MR. JEAN-WILLIAM DUPUY (Orcid ID : 0000-0002-2448-4797)

DR. ANNE-AURÉLIE RAYMOND (Orcid ID : 0000-0003-2697-9352)

Article type : Original

# Proteomic profiling of hepatocellular adenomas paves the way to new diagnostic and prognostic approaches

Cyril Dourthe[1,2], Céline Julien[1,8], Sylvaine Di Tommaso[1,2], Jean-William Dupuy[3], Nathalie Dugot-Senant[4], Alexandre Brochard[5], Brigitte Le Bail[1,6], Jean-Frédéric Blanc[1,7], Laurence Chiche[1,8], Charles Balabaud[1], Paulette Bioulac-Sage[1], Frédéric Saltel[1,2], Anne-Aurélie Raymond[1,2]

[1] Univ. Bordeaux, INSERM, BaRITOn, U1053, F-33000 Bordeaux, France
[2] Oncoprot Platform, TBM-Core US 005, F-33000 Bordeaux, France
[3] Univ. Bordeaux, Proteome Platform, F-33000 Bordeaux, France
[4] Histopathology Platform, TBM-Core US 005, F-33000 Bordeaux, France
[5] Neurocentre Magendie, INSERM U1215, Bordeaux, France
[6] Department of Pathology, Bordeaux University Hospital, F-33000 Bordeaux, France
[7] Department of Hepatology and Oncology, Bordeaux University Hospital, F-33000 Bordeaux, France
[8] Department of Digestive Surgery, Bordeaux University Hospital, F-33000 Bordeaux, France
These authors contributed equally to this work: Cyril Dourthe and Céline Julien
These authors jointly supervised this work: Frédéric Saltel and Anne-Aurélie Raymond.

**Author Contributions:** A.-A.R., F.S. were responsible for the study, design analysis and for obtaining funding. C.D., C.J, B.L.B, P.B.S, C.B, A.-A.R and F.S. contributed to interpretation of data and writing of the manuscript. C.B., P.B.S., B.L.B, C.J, L.C. and J.-F.B. were responsible for

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/HEP.31826

contributing to the HCA cohort. C.J., P.B.S. and C.B. collected clinical data from patients. S.D.T. was responsible for FFPE tissues preparation and laser microdissection. S.D.T and J.-W.D. have performed the sample preparation for mass spectrometry analysis. J-W.D. was responsible of proteomic analysis by mass spectrometry. J.-W.D. and A.-A.R. were responsible of processing of raw mass spectrometry. C.D., and A.B. were responsible for bioinformatic construction of HCA proteomic database. C.D. and A-A.R were responsible of bioinformatic and biostatistic treatment of proteomic data. N.D.S carried out the cuts, staining and histopathological labellings.

**Keywords**

Machine learning, clinical proteomics, profile matching, diagnosis, biopsies, hepatocellular adenomas, malignancy

**Address correspondence and reprint request to :**

Anne-Aurélie Raymond, Ph.D.

Oncoprot, TBM-Core US 005, INSERM UMR1053,

BaRITOn Bordeaux Research in Translational Oncology

146 rue Leo Saignat

33076 Bordeaux, France

E-mail: anne-aurelie.raymond@inserm.fr

Tel: +33 (0)5 57 57 17 71

or

Frédéric Saltel, Ph.D.

INSERM, UMR1053, BaRITOn Bordeaux Research inTranslational Oncology

146 rue Leo Saignat

33076 Bordeaux, France

E-mail: frederic.saltel@inserm.fr

Tel: +33 (0)5 57 57 17 07

**Abbreviations:** ASS1, argininosuccinate synthase; b-HCA, b-catenin mutated HCA; b-IHCA, b-catenin mutated and inflammatory HCA; BMI, body mass index; CRP, C-reactive protein; FFPE, formalin-fixed and paraffin-embedded; FNH, focal nodular hyperplasia; GPC3, Glypican 3; GS, glutamine synthetase; GSEA, Gene Set Enrichment Analysis; HCAs, hepatocellular adenomas; HCC, hepatocellular carcinoma; H-HCA, HNF1A mutated HCA; HNF1A, hepatocyte nuclear factor 1A; Hsp70, Heat Shock Protein 70; IHC, immunohistochemistry; IHCA, inflammatory HCA; LFABP, liver-type fatty acid binding protein; NT, nontumoral; PCA, Principal Component Analysis; SAA, serum amyloid A; shHCA, sonic hedgehog HCA; T, tumoral; TERT, Telomerase Reverse Transcriptase.

**Financial supports:** This work was supported by the Nouvelle Aquitaine Region (European FEDER Funds), the SIRIC BRIO, and the HepA association.

**Abstract**

**Background and aims:** Through an exploratory proteomic approach based on typical hepatocellular adenomas (HCA), we previously identified a new diagnostic biomarker for a distinctive subtype of HCA with high risk of bleeding, already validated on a multicenter cohort. We hypothesized that the whole protein expression deregulation profile could deliver much more informative data for tumors characterization. Therefore, we pursued our analysis with the characterization of HCAs proteomic profiles, evaluating their correspondence with the established genotype/phenotype classification and assessing whether they could provide added diagnosis and prognosis values.

**Approach & Results *:*** From a collection of 260 cases, we selected 52 typical cases of all different subgroups on which we built the first HCA proteomics database. Combining laser microdissection and mass spectrometry based proteomic analysis, we compared the relative protein abundances between tumoral (T) and non-tumoral (NT) liver tissues from each patient and we defined specific proteomic profile of each HCA sub-groups. Next, we built a matching algorithm comparing proteomic profile extracted from a patient with our reference HCA database. Proteomic profiles allowed HCA classification and made diagnosis possible, even for complexes cases with immunohistological or genomic analysis that did not lead to a formal conclusion. Despite a well-established pathomolecular classification, clinical practices have not substantially changed and HCA management link to the assessment of the malignant transformation risk

remains delicate for many surgeons. That's why we also identified and validated a proteomic profile that directly evaluate malignant transformation risk regardless of HCA subtype.

**Conclusions:** This pioneering work proposes a proteomic-based machine learning tool, operational on fixed biopsies, that can improve diagnosis and prognosis and therefore patient management for HCA.

## Introduction

Pathology diagnosis and management, especially for a tumor, require different clinical expertise. Today, medical imaging has made enormous progress and provides very effective results in terms of diagnosis and tumor prognosis[1,2]. However, one of the major challenges often remains to determine whether a tumor is malignant or not. If it is benign, it's important to know whether it still presents clinical risks, including malignant transformation into cancer. If a cancer is identified, the maximum of biological information must be obtained in order to propose the best possible patient management. Biopsy is an invasive procedure that is generally the first unavoidable step required to establish the diagnosis. The histological analysis of biopsies is an essential tool for diagnosis complementary to medical imaging. Pathologists can analyze the morphological features of tissues and visualize biomarker expression by immunohistochemistry (IHC) to determine the nature of tumor. Today, various biomarkers are currently available to support diagnosis and prognosis, and theranostics and pathologists use serial immunostainings to investigate and evaluate the abundance of an increasing number of biomarkers. Other molecular approaches, such as PCR (Polymerase Chain Reaction) and FISH (Fluorescence *in situ* hybridization), are routinely performed on human samples. Recent technological advances now make it possible to carry out very advanced genomic (next generation sequencing) and transcriptomic (RNA sequencing) analyses from frozen and even formalin-fixed and paraffin-embedded (FFPE) tissues. Nevertheless, the latter are not completely suited to such analyses given the degradation of genetic material. These molecular analyses can be used for mutation analysis or for detecting the presence of certain transcripts[3] in routine care samples. However, in addition to the IHC analyses performed, they often exhaust the biological sample. Downstream of translation, proteomic analysis allows the extraction of a considerable amount of information on the functional status of a disease, directly identifying and quantifying potential biomarker proteins[4] and pharmacological targets usable in clinical practice[5]. However, to date proteomics is not effective in diagnosing benign versus malignant tumors.

In-depth analysis of FFPE tissue proteomes is a technique now commonly used in translational research[6,7]. Experts in histoproteomics agree that proteomic analysis will integrate the surgical pathologist field in the near future[6,8,9]. Mass spectrometry already meets some specific

requirements for biomarker identification[10,11]. We have already combined FFPE tissue proteomics with laser microdissection to enable the precise selection of a tissue area of interest, and we have optimized this technique for compatibility with very small amounts of material, such as needle microbiopsies[4,7]. Indeed, a 1 $mm^2$ area of a 5 µm thick FFPE section is sufficient to compare deregulations in protein expression of 1000 proteins in pathological versus healthy tissues. Previously, using this method in an exploratory approach based on typical cases of HCA, we identified Argininosuccinate Synthase (ASS1) as diagnostic biomarker for a distinctive subtype of hepatocellular adenomas with high bleeding risk[7], which we then validated on an enlarged multicenter cohort[4]. We also showed that when in doubt of IHC staining interpretation, especially on small liver biopsies, HCA biomarker quantification by mass spectrometry could in fact be a valuable complement of information in diagnosis support[4]. Beyond a single biomarker, we hypothesized that the whole profile of protein expression deregulation could be diagnostic in a much more complete way. In this study, our goal was to develop proteomic profiling to improve diagnosis and therefore patient management. We pursued our analysis of HCA in order to characterize their overall proteomic profiles, evaluate their correspondence with the established molecular classification and assess whether they could provide added diagnostic value. HCA mainly affects women in their middle ages and men to a smaller extent. The main risk factor for HCA is hormonal exposure to estrogens or androgens, but metabolic, vascular, glycogen storage diseases, and some other rare genetic diseases have also been associated with the development of HCA[12,13]. Bleeding[14] and malignant transformation into hepatocellular carcinoma (HCC)[15] are the two major complications, both strongly related to the size of the adenoma and its pathomolecular subtype. A resection is usually recommended when the HCA reaches 5 cm[16] and when the β-Catenin pathway is activated. In order to specify the cases with significant clinical risk, HCA have been classified into four groups based on the identification of mutations and on corresponding clinical, histological, and immunohistological characteristics[17,18,15,19,20]: (1) H-HCA with inactivating mutations of *HNF1A*. L-FABP, whose expression is controlled by *HNF1A*, is not immunodetected in H-HCA tumor cells[19]; (2) Inflammatory HCAs (IHCA) exhibit various mutations (IL6ST, FRK, STAT3, JAK1, GNAS) or chromosome alterations that activate the JAK/STAT signaling pathway[18,21]. Neoplastic hepatocytes show strong and diffuse immunoreactivity for the acute phase inflammatory proteins SAA and C-reactive protein (CRP)[22,23]; (3) b-HCA with activating mutations of the *CTNNB1* gene encoding β-catenin (exon 3 at various hot spots, including S45 or exon 7/8)[24,25]. The *GLUL* gene, a β-catenin target, encoding Glutamine Synthetase (GS) is strongly detected in non-S45 exon 3 b-HCA tumors[15,24]. B-HCA exon 3 non-S45 tumors are at high risk of malignant transformation into HCC, particularly if they are associated with mutations in the Telomerase Reverse Transcriptase (TERT) promoter[26]. In

addition, half of β-catenin-mutated HCAs also exhibit inflammatory features (b-IHCA)[18,25]; (4) The last group corresponds to sh-HCA, defined by microdeletions fusing the INHBE promotor with the GLI gene. Sh-HCA is identified by the overexpression of ASS1, and has a high hemorrhage risk which is a major clinical issue[4,7,18].

The genotype/phenotype classification of HCA is currently considered as almost complete from a scientific point of view. However, the mechanisms underlying malignant transformation and bleeding remain unknown. In addition, the daily management of patients with HCA remains a challenge for many surgeons and hepatologists. The first challenge is the differentiation of HCA from other benign liver "tumors". These other tumors are characterized as hyperplastic in nature: Focal Nodular Hyperplasia (FNH). For some cases, differential diagnosis between HCA and FNH remains challenging by imaging and biopsy analysis, even for liver pathologist experts[27]. This is notably because GS is also upregulated in FNH. As FNH does not hemorrhage or undergo malignant transformation, resection is not indicated and monitoring is not mandatory. After HCA diagnosis, the second challenge lies with the determination of the risk of transformation into HCC and hemorrhage in order to propose resection when appropriate (transplantation in rare cases), taking into account the risk-benefit ratio associated with surgery.

With this study we prove that proteomic-based machine learning analysis is operational in routine clinical practice and can improve diagnosis and therefore patient management for HCA. This type of approach could be a breakthrough that is transposable to other pathologies.

## Material and Methods

### Patients

All patients gave informed consent and this study was approved by our local committee "Direction de la Recherche Clinique et de l'Innovation" of Bordeaux University Hospital, Bordeaux Liver Biobank BB-0033-00036.

### HCA database construction

The 260 resected HCA of our cohort (227 female/33 males), included for study between 1984 and May 2020, were routinely classified by histopathological analysis and IHC. Seventy one percent (185/260 cases) of cases were also analyzed by molecular biology primarily on resected specimen to identify genetic mutations. Clinical data and patient management history were collected for all cases.

We selected 52 samples that were surgical resections of HCA and representative of all the different subtypes as defined by the 2019 WHO classification[20] (6 H-HCA, 5 IHCA, 16 b-HCA (4

b-HCA exon 3 non S45, 5 b-HCA exon 3 S45, 5 b-HCA exon 7/8 plus two cases of b-HCA (1 exon 3 S45 and 1 exon 7/8), sent for second opinion, were added in this collection), 16 b-IHCA (6 b-IHCA exon 3 non S45, 4 b-IHCA exon 3 S45, 6 b-IHCA exon 7/8), 9 sh-HCA). The diagnosis of HCA and HCA subtype was considered as characteristic by the pathologist according to morphology and immunostainings, and results were always confirmed by genotyping. Main clinical data (age, sex, BMI, mode of discovery) were summarized in **Supplemental Table 1**. Five resected cases of FNH were also included as controls (**Supplemental Table 1**).

**Selection of cases for validation of proteomic profiles**

To validate the diagnostic applicability of the proteomic profiles we identified, we selected a panel of 11 cases of biopsies and 1 resection for sample classification (**Supplemental Table 2**) and 4 resections for malignancy status (**Supplemental Table 3**).

**Distance calculations and profile matching**

Analyzes were performed using the free R software. Hierarchical clusterings were represented *via* the R package and its "hclust" function. The "ward.D" method was used as the agglomeration method and the Euclidean distance was used for the distance calculation. For Principal Component Analysis (PCA), missing values were imputed using the k-Nearest Neighbor (KNN) imputation method from the VIM package. PCA data was calculated and formatted by the PCA and fviz_pca_ind functions in the factoextra and FactoMineR packages.

Machine Learning is a method of automatic learning by the computer through training with sets of data, in order to identify patterns, i.e. recurring patterns, in the data set. In our case, the training data sets correspond to the proteomic profiles for each of the HCA groups (**Supplemental Figure 1a**). The lists of proteins constituting the reference proteomic profiles of FNH and of each HCA subtype has been detailed in the **Supplemental Table 4**.

The second step of the process is to test our training algorithm with several methods to calculate the similarities between the profile of a biopsy and the profiles of each HCA group: (1) a Chi-square test to compare upregulated proteins (T/NT ratio≥1.5), down-regulated proteins (T/NT ratio≤0.67), and non-regulated proteins (T/NT ratio≥0.67 and ≤1.5) between the different reference proteomic profiles; (2) a calculation of Euclidean distance between each group from the PCA; (3) a Random Forest analysis on PCA-reduced datasets (**Supplemental Figure 1b**). The diagnostic tool learns and refines the reference proteomic profiles as the database is incremented.

Euclidean distances were calculated using the dist function and Chi-square test of independence using the chisq.test function (both within the R package). The Random Forests were generated

using the randomForest function in the randomForest package. The optimal parameters were calculated by the train function in the caret package. The Chi-square scoring method was not applicable for the malignant profile consisting of ten proteins.

Proteomics analysis mass spectrometry data processing, integrative biological analyses, and graphical representations were performed as described in **Supplemental Materials and Methods**.

## Results

### Creation of a reference HCA proteomic database

In order to construct a robust database from our collection of 260 cases, we selected 52 HCA surgical resection cases that were representative of all HCA subtypes and FNH. The diagnosis of these resections was considered typical by the pathologist and confirmed by genetic analysis (**Supplemental Table 1**).

We compared the relative protein abundances between tumoral (T) and non-tumoral (NT) liver tissues from each patient. To begin, we confirmed the specific and expected deregulation of GS, CRP/SAA, ASS1, and LFABP markers in the corresponding neoplasms: b-HCA/b-IHCA exon 3 non S45, IHCA/b-IHCA, sh-HCA and H-HCA respectively (**Figure 1a**), validating the robustness of our database.

We assumed that the protein signature for each HCA subtype would reflect the underlying mutation(s) that defined them. Therefore, we examined the proteins functionally associated to the underlying mutated genes defining each HCA subtype (*HNF1A*, *CTNNB1*, JAK/STAT signaling pathway). Surprisingly, with exception of the validated target proteins currently used as immunomarkers (L-FABP, GS, CRP/SAA), we did not observe major deregulation of the other proteins frequently found associated with each mutation, as illustrated with ALB[28], OAT[29], and MIF[30] (**Figure 1b, c and d**). In the same way, the proteins functionally associated with the GLI1 gene were no more modified in sh-HCA than in other subtypes as illustrated with MMP9[31] (**Supplemental Figure 2**). In addition, we explored the previously described typical/specific pathways for each HCA group from a transcriptomic approach[18]. We found all these pathways significantly deregulated in all HCA subtypes by means of unsupervised analysis (**Figure 1e**). We noted within these results that proteins associated with HCC development were significantly deregulated in all HCA subtypes, suggesting thus that all subtypes can potentially be transformed.

We then investigated significant differences between HCA and FNH (**Figure 2a**). The proteomic profiles of HCA and FNH were quite distinct (**Figure 2a**). Then, we tested whether it was possible to differentiate the different HCA subtype proteomic profiles. We first separated the most dissimilar profiles: H-HCA and sh-HCA **(Figure 2b).** Interestingly, the inflammatory protein expression profile was a strong signature shared by inflammatory HCA (IHCA and b-IHCA) which were distinct from the other HCA **(Figure 2c)**. The proteomic profile of b-HCA was not well distinguished from the other HCA by Principal Component Analysis (PCA) (**Figure 2d**), even following hierarchical clustering of b-HCA (**Supplemental Figure 3).** Yet, when IHCA and b-IHCA were specifically compared, their proteomic profiles were markedly different **(Figure 2e)**.

Moreover, their protein expression profiles were clearly grouped according to the type of *CTNNB1* mutation (exon 3 non S45, exon 3 S45, or exon 7/8) for both b-HCA and b-IHCA (**Figure 2f**).

These results demonstrated the excellent correlation between the proteomic profiles and the genotype/phenotype classification and revealed the existence of both dominant (sh-HCA, inflammatory H-HCA) and less obvious profiles (*CTNNB1* mutations).

We then examined the biological functions that were associated with each specific profile. Gene Set Enrichment Analysis (GSEA) revealed a major difference between FNH and HCA. This was the presence of mitochondrial disorders in HCA, including disruptions in amino acid metabolism that were not significant in FNH (**Figure 2g**). We found that most of the functional deregulations were shared by the different HCA subtypes (**Figure 2g**). As expected, the most significantly deregulated pathway for H-HCA was the fatty acid beta-oxidation pathway related to steatosis, which is characteristically observed in these tumors but was also found in inflammatory HCA. These data may be consistent with the fact that 23/80 cases of IHCA were steatotic in our collection. H-HCAs also showed strong deregulations in amino acid metabolism and xenobiotic metabolism, a significant deregulation shared with sh-HCA, for which we also found the previously described urea cycle deregulation[7].

The most significantly enriched pathways for inflammatory HCA (IHCA and b-IHCA) were, as anticipated, related to activation of the inflammatory response (FXR/RXR activation, acute-phase response signaling). Stress-related signals regulating mRNA translation associated with Eif2 signaling pathways were also significantly enriched in inflammatory HCA, similarly to the sh-HCA-specific proteomic profile (**Figure 2g**). The b-HCA subtype could not be distinguished by the enrichment of a specific pathway, including even the β-catenin pathway (**Figure 2g**).

This integrative analysis revealed that genetic mutations did not translate into a strictly clear functional classification at the protein level.

**A machine learning tool for HCA and FNH diagnosis**

Given our proteomic data enabled us to distinguish each different HCA subtype, we decided to set up a diagnostic tool. The first important step was to define whether the sample corresponded to a HCA or an FNH, and not another type of well-differentiated liver tumor. To examine this, we analyzed a low-grade dysplastic nodule in a cirrhotic liver and a well-differentiated HCC. For both cases, the similarity scores with HCA or FNH were very low, ruling out the possibility of erroneous attribution and indicating that tumors were neither HCA nor FNH (**Supplemental Table 2**).

Next, we built an algorithm that integrated first the main HCA groups identified (sh-HCA, H-HCA, Inflammatory (IHCA and b-IHCA) and then the secondary proteomic profiles (b-HCA according to the different *CTNNB1* mutations) (**Figure 3a)**. The principle being thus to sequentially compare

the HCA subtype reference proteomic profiles with new cases in order to diagnose their HCA subtype (**Figure 3b, c, and d**). We used three approaches to calculate the similarities: (1) a Chi-square test to compare T/NT deregulations with the different reference proteomic profiles; (2) a calculation of Euclidean distance between each group; (3) a Random Forest analysis on PCA-reduced datasets. We applied them to a panel of 11 biopsies and 1 resection of FNH and HCA (any subtype) to test the advantages and limitations of each analysis (**Supplemental Table 2**).

Three out of the 12 cases were typical b-HCA exon 3 non S45, sh-HCA and H-HCA (R_01, BP_01, and BP_02, respectively) and did not present any difficulty in classification (**Supplemental Table 2, Figure 3b, c, d and e**). The other nine cases consisted of a representative panel of biopsies diagnosed in our tertiary center (**Figure 3e, Supplemental Table 2**). HCA biomarkers were noncontributory to the subtyping of any of these cases (**Supplemental Table 2**). However, the proteomic profiles were consistent with the clinicopathological interpretation for three cases (BP_03, _04 and, _05). For two other cases (BP_06 and _07), Random Forest gave the correct diagnosis, one another test matched with the pathological diagnosis and the last one test gave a different result. These cases were less typical and did not perfectly match with the HCA in the database. Case BP_08 was a b-HCA exon 7/8 subtype subsequently identified by genetic analysis on resection. Random Forest and Euclidean distance firmly recognized it as a b-HCA exon 7/8 subtype, but Chi-square testing grouped it with the very similar b-HCA exon 3 S45 group. Finally, we analysed three biopsies without subsequent resection (BP_09, _10, and _11), subtyped as an IHCA or for which the pathologist could not conclude between a IHCA and a b-IHCA (exon 7/8 or exon 3 S45). These cases were found to group with b-IHCA exon 7/8 or exon 3 S45 profiles (**Figure 3e, Supplemental Table 2**).

In conclusion, these results demonstrate that proteomic profile matching can differentiate HCA from FNH and assign the molecular HCA subtype. In case of difficulties, in particular with non-contributory HCA biomarkers (**Supplemental Figure 4**), proteomic profiling can bring additional clues that support diagnosis.

**Identification of the proteomic profile of malignancy for transformed HCA**

Despite a well-established pathomolecular classification[20], clinical practices have not substantially changed and HCA management remains delicate for many surgeons. To illustrate it in our center, the molecular classification has not modified HCA management according to subgroup type. This is despite a slight tendency to perform more preoperative biopsies (**Supplemental Figure 5**).

As an indication, we reviewed HCA management in our center (Bordeaux, France) from the initial reports. Twenty-seven percent (70/260 cases) of cases had been biopsied prior to surgery (**Supplemental Figure 6a**). Among these 70 biopsies, 43 (61.4%) were contributory to the

diagnosis of HCA (**Supplemental Figure 6b**). When biopsy interpretation was not formal, it was either due to technical noncontributory factors, such as lack of interpretable material (9 cases), or to the impossibility of making a standard differential diagnosis with well-differentiated HCC (15 cases) or FNH (3 cases) using standard pathological tools (**Supplemental Figure 6b**).

These data, based on our experience, revealed that even for a well-characterized tumor such as HCA and expert pathologists in the field, an additional diagnostic support tool could have been useful, in particular to provide additional information on benign/malignant status.

Moreover, malignant transformation is a major complication of HCA modifying surgical management and we wanted to determine if our proteomic profile-based tool could help to address this issue.

The HCA cases in our collection can transformed into HCC regardless of the HCA subtype. (**Figure 4a**). Even if b-HCA exon 3 non S45 subtype showed the highest risk of malignant transformation, a non-negligible proportion of cases from other subgroups could be concerned (up to 13.6% for sh-HCA) (**Figure 4a**). It is noteworthy that the interpretation of histological features of malignancy can be challenging, depends on the expertise of the pathologist, and can be observer-dependent.

In half of these cases we noticed a reservation on the pathologist's behalf. This was described by the term "borderline" and defined as a neoplasm that was no longer characteristic of HCA, while at the same time presenting atypical foci lacking some characteristics of definite HCC (**Figure 4a and b**). In addition to the clinical and imaging context, the assertion of HCC diagnosis in histology is based on a set of features that include morphological criteria, such as cytonuclear atypia (**Figure 4c and d)**, but also immunohistochemical positivity (for example Glypican 3 (GPC3) and Heat Shock Protein 70 (HSP70))[32], and TERT promotor mutation[33]. These criteria are not completely specific and sensitive. Therefore, it appeared pragmatic to identify a proteomic signature of malignancy in order to improve patient management.

**Identification of the HCC developed on HCA malignancy profile**

To identify proteins related to malignancy in HCA, we selected six cases of surgical resection that presented definite HCC developed on HCA (we refer to this as "HCC/HCA") from the different HCA subtypes (2 b-HCA exon 3 non S45, 1 b-IHCA exon 3 S45, 1 UHCA, 2 sh-HCA) (**Supplemental Table 3**). As previously, we compared the T/NT ratios of six HCA and their corresponding HCC/HCA and isolated a significant proteomic profile consisting of ten proteins. These proteins allowed the perfect separation of malignant and benign tissues, both by hierarchical clustering (**Figure 5a)** and PCA (**Figure 5b**). Moreover, the protein expression deregulations within each group were homogeneous (**Figure 5c**).

In this malignant proteomic profile, half of the extracted functional annotations of the malignancy protein profiles were associated with immune response and immune-related cell activation (MPO, PRDX2, LCP1, PPIA, SERPINA1, PLD3), Interleukin-12 signaling (PPIA, LCP1), and neutrophil degranulation (PPIA, SERPINA1, MPO) (**Figure 5d**). Among these ten proteins, none were part of the known β-catenin functional environment.

We then wanted to validate this proteomic profile of HCA transformation with a set of new cases. The Chi-square scoring method was not applicable for a profile consisting of ten proteins. With Euclidean distance and Random Forest, we first tested our ten-protein proteomic profile of transformation on one positive control of HCC developed on HCA and 7 cases of benign HCA from our database.

**Validation with a control HCC developed on HCA**

Case 167 (**Supplemental Table 3)** was a HCC developed on HCA without any histological ambiguity and confirmed by molecular analysis. The proteomic profile was closer to the transformed cases, confirmed by the Random Forest test and the smaller Euclidean distance for the HCC group (5.18 for HCA vs 5.43 for HCC) (**Supplemental Table 3, Figure 6b and c**). Our proteomic malignant profile allowed confirmation of the HCC diagnosis.

**Validation with benign controls**

Next, we selected seven cases of HCA (cases 105, 116, 121, 135, 119, 83, 218) that were not questionable regarding their diagnosis of benign tumors by histology (**Figure 6b and c and Supplemental Table 1**). The proteomic profiles of all these tumors confirmed the benign nature of all of these cases (**Figure 6b and c**).

**Cases with doubt on the diagnosis of malignancy**

We then selected cases with a management history raising doubts concerning the diagnosis of malignancy. The first case (178) was suspected to be malignant at the time of biopsy analysis. The conclusive histological analysis of the surgical resection was clear on the benign nature of this HCA, which was later identified as a sh-HCA. In order to reproduce a clinical context, we analyzed the proteome of the preoperative biopsy that had initially raised the doubt. Unambiguously, the proteomic profile was closer benign HCA, confirmed by Random Forest testing and the smallest Euclidean distance for HCA (2.99 for HCA vs 5.01 for HCC) (**Supplemental Table 3, Figure 6b and c**).

Next, we selected a “borderline” case. Case 217 manifested intratumoral bleeding by imaging in a growing and unclassified HCA. The presence of an intratumoral hematoma made interpretation difficult and it was challenging to distinguish between HCA and HCC due to the presence of cytological atypias and abnormal sinusoids (CD34-positive and reduced reticulin staining). Molecular biology analysis found *a posteriori* diagnosis was a sh-HCA. The proteomic profile was very clearly associated with the HCC group, as confirmed by the Random Forest test. However, the Euclidean distance calculation did not allow a decision between the two groups (4.55 for HCA vs 4.65 for HCC). For this specific case, proteomics provided additional quantitative features that allowed the positioning of this neoplasm on a progressive scale between HCA and HCC (**Supplemental Table 3, Figure 6a, b and c**).

The last case (case 280, **Supplemental Table 3, Figure 6a, b and c**) was a female initially managed remotely for a benign 10 cm liver tumor. Histological analysis of the resection specimen was a challenge for the pathologists: unclassifiable HCA combined with diffuse cytological atypia and an abnormal reticulin network. This thus raised the diagnosis of probable HCC, but without HCC-favoring IHC markers (GPC3- and MIB1-negative) and a negative molecular analysis (TERT-negative). The proteomic profile clearly allowed the classification of this case with the HCC developed on HCA group (4.27 for HCA vs 4.99 for HCC).

In view of these results on this validation patient set, the proteomic malignant profile was effective for the determination of HCA engaged in a process of transformation.

Within this study we offer a complete tool for HCA diagnosis that can determine HCA subtype and a score reflecting its level of malignant transformation (**Supplemental Figure 7**).

**Discussion**

Proteomic analysis is the next omic step following on from genomics and transcriptomics to improve patient management. Until now, proteomic profiling of FFPE tissue had never been used in clinical practice, especially in the tumor pathology field.

We have carried out pioneering work using proteomic profile matching for diagnosis from FFPE tissues. The main principle of our tool is based on the comparison of a proteomic profile extracted from a biopsy or surgical specimen with a reference database composed of proteomic profiles from completely characterized patients. We have developed and adapted the analytical process for operational routine analysis. We use a small amount of tissue corresponding to 1 mm² on 3x5 µm section cuts, thus not depleting the sample and it can be reused. A complete analysis can be achieved within a week and so it remains compatible with the deadlines imposed by clinical practice. Moreover, the proteomic data generated remains a resource that can be re-exploited for further applications (prognosis, theranostics).

The main advantage of working on a profile of protein deregulation (tumoral vs non-tumoral) is the ability to directly access the tumor identity by reducing inter-individual variation, making the analysis efficient despite small patient collections. This also allows the association of profiles with functional biological pathways, the control of biological relevance, and identification of new targets. The additional advantage of working on a profile of protein expression deregulation versus spectral intensity profiles is that the construction of the reference database is not dependent on mass spectrometer type. Indeed, our method can be implemented on sites with different equipment, contrary to radiomics that encounter major obstacles concerning the heterogeneity between devices used for image acquisition[34].

In this study, we established our proof of concept on HCA. The molecular classification of HCA has been subject to many updates over the last ten years and is now considered almost established[18,35]. It must be noticed that while the enrichment of the molecular classification has led to a better understanding of HCA physiopathology, it has not modified the management of patients. We identified two main issues: (1) classification based on biopsies and (2) identification of the malignancy status. Indeed, for several reasons the diagnosis on biopsies using classical approaches is not always evident. Moreover, the fact that potentially all HCA can transform[36] can induce doubt at the time of diagnosis, especially for a pathologist who is not specialized in benign liver tumors. A suspicion of HCC significantly modifies the medical and surgical treatment and follow-up. Indeed, the surgical strategy is not thought-out in the same way; a simple enucleation of the tumor is sufficient when dealing with HCA. On the contrary, the surgical management of HCC consists, at best, in an anatomical resection (removing the corresponding portal branch), and when this it is not possible, the surgeon must strive to have an optimal margin of 2 cm.

For all these reasons, we first generated an HCA database for HCA subtype proteomic classification. Our study revealed such a significant disparity between genomic/transcriptomic data and protein expression profiles of HCAs, especially for the functional environment of β-catenin which is not particularly activated in b-HCA exon3 nonS45. Proteins constitute the functional elements whose expression depends not only on the mutation status of a tumor but also by many epigenetic parameters. This disparity can also be explained by the post-transcriptional regulations that classically lead to observe an overlap of barely half between transcriptome and proteome[37]. As an illustration, a positive glutamine synthetase rim has been reported with less or no expression in the tumor center in b-HCA S45 and exon7/8[38]. Secondly, we identified a protein profile that directly diagnoses malignant transformation regardless of HCA subtype. This profile was correlated with immune activation in response to cell degeneration, making sense for tumors transforming or differentiated tumors with active anti-tumor immunity[39,40]. It is likely that the β-catenin pathway did not emerge in our malignancy signature

given it was identified across all HCA subtypes. This validated signature allows us to define the HCA transformation status.

Machine learning is an emerging and promising disciplinary field for clinical data interpretation[41]. In our study, we tested different statistical analyses for calculations of similarity (difference score calculation: Chi-square test on simplified data, Euclidean distance calculation after PCA, Random Forest on data reduced by PCA). Without preconceived ideas, we wanted to use some of the analyses to provide evidence on the similarities of the proteomic profiles we were testing with respect to the reference groups. Although Random Forest was proven the most reliable, these methods were all successful. To interpret this, if the three mathematical approaches lead to the same conclusion, it is evidence of robustness. On the other hand, if they conflict with respect to atypical cases, it is necessary to be more cautious and consider the analyzes as additional features in diagnosis. Ongoing case addition to our HCA database will make the diagnosis more and more precise.

Thereafter, our methodology could be implemented to answer other critical clinical questions in the liver field, such as well-differentiated HCC for which differential diagnosis can be confusing in a common clinical context[42]. This could define whether the malignancy signature we identified is strictly related to HCA or is common to other more frequent etiologies and backgrounds.

From a broader point of view, the clinical recommendations for performing liver biopsies on liver tumors and peritumoral tissues are not systematic and must follow a logical benefit over risk assessment[43]. Indeed, this invasive procedure is not insignificant and can lead to complications, sometimes serious, such as hemorrhage and tumor cell spread. The value of biopsy performance is also under debate given advances in imaging. Indeed, radiologists are able to make increasingly accurate diagnoses, even by elimination. This already aids in management and especially in the decision for resection. For these reasons, the role of biopsy for patients with liver disease is one of the most important considerations among hepatologists. Nonetheless, information provided by the proteomic profile could reverse the benefit-risk ratio of performing liver biopsy, making a real difference in clinical practice and pave a way for personalized medicine. For example, malignancy diagnosis could indeed be very useful for dysplastic liver nodules found in cirrhotic livers because the malignancy conditions patient eligibility for transplantation[44]. In the future, proteomic patterns could also be used to identify elements of response or non-response to anti-cancer treatments, for example in patients with advanced HCC for which the therapeutic arsenal has recently been expanded[45].

In conclusion, we show that proteomic profiling could give biopsy a more important status in the patient management process. The diagnosis and subtyping of HCA and the evaluation of its malignant transformation is available for a transfer to routine clinical care in our center

(**Supplemental Figure 7**). This implies the integration of proteomic analysis into the clinics and the addition of proteomic results into patient pathology reports as part of the evidence-based care bundle. Proteomic analysis will bring pathology services into a new analytical era.

**References**


1. Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. CA Cancer J Clin 2019:caac.21552.

2. Hricak H. 2016 New Horizons Lecture: Beyond Imaging—Radiology of Tomorrow. Radiology 2018;286:764–775.

3. Bass BP, Engel KB, Greytak SR, et al. A Review of Preanalytical Factors Affecting Molecular, Protein, and Morphological Analysis of Formalin-Fixed, Paraffin-Embedded (FFPE) Tissue: How Well Do You Know Your FFPE Specimen? Arch Pathol Lab Med 2014;138:1520–1530.

4. Sala M, Gonzales D, Leste-Lasserre T, et al. ASS1 Overexpression: A Hallmark of Sonic Hedgehog Hepatocellular Adenomas; Recommendations for Clinical Practice. Hepatol Commun 2020;4:809–824.

5. Gustafsson OJR, Arentz G, Hoffmann P. Proteomic developments in the analysis of formalin-fixed tissue. Biochim Biophys Acta BBA - Proteins Proteomics 2015;1854:559–580.

6. Coscia F, Doll S, Bech JM, et al. A streamlined mass spectrometry-based proteomics workflow for large-scale FFPE tissue analysis. J Pathol 2020;251:100–112.

7. Henriet E, Abou Hammoud A, Dupuy J-W, et al. Argininosuccinate synthase 1 (ASS1): A marker of unclassified hepatocellular adenoma and high bleeding risk. Hepatol Baltim Md 2017;66:2016–2028.

8. Ahmed M, Broeckx G, Baggerman G, et al. Next-generation protein analysis in the pathology department. J Clin Pathol 2020;73:1–6.

9. Jin P, Lan J, Wang K, et al. Pathology, proteomics and the pathway to personalised medicine. Expert Rev Proteomics 2018;15:231–243.

10. Ahmed M, Broeckx G, Baggerman G, et al. Next-generation protein analysis in the pathology department. J Clin Pathol 2020;73:1–6.

11. Longuespée R, Casadonte R, Schwamborn K, et al. Proteomics in Pathology. PROTEOMICS 2018;18:1700361.

12. Edmondson HA, Henderson B, Benton B. Liver-Cell Adenomas Associated with Use of Oral Contraceptives. N Engl J Med 1976;294:470–472.

13. Nault J, Bioulac–Sage P, Zucman–Rossi J. Hepatocellular Benign Tumors—From Molecular Classification to Personalized Clinical Care. Gastroenterology 2013;144:888–902.

14. Aalten SM van, Man RA de, IJzermans JNM, et al. Systematic review of haemorrhage and rupture of hepatocellular adenomas. Br J Surg 2012;99:911–916.

15. Stoot JHMB, Coelen RJS, De Jong MC, et al. Malignant transformation of hepatocellular adenomas into hepatocellular carcinomas: a systematic review including more than 1600 adenoma cases. HPB 2010;12:509–522.

16. Thomeer MG, Broker M, Verheij J, et al. Hepatocellular adenoma: when and how to treat? Update of current evidence. Ther Adv Gastroenterol 2016;9:898–912.

17. Zucman-Rossi J, Villanueva A, Nault J-C, et al. Genetic Landscape and Biomarkers of Hepatocellular Carcinoma. Gastroenterology 2015;149:1226-1239.e4.

18. Nault J-C, Couchy G, Balabaud C, et al. Molecular Classification of Hepatocellular Adenoma Associates With Risk Factors, Bleeding, and Malignant Transformation. Gastroenterology 2017;152:880-894.e6.

19. Bioulac-Sage P, Rebouissou S, Thomas C, et al. Hepatocellular adenoma subtype classification using molecular markers and immunohistochemistry. Hepatology 2007;46:740–748.

20. Bioulac Sage P. Hepatocellular Adenoma. WHO Classification of Tumours. 5th Ed Geneva Switz Int Agency Res Cancer IARC Publ 2019:224–228.

21. Bayard Q, Caruso S, Couchy G, et al. Recurrent chromosomal rearrangements of *ROS1* , *FRK* and *IL6* activating JAK/STAT pathway in inflammatory hepatocellular adenomas. Gut 2020;69:1667–1676.

22. Rebouissou S, Amessou M, Couchy G, et al. Frequent in-frame somatic deletions activate gp130 in inflammatory hepatocellular tumours. Nature 2009;457:200–204.

23. Pilati C, Zucman-Rossi J. Mutations leading to constitutive active gp130/JAK1/STAT3 pathway. Cytokine Growth Factor Rev 2015;26:499–506.

24. Bioulac-Sage P, Rebouissou S, Thomas C, et al. Hepatocellular adenoma subtype classification using molecular markers and immunohistochemistry. Hepatology 2007;46:740–748.

25. Rebouissou S, Franconi A, Calderaro J, et al. Genotype-phenotype correlation of *CTNNB1* mutations reveals different ß-catenin activity associated with liver tumor progression: Hepatology, Vol. XX, No. X, 2016 Rebouissou, Franconi, et al. Hepatology 2016;64:2047–2061.

26. Nault JC, Mallet M, Pilati C, et al. High frequency of telomerase reverse-transcriptase promoter somatic mutations in hepatocellular carcinoma and preneoplastic lesions. Nat Commun 2013;4:2218.

27. Balabaud C, Al-Rabih WR, Chen P-J, et al. Focal Nodular Hyperplasia and Hepatocellular Adenoma around the World Viewed through the Scope of the Immunopathological Classification. Int J Hepatol 2013;2013:1–12.

28. Patitucci C, Couchy G, Bagattin A, et al. Hepatocyte nuclear factor 1α suppresses steatosis-associated liver cancer by inhibiting PPARγ transcription. J Clin Invest 2017;127:1873–1888.

29. Cadoret A, Ovejero C, Terris B, et al. New targets of β-catenin signaling in the liver are involved in the glutamine metabolism. Oncogene 2002;21:8293–8301.

30. Ohta S, Misawa A, Fukaya R, et al. Macrophage migration inhibitory factor (MIF) promotes cell survival and proliferation of neural stem/progenitor cells. J Cell Sci 2012;125:3210–3220.

31. Chen J-S, Li H-S, Huang J-Q, et al. Down-regulation of Gli-1 inhibits hepatocellular carcinoma cell migration and invasion. Mol Cell Biochem 2014;393:283–291.

32. Di Tommaso L, Destro A, Seok JY, et al. The application of markers (HSP70 GPC3 and GS) in liver biopsies is useful for detection of hepatocellular carcinoma. J Hepatol 2009;50:746–754.

33. Nault J-C, Zucman-Rossi J. TERT promoter mutations in primary liver tumors. Clin Res Hepatol Gastroenterol 2016;40:9–14.

34. Rizzo S, Botta F, Raimondi S, et al. Radiomics: the facts and the challenges of image analysis. Eur Radiol Exp 2018;2:36.

35. Bioulac-Sage P, Sempoux C, Frulio N, et al. Snapshot summary of diagnosis and management of hepatocellular adenoma subtypes. Clin Res Hepatol Gastroenterol 2019;43:12–19.

36. Putra J, Ferrell LD, Gouw ASH, et al. Malignant transformation of liver fatty acid binding protein-deficient hepatocellular adenomas: histopathologic spectrum of a rare phenomenon. Mod Pathol 2020;33:665–675.

37. Haider S, Pal R. Integrated analysis of transcriptomic and proteomic data. Curr Genomics 2013;14:91–110.

38. Sempoux C, Gouw ASH, Dunet V, et al. Predictive Patterns of Glutamine Synthetase Immunohistochemical Staining in CTNNB1-mutated Hepatocellular Adenomas. Am J Surg Pathol 2021.

39. Lasek W, Zagożdżon R, Jakobisiak M. Interleukin 12: still a promising candidate for tumor immunotherapy? Cancer Immunol Immunother 2014;63:419–435.

40. Mollinedo F. Neutrophil Degranulation, Plasticity, and Cancer Metastasis. Trends Immunol 2019;40:228–242.

41. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J 2019;6:94–98.

42. Quaglia A. Hepatocellular carcinoma: a review of diagnostic challenges for the pathologist. J Hepatocell Carcinoma 2018;Volume 5:99–108.

43. Tommaso LD, Spadaccini M, Donadon M, et al. Role of liver biopsy in hepatocellular carcinoma. World J Gastroenterol 2019;25:6041–6052.

44. Park HJ, Choi BI, Lee ES, et al. How to Differentiate Borderline Hepatic Nodules in Hepatocarcinogenesis: Emphasis on Imaging Diagnosis. Liver Cancer 2017;6:189–203.

45. Faivre S, Rimassa L, Finn RS. Molecular therapies for HCC: Looking outside the box. J Hepatol 2020;72:342–352.

**Acknowledgments:** We wish to thank David Bouyssie (IPBS Toulouse, France) for his help in using the Proline software and we thank the Cancer Biobank of Bordeaux University Hospital: BRIF BB-0033-00036.

**Legends of figures and tables**

**Figure 1: HCA database validation and linking of proteomic profiles with mutations defining HCA subtypes. a** Deregulation of HCA biomarkers (GS, CRP/SAA, ASS1 and L-FABP) for each subtype in the 57 cases constituting the proteomic database**.** Each point represents the Tumoral vs Non-Tumoral (T/NT) ratio of a case. The median of T/NT ratios is indicated for each subtype. The red line indicates identical expression level between the T and NT liver (T/NT=1). **b to d** Deregulation of the proteins frequently functionally associated with the mutated gene in each HCA subtype (*HNF1A*, *CTNNB1*, and genes of the JAK/STAT signaling pathway). Immunomarkers of each subtype are highlighted in red and orange. Examples of proteins with expression linked to mutations in *HNF1A* or *CTNNB1* (ALB and OAT respectively) or proinflammatory cytokines linked to the JAK/STAT signaling pathway (MIF) are highlighted in blue. **e** Gene set enrichment analysis (GSEA) performed from protein expression data identified by proteomics for the pathways and dysfunctions listed by Nault *et al* [18] used to characterize each HCA molecular subtype.

**Figure 2: a to f** Principal Component Analysis (PCA) comparing the profiles of protein expression deregulation in HCA versus FNH and between the different HCA subtypes. **g** Biological pathways associated with HCA subtype-specific profiles. Gene Set Enrichment analyses (GSEA) was carried out on the Ingenuity Pathway Analysis (IPA) database (Canonical Pathways). The most significantly enriched pathways are represented for HCA subtypes and compared to each other. Colors and circle size correspond to the level of significance.

**Figure 3**: **a** Diagnostic algorithm based on sequential proteomic profile matching. **b to d** Example of results that can be obtained using proteomic profile matching (**b** (R_01) corresponds to b-IHCA

exon 3 non S45, **c** (BP_01) to sh-HCA, **d** (BP_02) to H-HCA). The location of each case on the PCA from the reference database is indicated by a red arrow. **e** Validation results on a panel of biopsies and resections representative of cases analyzed by our expert center. Results that are inconsistent with the histopathological analysis are in italics.

**Figure 4: a** Identification of transformed HCA cases classified by molecular subtype. **b** Percentage of HCC/HCA or "borderline" developed on HCA diagnosed in men versus women in our collection of 260 cases at Bordeaux. **c and d** Percentage of HCA cases in our Bordeaux collection with cytological atypia and/or architectural anomalies extracted from pathological reports (**d**) according to molecular subtype (**c**).

**Figure 5:** Hierarchical clustering (**a**) and Principal Component Analysis (PCA) comparing the profiles of protein expression deregulation for HCC developed on HCA (HCC/HCA) and the corresponding HCA. Ten proteins were identified and constitute a signature for HCA malignancy. Their levels of deregulation (T/NT ratios) are represented in the form of a heatmap (**c**). The functional annotations extracted from the Reactome database are listed in table **d**.

**Figure 6:** Malignancy signature validation on 12 cases. **a** Histological hematoxylin eosin (H&E) staining of 2 cases presenting suspected malignancy : case 217 mild cytological atypia associated with foci of decreased reticulin network (not shown here) leading to the diagnosis of "borderline HCA" and case 280 suspected to be a well differentiated HCC with obvious cytological and architectural atypia; these 2 cases are located on principal component analysis (PCA) in g and h respectively. **b** Example locations on PCA for validation cases. The location of the tested cases on the PCA is indicated by a green arrow. Table (**c**) summarized the results of Random Forest testing and calculation of the Euclidian distance.

## Figure 1

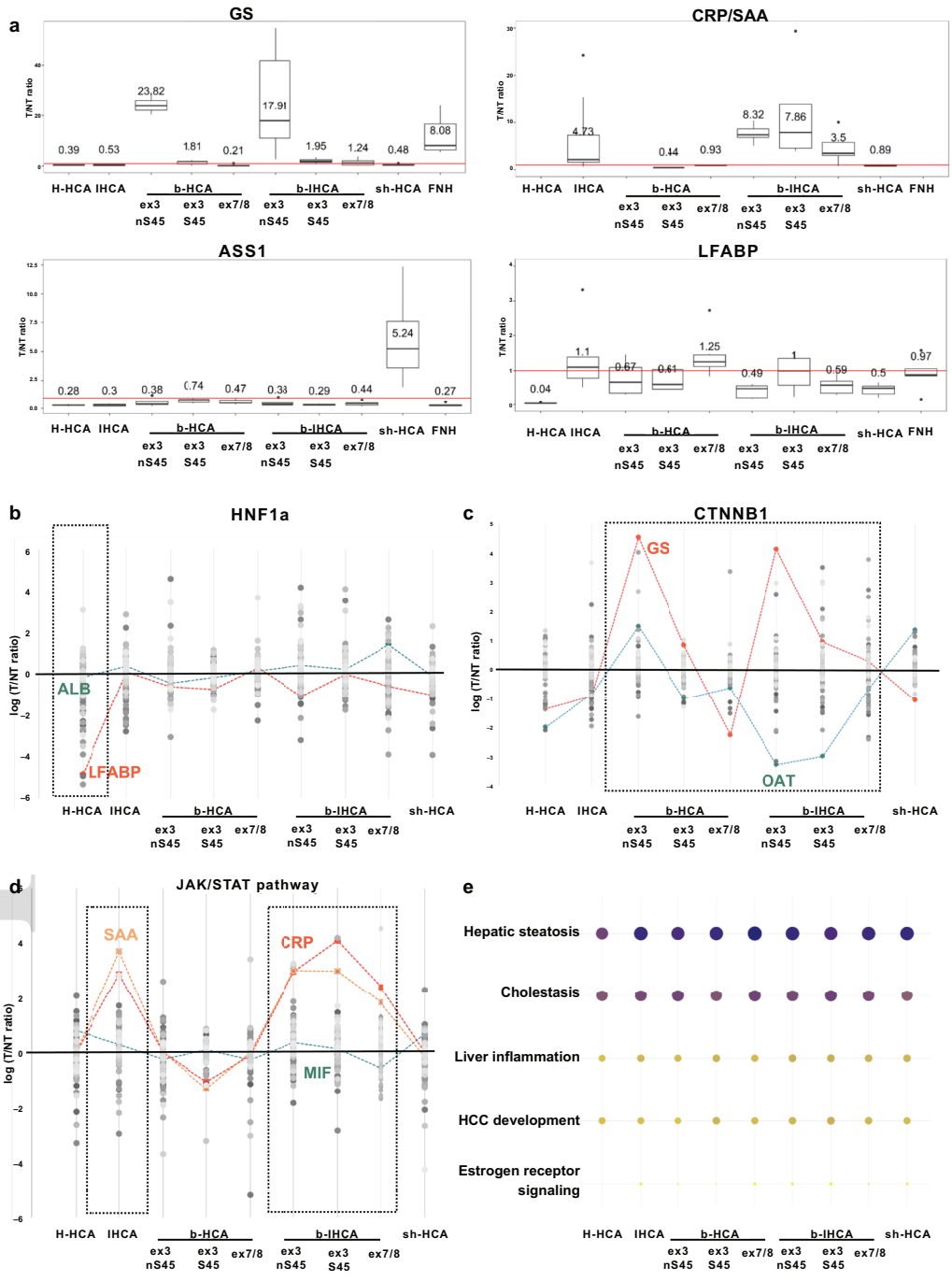
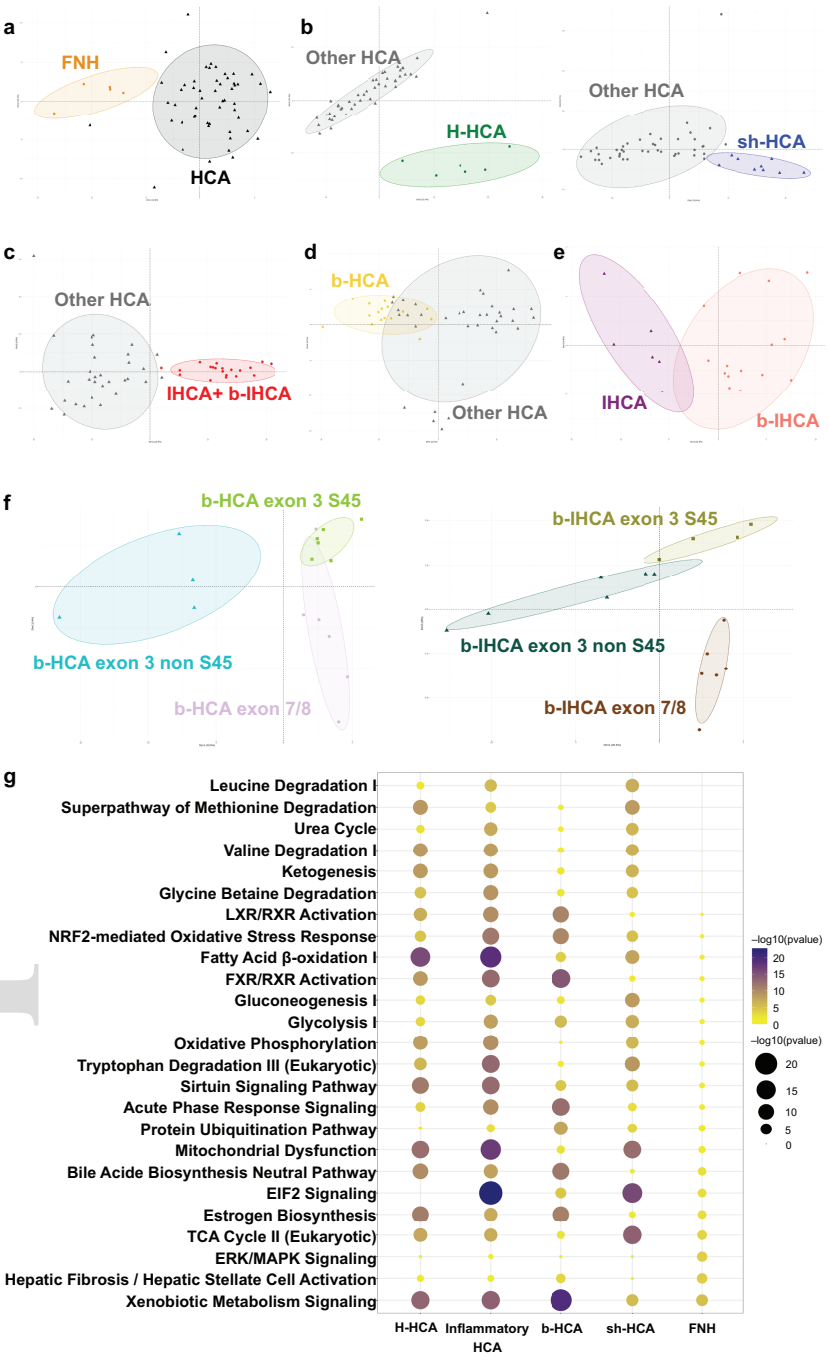
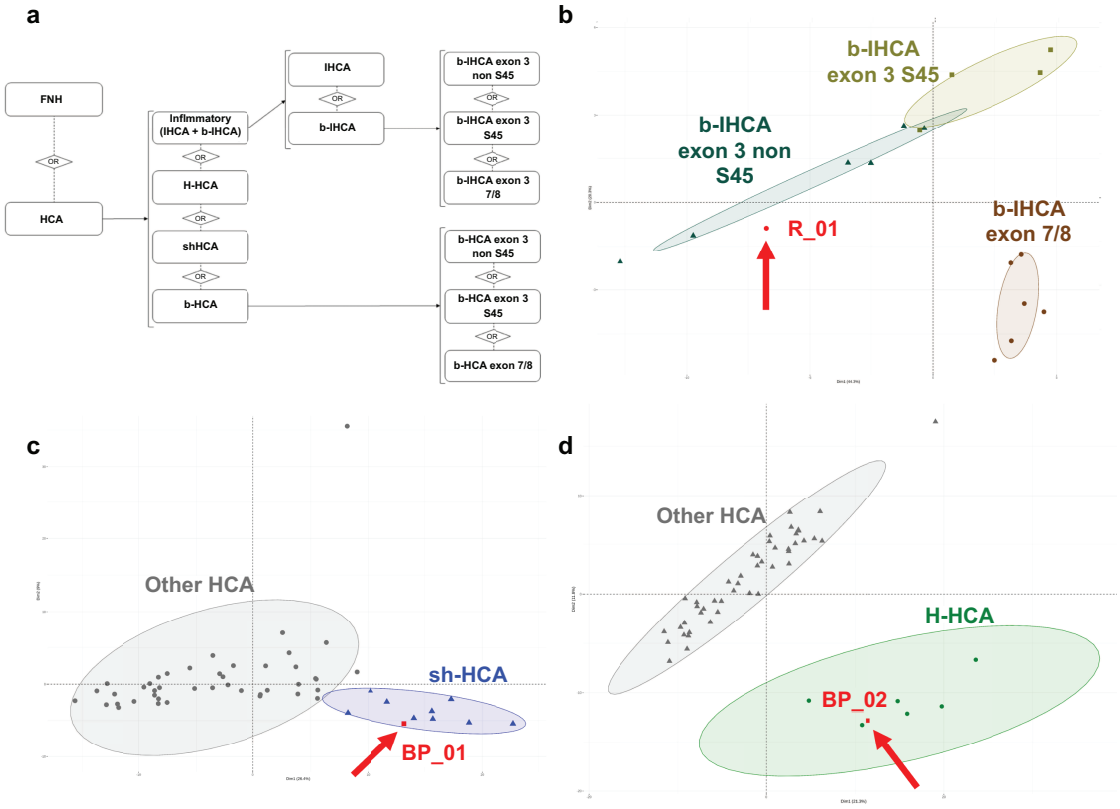


Figure 2



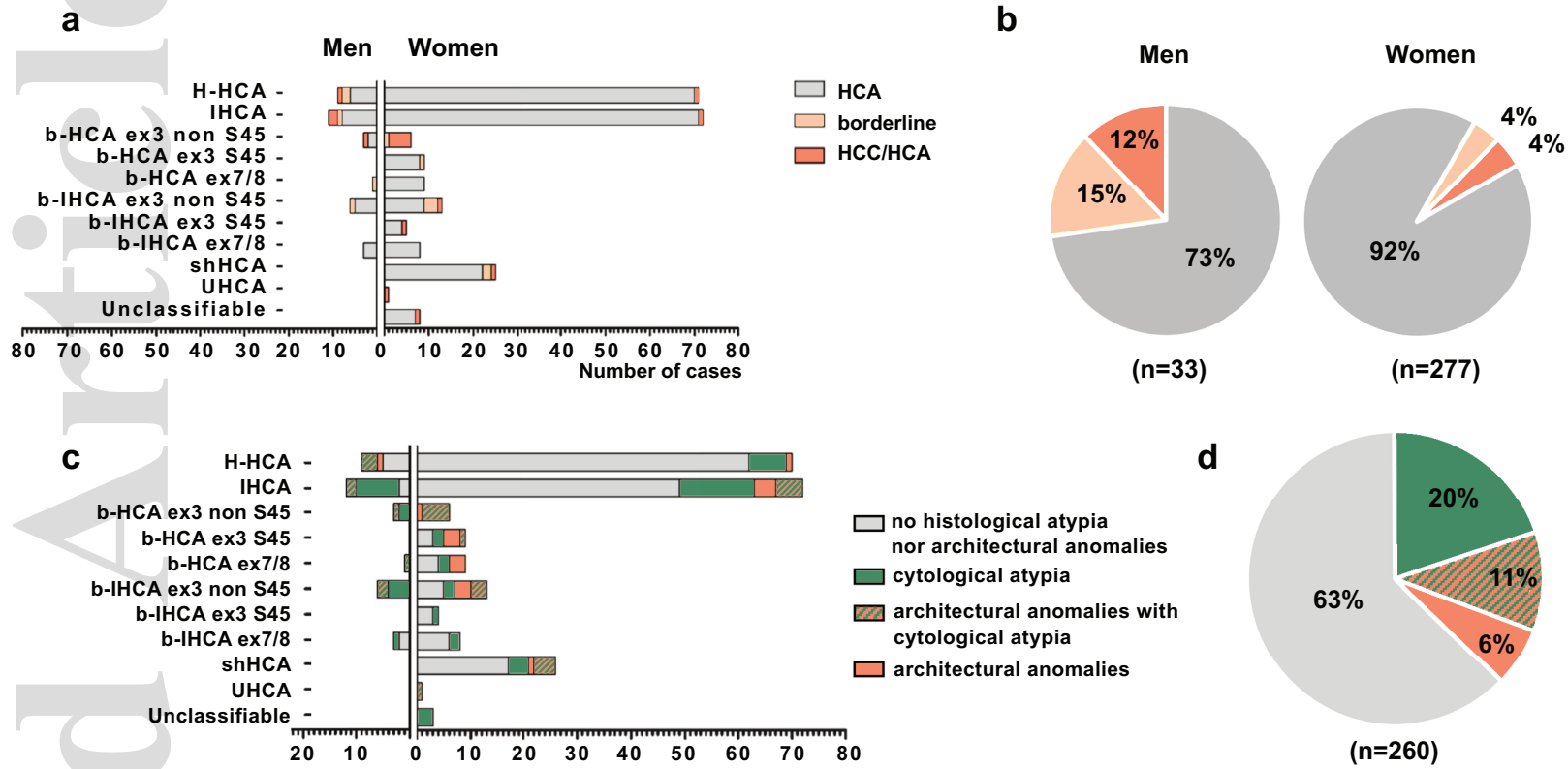
## Figure 3




e

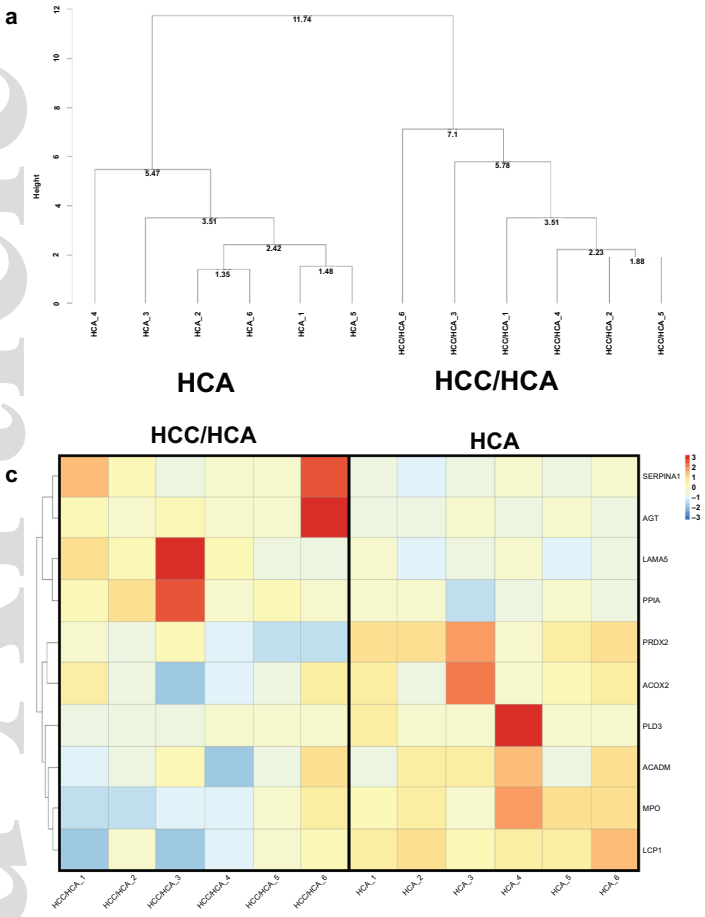
| cases | biopsy/resection | pathological diagnosis | Random Forest (with PCA reduction) | Euclidian distance calculation | Chi-square test |
|---|---|---|---|---|---|
| R_01 | resection | b-IHCA exon 3 non S45 | b-IHCA exon 3 non S45 | b-IHCA exon 3 non S45 | b-IHCA exon 3 non S45 |
| BP_01 | biopsy | sh-HCA | sh-HCA | sh-HCA | sh-HCA |
| BP_02 | biopsy | H-HCA | H-HCA | H-HCA | H-HCA |
| BP_03 | biopsy | H-HCA | H-HCA | H-HCA | H-HCA |
| BP_04 | biopsy | FNH | FNH | FNH | FNH |
| BP_05 | biopsy | FNH | FNH | FNH | FNH |
| BP_06 | biopsy | sh-HCA | sh-HCA | sh-HCA | *FNH* |
| BP_07 | biopsy | H-HCA | H-HCA | *b-HCA* | H-HCA |
| BP_08 | biopsy | b-HCA exon 7/8 | b-HCA exon 7/8 | b-HCA exon 7/8 | *b-HCA exon 3 S45* |
| BP_09 | biopsy | IHCA | *b-IHCA exon 7/8* | *b-IHCA exon 7/8* | *b-IHCA exon 7/8* |
| BP_10 | biopsy | IHCA | *b-IHCA exon 3 S45* | *b-iHCA exon 3 S45* | IHCA |
| BP_11 | biopsy | IHCA or b-IHCA (exon 3 S45 or exon 7/8) | b-IHCA exon 7/8 | b-IHCA exon 7/8 | b-IHCA exon 7/8 |

## Figure 4




hep_31826_f4.eps

**Figure 5**



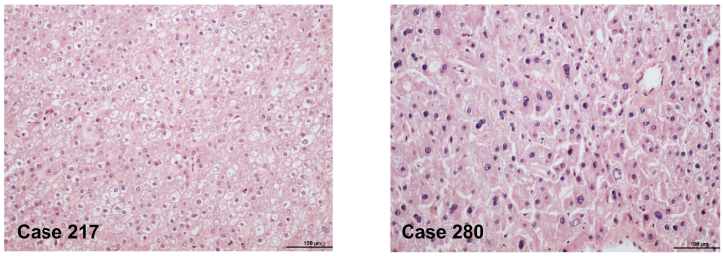

d

| | PPIA | LCP1 | SERPINA 1 | MPO | LAMA5 | PLD3 | ACADM | AGT | ACOX2 |
|---|---|---|---|---|---|---|---|---|---|
| interleukin 12 signaling | ■ | ■ | | | | | | | |
| innate immune system | ■ | | ■ | ■ | | ■ | | | |
| signaling by interleukins | ■ | ■ | | | ■ | | | | |
| neutrophil degranulation | ■ | | ■ | ■ | | | | | |
| metabolism of lipids | | | | | | ■ | ■ | ■ | ■ |

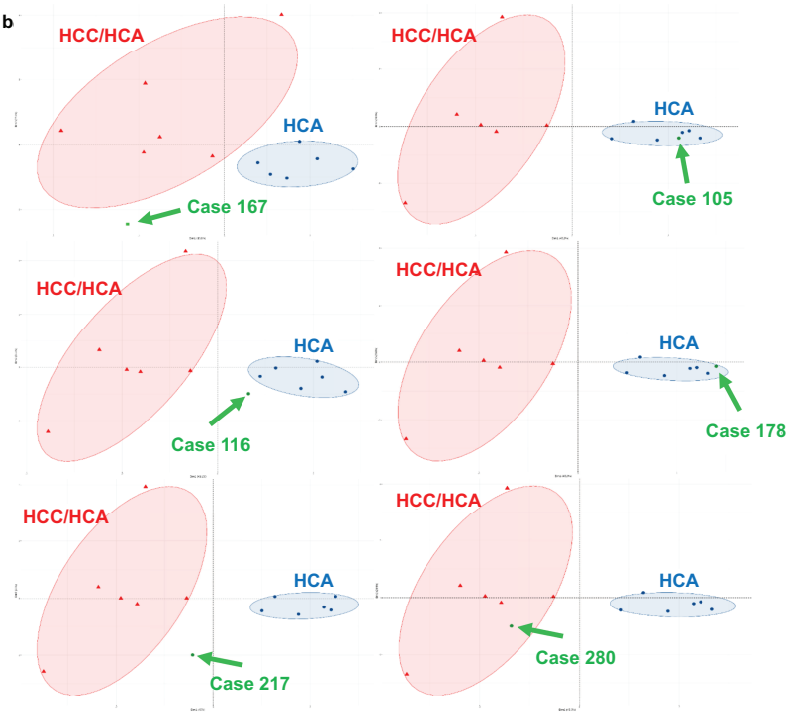
hep_31826_f5.eps

**Figure 6**

a




b




c

| Case | Euclidean distance (lowest result) | Random Forest (with PCA reduction) |
|---|---|---|
| Case 167 | HCC/HCA =5.18 ; HCA =5.43 | HCC/HCA |
| Case 105 | HCC/HCA = 5.01 ; HCA = 2.99 | HCA |
| Case 116 | HCC/HCA = 5.16 ; HCA = 4.26 | HCA |
| Case 121 | HCC/HCA = 5.23 ; HCA = 3.42 | HCA |
| Case 135 | HCC/HCA = 7.11 ; HCA = 4.86 | HCA |
| Case 119 | HCC/HCA = 4.87 ; HCA = 3.18 | HCA |
| Case 83 | HCC/HCA = 6.34 ; HCA = 4.90 | HCA |
| Case 218 | HCC/HCA = 4.96 ; HCA = 4.60 | HCA |
| Case 178 | HCC/HCA = 5.44 ; HCA = 3.88 | HCA |
| Case 217 | HCC/HCA = 4.65 ; HCA = 4.55 | HCC/HCA |
| Case 280 | HCC/HCA = 4.27 ; HCA =4.98 | HCC/HCA |