DNA METHYLATION AS A KEY EPIGENETICS PLAYER FOR HEPATOBLASTOMA STRATIFICATION

Maria Rivas*[1], Talita Aguiar*[2], Gustavo Fernandes[2], Luiz Caires-Júnior[1], Ernesto Goulart[1], Kayque Telles-Silva[1], Mariana Maschietto[3], Monica Cypriano[4], Silvia Toledo[5], Dirce Carraro[6], Isabela Werneck[7], Cecilia Da Costa[8], Carla Rosenberg[2], and Ana Krepischi[2]

[1]Institute of Biosciences, University of São Paulo
[2]Institute of Biosciences, University of Sao Paulo
[3]Boldrini Children's Hospital
[4]Institute of Pediatric Oncology
[5]Genetics Laboratory, Pediatric Oncology Institute GRAACC/UNIFESP
[6]A. C. Camargo Cancer Center
[7]Rede D'OR-São Luiz
[8]A.C.Camargo Cancer Center

June 1, 2020

Abstract

Background: Hepatoblastoma (HB) is a rare embryonal liver tumor occurring in the pediatric population, and although intrinsic biological differences between tumors can impact HBs prognosis, few groups have studied this aspect. Given the crescent relevance of epigenetic mechanisms in the genesis and progression of these tumors, we aim to classify HB samples according to the different stages of liver development as well as DNA methylation machinery. Procedures: Using bioinformatics tools, we evaluate the expression of 24 genes associated with epigenetics and stages of hepatocyte differentiation as well as global DNA methylation to propose a stratification model for HB. Results: Based on the gene expression profiles of DNA methylation machinery and hepatocyte differentiation markers, HBs were clustered into three groups. Besides reinforcing the molecular heterogeneity of these embryonal tumors, our data propose that a panel of 13 genes can be used for HB stratification (TET1, TET2, TET3, DNMT1, DNMT3A, UHRF1, ALB, CYP3A4, TDO2, UGT1A1, AFP, HNF4A, and FOXA2). DNA methylation machinery exerts a key role in the characterization of HBs, directly reflected in diverse DNA methylation content. Moreover, we suggest that the group of HBs presenting similarity with differentiated livers, such as upregulation of mature hepatocyte markers, decreased expression of DNA methylation enzymes and higher global methylation levels are associated with a worse prognosis. Conclusions: HBs are heterogeneous tumors and our findings point out the role of the epigenetic machinery in the samples characterization, suggesting the need of new stratification system for patients with HBs.

INTRODUCTION

Hepatoblastoma (HB) is a rare liver tumor occurring in childhood[1], hypothetically arising from hepatocyte precursors[2]. There is a reported global trend towards an increasing incidence of HB over the last years [3] which could be caused by the rising of premature children, associated with low birthweight, 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 low survival [7–10]. HB are heterogeneous with several histologic types [2], frequently displaying combinations of epithelial, stromal, mesenchymal and/or undifferentiated cells[11].Despite this clinical heterogeneity, a few recurrent genetic lesions were reported in HBs, that present

the lowest reported burden of somatic mutations [12]. Epigenetic mechanisms seem to be crucial factors to be considered in the genesis and progression of this tumor. In our previous work[13], the evaluation of HB methylomes disclosed a lower level of DNA methylation is non-repetitive sequences compared to non-tumoral livers, a finding corroborated by others[14], in addition to hypermethylation in CpG islands mapped in genes associated with tumor suppression, lipid metabolism, and liver differentiation. Moreover, we unraveled 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, for which the main hypothesis of origin proposes that cancer arises from a failure of cell differentiation, leading to the rupture of the normal process of organ development [16]. Therefore, some childhood tumors could mimic the activity of embryonic cells; as an example, the α-fetoprotein, only produced by the fetal liver, is increased in most HBs [17]. Although intrinsic biological differences between tumors can impact HBs prognosis, few groups have studied this aspect. Here, we focus on the model of the origin of embryonal tumors to classify HB samples according to the different stages of liver development, based on gene expression analysis of markers of hepatocyte differentiation as well as enzymes of the DNA methylation machinery. Using these data, we provide a framework to stratify HBs, correlating the molecular biomarkers with clinical features.

MATERIAL AND METHODS

Patients

Fresh-frozen HBs (n=21) and matched non-tumoral liver tissues (n=9) were recovered from biobanks. All patients received pre-surgery chemotherapy according to either SIOPEL (*http://www.siopel.org* /) or COG (*https://www.childrensoncologygroup.org* /) protocols, and they were followed by clinical examination, imaging tests, and alpha-fetoprotein dosage. This study was approved by the Ethics Committees of the A.C. Camargo Cancer Center and GRAACC. Biological samples were collected after signed informed consent obtained from parents.

Cell cultures

Induced pluripotent stem cell lines (iPSC) derived from human peripheral blood of three control individuals were obtained following the procedure reported by Okita et al. [18]. Hepatic differentiation was performed using the iPSC lines as described [19] to collect cells in different stages of differentiation (definitive endoderm, hepatoblasts, and precursor hepatocyte cells).

RNA isolation

RNA samples were obtained 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 manufacturer's recommendations. The quality of RNA samples was verified by 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). The cDNA was synthesized using the High Capacity RNA-to-cDNA kit (Applied Biosystems, EUA). 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[20] after expression analysis of *ACTB* ,*GAPDH* , *B2M* genes and 18S ribosomal RNA (*18S rRNA*). Data were normalized using the expression values of the housekeeping genes *18S rRNA* and *B2M* , and all reactions were performed with three technical replicates. The delta-delta Ct ([?][?]Ct) method was used for data analysis [21].

Statistical analysis

2

Statistical analysis of gene expression data was applied to $\log_2$-transformed values [22]. 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 prior to clustering procedure. Differentially expressed genes were identified by analysis of variance (ANOVA) [23] between sample sets 1, 2 and 3. All analyses were performed using R (version 3.4.4), with the following packages: STRINGR, GGPLOT2, GRID, gridExtra, and GGBIPLOT.

Global methylation level

The global mean methylation level was calculated as the average of all normalized B-values of the HM450K Methylation array (Illumina)[13] for each sample as this platform mainly reflects the content of CpG sites mapped to regulatory sequences.

In silico analysis

To explore the pathways in which the genes are involved with, we used Reactome Pathway Database (https://reactome.org/ - Version 72 Released). A binomial test is used to calculate the probability for each result and the p-values are corrected for the multiple testing (Benjamini–Hochberg procedure) that arises from evaluating the submitted list of identifiers against every pathway [24,25]. We considered significant, those p-values <0.05. Data were crossed with other resources such as NCBI, Ensembl, UniProt, KEGG (Gene and Compound), ChEBI, PubMed and GO.

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

RESULTS

Gene expression of differentiation markers was evaluated for specific stages: stem cells (*NANOG* , *POU5F1* , *SOX2*), definitive endoderm (*CXCR4* , *FOXA2* , *SOX17*), hepatoblasts (*AFP* , *CK19* , *HNF4A)* , hepatocyte (*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 correlated expression patterns (**Figure 1**): (1) *AFP, CK19, UHRF1, DNMT3B, HNF1A, FOXA2, HNF4A* , *SOX2* ; (2) *DNMT3A, TET3, TET1, DNMT1* , *TET2* ; and (3) *ALB, TDO2, CYP3A4, CK18* , and *UGT1A1* .

Based on the (**Supporting Information Figure S1A and S1B**) and K-means (**Supporting Information Figure S2)** The Set-1 cluster (in black) was composed of hepatoblasts and hepatocyte-like cells and only three HBs (33T, 35T, and 43T), suggesting a gene expression profile similar to transitional phases of the hepatocyte differentiation. A second group (Set-2, in red) contained nine HB samples (18T, 28T 32T, 37T, 38T, 40T, 42T, 44T, and 45T). The larger group (Set-3, in green) clustered all 9 non-tumoral liver samples plus nine HBs (15T, 17T, 30T, 31T, 34T, 36T, 39T, 41T and 46T), showing that these tumors exhibited a gene expression profile similar to differentiated livers. The Set-2 cluster (red) presented a partial overlap to Set-1 and Set-3, pointing to mixed characteristics of gene expression, in a more advanced phase of hepatocyte differentiation when compared to the Set-1 group. Finally, an isolated group consisting of only the iPSC and the definitive endoderm cells was detected (Set-4, in blue), exhibiting the most distinctive gene expression pattern when compared to the other groups, as expected.

Differential expression analysis indicated that 13 out the 24 genes showed differences between Set-1, Set-2 and Set-3 groups (**Supporting Information Figure S3)** . From these thirteen genes, six genes were associated with the epigenetic machinery (*TET1, TET2, TET3, DNMT1, DNMT3A,* and *UHRF1*), a definitive endoderm marker (*FOXA2*), two hepatoblast markers (*AFP* and*HNF4* A), and four hepatocyte mark-

ers (*ALB, CYP3A4, TDO2* , and *UGT1A1*). The gene expression analysis evidenced that Set-1 tumors have upregulation of hepatoblast markers (**Figure 3A-C**), and high expression of the DNA methylation genes, mainly *UHRF1* and *TET1* (**Figure 3D-E**). The Set-2 tumors exhibit upregulation of the DNA methylation genes (**Figure 3D-I**) and low expression of markers of both the intermediary stages of differentiation (**Figure 3A-C**) and mature hepatocyte genes (**Figure 3J-M**). Finally, in Set-3, samples show down-regulation of the DNA methylation genes (**Figure 3D-I**), associated with high expression of mature hepatocyte marker genes (**Figure 3J-K**). With the determination of these 13 genes derived from differential expression, the analysis of gene distance showed that these genes have three groups with a similar expression pattern (*UHRF1, AFP, FOXA2* and *HNF4A* ; *DNMT3A, TET3, TET1, DNMT1* , and *TET2* ;*ALB, TDO2, CYP3A4,* and *UGT1A1*). Besides, the analysis showed that in HB genes of epigenetic machinery have an inverse expression pattern to the mature hepatocyte's markers (**Supporting Information Figure S4**).

The thirteen genes were submitted to Reactome Pathways analysis, searching for 1) pathways related to biological mechanisms and 2) pathways that are specific for cancer. From the 158 pathways ($p<0.05$) (**Supporting Information Table S1**), the top five were mRNA splicing - major pathway (R-HSA-72163), mRNA splicing (R-HSA-72172), processing of capped intron-containing pre-mRNA (R-HSA-72203), apoptosis (R-HSA-109581) and protein localization (R-HSA-9609507). Considering cancer pathways, from 33 pathways ($p<0.05$) (**Supporting Information Table S2**), the top five were endosomal/vacuolar pathway (R-HSA-1236977), *TP53* regulates transcription of DNA repair genes (R-HSA-6796648), signaling by *BRAF* and *RAF* fusions (R-HSA-6802952), constitutive signaling by aberrant *PI3K* in cancer (R-HSA-2219530) and oncogenic MAPK signaling (R-HSA-6802957).

Comparison of the analysis of global DNA methylation between the three groups (**Figure 4**) showed that the Set-3 presented an overall methylation level similar to non-tumoral liver samples, in agreement with expression data. The samples from the Set-1, that presented expression profile similar to embryonic stages of the hepatocyte differentiation, had a lower level of global DNA methylation than the other two groups. The Set-2 samples, that contain HBs with transitional gene expression profile when compared to the others, also had an intermediate mean level of DNA methylation.

Table 1 presents the clinical characteristics of the HB cases, clustered according to the Set in which they were classified based on their gene expression profile. Moreover, the somatic driver mutations previously detected in some HBs, as well as the results of the immunohistochemistry analysis of the beta-catenin are presented (data derived from Aguiar et al. 2020 [26]). Set-1 was composed of intermediate-risk HBs, according to the CHIC criteria[27,28]. It is interesting to note that three out of nine patients of the Set-3 died, three developed pulmonary metastases, and one was characterized as HB/HCC features (carrying a *TERT* promoter mutation; **Supporting Information Figure S5 − Kaplan-Meier survival curve**). The Set-2 contains tumors that are clinically heterogeneous, with two cases of late diagnosis, one of them carrying a somatic mutation in the *TERT* promoter.

Despite the number of samples in each set, it is possible to notice a trend in relation to histology. In Set-1 there is a predominance of epithelial embryonal, Set-2 exhibits a predominance of epithelial and mesenchymal mixed, and in Set-3, epithelial fetal samples.**Figure 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 molecular subclassifications, including gene signatures, have been proposed [2,29,30]. Expression pattern of 16 genes grouped HBs according to tumor aggressiveness in two groups, called C1 and C2 signatures where β-catenin is upregulated in both signatures with genes being related to hepatocyte markers (*GLUL* , *CYP2E1* , *CYP1A1,* and *AQP9)*), cell cycle regulation (*E2F5* , *BUB1* ,*DLG7*), hepatic stem/progenitor markers (*DUSP9* , *AFP* ,*CK19* , and *TACSTD1)* , and metabolism (*APOC4, ALDH2*)[29]. The differential activation of hepatic stem/progenitor markers (*LIN28B, SALL4* , and *AFP*), genes related to cancer pathways (*TERT, TP53*), metabolism (*NFE2L2*) and hepatobiliary

system (*NOCHT1* , *HNF*) were also used to stratify HBs in molecular subtypes according to risk[30].

Here, we focused on epigenetic markers (expression of DNMTs and TETs as well as global DNA methylation) in addition to genes related to the stages of liver differentiation, in an attempt to perform HB stratification. Hepatocytes are the major cell type of the liver, accounting for ˜ 70% of the mass of the adult organ. 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 [31,32], and liver growth and hepatocyte maturation are processes regulated by genes acting intrinsically in these cells [33]. HB is postulated to develop from hepatoblasts, the precursor cells of hepatocytes [2]; however, microscopically, HBs are heterogeneous, rarely composed of only one cell type, often exhibiting combinations of epithelial, stromal, mesenchymal and/or undifferentiated cells components[11].

Genes associated with DNA methylation exhibited in HBs a pattern of expression directly correlated, which is inverse to the expression pattern of mature hepatocyte markers, highlighting the central role of this specific epigenetic pathway in liver differentiation and as well as in HB stratification. Our data also disclosed a direct correlation of *UHRF1* and *AFP* expression in tumors, which was previously described for HCC [34], possibly representing a new finding to be explored in HBs. Additionally, there is a similar expression pattern of *HFN4A* and *FOXA2* , which are nuclear factors that present tissue co-expression and cooperate for hepatic pathway cell commitment[35], emphasizing the key role of the hepatic differentiation blockage in HB development.

Besides reinforcing the molecular heterogeneity of these embryonal tumors, we propose a panel of 13 genes for this HB stratification (*TET1, TET2, TET3, DNMT1, DNMT3A, UHRF1, ALB, CYP3A4, TDO2* , *UGT1A1, AFP* , *HNF4* A, and *FOXA2*). Remarkably, our data showed that the DNA methylation machinery exerts a key role in the characterization of HBs, directly reflected in diverse DNA methylation content. Six out of the eight genes associated with DNA methylation were determinants for HB stratification, evidencing the importance of the epigenetic machinery for the biology of this embryonal liver cancer.

At the beginning of the embryo development, the epigenetic machinery acts by decreasing the overall level of DNA methylation, allowing the expression of genes associated with pluripotency [36]. During the maturation of the organ, there is an increase in the DNA methylation, a process that will culminate in the selective expression of tissue-specific genes, associated with the promotion of cell differentiation. The iPSC and definitive endoderm cell lines clustered separately (Set-4), reinforcing that HBs are composed of cells already compromised with the hepatocyte differentiation. A small group of intermediate-risk tumors exhibited high expression of the epigenetic machinery genes as well as hepatoblast markers (Set-1), a profile suggestive of a precursor stage of hepatocyte differentiation. This HB group present the highest expression of *TET1* , *TET3* , and *UHRF1* among all tumors; in a recent work, we proposed an active demethylation process in HBs, particularly associated with upregulation of *TETs* and *UHRF1* [15] and, accordingly, Set-1 tumors exhibited marked global DNA hypomethylation. In the other hand, the Set-3 presented downregulation of the DNA methylation genes and high expression of mature hepatocyte markers, similarly to non-tumoral control livers; they were likely derived from cells compromised with more advanced stages of hepatocyte differentiation, and their global DNA methylation content is also similar to non-tumoral livers, corroborating the hypothesis of origin in late stages of hepatocyte differentiation. Considering that Set-3 comprised HBs from patients who are deceased and/or developed metastases, as well as one sample characterized as HB/HCC features, the major markers of hepatocyte (*UGT1A1* , *TDO2* , and *CYP3A4*) that are differentially expressed in this group, deserve further investigation as biomarkers of worse prognosis. Finally, 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 hepatoblast and mature hepatocyte markers. The Set-2 has a global DNA methylation level intermediary between Set-1 and Set-3, which is in accordance with the hypothesis that these tumor samples have a molecular profile transitional between hepatoblasts and mature hepatocytes.

In a recent analysis involving a large group of 113 samples, HBs presented two epigenetic signatures (Epi-CA and Epi-CB), according to the degree of sample hypomethylation. Moreover, amplification of the oncogenic

14q32 *DLK1-DIO3* locus in part of the HBs delineated signatures of moderate or strong expression of genes mapped to this region. Using both findings, a molecular risk stratification of three categories was proposed (MRS-HB) [37]. Despite using a smaller cohort of 21 HB samples in our study, we also showed that DNA methylation is a strong biomarker for HB stratification, pinpointing to specific DNA methylation genes as important players.

Precision medicine brings the promise of improvement in diagnosis and matching patients to personalized targeted therapies using genomics, epigenomics, metabolomics and proteomics. One of the major barriers to the implementation of this approach in clinical settings is the intra/inter- tumor heterogeneity, in addition to the rarity in the case of pediatric cancer. In particular, HB is a tumor with a genetic background with low mutational background and few cytogenetic alterations[12,30,38–41], in addition to DNA methylation changes[13,42].

The present work describes the possibility of molecular stratification of HBs according to markers of hepatocyte differentiation and DNA methylation. Tumor heterogeneity could be overcome using molecular signatures that are linked to clinical data. Here, we evidenced a HB group (Set-3) with a similarity of late stages of hepatocyte differentiation and lower DNA methylation genes that present a tendency to a worse prognosis. Although the other two sets do not have clinical characteristics that are divided evenly between groups, the association of Set-3 with a worse prognosis is in concordance with the literature that have shown metastasis at diagnosis, advanced age of diagnosis[27,28,43], mutations at *TERT gene*[30,39] associated with high mortality.

Most drugs based on the mechanisms of anticancer resistance in HB will act on pathways such as p53, tyrosine kinase, cell cycle control, and transcriptional and translational events. A specific gene, *CYP3A4*, upregulated in Set-3, is related to five drugs used in the treatment of HBs [44], particularly etoposide (a drug used in neoadjuvant chemotherapy), which causes demethylation in liver cells [45,46], and Sorafenib, a kinase inhibitor mainly oxidized by *CYP3A4* , with the collaboration of *CYP1A1* and *CYP1B1* [47]. Vincristine, Cyclophosphamide and Isofosfamide are also metabolized by *CYP3A4*[48–51]. These findings suggest that changes in this gene could lead to resistance to treatment, maybe explaining the worse clinical signs observed in Set-3, with *CYP3A4* levels similar to mature hepatocytes.

Together, data here showed the possibility of HB stratification in three groups considering the epigenetic machinery and markers of liver differentiation stages. These groups have levels of DNA methylation according to their stage of cell differentiation, which indicates that the clinical heterogeneity widely described for this tumor, also occurs at the epigenetic level. Therefore, we demonstrate that epigenetics can be an important tool that must be considered in stratification of HB.

CONFLICTS OF INTEREST STATEMENT

We declare that we have no conflicts of interest.

ACKNOWLEDGMENTS

REFERENCES

1. Heck JE, Meyers TJ, Lombardi C, et al. Case-control study of birth characteristics and the risk of hepatoblastoma. *Cancer Epidemiol* . 2013;37(4):390-395.

2. Lopez-Terrada D, Alaggio R, de Davila MT, et al. Towards an international pediatric liver tumor consensus classification: proceedings of the Los Angeles COG liver tumors symposium. *Mod Pathol* . 2014;27(3):472-491.

3. Howlader N, Noone A, Krapcho M, et al. SEER Cancer Statistics Review 1975-2013. *Natl Cancer Inst* . 2013.

4. Turcotte LM, Georgieff MK, Ross JA, et al. Neonatal medical exposures and characteristics of low birth weight hepatoblastoma cases: a report from the Children's Oncology Group. *Pediatr Blood Cancer* . 2014;61(11):2018-2023.

5. Perilongo G, Malogolowkin M, Feusner J. Hepatoblastoma clinical research: lessons learned and future challenges. *Pediatr Blood Cancer* . 2012;59(5):818-821.

6. Trobaugh-Lotrario AD, Katzenstein HM. Chemotherapeutic approaches for newly diagnosed hepatoblastoma: past, present, and future strategies.*Pediatr Blood Cancer* . 2012;59(5):809-812.

7. 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(22):3819-3828.

8. Douglass EG, Reynolds M, Finegold M et al. Cisplatin, vincristine, and fluorouracil therapy for hepatoblastoma: A pediatric oncology group study. *J Clin Oncol* . 1993;11(1):96-9.

9. Ortega JA, Douglass EC, Feusner JH, et al. Randomized comparison of cisplatin/vincristine/fluorouracil and cisplatin/continuous infusion doxorubicin for treatment of pediatric hepatoblastoma: A report from the Children's Cancer Group and the Pediatric Oncology Group. *J Clin Oncol* . 2000;18(14):2665-2675.

10. Meyers RL, Czauderna P, Otte JB. Surgical treatment of hepatoblastoma. *Pediatr Blood Cancer* . 2012;59(5):800-8

11. Weinberg AG, Finegold MJ. Primary hepatic tumors of childhood.*Hum Pathol* . 1983;14(6):512-537.

12. 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(7696):371-376.

13. 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; 17; 8(58): 97871–97889.

14. Yamada N, Yasui K, Dohi O, et al. Genome-wide DNA methylation analysis in hepatocellular carcinoma. *Oncol Rep* . 2016;35(4):2228-2236.

15. Rivas MP, Aguiar TFM, Fernandes GR, et al. TET Upregulation Leads to 5-Hydroxymethylation Enrichment in Hepatoblastoma. *Front Genet* . 2019;10:1-7.

16. Maris JM, Denny CT. Focus on embryonal malignancies. *Cancer Cell* . 2002;2(6):447-450.

17. Kim EH, Koh KN, Park M, et al. Clinical features of infantile hepatic hemangioendothelioma. *Korean J Pediatr* . 2011; 54(6):260-6.

18. 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.

19. 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.

20. 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).

21. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.*Methods* . 2001;25(4):402-408.

22. Edmunds RC, McIntyre JK, Luckenbach JA, Baldwin DH, Incardona JP. Toward enhanced MIQE compliance: reference residual normalization of qPCR gene expression data. *J Biomol Tech* . 2014;25(2):54-60.

23. Hardeo S, Mohammed AI. *The Analysis of Variance: Fixed, Random and Mixed Models* . Birkhauser; 2000.

24. Fabregat A, Sidiropoulos K, Viteri G, et al. Reactome pathway analysis: A high-performance in-memory approach. *BMC Bioinformatics* . 2017;18(1):142.

25. Fabregat A, Sidiropoulos K, Garapati P, et al. The reactome pathway knowledgebase. *Nucleic Acids Res* . 2016;44(D1):D481-D487.

26. Aguiar TFM, Rivas MP, Costa S, et al. Insights Into the Somatic Mutation Burden of Hepatoblastomas From Brazilian Patients. *Front Oncol* . 2020;10:556.

27. 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(1):122-131.

28. 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.

29. 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(6):471-484.

30. Sumazin P, Chen Y, Trevino LR, et al. Genomic analysis of hepatoblastoma identifies distinct molecular and prognostic subgroups.*Hepatology* . 2017;65(1):104-121.

31. Houssaint E. Differentiation of the mouse hepatic primordium. I. An analysis of tissue interactions in hepatocyte differentiation.*Cell Differ* . 1980; 9(5):269-79

32. Medlock ES, Haar JL. The liver hemopoietic environment: I. Developing hepatocytes and their role in fetal hemopoiesis. *Anat Rec* . 1983; 207(1):31-41.

33. Zorn. Liver Development. *StemBook* . 2008.

34. Mudbhary R, Hoshida Y, Chernyavskaya Y, et al. UHRF1 Overexpression Drives DNA Hypomethylation and Hepatocellular Carcinoma. *Cancer Cell* . 2014;25(2):196-209.

35. 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(5):1033-1048.

36. Combes AN, Whitelaw E. Epigenetic reprogramming: Enforcer or enabler of developmental fate? *Dev Growth Differ* . 2010; 52(6):483-91

37. Carrillo-Reixach J, Torrens L, Simon-Coma M, et al. Epigenetic footprint enables molecular risk stratification of hepatoblastoma with clinical implications. *J Hepatol* . March 2020; 30;S0168-8278(20)30187-2.

38. Aguiar TFM, Carneiro TN, da Costa CML, Rosenberg C, da Cunha IW, Krepischi ACV. The genetic and epigenetic landscapes of hepatoblastomas.*Appl Cancer Res* . 2017; (37),20.

39. Eichenmüller M, Trippel F, Kreuder M, et al. The genomic landscape of hepatoblastoma and their progenies with HCC-like features. *J Hepatol* . 2014;61(6):1312-1320.

40. Rodrigues TC, Fidalgo F, da Costa CML, et al. Upregulated genes at 2q24 gains as candidate oncogenes in hepatoblastomas. *Future Oncol* . 2014;10(15):2449-2457.

41. Gröbner SN, Worst BC, Weischenfeldt J, et al. The landscape of genomic alterations across childhood cancers. *Nature* . 2018;555(7696):321-327.

42. Cui X, Liu B, Zheng S, Dong K, Dong R. Genome-wide analysis of DNA methylation in hepatoblastoma tissues. *Oncol Lett* . 2016;12(2):1529-1534.

43. Czauderna P, Lopez-Terrada D, Hiyama E, Haberle B, Malogolowkin MH, Meyers RL. Hepatoblastoma state of the art: pathology, genetics, risk stratification, and chemotherapy. *Curr Opin Pediatr* . 2014;26(1):19-28.

44. Marin JJG, Cives-Losada C, Asensio M, Lozano E, Briz O, Macias RIR. Mechanisms of anticancer drug resistance in hepatoblastoma.*Cancers (Basel)* . 2019;11(3).

45. Kawashiro T, Yamashita K, Zhao XJ, et al. A study on the metabolism of etoposide and possible interactions with antitumor or supporting Agents by Human Liver Microsomes. *J Pharmacol Exp Ther* . 1998;286(3):1294-1300.

46. Lagas JS, Fan L, Wagenaar E, et al. P-glycoprotein (P-gp/Abcb1), Abcc2, and Abcc3 determine the pharmacokinetics of etoposide. *Clin Cancer Res* . 2010;16(1):130-140.

47. Ghassabian S, Rawling T, Zhou F, et al. Role of human CYP3A4 in the biotransformation of sorafenib to its major oxidized metabolites.*Biochem Pharmacol* . 2012;84(2):215-223.

48. Martino E, Casamassima G, Castiglione S, et al. Vinca alkaloids and analogues as anti-cancer agents: Looking back, peering ahead.*Bioorganic Med Chem Lett* . 2018;28(17):2816-2826.

49. Sarper N, Çorapçoğlu F, Ank Y, Ural D, Yldz K, Tugay M. Unresectable multifocal hepatoblastoma with cardiac extension: Excellent response with HB-94 chemotherapy protocol. *J Pediatr Hematol Oncol* . 2006;28(6):386-390.

50. Hiyama E, Hishiki T, Watanabe K, et al. Resectability and tumor response after preoperative chemotherapy in hepatoblastoma treated by the Japanese Study Group for Pediatric Liver Tumor (JPLT)-2 protocol.*J Pediatr Surg* . 2016;51(12):2053-2057.

51. Cacciavillano WD, Brugières L, Childs M, et al. Phase II study of high-dose cyclophosphamide in relapsing and/or resistant hepatoblastoma in children: A study from the SIOPEL group. *Eur J Cancer* . 2004;40(15):2274-2279.

LEGENDS OF FIGURES

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 3 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.

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 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; samples grouped in Set-2 are in red; in green, samples grouped in Set-3; in blue, samples grouped in Set-4.

Figure 3. Analysis of gene expression of 13 genes in hepatoblastoma groups according to their clustering. Set-1 presents the upregulation of hepatoblast markers *AFP* , *HNF4* A,*FOX2A* (**A-C**), as well as the high expression level of the epigenetic machinery genes, mainly *UHRF1* and *TET1*(**D-I**). The Set-2 cluster samples revealed the upregulation of the epigenetic machinery genes, highlighting *TET2* (**D-I**), and low expression of markers of both the intermediary stages of differentiation (*AFP* ,*HNF4A* and *FOXA2*) (**A-C**) and mature hepatocyte genes as *ALB* , *UGT1A1* , and *TDO2* (**J-I; M**). Set-3 group the samples showed down-regulation of the epigenetic machinery genes (*TET1, TET2, TET3, DNMT1, DNMT3A, UHRF1*)

(**D-I**), associated with high expression of mature hepatocyte marker genes (*ALB* , *UGT1A1, CYP3A4* , and *TDO2*) (**J-M**).

Figure 4. Global methylation analysis of the three different clusters of hepatoblastomas and liver control samples. On the y-axis, the percentage of methylation is indicated for each sample. 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.

Figure 5. Hepatoblastoma stratification according to global content of DNA methylation and expression profile of DNA methylation enzymes and markers of hepatocyte differentiation. In the middle is 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.

Hosted file

Table1.docx available at https://authorea.com/users/328344/articles/455683-dna-methylation-as-a-key-epigenetics-player-for-hepatoblastoma-stratification

10

Gene Expression – PCA
kmeans grouping