)
RP5-828K20.1	3,006313105	0,024667718	Processed pseudogene; ENSG00000225170 (Ensembl)
RP11-433O3.1	2,713532674	0,048123024	Processed pseudogene; ENSG00000225356 (Ensembl)
LINC00475	2,794346744	0,037262712	lincRNA (GeneCards) |Upregulated in glioma (https://doi.org/10.1111/jcmm.16100) and gastric cancer (https://doi.org/10.1002/jcla.22692)
AC026150.5	2,802164001	0,037436731	No information
AC006037.2	2,792525398	0,031825939	TEC; ENSG00000279218 (Ensembl) | Protein-tyrosine kinase (GeneCards) | GO:0002250: Adaptive immune response; GO:0006468: Protein phosphorylation; GO:0007229: I integrin-mediated signaling pathway; GO:0010543: Regulation of platelet activation | Reactome R-HSA-1280215.5: Cytokine signaling in immune system | KEGG hsa04660: T cell receptor signaling pathway
RP11-365D9.1	3,116084971	0,018316562	No information
VN1R51P	2,827433049	0,036593432	VN1R51 pseudogene (GeneCards)
GLUDP5	2,983266439	0,024069603	GLUD1 pseudogene (GeneCards)

LEF1-AS1	3,480239637	0,006688035	LEF1 antisense RNA (GeneCards) | Interacts with miRNAs and is involved in the development of colorectal cancer (https://doi.org/10.1016/j.bbrc.2020.03.153), glioblastoma (https://doi.org/10.1016/j.brainres.2020.146781), ovarian cancer (https://doi.org/10.2147/cmar.s227652), retinoblastoma (https://doi.org/10.1111/1440-1681.13263), among others
RP11-187C18.2	3,270399127	0,010449378	lincRNA; ENSG00000232116 (Ensembl)
RP11-154D17.1	3,446546609	0,00637598	lncRNA; Uncharacterized LOC101927502 (GeneCards)
U91328.2	2,712155474	0,041785349	lncRNA; ENSG00000272558 (GeneCards)
RP11-32F11.2	2,93011561	0,023948042	lincRNA; LINC01231 (GeneCards)
RP4-668G5.1	2,639814765	0,04827435	lncRNA; ENSG00000237928; NFIA-AS2 (GeneCards) | Overexpressed in gliomas (https://doi.org/10.1007/s13577-020-00408-9)
RP11-431J24.2	4,368201534	0,000256996	Antisense RNA; ENSG00000238178 (Ensembl) | Antisense To MAGEB17 (GeneCards)
RPS27AP1	2,885833945	0,031706594	RPS27A pseudogene (GeneCards)
RPL3P3	3,27247075	0,011472422	RPL3 pseudogene (GeneCards)
RP5-839B4.8	3,175676038	0,012707036	lincRNA; ENSG00000243961 (Ensembl)
CTB-73N10.1	2,947198491	0,028584338	Uncharacterized LOC105378211 (GeneCards)
RP11-184M15.1	2,75805408	0,04377416	lincRNA; ENSG00000248187 (Ensembl)
LINC00616	2,670312333	0,002078894	lincRNA (GeneCards) | Targets ABI2, CCDC85C, EMC10, FAM98B, MLYCD, PARVG, RP11-573D15.8, RP3-323A16.1, RPGR, SNX8 (https://doi.org/10.1186/s12958-020-00660-6)
RP13-644M16.5	2,659072378	0,035355164	UGT2B11 pseudogene; ENSG00000248763 (GeneCards)
RP11-798M19.3	3,04651749	0,021472365	Antisense RNA; ENSG00000272870 (Ensembl)
GYPB	2,744568968	0,044592269	GO:0050900: Leukocyte migration | KEGG hsa05144: Malaria | Reactome R-HSA-202733.5: Cell surface interactions at the vascular wall; Reactome R-HSA-109582.2: Hemostasis | Fusion transcript with SAV1 identified in acute myeloid leukemia (https://doi.org/10.3390/cancers11121951) | Upregulated in prostate cancer (https://doi.org/10.3390/ijms17020250)
RP11-436H11.3	3,078450947	0,017716768	lncRNA; High expression levels in renal cell carcinoma (https://doi.org/10.1186/s12943-017-0735-3)
RP11-427M20.1	2,867829748	0,032730799	lincRNA; ENST00000507844.1 (lncRNome)
RP11-724M22.1	3,056450945	0,019331773	lncRNA; ENST00000505408.1 (Ensembl)
RP11-501C14.7	2,960205459	0,025988622	CMC1 pseudogene; ENSG00000251550 (Ensembl)
SNORD112	4,128417672	0,000723825	snoRNA (GeneCards) | GO:0006396: RNA processing | Downregulated in acute leukemia (https://doi.org/10.1038/leu.2012.111)
SNORD112	3,553844072	0,005457504	snoRNA (GeneCards) | GO:0006396: RNA processing | Downregulated in acute leukemia (https://doi.org/10.1038/leu.2012.111)

RP11-150O12.1	3,307428764	0,009979028	lincRNA; ENSG00000253161 (Ensembl) | LINC01605 (GeneCards) | Promotes tumor cells proliferation, migration and metastasis and is upregulated in colorectal cancer (https://doi.org/10.26355/eurrev_202102_24837), laryngeal squamous cell carcinoma (https://doi.org/10.26355/eurrev_201912_19677) and bladder cancer (https://doi.org/10.1042/bsr20180562)
RP11-150O12.6	3,037830524	0,021065846	lncRNA; ENSG00000253414 (Ensembl) | LINC01605 (GeneCards) | Upregulated in esophageal squamous cell carcinoma (https://doi.org/10.1007/s13277-016-5227-3)
RP11-396O20.1	3,164459819	0,015420204	lincRNA; ENSG00000254695 (Ensembl)
ENPP7P8	2,685824095	0,045737148	ENPP7 pseudogene (GeneCards)
TRAV1-2	2,895280711	0,031961331	T cell receptor (GeneCards) | GO:0002250: Adaptive immune response; GO:0002376: Immune system process
AC005477.1	2,796302802	0,040521853	Protein coding; ENSG00000257040 (Ensembl)
RP11-76E16.2	2,767045771	0,040640466	IGF2BP2 processed pseudogene; ENSG00000257674 (Ensembl)
RP1-97G4.1	2,698262966	0,048685684	lncRNA; ENSG00000257835 (Ensembl)
RP11-18J9.3	2,873364936	0,032254938	lncRNA; ENSG00000258178 (Ensembl)
RP11-16B13.1	3,426768547	0,00747266	lncRNA; ENSG00000258553 (Ensembl)
RP11-242P2.1	2,769221651	0,041673015	Antisense RNA; ENSG00000259106 (Ensembl) | Antisense To NRXN3 (GeneCards)
RP11-142G1.2	2,71034093	0,043274787	lncRNA; Cancer Susceptibility 22 (CASC22); ENSG00000260887 (Ensembl)
RP11-395N3.1	3,73169697	0,003139851	lncRNA; ENSG00000261379 (Ensembl)
AL132709.3	3,751635519	0,002862181	snoRNA; ENSG00000201500 (Ensembl)
LINC00906	2,198823599	0,014755305	lincRNA (GeneCards)
CTD-2132N18.4	2,940584112	0,024072481	Zinc Finger Protein 385C (ZNF385C) (Gene Cards) | Involved in craniofacial development (ossification) (https://doi.org/10.1016/j.ydbio.2015.01.011)
CTC-203F4.2	2,710432515	0,046922297	No information
RP11-435O5.4	2,757894542	0,040449328	Antisense RNA; ENSG00000271659 (Ensembl) | Antisense To PTCH1 (GeneCards)
SNORD113-5	3,686997412	0,003551869	snoRNA (GeneCards) | GO:0006396: RNA processing

Capítulo VIII. Considerações Finais

Apresentando a menor taxa de mutação somática dentre os tumores sólidos, o hepatoblastoma é um câncer que ainda tem sua origem e mecanismos moleculares bastante debatidos na literatura científica. Há reconhecido link entre epigenética, desenvolvimento fetal e tumores pediátricos e trabalho anterior do grupo evidenciou hipometilação global de sequências não-repetitivas em hepatoblastomas em relação ao tecido hepático fetal e diferenciado[101], milhares de sítios diferencialmente metilados, resultando em desregulação de vias biológicas específicas, como diferenciação celular, desenvolvimento embrionário e metabolismo. Nos últimos anos, o estudo com foco na epigenética de hepatoblastomas ganhou maior importância como um mecanismo chave; o que foi proposto inicialmente de maneira pioneira por nosso grupo, recentemente foi revelado por grandes estudos internacionais como importante marcador de assinaturas de prognóstico tumoral.

Estruturamos esta tese objetivando compor um quadro acerca do papel e mecanismos de atuação da metilação de DNA e consequente alteração no padrão de expressão gênica na biologia dos hepatoblastomas. A abordagem foi baseada em um conjunto de 30 amostras de hepatoblastomas, além de tecido hepático não tumoral, coletados com apoio e colaboração de pesquisadores e médicos de três grandes centros de câncer infantil da cidade de São Paulo (AC Camargo Cancer Center, GRAAC e ITACI). Também utilizamos linhagens celulares hepáticas e células pluripotentes induzidas (IPSCs) diferenciadas em hepatoplastos e hepatócitos-*like*. O foco foi **(a)** identificar o mecanismo mais provável de hipometilação (desmetilação ativa ou passiva) pela caracterização do perfil de expressão de genes reguladores da metilação e hidroximetilação; **(b)** determinar o nível de 5-hidroximetilcitosina (5hmC) em comparação ao de 5-metilcitosina (5mC); **(c)** caracterizar as amostras de hepatoblastoma de acordo com as etapas de diferenciação celular de hepatócitos; **(d)** avaliar o impacto das alterações de metilação, previamente evidenciadas pelo grupo, nas vias biológicas de hepatoblastomas, pela análise do transcriptoma (RNA-seq) e de genes específicos.

Desta formal, mostramos neste estudo o perfil de expressão de genes associados a maquinaria de metilação em HBL (*TET1, TET2, TET3, DNMT1, DNMT3A, DNMT3B e DNMT3L*) [138], indicando que a hipometilação descrita em trabalho prévio provavelmente

ocorre de forma ativa, sob influência dos genes da família *TET* e *UHRF1*, com consequenteme aumento do seu produto, 5-hidroximetilcitonisa. Esse aumento de 5-hidroximetilcitosina em HBL foi associado a uma possível redução da sobrevida global neste trabalho, destacando que o conteúdo de 5hmC é biomarcador potencial de risco para HBL.

Sendo um tumor embrionário, a principal hipótese de origem do HBL preconiza a ocorrência de eventos que geram o bloqueio do processo de diferenciação normal de hepatócitos. Analisamos a expressão de 24 genes, 08 associados à maquinaria epigenética e 16 marcadores de diferenciação hepática, assim como o nível global de metilação dos tumores, para avaliar em qual fase da diferenciação de hepatócitos os hepatoblastomas devem se originar [202]. Considerando estes fatores, observamos que amostras tumorais podem ser agrupadas em três grupos, de acordo com o perfil de expressão do conjunto de genes estudados: similares a hepatoblastos, similares a hepatócitos e um terceiro grupo, intermediários entre os anteriores. Este resultado mostra que, de maneira geral, os hepatoblastomas são de fato tumores que, em sua maioria, são originários de estágios da diferenciação de hepatócitos, embora também revele que há heterogeneidade dentre os tumores. Este estudo também evidenciou que o grupo de tumores com perfil de expressão gênica mais diferenciado estaria associado a pior prognóstico, novamente apontando a importância da epigenética para classificação destes tumores, uma vez que dos 13 genes determinantes para a estraficação tumoral obtida, seis codificam proteínas relacionadas à metilação de DNA.

Inúmeros genes surgiram como candidatos após a análise dos dados de metilação de HBL de pacientes brasileiros, com evidência de hipermetilação em suas regiões promotoras [101]. Nesse cenário, o gene *NNMT* era uma proteína interessante a ser estudada, uma vez que é central para o metabolismo de hepatócitos e adipócitos. Havíamos identificado hipermetilação no promotor deste gene nos tumores e, com o aprofundamento do estudo, verificamos que ocorre redução de expressão e consequente diminuição no nível proteíco [142]. Para avaliar o impacto funcional da redução de NNMT, foi feita uma análise de metaboloma de algumas amostras de HBL, na qual foi revelado a redução de classes específicas de lipídios, possivelmente associada à ação de NNMT, o que resta ser provado por meio de análise funcional futura. Estes dados corroboraram também a alteração de vias metabólicas em hepatoblastomas, como inicialmente detectamos nos dados de metiloma.

Por fim, a análise do transcriptoma de HBLs, embora ainda em processo de finalização, claramente indica a existência de dois eixos centrais na biologia destes tumores: metabolismo disfuncional em relação a células hepáticas normais e regulação transcricional fortemente associada a genes ncRNAs. Em uma análise exploratória do transcriptoma, vias metabólicas reduzidas nos tumores são associadas a nicotinamidas e lipídios, dando suporte a nossos achados anteriores, assim como há aumento da atividade de vias relacionadas ao desenvolvimento embrionário, o que corrobora a hipótese de bloqueio/disfunção do processo de diferenciação celular como gênese do HBL. De maneira interessante, aprofundando o olhar para a rede de genes miRNA, foi possível detectar que os dois eixos centrais, metabolismo e ncRNA, estão conectados, com aumento de expressão de ncRNAs que atuam em vias metabólicas alteradas nos tumores, com destaque para quatro miRNAs (miR-186, miR-214, miR-377, and miR-494). Por fim, destacamos dois genes candidatos (*THRSP* e *GCK*) para futuras análises de validação funcional, também associados a metabolismo, o primeiro deles com detecção prévia de possível regulado por metilação [101]. Mais análises serão realizadas buscando explorar o papel da classe de lncRNAs no desenvolvimento de hepatoblastomas.

Em resumo, o presente trabalho evidenciou a importância de mecanismos epigenéticos, como metilação de DNA e RNAs não codificadores, na gênese e desenvolvimento do hepatoblastoma. Adicionalmente, diversas linhas de evidência de nossos dados sugerem que alterações metabólicas são frequentes nestes tumroes, em especial, redução de lipídios. Os achados são inéditos e trazem novas perguntas a serem exploradas.

Capítulo IV. Resumo

Tumores hepáticos em crianças são raros, representando 1 - 4% de todos os tumores sólidos pediátricos, sendo o hepatoblastoma o mais frequente. Hepatoblastomas são tumores embrionários, com idade muito precoce de manifestação e apresentação histológica característica, que recapitula diferentes fases da organogênese do fígado. Em decorrência de sua raridade, os mecanismos moleculares subjacentes ao desenvolvimento e progressão deste tipo tumoral ainda são pouco explorados e praticamente inexistem biomarcadores para estratificação de risco. Apresentando a menor taxa de mutação somática dentre os tumores sólidos, o hepatoblastoma é um câncer que ainda tem sua origem e mecanismos moleculares debatidos na literatura científica. Há reconhecido link entre epigenética, desenvolvimento fetal e tumores pediátricos e, nos últimos anos, o estudo com foco na epigenética de hepatoblastomas ganhou maior importância como um mecanismo chave.

Este estudo teve como objetivo central a avaliação do papel da metilação de DNA na alteração do padrão de expressão gênica de hepatoblastomas. Foram avaliadas no total um conjunto de 30 amostras de hepatoblastomas e 11 amostras de tecido hepático não tumoral, derivadas de colaboração com três centros de câncer infantil da cidade de São Paulo (AC Camargo Cancer Center, GRAAC e ITACI). Também foram utilizadas linhagens celulares de tumores hepáticos e células pluripotentes induzidas (IPSCs) diferenciadas em hepatoblastos e hepatócitos-like. Utilizando técnicas como qRT-PCR, western blot, imunohistoquímica, metabolômica, RNASeq e nível global de hidroximetilação, os objetivos foram (a) identificar o mecanismo de hipometilação de DNA em hepatoblastomas e determinar o nível de 5-hidroximetilcitosina; (b) caracterizar as amostras de hepatoblastoma de acordo com as etapas de diferenciação celular de hepatócitos; (c) avaliar o impacto de alterações de metilação previamente identificadas no transcriptoma tumoral (genes codificadores e não codificadores).

Estudamos o padrão de expressão de genes da maquinaria de metilação de DNA (*DNMT1, DNMT3A, DNMT3B, DNMT3L, UHRF1, TET1, TET2* e *TET3*) nestes tumores e o nível global de 5hmC. Detectamos elevada expressão dos genes da maquinaria epigenética, principalmente *UHRF1*, *TET1* e *TET2*, em associação com enriquecimento do conteúdo de 5hmC. Estes resultados indicam a ocorrência em hepatoblastomas de desmetilação ativa do DNA, mediada por TETs, provavelmente durante os estágios iniciais do desenvolvimento do

fígado. Além disso, níveis mais baixos de 5hmC em hepatoblastomas podem estar associados à redução da sobrevida global.

Em seguida, avaliamos a utilização do padrão de expressão de genes da maquinaria epigenética e de diferenciação hepática como elementos-chave para a estratificação de hepatoblastomas. Foi possível a estratificação das amostras tumorais em três grupos, destacando a importância de 13 genes (*TET1, TET2, TET3, DNMT1, DNMT3A, UHRF1, ALB, CYP3A4, TDO2, UGT1A1, AFP, HNF4A* e *FOXA2*) para o estabelecimento desta classificação. Os resultados permitiram correlacionar os grupos identificados com distintos níveis de metilação global de DNA, que estão de acordo com a fase de diferenciação hepática em que as células tumorais se encontrariam, o que corrobora a hipótese de bloqueio de diferenciação do órgão de origem. Deste modo, propusemos assinaturas tumorais associadas a diferentes fases de diferenciação hepática com base em expressão gênica e em nível de metilação de DNA.

Evidenciamos redução significativa de expressão gênica e proteica de NNMT em hepatoblastomas. Níveis mais elevados de NNMT foram estatisticamente associados ao diagnóstico tardio em hepatoblastoma, uma reconhecida variável clínica de pior prognóstico. A análise metabolômica não direcionada detectou metabolismo lipídico aberrante em hepatoblastomas, sugerindo que a diminuição de NNMT pode estar associada à redução detectada do conteúdo lipídico em amostras de hepatoblastoma, possivelmente relacionada a concentrações alteradas de metabólicos ligados a função do NNMT ou precursores de NAD+. Os dados apresentados aqui mostraram pela primeira vez que a redução de NNMT ocorre em hepatoblastomas, fornecendo apoio adicional à teoria de que sua regulação negativa é um fenômeno geral em câncer de fígado.

Com o objetivo de compreender melhor os processos biológicos alterados, realizamos análises de bioinformática a partir de dados de RNASeq de hepatoblastomas e amostras de fígado normais. Foi identificado um grupo de 1492 genes com expressão diferencial entre tumores e amostras controle (Differentially Expressed Genes - DEGs), 1.031 deles regulados positivamente e 461 com diminuição de expressão. A lista de DEGs contém genes codificadores de proteínas e ncRNAs, estes principalmente regulados positivamente. Oitenta e cinco dos genes com expressão diferencial são expressos exclusivamente em amostras tumorais, sendo aproximadamente 55% deles ncRNAs. Dentre os 50 principais genes, aqui definidos como aqueles com a maior alteração de expressão, 42% correspondem a ncRNAs; três (*MAGED4*, *SNO144-16* e *RP11-431J24.2*) com expressão detectável exclusivamente em

tumores. Adicionalmente, 16 destes genes com maior diferença de expressão entre tumores e fígado normal já haviam sido identificados anteriormente como alterados neste grupo tumoral (*GCK, HMGA2, INS-IGF2, DKK1, CPA6, TTC36, LIX1, PART1, DKK4, CDCA7, GNG4, THRSP, CNDP1, CYP2B7P, ASPG* e *HAO2*). Os processos biológicos foram enriquecidos em regulação negativa do metabolismo, principalmente de amina, nicotinamida e lipídios/ácidos graxos; também há alteração de mecanismos epigenéticos, como metilação do DNA, ncRNAs e silenciamento de genes por miRNA; citotoxicidade e vias de sinalização, incluindo mediada por citocinas. Uma rede de interação miRNA-mRNA foi construída, identificando quatro genes (*miR-186, miR-214, miR-377* e *miR-494*), cada um deles conectado a mais de 10 mRNAs. Os principais processos biológicos relacionados a esta rede miRNA-mRNA foram associados também a metabolismo e reações de oxidação de lipídios e carboidratos. Em conclusão, a análise de RNASeq evidenciou grande alteração de vias metabólicas em hepatoblastomas, incluindo lipídios, aminas e nicotinamidas, o que retoma resultados anteriores de nosso grupo. Também revelou grande modificação na rede de expressão de ncRNAs neste tipo de tumores.

A detecção de processos biológicos alterados fornece pistas para o entendimento das vias de tumorigênese em hepatoblastomas e os achados de genes centrais indicam candidatos a futuras análises funcionais.

Capítulo X. *Abstract*

Liver tumors in children are rare, representing 1 - 4% of all pediatric solid tumors, with hepatoblastoma being the most frequent. Hepatoblastoma is an embryonic tumor with a very early age of manifestation and characteristic histological presentation that recapitulates different phases of liver organogenesis. Due to its rarity, the molecular mechanisms underlying the development and progression of this tumor type are still poorly explored, and there are few biomarkers for risk stratification. With the lowest rate of somatic mutation among solid tumors, the origin and molecular mechanisms of hepatoblastoma are still discussed in the literature. There is a recognized link between epigenetics, fetal development, and pediatric tumors and, in recent years, the study focused on the epigenetics of hepatoblastomas has gained greater importance as a key mechanism.

This study had as its central objective the evaluation of the role of DNA methylation in altering the pattern of gene expression of hepatoblastomas. A total of 30 samples of hepatoblastomas and 11 samples of non-tumor liver tissue were evaluated, derived from collaboration with three pediatric cancer centers in São Paulo (AC Camargo Cancer Center, GRAAC, and ITACI). Liver tumor cell lines and induced pluripotent cells (IPSCs) differentiated into hepatoblasts and hepatocytes-like were also used. Using techniques such as qRT-PCR, western blot, immunohistochemistry, metabolomics, RNASeq, and global level of hydroxymethylation, the objectives were (a) to identify the mechanism of DNA hypomethylation in hepatoblastomas and to determine the level of 5-hydroxymethylcytosine; (b) characterize the samples of hepatoblastoma according to the steps of cell differentiation from hepatocytes; (c) assess the impact of previously identified changes in the tumor methylome on the total transcriptome (coding and non-coding genes).

We studied the gene expression pattern of DNA methylation machinery (*DNMT1*, *DNMT3A*, *DNMT3B*, *DNMT3L*, *UHRF1*, *TET1*, *TET2,* and *TET3*) in these tumors and investigated the overall level of 5hmC in hepatoblastoma. We detected high expression of epigenetic machinery genes, mainly *UHRF1, TET1*, and *TET2*, in association with enrichment of the content of 5hmC. These results indicate the occurrence of active DNA demethylation in hepatoblastomas, mediated by TETs, probably during the early stages of liver development. In addition, lower levels of 5hmC in hepatoblastomas may be associated with reduced overall survival.

In addition, we evaluated the use of the gene expression pattern of epigenetic machinery and liver differentiation as pivotal elements for hepatoblastoma stratification. It was possible to stratify tumor samples into three groups, highlighting the importance of 13 (*TET1, TET2, TET3, DNMT1, DNMT3A, UHRF1, ALB, CYP3A4, TDO2, UGT1A1, AFP, HNF4A,* and *FOXA2*) genes for the establishment of this classification. The results made it possible to correlate the identified groups with different levels of global DNA methylation, which are in accordance with the phase of liver differentiation in which the tumor cells would meet, which corroborates the hypothesis of blocking the differentiation of the source organ. Thus, we proposed tumor signatures associated with different stages of liver differentiation based on gene expression and DNA methylation level.

We evidenced a significant reduction in NNMT gene and protein expression in hepatoblastomas. The higher expression of NNMT was statistically associated with a late diagnosis of hepatoblastoma, a recognized clinical variable with a worse prognosis. The undirected metabolomic analysis detected aberrant lipid metabolism in hepatoblastoma, suggesting that the decrease in NNMT may be associated with the detected reduction in lipid content in tumor samples, possibly related to altered concentrations of metabolic linked to NNMT function or NAD+ 123 precursors. The data presented here showed for the first time that the reduction of NNMT occurs in hepatoblastomas, providing additional support for the theory that the negative regulation of this protein is a general phenomenon in liver cancer.

To better understand the altered biological processes, we performed bioinformatics analyzes based on data from RNASeq of normal hepatoblastomas and liver. A group of 1492 genes with differential expression between tumors and control (Differentially Expressed Genes - DEGs) was identified, 1,031 of them regulated positively and 461 with decreased expression. A list of DEGs contains genes encoding proteins and ncRNAs, which are mainly upregulated. Eighty-five of the genes with differential expression are expressed exclusively in our tumors, with approximately 55% of them being ncRNAs. Among the 50 main genes, defined here as those with the greatest expression alteration, 42% correspond to ncRNAs; three (*MAGED4*, *SNO144-16* and *RP11-431J24.2*) had detectable expression exclusively in tumors and, in addition, 16 of these genes with greater difference in expression between tumors and normal liver had previously been identified as altered in this tumor group (*GCK, HMGA2, INS-IGF2*, *DKK1, CPA6, TTC36, LIX1, PART1, DKK4, CDCA7, GNG4, THRSP, CNDP1, CYP2B7P*, *ASPG* and *HAO2*). The biological processes were enriched in the sense of negative regulation of metabolism, mainly of amine, nicotinamide and lipids/fatty acids; there is also a change

in epigenetic mechanisms, such as DNA methylation, ncRNAs and gene silencing by miRNA; cytotoxicity and signaling pathways, including cytokine-mediated signaling pathway. A miRNA-mRNA interaction network was built, identifying four genes (*miR-186*, *miR-214*, *miR-377* and *miR-494*) where each is connected to more than 10 mRNAs. The main biological processes related to this miRNA-mRNA network were also associated with metabolism and oxidation reactions of lipids and carbohydrates. In conclusion, the analysis of RNASeq showed great alteration of metabolic pathways in hepatoblastomas, including lipids, amines and nicotinamides, which resumes previous results of our group. It also revealed major changes in the ncRNA expression network in this type of tumor.

The detection of altered biological processes provides clues for the understanding of tumorigenesis pathways in hepatoblastoma and the findings of key genes indicate candidates for future functional analyzes.

Capítulo XI. Biografia

- Adriano Bonaldi, André Y. Kashiwabara, Érica S.S. de Araújo, Lygia V. Pereira, Alexandre R. Paschoal, Mayra B. Andozia, Darine Villela, **Maria P. Rivas**, Claudia K. Suemoto, Carlos A. Pasqualucci, Lea T. Grinberg, Helena P. Brentani, Silvya S. Maria-Engler, Dirce M. Carraro, Angela M. Vianna-Morgante, Carla Rosenberg, Ana Cristina Krepischi, Luciana R. Vasques; MINING NOVEL CANDIDATE IMPRINTED GENES USING GENOME-WIDE METHYLATION SCREENING. *Epigenomes*, (2017) 1(2), 13; https://doi.org/10.3390/epigenomes1020013

- Dimitrius T. Pramio, André Y. Kashiwabara, Paula C. Pennacchi, **Maria P. Rivas**, Silvya S. Maria-Engler, Antônio H. J. F. M. Campos, João P. Duprat, Dirce M. Carraro, Ana Cristina Krepischi. EPIGENETIC SIGNATURE OF DIFFERENTIALLY METHYLATED GENES IN CUTANEOUS MELANOMA. *Applied Cancer Res*, (2017) 37: 34. https://doi.org/10.1186/s41241-017-0041-9

- **Maria P. Rivas**, Anne C. Teixeira, Ana Cristina Krepischi. EPIGENÉTICA: CONCEITO, MECANISMOS E IMPACTO EM DOENÇAS HUMANAS. *Revista Genética na Escola*, (2018) v. 14, p. 14-25.

- **Maria P. Rivas**, Talita M. Aguiar, Gustavo R. Fernandes, Luiz C. Caires-Júnior, Ernesto Goulart, Kayque A. Telles-Silva, Monica Cypriano, Silvia R. Toledo, Carla Rosenberg, Dirce M. Carraro, Cecilia M. da Costa, Isabela W. da Cunha, Ana C. Krepischi. TET GENES UPREGULATION LEADS TO 5-HYDROXYMETHYLATION ENRICHMENT IN HEPATOBLASTOMA. *Frontiers Genetics*, (2019) 10: 553. doi: 10.3389/fgene.2019.00553

- **Maria P. Rivas**, Sara F. Pires, Talita M. Aguiar. CRIANÇA TAMBÉM TEM CÂNCER? *Revista Genética na Escola*, (2019) v. 14, p. 98-107.

- Talita M. Aguiar, **Maria P. Rivas**, Silvia Souza, Tatiane Rodrigues, Juliana Sobral, Anne C. Barbosa, Mariana Maschietto, Renan Valieris, Gustavo R. Fernandes, Monica Cypriano, Silvia R. Toledo, Israel Tojal, Maria de Pinho, Dirce M. Carraro, Carla Rosenberg, Cecilia M. da Costa, Isabela W. da Cunha, Ana C. Krepischi. MUTATIONAL BURDEN OF HEPATOBLASTOMAS: A ROLE FOR THE CX3CL1/CX3CR1 CHEMOKINE SIGNALING PATHWAY. *Frontiers Oncology*, (2020): 10.3389/fonc.2020.00556.

- **Maria Prates Rivas**, Talita Marques Aguiar, Mariana Maschietto, Luiz Carlos Caires-Júnior, Ernesto Goulart, Kayque Alves Telles-Silva, Estela Novak, Lilian Maria Cristofani,

Vicente Odone, Monica Cypriano, Silvia Regina Caminada de Toledo, Dirce Maria Carraro, Melissa Quintero Escobar, Hana Leeh, Michael Johnston, Cecilia Maria Lima da Costa, Isabela Werneck da Cunha, Ljubica Tasic, Peter L. Pearson, Carla Rosenberg, Nikolai Timchenko, Ana Cristina Victorino Krepischi. HEPATOBLASTOMAS EXHIBIT MARKED NNMT DOWN-REGULATION DRIVEN BY PROMOTER DNA HYPERMETHYLATION. *Tumor Biology,* (2020): 10.1177/1010428320977124

- Amber D'Souza, Ashley Cast, Meenasri Kumbaji, **Maria Rivas**, Ruhi Gulati, Michael Johnston, David Smithrud, James Geller and Nikolai Timchenko. SMALL MOLECULE CJOC42 IMPROVES CHEMO-SENSITIVITY AND INCREASES LEVELS OF TUMOR SUPPRESSOR PROTEINS IN MICE BY INHIBITING ONCOGENE GANKYRIN. *Frontiers in Pharmacology,* (2021): 10.3389/fphar.2021.580722

- **Maria Rivas**, Talita Aguiar, Gustavo Fernandes, Renan Lemes, Luiz Caires-Júnior, Ernesto Goulart, Kayque Telles-Silva, Mariana Maschietto, Monica Cypriano, Silvia de Toledo, Dirce Carraro, Isabela da Cunha, Cecilia da Costa, Carla Rosenberg, Ana Krepischi. DNA METHYLATION AS A KEY EPIGENETICS PLAYER FOR HEPATOBLASTOMA CHARACTERIZATION. *Clinics and Research in Hepatology and Gastroenterology* (2021): 10.1016/j.clinre.2021.101684

- Michael Edward Johnston, **Maria Prates Rivas**, Delphine Nicolle, Aurore Gorse, Ruhi Gulati, Meena Kumbaji, Matthew T. Weirauch, Alexander Bondoc, Stefano Cairo, James Geller, Gregory Tiao, Nikolai Timchenko. OLAPARIB INHIBITS TUMOR GROWTH OF HEPATOBLASTOMA IN PATIENT DERIVED XENOGRAFT MODELS. *Artigo em revisão na Hepatology*.

- **Maria Rivas**, Michael Johnston, Ruhi Gulati, Talita Aguiar, Meenasri Kumbaji, Ashley Cast, Ana Krepischi, Alexander Bondoc, Gregory Tiao, James Geller, Nikolai Timchenko. HDAC1-DEPENDENT REPRESSION OF MARKERS OF HEPATOCYTES AND P21 IS INVOLVED IN DEVELOPMENT OF PEDIATRIC LIVER CANCER. *Artido em revisão na Cellular and Molecular Gastroenterology and Hepatology*.

Referências bibliográficas

1. Feliciano SVM, Santos M de O, Pombo-de-Oliveira MS. Incidência e Mortalidade por Câncer entre Crianças e Adolescentes: uma Revisão Narrativa. *Rev Bras Cancerol*. Epub ahead of print 2019. DOI: 10.32635/2176-9745.rbc.2018v64n3.45.
2. INCA. Estimativa 2020: incidência de câncer no Brasil | INCA - Instituto Nacional de Câncer José Alencar Gomes da Silva. *INCA*.
3. Santos M de O. 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. *Rev Bras Cancerol*. Epub ahead of print 2019. DOI: 10.32635/2176-9745.rbc.2018v64n3.56.
4. Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. *Science*. Epub ahead of print 2013. DOI: 10.1126/science.1235122.
5. Anderson LM. Environmental genotoxicants/carcinogens and childhood cancer: bridgeable gaps in scientific knowledge. *Mutat Res* 2006; 608: 136–156.
6. Ma X, Liu Y, Liu Y, et al. Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature* 2018; 555: 371–376.
7. Maris JM, Denny CT. Focus on embryonal malignancies. *Cancer Cell* 2002; 2: 447–450.
8. Kim EH, Koh KN, Park M, et al. Clinical features of infantile hepatic hemangioendothelioma. *Korean J Pediatr*. Epub ahead of print 2011. DOI: 10.3345/kjp.2011.54.6.260.
9. Pahlman S, Stockhausen M-T, Fredlund E, et al. Notch signaling in neuroblastoma. *Semin Cancer Biol* 2004; 14: 365–373.
10. Zorn. Liver Development. *StemBook*. Epub ahead of print 2008. DOI: 10.3824/stembook.1.25.1.
11. Zhao R, Duncan SA. Embryonic development of the liver. *Hepatology* 2005; 41: 956–967.
12. Si-Tayeb K, Lemaigre FP, Duncan SA. Organogenesis and Development of the Liver. *Dev Cell* 2010; 18: 175–189.
13. Houssaint E. Differentiation of the mouse hepatic primordium. I. An analysis of tissue interactions in hepatocyte differentiation. *Cell Differ*. Epub ahead of print 1980. DOI: 10.1016/0045-6039(80)90026-3.

14. Medlock ES, Haar JL. The liver hemopoietic environment: I. Developing hepatocytes and their role in fetal hemopoiesis. *Anat Rec*. Epub ahead of print 1983. DOI: 10.1002/ar.1092070105.
15. Clotman F, Jacquemin P, Plumb-Rudewiez N, et al. Control of liver cell fate decision by a gradient of TGF beta signaling modulated by Onecut transcription factors. *Genes Dev* 2005; 19: 1849–1854.
16. Tan X, Apte U, Micsenyi A, et al. Epidermal growth factor receptor: a novel target of the Wnt/beta-catenin pathway in liver. *Gastroenterology* 2005; 129: 285–302.
17. Tan X, Yuan Y, Zeng G, et al. β-catenin deletion in hepatoblasts disrupts hepatic morphogenesis and survival during mouse development. *Hepatology*. Epub ahead of print 2008. DOI: 10.1002/hep.22225.
18. Michalopoulos GK. Liver regeneration. *Journal of Cellular Physiology*. Epub ahead of print 2007. DOI: 10.1002/jcp.21172.
19. Messerschmidt DM, Knowles BB, Solter D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev* 2014; 28: 812–828.
20. Mikkelsen TS, Ku M, Jaffe DB, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*. Epub ahead of print 2007. DOI: 10.1038/nature06008.
21. Luks FI, Yazbeck S, Brandt ML, et al. Benign liver tumors in children: A 25-year experience. *J Pediatr Surg*. Epub ahead of print 1991. DOI: 10.1016/0022-3468(91)90612-W.
22. Julio C. Wiederkehr, Izabel M. Coelho, Sylvio G. Avilla BAW and HAW. Liver tumors in infancy. In: *Hepatic Surgery*. 2012, pp. 423–459.
23. López-Terrada D, Alaggio R, De Dávila MT, et al. Towards an international pediatric liver tumor consensus classification: Proceedings of the Los Angeles COG liver tumors symposium. *Mod Pathol* 2014; 27: 472–491.
24. Oue T, Kubota A, Okuyama H, et al. Hepatoblastoma in children of extremely low birth weight: A report from a single perinatal center. In: *Journal of Pediatric Surgery*. 2003. Epub ahead of print 2003. DOI: 10.1053/jpsu.2003.50027.
25. Teplick A, Kowalski M, Biegel JA, et al. Educational paper: screening in cancer predisposition syndromes: guidelines for the general pediatrician. *Eur J Pediatr* 2011; 170: 285–294.
26. Ranganathan S, Lopez-Terrada D, Alaggio R. Hepatoblastoma and Pediatric

Hepatocellular Carcinoma: An Update. *Pediatric and Developmental Pathology*. Epub ahead of print 2019. DOI: 10.1177/1093526619875228.

27. Hermann RE, Lonsdale D. Chemotherapy, radiotherapy, and hepatic lobectomy for hepatoblastoma in an infant: report of a survival. *Surgery*. Epub ahead of print 1970. DOI: 10.1016/s0022-3468(70)80081-1.
28. Germain T, Favelier S, Cercueil JP, et al. Liver segmentation: Practical tips. *Diagnostic and Interventional Imaging*. Epub ahead of print 2014. DOI: 10.1016/j.diii.2013.11.004.
29. Triviño T, Abib S de CV. Anatomia cirúrgica do fígado. *Acta Cir Bras*. Epub ahead of print 2003. DOI: 10.1590/s0102-86502003000500006.
30. Pritchard J, Brown J, Shafford E, et al. Cisplatin, doxorubicin, and delayed surgery for childhood hepatoblastoma: A successful approach - Results of the first prospective study of the International Society of Pediatric Oncology. *J Clin Oncol* 2000; 18: 3819–3828.
31. Meyers RL, Czauderna P, Otte JB. Surgical treatment of hepatoblastoma. *Pediatric Blood and Cancer*. Epub ahead of print 2012. DOI: 10.1002/pbc.24220.
32. Sidebotham E. Abdominal and pelvic tumours in children. *Surgery (United Kingdom)*. Epub ahead of print 2019. DOI: 10.1016/j.mpsur.2019.02.003.
33. Czauderna P, Haeberle B, Hiyama E, et al. The Children's Hepatic tumors International Collaboration (CHIC): Novel global rare tumor database yields new prognostic factors in hepatoblastoma and becomes a research model. *Eur J Cancer* 2016; 52: 92–101.
34. Aronson DC, Czauderna P, Maibach R, et al. The treatment of hepatoblastoma: Its evolution and the current status as per the SIOPEL trials. *Journal of Indian Association of Pediatric Surgeons*. Epub ahead of print 2014. DOI: 10.4103/0971-9261.142001.
35. Meyers RL, Maibach R, Hiyama E, et al. Risk-stratified staging in paediatric hepatoblastoma: a unified analysis from the Children's Hepatic tumors International Collaboration. *Lancet Oncol* 2017; 18: 122–131.
36. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature*. Epub ahead of print 2009. DOI: 10.1038/nature07943.
37. Ding L, Ellis MJ, Li S, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature*. Epub ahead of print 2010. DOI: 10.1038/nature08989.

38. Kan Z, Jaiswal BS, Seshagiri S. Diverse somatic mutation patterns and pathway alterations in human cancers. *Genome Biol*. Epub ahead of print 2010. DOI: 10.1186/gb-2010-11-s1-p37.
39. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. *Nature Reviews Genetics*. Epub ahead of print 2010. DOI: 10.1038/nrg2841.
40. Zhang J, Walsh MF, Wu G, et al. Germline Mutations in Predisposition Genes in Pediatric Cancer. *N Engl J Med* 2015; 373: 2336–2346.
41. Eichenmüller M, Trippel F, Kreuder M, et al. The genomic landscape of hepatoblastoma and their progenies with HCC-like features. *J Hepatol* 2014; 61: 1312–1320.
42. Tomlinson GE, Kappler R. Genetics and epigenetics of hepatoblastoma. *Pediatric Blood and Cancer*. Epub ahead of print 2012. DOI: 10.1002/pbc.24213.
43. Gröbner SN, Worst BC, Weischenfeldt J, et al. The landscape of genomic alterations across childhood cancers. *Nature*. Epub ahead of print 2018. DOI: 10.1038/nature25480.
44. Rodrigues TC, Fidalgo F, Da Costa CML, et al. Upregulated genes at 2q24 gains as candidate oncogenes in hepatoblastomas. *Futur Oncol*. Epub ahead of print 2014. DOI: 10.2217/fon.14.149.
45. Barros JS de. *Alterações de número de cópias genômicas em hepatoblastomas: microarranjos genômicos e sequenciamento de nova geração*. Universidade de São Paulo, 2019.
46. Armengol C, Cairo S, Fabre M, et al. Wnt signaling and hepatocarcinogenesis: The hepatoblastoma model. *Int J Biochem Cell Biol* 2011; 43: 265–270.
47. Ozen C, Yildiz G, Dagcan AT, et al. Genetics and epigenetics of liver cancer. *N Biotechnol*. Epub ahead of print 2013. DOI: 10.1016/j.nbt.2013.01.007.
48. Aguiar TFM. *ESTUDO GENÔMICO DO TUMOR EMBRIONÁRIO HEPATOBLASTOMA*. Fundação Antônio Prudente, 2019.
49. Aguiar TFM, Rivas MP, Costa S, et al. Insights Into the Somatic Mutation Burden of Hepatoblastomas From Brazilian Patients. *Front Oncol* 2020; 10: 556.
50. Cairo S, Armengol C, Maibach R, et al. A combined clinical and biological risk classification improves prediction of outcome in hepatoblastoma patients. *Eur J Cancer*. Epub ahead of print 2020. DOI: 10.1016/j.ejca.2020.09.026.
51. Buendia MA. Unravelling the genetics of hepatoblastoma: Few mutations, what

else? *J Hepatol* 2014; 61: 1202–1204.

52. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. *Nature Reviews Genetics*. Epub ahead of print 2019. DOI: 10.1038/s41576-019-0150-2.
53. Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics*. Epub ahead of print 2009. DOI: 10.1038/nrg2484.
54. Ozsolak F, Milos PM. RNA sequencing: Advances, challenges and opportunities. *Nature Reviews Genetics*. Epub ahead of print 2011. DOI: 10.1038/nrg2934.
55. Hooks KB, Audoux J, Fazli H, et al. New insights into diagnosis and therapeutic options for proliferative hepatoblastoma. *Hepatology*. Epub ahead of print 2018. DOI: 10.1002/hep.29672.
56. Hong M, Tao S, Zhang L, et al. RNA sequencing: new technologies and applications in cancer research. *Journal of Hematology and Oncology*. Epub ahead of print 2020. DOI: 10.1186/s13045-020-01005-x.
57. Carrillo-Reixach J, Torrens L, Simon-Coma M, et al. Epigenetic footprint enables molecular risk stratification of hepatoblastoma with clinical implications. *J Hepatol*. Epub ahead of print March 2020. DOI: 10.1016/j.jhep.2020.03.025.
58. Sekiguchi M, Seki M, Kawai T, et al. Integrated multiomics analysis of hepatoblastoma unravels its heterogeneity and provides novel druggable targets. *npj Precis Oncol*. Epub ahead of print 2020. DOI: 10.1038/s41698-020-0125-y.
59. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. *Carcinogenesis* 2009; 31: 27–36.
60. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. *Nat Rev Genet* 2016; 17: 487–500.
61. Tchurikov NA. Molecular mechanisms of epigenetics. *Biochemistry (Mosc)* 2005; 70: 406–423.
62. Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. *Nature Reviews Genetics*. Epub ahead of print 2006. DOI: 10.1038/nrg1748.
63. Sandoval, J., Esteller M. Cancer epigenomics: beyond genomics. *Curr Opin Genet Dev* 2012; 22: 50–5.
64. You JS, Jones PA. Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell* 2012; 22: 9–20.
65. Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. *Science* 1975; 187: 226–232.
66. Das PM, Singal R. DNA methylation and cancer. *Journal of Clinical Oncology*.

Epub ahead of print 2004. DOI: 10.1200/JCO.2004.07.151.

67. Lister R, Ecker JR. Finding the fifth base: Genome-wide sequencing of cytosine methylation. *Genome Research*. Epub ahead of print 2009. DOI: 10.1101/gr.083451.108.
68. Turek-Plewa J, Jagodziński PP. The role of mammalian DNA methyltransferases in the regulation of gene expression. *Cell Mol Biol Lett*.
69. Jones PA. The DNA methylation paradox. *Trends Genet* 1999; 15: 34–37.
70. Barau J, Teissandier A, Zamudio N, et al. The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science (80-)*. Epub ahead of print 2016. DOI: 10.1126/science.aah5143.
71. Goll MG, Kirpekar F, Maggert KA, et al. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science (80-)*. Epub ahead of print 2006. DOI: 10.1126/science.1120976.
72. Jeltsch A, Ehrenhofer-Murray A, Jurkowski TP, et al. Mechanism and biological role of Dnmt2 in Nucleic Acid Methylation. *RNA Biology*. Epub ahead of print 2017. DOI: 10.1080/15476286.2016.1191737.
73. Neri F, Krepelova A, Incarnato D, et al. XDnmt3L antagonizes DNA methylation at bivalent promoters and favors DNA methylation at gene bodies in ESCs. *Cell* 2013; 155: 121.
74. Jang HS, Shin WJ, Lee JE, et al. CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function. *Genes (Basel)*; 8. Epub ahead of print May 2017. DOI: 10.3390/genes8060148.
75. Ramsahoye BH, Biniszkiewicz D, Lyko F, et al. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A* 2000; 97: 5237–5242.
76. Laurent L, Wong E, Li G, et al. Dynamic changes in the human methylome during differentiation. *Genome Res* 2010; 20: 320–331.
77. Lienert F, Wirbelauer C, Som I, et al. Identification of genetic elements that autonomously determine DNA methylation states. *Nat Genet*. Epub ahead of print 2011. DOI: 10.1038/ng.946.
78. Jones PA. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*. Epub ahead of print 2012. DOI: 10.1038/nrg3230.
79. Cokus SJ, Feng S, Zhang X, et al. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature*. Epub ahead of print 2008.

DOI: 10.1038/nature06745.

80. Lister R, Pelizzola M, Dowen RH, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. Epub ahead of print 2009. DOI: 10.1038/nature08514.
81. Rodríguez-Paredes M, Esteller M. Cancer epigenetics reaches mainstream oncology. *Nature Medicine*. Epub ahead of print 2011. DOI: 10.1038/nm.2305.
82. Łuczak MW, Jagodziński PP. The role of DNA methylation in cancer development. *Folia Histochemica et Cytobiologica*. Epub ahead of print 2006. DOI: 10.5603/4561.
83. Lopez-Serra P, Esteller M. DNA methylation-associated silencing of tumor-suppressor microRNAs in cancer. *Oncogene* 2012; 31: 1609–1622.
84. Schofield MJ, Hsieh P. DNA mismatch repair: molecular mechanisms and biological function. *Annu Rev Microbiol* 2003; 57: 579–608.
85. Münzel M, Globisch D, Carell T. 5-hydroxymethylcytosine, the sixth base of the genome. *Angewandte Chemie - International Edition*. Epub ahead of print 2011. DOI: 10.1002/anie.201101547.
86. Mellen M, Ayata P, Dewell S, et al. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* 2012; 151: 1417–1430.
87. Wu H, Zhang Y. Mechanisms and functions of Tet proteinmediated 5-methylcytosine oxidation. *Genes and Development*. Epub ahead of print 2011. DOI: 10.1101/gad.179184.111.
88. Wu X, Zhang Y. TET-mediated active DNA demethylation: Mechanism, function and beyond. *Nature Reviews Genetics*. Epub ahead of print 2017. DOI: 10.1038/nrg.2017.33.
89. An J, Rao A, Ko M. TET family dioxygenases and DNA demethylation in stem cells and cancers. *Experimental and Molecular Medicine*. Epub ahead of print 2017. DOI: 10.1038/emm.2017.5.
90. Pfeifer GP, Kadam S, Jin SG. 5-Hydroxymethylcytosine and Its Potential Roles in Development and Cancer. *Epigenetics and Chromatin* 2013; 6: 1–9.
91. Ito S, Dalessio AC, Taranova O V., et al. Role of tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 2010; 466: 1129–1133.
92. Ecsedi S, Rodríguez-Aguilera J, Hernandez-Vargas H. 5-Hydroxymethylcytosine (5hmC), or How to Identify Your Favorite Cell. *Epigenomes*. Epub ahead of print 2018. DOI: 10.3390/epigenomes2010003.

93. Ficz G, Gribben JG. Loss of 5-hydroxymethylcytosine in cancer: Cause or consequence? *Genomics* 2014; 104: 352–357.
94. Kudo Y, Tateishi K, Yamamoto K, et al. Loss of 5-hydroxymethylcytosine is accompanied with malignant cellular transformation. *Cancer Sci*. Epub ahead of print 2012. DOI: 10.1111/j.1349-7006.2012.02213.x.
95. Couronné L, Bastard C, Bernard OA. TET2 and DNMT3A mutations in human T-Cell lymphoma. *New England Journal of Medicine*. Epub ahead of print 2012. DOI: 10.1056/NEJMc1111708.
96. Bezerra Salomão K, Cruzeiro GAV, Bonfim-Silva R, et al. Reduced hydroxymethylation characterizes medulloblastoma while TET and IDH genes are differentially expressed within molecular subgroups. *J Neurooncol*. Epub ahead of print 2018. DOI: 10.1007/s11060-018-2845-1.
97. Nasim Azizgolshani, Curtis L Petersen, Youdinghuan Chen, Lucas A Salas, Laurent Perreard, Lananh N Nguyen VOPC. *DNA hypohydroxymethylation in pediatric central nervous system tumors is associated with CTCF binding sites and reduced survival*. 2020. Epub ahead of print 2020. DOI: https://doi.org/10.1101/2020.06.20.20136184.
98. Rumbajan JM, Maeda T, Souzaki R, et al. Comprehensive analyses of imprinted differentially methylated regions reveal epigenetic and genetic characteristics in hepatoblastoma. *BMC Cancer*. Epub ahead of print 2013. DOI: 10.1186/1471-2407-13-608.
99. Mah WC, Lee CGL. DNA methylation: Potential biomarker in Hepatocellular Carcinoma. *Biomarker Research*. Epub ahead of print 2014. DOI: 10.1186/2050-7771-2-5.
100. Mudbhary R, Hoshida Y, Chernyavskaya Y, et al. UHRF1 Overexpression Drives DNA Hypomethylation and Hepatocellular Carcinoma. *Cancer Cell*. Epub ahead of print 2014. DOI: 10.1016/j.ccr.2014.01.003.
101. Maschietto M, Rodrigues TC, Kashiwabara AY, et al. DNA methylation landscape of hepatoblastomas reveals arrest at early stages of liver differentiation and cancer-related alterations. *Oncotarget*. Epub ahead of print 2016. DOI: 10.18632/oncotarget.14208.
102. Honda S, Chatterjee A, Leichter AL, et al. A MicroRNA Cluster in the DLK1-DIO3 Imprinted Region on Chromosome 14q32.2 Is Dysregulated in Metastatic Hepatoblastomas. *Front Oncol*. Epub ahead of print 2020. DOI:

10.3389/fonc.2020.513601.

103. Han YG, Kim HJ, Dlugosz AA, et al. Dual and opposing roles of primary cilia in medulloblastoma development. *Nat Med*. Epub ahead of print 2009. DOI: 10.1038/nm.2020.
104. Schüller U, Heine VM, Mao J, et al. Acquisition of Granule Neuron Precursor Identity Is a Critical Determinant of Progenitor Cell Competence to Form Shh-Induced Medulloblastoma. *Cancer Cell*. Epub ahead of print 2008. DOI: 10.1016/j.ccr.2008.07.005.
105. Li CM, Guo M, Borczuk A, et al. Gene expression in Wilms' tumor mimics the earliest committed stage in the metanephric mesenchymal-epithelial transition. *Am J Pathol*. Epub ahead of print 2002. DOI: 10.1016/S0002-9440(10)61166-2.
106. Li W, Kessler P, Williams BRG. Transcript profiling of Wilms tumors reveals connections to kidney morphogenesis and expression patterns associated with anaplasia. *Oncogene*. Epub ahead of print 2005. DOI: 10.1038/sj.onc.1208228.
107. Rodrigues TC. *The common set of DMSs displayed a non-random distribution regarding different categories of genomic sequences being mostly hypermethylated at gene bodies and CpG islands, and hypomethylated at intergenic regions*. Universidade de São Paulo, 2014.
108. Hay DC, Zhao D, Fletcher J, et al. Efficient Differentiation of Hepatocytes from Human Embryonic Stem Cells Exhibiting Markers Recapitulating Liver Development In Vivo. *Stem Cells*. Epub ahead of print 2008. DOI: 10.1634/stemcells.2007-0718.
109. Okita K, Yamakawa T, Matsumura Y, et al. An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells. *Stem Cells* 2013; 31: 458–466.
110. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*.
111. Bustin SA, Benes V, Garson JA, et al. The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin Chem*. Epub ahead of print 2009. DOI: 10.1373/clinchem.2008.112797.
112. Costa TBBC, Lacerda ALT, Mas CD, et al. Insights into the Effects of Crack Abuse on the Human Metabolome Using a NMR Approach. *J Proteome Res* 2019; 18: 341–348.

113. Tian Y, Xu T, Huang J, et al. Tissue metabonomic phenotyping for diagnosis and prognosis of human colorectal cancer. *Sci Rep*; 6. Epub ahead of print February 2016. DOI: 10.1038/srep20790.
114. Alhosin M, Omran Z, Zamzami MA, et al. Signalling pathways in UHRF1-dependent regulation of tumor suppressor genes in cancer. *Journal of Experimental and Clinical Cancer Research*. Epub ahead of print 2016. DOI: 10.1186/s13046-016-0453-5.
115. Scourzic L, Mouly E, Bernard OA. TET proteins and the control of cytosine demethylation in cancer. *Genome Med* 2015; 7: 1–16.
116. Aguiar TFM, Carneiro TN, da Costa CML, et al. The genetic and epigenetic landscapes of hepatoblastomas. *Appl Cancer Res*. Epub ahead of print 2017. DOI: 10.1186/s41241-017-0021-0.
117. Cairo S, Armengol C, De Reynies A, et al. Hepatic stem-like phenotype and interplay of Wnt/beta-catenin and Myc signaling in aggressive childhood liver cancer. *Cancer Cell* 2008; 14: 471–484.
118. Sumazin P, Chen Y, Treviño LR, et al. Genomic analysis of hepatoblastoma identifies distinct molecular and prognostic subgroups. *Hepatology* 2017; 65: 104–121.
119. Bogan KL, Brenner C. Nicotinic Acid, Nicotinamide, and Nicotinamide Riboside: A Molecular Evaluation of NAD $^{+}$ Precursor Vitamins in Human Nutrition. *Annu Rev Nutr* 2008; 28: 115–130.
120. Hong S, Moreno-Navarrete JM, Wei X, et al. Nicotinamide N-methyltransferase regulates hepatic nutrient metabolism through Sirt1 protein stabilization. *Nat Med* 2015; 21: 887–94.
121. Kim J, Hong SJ, Lim EK, et al. Expression of nicotinamide N-methyltransferase in hepatocellular carcinoma is associated with poor prognosis. *J Exp Clin Cancer Res* 2009; 28: 1–9.
122. Pissios P. Nicotinamide N-Methyltransferase: More Than a Vitamin B3 Clearance Enzyme. *Trends Endocrinol Metab* 2017; 28: 340–353.
123. Shabalin K, Nerinovski K, Yakimov A, et al. NAD metabolome analysis in human cells using 1 H NMR spectroscopy. *Int J Mol Sci*; 19. Epub ahead of print December 2018. DOI: 10.3390/ijms19123906.
124. Xu J, Capezzone M, Xu X, et al. Activation of nicotinamide N-methyltransferase gene promoter by hepatocyte nuclear factor-1beta in human papillary thyroid cancer

cells. *Mol Endocrinol* 2005; 19: 527–39.

125. Xu J, Moatamed F, Caldwell JS, et al. Enhanced Expression of Nicotinamide N-Methyltransferase in Human Papillary Thyroid Carcinoma Cells. *J Clin Endocrinol Metab*. Epub ahead of print 2003. DOI: 10.1210/jc.2002-021843.
126. Yamada K, Miyazaki T, Hara N, et al. Interferon-γ elevates nicotinamide N-Methyltransferase activity and nicotinamide level in human glioma cells. *J Nutr Sci Vitaminol (Tokyo)*. Epub ahead of print 2010. DOI: 10.3177/jnsv.56.83.
127. Zhang J, Wang Y, Li G, et al. Down-regulation of nicotinamide N-methyltransferase induces apoptosis in human breast cancer cells via the mitochondria-mediated pathway. *PLoS One*; 9. Epub ahead of print 2014. DOI: 10.1371/journal.pone.0089202.
128. Ulanovskaya OA, Zuhl AM, Cravatt BF. NNMT promotes epigenetic remodeling in cancer by creating a metabolic methylation sink. *Nat Chem Biol* 2013; 9: 300–306.
129. Shin JH, Park CW, Yoon G, et al. NNMT depletion contributes to liver cancer cell survival by enhancing autophagy under nutrient starvation. *Oncogenesis*. Epub ahead of print 2018. DOI: 10.1038/s41389-018-0064-4.
130. Jin D, Song Y, Chen Y, et al. Identification of Three lncRNAs as Potential Predictive Biomarkers of Lung Adenocarcinoma. *Biomed Res Int*. Epub ahead of print 2020. DOI: 10.1155/2020/7573689.
131. Xie W, Yuan S, Sun Z, et al. Long noncoding and circular RNAs in lung cancer: Advances and perspectives. *Epigenomics*. Epub ahead of print 2016. DOI: 10.2217/epi-2016-0036.
132. Ferreira HJ, Esteller M. Non-coding RNAs, epigenetics, and cancer: tying it all together. *Cancer Metastasis Rev*. Epub ahead of print 2018. DOI: 10.1007/s10555-017-9715-8.
133. Cristóbal I, Sanz-Álvarez M, Luque M, et al. The role of microRNAs in hepatoblastoma tumors. *Cancers*. Epub ahead of print 2019. DOI: 10.3390/cancers11030409.
134. Lupski JR, Stankiewicz P. Genomic disorders: Molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genetics*. Epub ahead of print 2005. DOI: 10.1371/journal.pgen.0010049.
135. Villela D, de Barros JS, da Costa SS, et al. Detection of mosaicism for segmental and whole chromosome imbalances by targeted sequencing. *Ann Hum Genet*. Epub ahead of print 2021. DOI: 10.1111/ahg.12402.

136. Tomlinson GE, Douglass EC, Pollock BH, et al. Cytogenetic evaluation of a large series of hepatoblastomas: Numerical abnormalities with recurring aberrations involving 1q12-q21. *Genes Chromosom Cancer*. Epub ahead of print 2005. DOI: 10.1002/gcc.20227.
137. Arai Y, Honda S, Haruta M, et al. Genome-wide analysis of allelic imbalances reveals 4q deletions as a poor prognostic factor and MDM4 amplification at 1q32.1 in hepatoblastoma. *Genes Chromosom Cancer*. Epub ahead of print 2010. DOI: 10.1002/gcc.20770.
138. Rivas MP, Aguiar TFM, Fernandes GR, et al. TET Upregulation Leads to 5-Hydroxymethylation Enrichment in Hepatoblastoma. *Front Genet* 2019; 10: 1–7.
139. Zhang S, Wei JS, Khan J. The Significance of Transcriptome Sequencing in Personalized Cancer Medicine. In: *Cancer Genomics: From Bench to Personalized Medicine*. 2013. Epub ahead of print 2013. DOI: 10.1016/B978-0-12-396967-5.00004-9.
140. Gomez-Casati DF, Grisolía M, Busi M V. The Significance of Metabolomics in Human Health. In: *Medical and Health Genomics*. 2016. Epub ahead of print 2016. DOI: 10.1016/B978-0-12-420196-5.00007-1.
141. Cieślik M, Chinnaiyan AM. Cancer transcriptome profiling at the juncture of clinical translation. *Nature Reviews Genetics*. Epub ahead of print 2018. DOI: 10.1038/nrg.2017.96.
142. Rivas MP, Aguiar TFM, Maschietto M, et al. Hepatoblastomas exhibit marked NNMT downregulation driven by promoter DNA hypermethylation . *Tumor Biol*. Epub ahead of print 2020. DOI: 10.1177/1010428320977124.
143. Li H. [Heng Li - Compares BWA to other long read aligners like CUSHAW2] Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv Prepr arXiv*.
144. Dobin A, Davis CA, Schlesinger F, et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*. Epub ahead of print 2013. DOI: 10.1093/bioinformatics/bts635.
145. Anders S, Pyl PT, Huber W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*. Epub ahead of print 2015. DOI: 10.1093/bioinformatics/btu638.
146. R: a language and environment for statistical computing, https://www.gbif.org/pt/tool/81287/r-a-language-and-environment-for-statistical-

computing (accessed 21 April 2021).

147. Alboukadel Kassambara and Fabian Mundt. factoextra : Extract and Visualize the Results of Multivariate Data Analyses. *R package version 1.0.7*, https://cran.r-project.org/package=factoextra (2020).
148. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. Epub ahead of print 2014. DOI: 10.1186/s13059-014-0550-8.
149. Love M, Anders S, Huber W. Analyzing RNA-seq data with DESeq2. *Bioconductor*.
150. Graffelman J, Van Eeuwijk F. Calibration of multivariate scatter plots for exploratory analysis of relations within and between sets of variables in genomic research. *Biometrical J*. Epub ahead of print 2005. DOI: 10.1002/bimj.200510177.
151. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016; 32: 2847–2849.
152. Gel B, Serra E. KaryoploteR: An R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* 2017; 33: 3088–3090.
153. Yu G, Wang LG, Han Y, et al. ClusterProfiler: An R package for comparing biological themes among gene clusters. *Omi A J Integr Biol* 2012; 16: 284–287.
154. Walter W, Sánchez-Cabo F, Ricote M. GOplot: An R package for visually combining expression data with functional analysis. *Bioinformatics* 2015; 31: 2912–2914.
155. Maere S, Heymans K, Kuiper M. BiNGO: A Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics*. Epub ahead of print 2005. DOI: 10.1093/bioinformatics/bti551.
156. Csardi G, Nepusz T. The igraph software package for complex network research, InterJournal, Complex Systems. *InterJournal*.
157. Chou CH, Shrestha S, Yang CD, et al. MiRTarBase update 2018: A resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res* 2018; 46: D296–D302.
158. Agarwal V, Bell GW, Nam JW, et al. Predicting effective microRNA target sites in mammalian mRNAs. *Elife*; 4. Epub ahead of print 12 August 2015. DOI: 10.7554/eLife.05005.
159. Meyers RL, Maibach R, Hiyama E, et al. Risk-stratified staging in paediatric hepatoblastoma: a unified analysis from the Children's Hepatic tumors International Collaboration. *Lancet Oncol* 2017; 18: 122–131.

160. Warner JP, Leek JP, Intody S, et al. Human glucokinase regulatory protein (GCKR): cDNA and genomic cloning, complete primary structure, and chromosomal localization. *Mamm Genome*. Epub ahead of print 1995. DOI: 10.1007/BF00356171.
161. Lee CT, Zhang L, Mounajjed T, et al. High mobility group AT-hook 2 is overexpressed in hepatoblastoma. *Hum Pathol*. Epub ahead of print 2013. DOI: 10.1016/j.humpath.2012.08.003.
162. Davies SM. Maintenance of Genomic Imprinting at the IGF2 Locus in Hepatoblastoma. *Cancer Res*.
163. Wirths O, Waha A, Weggen S, et al. Overexpression of human Dickkopf-1, an antagonist of wingless/WNT signaling, in human hepatoblastomas and Wilms' tumors. *Lab Investig*. Epub ahead of print 2003. DOI: 10.1097/01.LAB.0000059926.66359.BD.
164. Vastrad B, Vastrad C, Kottursheti I. Identification of potential core genes in hepatoblastoma via bioinformatics analysis. *medRxiv*. Epub ahead of print 2020. DOI: 10.1101/2020.12.22.20248756.
165. Czauderna P, Garnier H. Hepatoblastoma: Current understanding, recent advances, and controversies. England, 2018.
166. Lu R, Fan C, Shangguan W, et al. Neurons generated from carcinoma stem cells support cancer progression. *Signal Transduct Target Ther*. Epub ahead of print 2017. DOI: 10.1038/sigtrans.2016.36.
167. Crippa S, Ancey P, Vazquez J, et al. Mutant CTNNB 1 and histological heterogeneity define metabolic subtypes of hepatoblastoma . *EMBO Mol Med*. Epub ahead of print 2017. DOI: 10.15252/emmm.201707814.
168. Wang J, Tian R, Shan Y, et al. Metabolomics study of the metabolic changes in hepatoblastoma cells in response to NTCP/SLC10A1 overexpression. *Int J Biochem Cell Biol* 2020; 125: 105773.
169. Alexander RP, Fang G, Rozowsky J, et al. Annotating non-coding regions of the genome. *Nature Reviews Genetics*. Epub ahead of print 2010. DOI: 10.1038/nrg2814.
170. Li J, Liu C. Coding or noncoding, the converging concepts of RNAs. *Frontiers in Genetics*. Epub ahead of print 2019. DOI: 10.3389/fgene.2019.00496.
171. Qiao HP, Gao WS, Huo JX, et al. Long non-coding RNA GAS5 Functions as a tumor suppressor in renal cell carcinoma. *Asian Pacific J Cancer Prev*. Epub ahead of print 2013. DOI: 10.7314/APJCP.2013.14.2.1077.

172. Pan YF, Qin T, Feng L, et al. Expression profile of altered long non-coding RNAs in patients with HBV-associated hepatocellular carcinoma. *J Huazhong Univ Sci Technol - Med Sci*. Epub ahead of print 2013. DOI: 10.1007/s11596-013-1078-y.
173. Sun W, Wu Y, Yu X, et al. Decreased expression of long noncoding RNA AC096655.1-002 in gastric cancer and its clinical significance. *Tumor Biol*. Epub ahead of print 2013. DOI: 10.1007/s13277-013-0821-0.
174. Ge X, Chen Y, Liao X, et al. Overexpression of long noncoding RNA PCAT-1 is a novel biomarker of poor prognosis in patients with colorectal cancer. *Med Oncol*. Epub ahead of print 2013. DOI: 10.1007/s12032-013-0588-6.
175. Magrelli, A., Azzalin, G., Salvatore, M., Viganotti, M., Tosto, F., Colombo, T., Devito, R., Di Masi, A., Antoccia, A., Lorenzetti, S., Maranghi, F., Mantovani, A., Tanzarella, C., Macino, G., & Taruscio D. Altered microRNA Expression Patterns in Hepatoblastoma Patients. *Transl Oncol* 2009; 2: 157–163.
176. Kong YW, Ferland-McCollough D, Jackson TJ, et al. MicroRNAs in cancer management. *The Lancet Oncology*. Epub ahead of print 2012. DOI: 10.1016/S1470-2045(12)70073-6.
177. Cui X, Wang Z, Li J, et al. Cross talk between RNA N6-methyladenosine methyltransferase-like 3 and miR-186 regulates hepatoblastoma progression through Wnt/β-catenin signalling pathway. *Cell Prolif*; 53. Epub ahead of print 1 March 2020. DOI: 10.1111/cpr.12768.
178. Gyugos M, Lendvai G, Kenessey I, et al. MicroRNA expression might predict prognosis of epithelial hepatoblastoma. *Virchows Arch* 2014; 464: 419–427.
179. Wu JF, Ho MC, Ni YH, et al. Dysregulation of liver developmental microRNA contribute to hepatic carcinogenesis. *J Formos Med Assoc* 2020; 119: 1041–1051.
180. Chuang KH, Whitney-Miller CL, Chu CY, et al. MicroRNA-494 is a master epigenetic regulator of multiple invasion-suppressor microRNAs by targeting ten eleven translocation 1 in invasive human hepatocellular carcinoma tumors. *Hepatology* 2015; 62: 466–480.
181. Liu K, Liu S, Zhang W, et al. miR-494 promotes cell proliferation, migration and invasion, and increased sorafenib resistance in hepatocellular carcinoma by targeting PTEN. *Oncol Rep* 2015; 34: 1003–1010.
182. Lieven O, Knobloch J, Rüther U. The regulation of Dkk1 expression during embryonic development. *Dev Biol* 2010; 340: 256–268.
183. Wang RN, Green J, Wang Z, et al. Bone Morphogenetic Protein (BMP) signaling in

development and human diseases. *Genes and Diseases* 2014; 1: 87–105.

184. Camilleri S, McDonald F. Runx2 and dental development. *European Journal of Oral Sciences* 2006; 114: 361–373.
185. Valdivia LE, Young RM, Hawkins TA, et al. Lef1-dependent Wnt/β-catenin signalling drives the proliferative engine that maintains tissue homeostasis during lateral line development. *Development* 2011; 138: 3931–3941.
186. Nunes FD, de Almeida FC, Tucci R, et al. Homeobox genes: a molecular link between development and cancer. *Pesquisa odontológica brasileira = Brazilian oral research* 2003; 17: 94–98.
187. Ogawa Y, Oses-Prieto J, Kim MY, et al. ADAM22, a Kv1 channel-interacting protein, recruits membrane-associated guanylate kinases to juxtaparanodes of myelinated axons. *J Neurosci* 2010; 30: 1038–1048.
188. Goldsmith AP, Gossage SJ, Ffrench-Constant C. ADAM23 is a cell-surface glycoprotein expressed by central nervous system neurons. *J Neurosci Res* 2004; 78: 647–658.
189. Chou CW, Huang YK, Kuo TT, et al. An overview of ADAM9: Structure, activation, and regulation in human diseases. *International Journal of Molecular Sciences* 2020; 21: 1–22.
190. Fatima S, Lee NP, Tsang FH, et al. Dickkopf 4 (DKK4) acts on Wnt/B-catenin pathway by influencing B-catenin in hepatocellular carcinoma. *Oncogene* 2012; 31: 4233–4244.
191. Cai X, Yao Z, Li L, et al. Role of DKK4 in tumorigenesis and tumor progression. *International Journal of Biological Sciences* 2018; 14: 616–621.
192. Li L, Zhuo Z, Yang Z, et al. HMGA2 polymorphisms and hepatoblastoma susceptibility: A five-center case-control study. *Pharmgenomics Pers Med* 2020; 13: 51–57.
193. Song L, Guo X, Zhao F, et al. TTC36 inactivation induce malignant properties via Wnt-β-catenin pathway in gastric carcinoma. *J Cancer* 2021; 12: 2598–2609.
194. Těšínský M, Šimčíková D, Heneberg P. First evidence of changes in enzyme kinetics and stability of glucokinase affected by somatic cancer-associated variations. *Biochim Biophys Acta - Proteins Proteomics* 2019; 1867: 213–218.
195. Kanda M, Tanaka H, Shimizu D, et al. SYT7 acts as a driver of hepatic metastasis formation of gastric cancer cells. *Oncogene* 2018; 37: 5355–5366.
196. Jin H, Xu G, Zhang Q, et al. Synaptotagmin-7 is overexpressed in hepatocellular

carcinoma and regulates hepatocellular carcinoma cell proliferation via Chk1–p53 signaling. *Onco Targets Ther* 2017; 10: 4283–4293.

197. Kim JH, Kim TW, Kim SJ. Downregulation of ARFGEF1 and CAMK2B by promoter hypermethylation in breast cancer cells. *BMB Rep* 2011; 44: 523–528.
198. Noh JH, Jung KH, Kim JK, et al. Aberrant Regulation of HDAC2 Mediates Proliferation of Hepatocellular Carcinoma Cells by Deregulating Expression of G1/S Cell Cycle Proteins. *PLoS One* 2011; 6: e28103.
199. Hagelkruys A, Sawicka A, Rennmayr M, et al. The biology of HDAC in cancer: The nuclear and epigenetic components. *Handb Exp Pharmacol*. Epub ahead of print 2011. DOI: 10.1007/978-3-642-21631-2_2.
200. Beck A, Eberherr C, Hagemann M, et al. Connectivity map identifies HDAC inhibition as a treatment option of high-risk hepatoblastoma. *Cancer Biol Ther*. Epub ahead of print 2016. DOI: 10.1080/15384047.2016.1235664.
201. Pakakasama S, Chen TTL, Frawley W, et al. CCND1 polymorphism and age of onset of hepatoblastoma. *Oncogene* 2004; 23: 4789–4792.
202. Rivas M, Aguiar T, Fernandes G, et al. DNA methylation as a key epigenetic player for hepatoblastoma characterization. *Clin Res Hepatol Gastroenterol* 2021; 45: 101684.