



Article

Combined Salivary Proteome Profiling and Machine Learning Analysis Provides Insight into Molecular Signature for Autoimmune Liver Diseases Classification

Giulia Guadalupi ¹, Cristina Contini ¹, Federica Iavarone ^{2,3}, Massimo Castagnola ⁴, Irene Messana ⁵, Gavino Faa ⁶, Simona Onali ⁷, Luchino Chessa ⁷, Rui Vitorino ^{8,9}, Francisco Amado ¹⁰, Giacomo Diaz ¹¹, Barbara Manconi ^{1,*}, Tiziana Cabras ^{1,†} and Alessandra Olianas ^{1,†}

- ¹ Dipartimento di Scienze della Vita e dell'Ambiente, Università di Cagliari, 09124 Cagliari, Italy; giulia.guadalupi@unica.it (G.G.); cristina.contini93@unica.it (C.C.); tcabras@unica.it (T.C.); olianas@unica.it (A.O.)
 - ² Fondazione Policlinico Universitario IRCCS "A. Gemelli", 00168 Rome, Italy; federica.iavarone@unicatt.it
 - ³ Dipartimento di Scienze Biotecnologiche di Base, Cliniche Intensivologiche e Perioperatorie, Università Cattolica del Sacro Cuore, 00168 Rome, Italy
 - ⁴ Laboratorio di Proteomica, Centro Europeo di Ricerca sul Cervello, IRCCS Fondazione Santa Lucia, 00168 Rome, Italy; maxcastagnola@outlook.it
 - ⁵ Istituto di Scienze e Tecnologie Chimiche "Giulio Natta", Consiglio Nazionale delle Ricerche, 00168 Rome, Italy; imessana53@gmail.com
 - ⁶ Division of Pathology, Department of Medical Sciences and Public Health, University Hospital, 09124 Cagliari, Italy; gavinofaa@gmail.com
 - ⁷ Liver Unit, University Hospital of Cagliari, 09124 Cagliari, Italy; simona.onali@yahoo.it (S.O.); luchinochessa@unica.it (L.C.)
 - ⁸ iBiMED, Department of Medical Science, University of Aveiro, 3810-193 Aveiro, Portugal; rvitorino@ua.pt
 - ⁹ UnIC@RISE, Department of Surgery and Physiology, Faculty of Medicine of the University of Porto, 4200-319 Porto, Portugal
 - ¹⁰ LAQV/REQUIMTE, Department of Chemistry, University of Aveiro, 3810-193 Aveiro, Portugal; famado@ua.pt
 - ¹¹ Dipartimento di Scienze Biomediche, Università di Cagliari, 09124 Cagliari, Italy; gdiaz@unica.it
- * Correspondence: bmanconi@unica.it
† These authors contributed equally to this work.



Citation: Guadalupi, G.; Contini, C.; Iavarone, F.; Castagnola, M.; Messana, I.; Faa, G.; Onali, S.; Chessa, L.; Vitorino, R.; Amado, F.; et al. Combined Salivary Proteome Profiling and Machine Learning Analysis Provides Insight into Molecular Signature for Autoimmune Liver Diseases Classification. *Int. J. Mol. Sci.* **2023**, *24*, 12207. <https://doi.org/10.3390/ijms241512207>

Academic Editor: Chulso Moon

Received: 15 June 2023
Revised: 26 July 2023
Accepted: 28 July 2023
Published: 30 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Autoimmune hepatitis (AIH) and primary biliary cholangitis (PBC) are autoimmune liver diseases that target the liver and have a wide spectrum of presentation. A global overview of quantitative variations on the salivary proteome in presence of these two pathologies is investigated in this study. The acid-insoluble salivary fraction of AIH and PBC patients, and healthy controls (HCs), was analyzed using a gel-based bottom-up proteomic approach combined with a robust machine learning statistical analysis of the dataset. The abundance of Arginase, Junction plakoglobin, Desmoplakin, Hexokinase-3 and Desmocollin-1 decreased, while that of BPI fold-containing family A member 2 increased in AIHp compared to HCs; the abundance of Gelsolin, CD14, Tumor-associated calcium signal transducer 2, Clusterin, Heterogeneous nuclear ribonucleoproteins A2/B1, Cofilin-1 and BPI fold-containing family B member 2 increased in PBCp compared to HCs. The abundance of Hornerin decreased in both AIHp and PBCp with respect to HCs and provided an area under the ROC curve of 0.939. Machine learning analysis confirmed the feasibility of the salivary proteome to discriminate groups of subjects based on AIH or PBC occurrence as previously suggested by our group. The topology-based functional enrichment analysis performed on these potential salivary biomarkers highlights an enrichment of terms mostly related to the immune system, but also with a strong involvement in liver fibrosis process and with antimicrobial activity.

Keywords: autoimmune hepatitis; autoimmune liver diseases; bottom-up proteomics; Cofilin-1; mass spectrometry; hornerin; primary biliary cholangitis; saliva; random forest

1. Introduction

Autoimmune hepatitis (AIH) and primary biliary cholangitis (PBC) are two types of autoimmune liver diseases (AILDs) that target the liver and are characterized by a wide spectrum of presentation.

AIH refers to chronic and progressive inflammation of the liver from an unknown cause but is thought to be a combination of genetic predisposition, environmental triggers and failure of the native immune system, which results in chronic inflammation of the liver, necrosis of hepatocyte and subsequent fibrosis. AIH is a rare worldwide disease predominant in women by a ratio of 4:1 with respect to men [1,2]. There are two known types of AIH: the most common Type 1 involves anti-smooth muscle antibodies (ASMA) with or without anti-nuclear antibodies (ANA). The more rare and often more severe Type 2 tends to appear earlier, usually during childhood, and progresses faster than type 1; it involves anti-liver/anti-kidney microsomal (anti-LKM) type 1 antibodies targeting cytochrome P450-2D6 and anti-liver cytosol (anti-LC) type 1 antibodies. Management frequently consists of lifelong nonspecific immunosuppression with azathioprine or other salvage therapies [3,4].

PBC usually affects women aged 40 to 60, with a male to female ratio of 1:10. It is characterized by the attack of the immune system towards the bile ducts and their epithelial cells, leading to progressive destruction of the intrahepatic bile ducts, and if not diagnosed and adequately treated it develops toward fibrosis, cirrhosis and finally liver failure [5]. Management consists of lifelong administration of ursodeoxycholic acid (UDCA) [6].

PBC and AIH can be difficult to distinguish clinically at early stages. There are still critical issues concerning early diagnosis, risk stratification of disease progression and identification of response to therapy predictors. Moreover, some patients may have “overlap syndromes” with characteristics of PBC or primary sclerosing cholangitis (PSC) in combination with AIH and cannot be assimilated into classical diagnostic categories. The most common AIH–PBC overlap syndrome [7] can be diagnosed using the Paris criteria, while so far there are no standardized diagnostic criteria for the other types of overlap syndromes [8]. In recent years, the definition of “overlap syndrome” has been substituted by the new definition of “variant syndrome of AIH and PBC” [9]. The variant syndrome shares diagnostic clinical and histological features of both entities [10]. Another challenge in AIH/PBC diagnosis is represented by the coexistence of comorbidities which may affect clinical phenotype at presentation. The most common association was found with other concurrent extrahepatic autoimmune disorders (CEHAID), mainly with autoimmune thyroid disease, but also with Sjögren’s syndrome, rheumatoid arthritis, nondestructive polyarthropathy, type 1 diabetes, vitiligo, ulcerative colitis and psoriasis [11,12].

AIH and PBC can be differentiated using clinical, biochemical, serologic and histologic findings. Blood testing is often the first step for the diagnosis of AILDs because many patients do not show symptoms until the disease has progressed to cirrhosis or liver failure. Serum autoantibodies ANA/SMA and LKM-1 are generally considered the diagnostic hallmarks of AIH [2]; however, a recent review reported a very poor diagnostic accuracy for ANA, SMA and LKM-1 if detected alone, with accuracy increasing only in the presence of both ANA and SMA [13]. Moreover, ANA tests also detect antigen specificities associated exclusively with PBC, including autoantibodies to Sp100-containing nuclear bodies (NBs) or gp210 protein [14]. Among candidate autoantibodies that may aid in the diagnosis of AIH, the most promising is α -actinin, a ubiquitous cytoskeletal cross-linking protein within the family of filamentous actin (F-actin) [15]. Nevertheless, a minority of patients with AIH do not show detectable autoantibodies at presentation and may express them only later and sporadically [2]. As a result, such patients must be scored using revised diagnostic criteria (RDC) or should undergo liver biopsy. Current guidelines recommend liver biopsy as a prerequisite for the diagnosis of AIH, to determine disease severity and to discriminate acute and chronic forms [16]. Regarding the diagnosis of PBC, liver biopsy is required only in specific anti-mitochondrial antibodies (AMA)-negative patients [17].

In this scenario, while the possible use of saliva as a diagnostic fluid has been largely investigated for oral and systemic diseases [18–22], it has been only marginally used in autoimmune liver diseases, mainly to investigate the role played by oral microbiota in their pathogenesis [23]. As a mirror of oral and systemic health, saliva provides valuable information because it contains not only proteins specifically secreted by the salivary glands [24], but also proteins from the gingival crevicular fluid [25,26], from oral microflora [27] and plasmatic proteins transported from blood to saliva by both intra- and extracellular pathways.

In a recent work, performed on the acid-soluble fraction of saliva from AIH and PBC patients explored by a top-down mass-spectrometry pipeline, our group evidenced qualitative and quantitative variations of some naturally occurring proteins/peptides belonging to the main families of salivary proteins, either secreted or not secreted by salivary glands [28]. These proteins/peptides were acid proline-rich proteins; statherin and P-B peptide; histatins; salivary cystatins; cystatins A, B, C and D; α -defensins; antileukoproteinase; S100A7, S100A8, S100A9 and S100A12 proteins. Supervised machine learning analysis of Mass Spectrometry (MS) data revealed for the first time the feasibility of selected salivary proteins to discriminate groups of subjects based on AIH or PBC occurrence. The acid-soluble salivary fraction did not contain, however, protein prone to precipitation under acidic conditions which may be influenced, in their expression, by the pathological condition. With the aim of extending our previous finding on salivary proteome in patients affected by AIH and PBC, and therefore to obtain a global overview of quantitative variations on the salivary proteome in presence of these two pathologies, we explored the acid-insoluble salivary fraction using a gel-based bottom-up proteomic approach combined with a robust machine learning statistical analysis of the dataset. Mass spectrometry data were analyzed by exact Mann–Whitney and Kruskal–Wallis tests to provide plausible salivary biomarkers related to AIH or PBC while Random Forest (RF) and multidimensional scaling (MDS) were used to individuate a panel of salivary proteins able to accurately classify the subjects based on AIH or PBC manifestation. Finally, a topology-based functional enrichment analysis has been performed for gaining insights into biological and functional pathways of both proteins commonly found and proteins showing varied levels among the three comparison groups.

2. Results

2.1. Characteristics of the Participants

A total of 17 AIH patients (in the following indicated as AIHp) 60.6 ± 12.8 (mean \pm standard deviation) years old, males $n = 2$, females $n = 15$, and 17 PBC patients (PBCp) 63.0 ± 9.4 years old, males $n = 0$, females $n = 17$ were recruited. The control group (HCs) included 17 age- and sex-matched healthy volunteers (58.9 ± 13.4 years old, males $n = 0$, females $n = 17$).

The detailed demographic data of HCs, AIHp and PBCp are reported in Table S1. Demographic characteristics, including age and gender, were matched between the AIHp and HCs and between PBCp and HCs ($p > 0.05$). Clinical and pharmacological features of AIHp and PBCp, collected at the same time of saliva sampling, are reported in Table 1. They comprise a panel of serum markers of liver dysfunction, including markers of hepatocellular dysfunction alanine transaminase (ALT) and/or aspartate aminotransferase (AST); markers of biliary disease such as alkaline phosphatase (ALP); markers of parenchymal liver dysfunction or biliary obstruction; or total bilirubin (TB) and γ -glutamyltransferase (GGT) and albumin, useful in assessing hepatic synthetic function. Patients were also tested for antinuclear antibodies (ANAs), smooth muscle antibodies (SMAs) and renal microsomal antigen antibodies (LKMs).

Table 1. Clinical features and pharmacological treatments of AIHp and PBCp included in the study measured at the time of saliva sampling.

Parameters		AIHp	PBCp
Age, average (range)	Years	60.6 (40–83)	63.0 (52–83)
Gender, <i>n</i> (%)	Female	15 (88.3%)	17 (100%)
BMI, average (range)	Kg/m ²	26.46 (17.58–36.13)	24.63 (19.33–30.27)
Cirrhosis, <i>n</i> (%)		4 (23.5%)	4 (23.5%)
Histological stage <i>n</i> (%)	I–II	5 (29.4%)	11 (64.7%)
	III–IV	12 (70.6%)	6 (35.3%)
Positivity to autoantibodies, <i>n</i> (%)	ANA	12 (70.6%)	8 (47.1%)
	SMA	10 (58.8%)	3 (17.6%)
	LKM	1 (5.9%)	2 (11.8%)
	AMA	0 (0%)	16 (94.1%)
AST, median (range)	IU/L	25.0 (14–57)	28.0 (17–65)
ALT, median (range)	IU/L	20.0 (8–46)	28.0 (17–133)
GGT, median (range)	IU/L	30.0 (0–138)	68.0 (14–452)
ALP, median (range)	IU/L	64.5 (28–216)	118.0 (65–415)
IgG, median (range)	g/dL	1.4 (0.69–2.51)	1.4 (1.1–2.3)
Albumine, median (range)	g/dL	3.9 (3.24–4.76)	3.9 (2.8–4.4)
Prothrombin time, median (range)	INR	0.93 (0.92–1.03)	1.0 (0.91–1.61)
TB, median (range)	mg/dL	0.69 (0.3–1.46)	0.61 (0.28–2.95)
Platelets, median (range)	10 ⁹ /L	199.0 (91–402)	223 (46–418)
Pharmacological treatment (% treated)	AZA (only or + UDCA and/or Steroids)	58.8%	
	Steroids	11.8%	
	UDCA (only or + steroids)	23.5%	100%
	Naïve	5.9%	

Range refers to total range; BMI: body mass index; ANA: antinuclear antibodies, SMA: smooth muscle antibodies; LKM: renal microsomal antigen antibodies; AMA: anti-mitochondrial autoantibodies; AST: aspartate aminotransferase; ALT: alanine aminotransferase; GGT: glutamyl transferase; ALP: alkaline phosphatase, IgG: immunoglobulin G; TB: total bilirubin; AZA: Azathioprine; UDCA: Ursodeoxycholic Acid.

Regarding pharmacological therapies, 58.8% of AIHp were under Azathioprine only or in combination with Steroids and/or UDCA, 11.8% under Steroid only, 23.5% under UDCA only or with Steroids and 5.9% were without therapy (naïve); PBCp were 100% under UDCA treatment.

The presence of other concurrent autoimmune diseases was investigated in patients; five AIHp and six PBCp presented Hashimoto's thyroiditis, one AIHp and one PBCp presented rheumatoid arthritis.

2.2. Saliva Sampling

To unravel the best approach for protein solubilization, the acid-insoluble fraction of saliva samples was subjected to three different solubilization solutions before bottom-up analysis. Based on bicinchoninic acid (BCA) assay, it was possible to demonstrate that 2% sodium dodecyl sulfate (SDS), 0.5 mM Dithiothreitol (DTT), 30 mM Tris-HCl pH 6.8 (Solution 3) allowed the retrieval of the best quantity of acid-insoluble proteins with a total protein concentration (TPC) of 2.5 mg/mL, approximately double the recovery of proteins compared to Solution 2 (TPC = 1.3 mg/mL) and more than 10 times the recovery obtained with Solution 1 (TPC = 0.2 mg/mL). Therefore, solution 3 was chosen to treat all AIHp, PBCp and HC samples.

Figure S1 reports the stained SDS-PAGE gel of AIHp, PBCp and HC pellets. For each lane, three portions of the gel were manually excised and submitted separately to trypsin digestion. This approach was preferred over total lane digestion to minimize the amount of gel to be submitted to in-gel digestion and therefore improve protein identification.

2.3. RP-nanoHPLC-High Resolution ESI-MS and MS/MS Analysis and Protein Identification

MS analysis of tryptic peptides resulted in 918, 891 and 787 high-confidence proteins for AIHp, PBCp and HCs, respectively, of which 467 were quantified in all groups and used for statistical analysis. Mann–Whitney and Kruskal–Wallis tests were used to identify possible protein abundance variations among AIHp vs. PBCp, AIHp vs. HCs and PBCp vs. HCs. In this way, non-parametric tests highlighted 14 varied proteins out of 467 with significant changes (p -values < 0.05 with and FDR < 10%). The results are shown in Table 2, reporting: (i) six proteins decreased in AIHp compared to HCs, namely Arginase (P05089), Junction plakoglobin (P14923), Desmoplakin (P15924), Hexokinase-3 (P52790), Desmocollin-1 (Q08554) and Hornerin (Q86YZ3); (ii) one protein increased in AIHp compared to HCs: BPI fold-containing family A member 2 (Q96DR5); (iii) seven proteins increased in PBCp compared to HCs, namely Gelsolin (P06396), Monocyte differentiation antigen (P08571), Tumor-associated calcium signal transducer 2 (P09758), Clusterin (P10909), Heterogeneous nuclear ribonucleoproteins A2/B1 (P22626), Cofilin-1 (P23528) and BPI fold-containing family B member 2 (Q8N4F0); (iv) one protein decreased in PBCp compared to HCs: Hornerin (Q86YZ3). Interestingly, only Hornerin was found to be simultaneously reduced in both pathologies. No significant changes were found in the comparison between AIHp and PBCp.

Table 2. Pairwise Mann–Whitney comparisons and Kruskal–Wallis test between HCs, AIHp and PBCp. Significant p -values < 0.05 with and FDR < 10% are highlighted. Direction of significant changes is also shown.

Proteins		HCs vs. AIHp	HCs vs. PBCp	AIHp vs. PBCp	HCs vs. AIHp vs. PBCp
Uniprot Code	Description	Mann Whitney	Mann Whitney	Mann Whitney	Kruskal Wallis
		p -Value	p -Value	p -Value	p -Value
P05089	Arginase *	0.0003 AIH < HC	0.0730	0.0061	0.0006
P06396	Gelsolin	0.2313	0.0015 PBC > HC	0.0653	0.0077
P08571	Monocyte differentiation antigen *	0.0256	0.0001 PBC > HC	0.1474	0.0005
P09758	Tumor-associated calcium signal transducer 2	0.0187	0.0010 PBC > HC	0.5352	0.0035
P10909	Clusterin *	0.0540	0.0002 PBC > HC	0.0777	0.0009
P14923	Junction plakoglobin	0.0007 AIH < HC	0.2313	0.0287	0.0034
P15924	Desmoplakin *	0.0007 AIH < HC	0.3057	0.0177	0.0030
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 *	0.0036	0.0006 PBC > HC	0.1899	0.0009
P23528	Cofilin-1 *	0.0679	0.0007 PBC > HC	0.5177	0.0062
P52790	Hexokinase-3	0.0013 AIH < HC	0.0987	0.0819	0.0053
Q08554	Desmocollin-1 *	<0.0001 AIH < HC	0.0064	0.0114	<0.0001

Table 2. Cont.

Proteins		HCs vs. AIHp	HCs vs. PBCp	AIHp vs. PBCp	HCs vs. AIHp vs. PBCp
Uniprot Code	Description	Mann Whitney	Mann Whitney	Mann Whitney	Kruskal Wallis
		<i>p</i> -Value	<i>p</i> -Value	<i>p</i> -Value	<i>p</i> -Value
Q86YZ3	Hornerin *	<0.0001 AIH < HC	<0.0001 PBC < HC	0.4824	<0.0001
Q8N4F0	BPI fold-containing family B member 2	0.0145	0.0007 PBC > HC	0.5629	0.0035
Q96DR5	BPI fold-containing family A member 2	0.0007 AIH > HC	0.0216	0.5861	0.0044

* Proteins selected by Boruta algorithm for RF classification.

2.4. Subjects Classification with Random Forest (RF) Analysis

RF classification among AIHp, PBCp and HCs was applied to a subset of 17 proteins selected according to the Boruta method. Figure 1A shows the relative importance of the selected proteins for classification, evaluated by the mean decreased Gini purity score (*x*-axis) and the mean decreased accuracy (*y*-axis). These parameters account for how much the purity and accuracy of classification decrease by excluding the specific protein from classification. Therefore, the greater the value of the parameter the greater the importance of the protein in the model. In Figure 1, Hornerin (Q86YZ3) appears to be the most important protein, consistent with the fact that it is the only protein showing significant changes in both comparisons: AIHp vs. HCs and PBCp vs. HCs (Table 2). Clusterin (P10909) was also a good discriminant component.

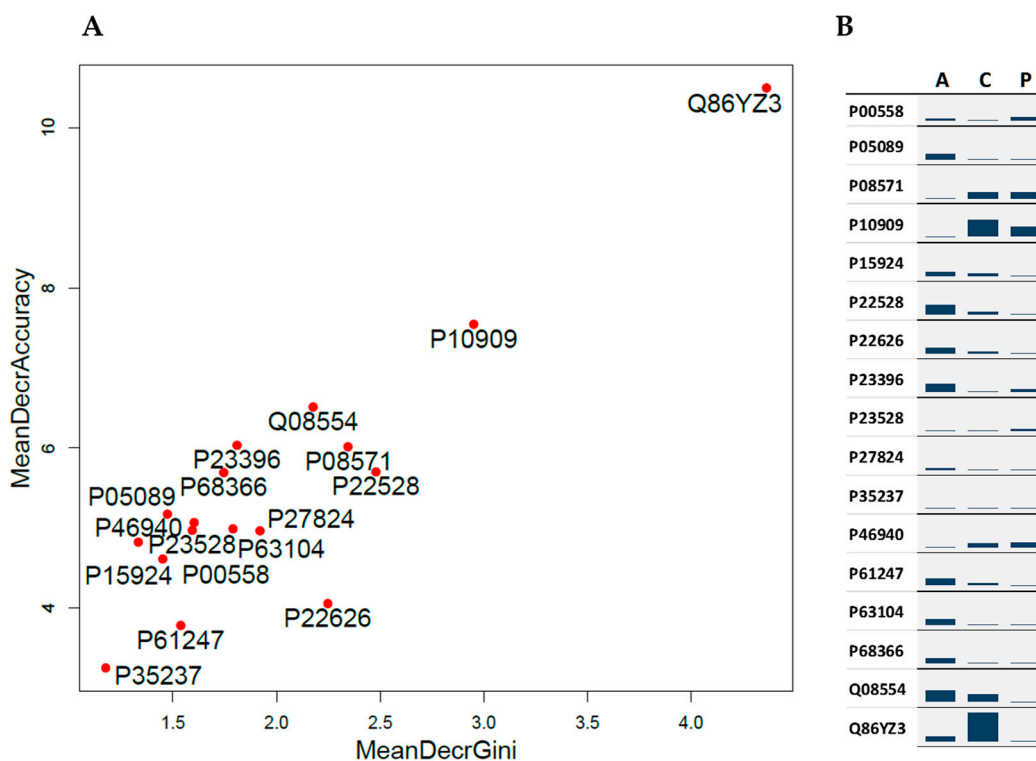


Figure 1. Variable importance for RF classification (A) Plot of the mean decreased Gini purity score (*x*-axis) and the mean decreased accuracy (*y*-axis). (B) Plot of the mean decreased Gini purity score (relative scale) calculated for each group. A, C and P are abbreviations of AIHp, HCs and PBCp, respectively.

The most important proteins for the classification of each group are shown in Figure 1B. As evidenced from this figure, Arginase (P05089), Desmoplakin (P15924), Cornifin-B (P22528), Heterogeneous nuclear ribonucleoproteins A2/B1 (P22626), 40S ribosomal protein S3 (P23396), 40S ribosomal protein S3a (P61247), 14-3-3 protein zeta/delta (P63104), Tubulin alpha-4A chain (P68366), Desmocollin-1 (Q08554) and Calnexin (P27824) were the most important proteins for AIHp classification, while Phosphoglycerate kinase 1 (P00558), Monocyte differentiation antigen (P08571), Clusterin (P10909) and Ras GTPase-activating-like protein (P46940) were selected for PBCp classification.

This subset of proteins provided a consistent increase in classification accuracy. The out-of-bag (OOB) classification error for AIHp and PBCp was 6% and 12%, respectively. When grouping AIHp and PBCp, the classification error of patients versus HCs was 0% (Table 3).

Table 3. Confusion matrix of the classification of the three groups of subjects.

		Predicted Class			OOB Error (%)
		HCs	AIHp	PBCp	
True class	HCs	17	0	0	0
	AIHp	0	16	1	6
	PBCp	0	2	15	12

Figure 2 shows the multidimensional scaling (MDS) diagram obtained by RF classification analysis of the three compared groups. The MDS diagram highlights a complete separation between HCs and both AIHp and PBCp in accordance with 100% accuracy in the classification shown in Table 3. A 3D video showing the first three MDS axes is reported in Supplemental materials (Video S1).

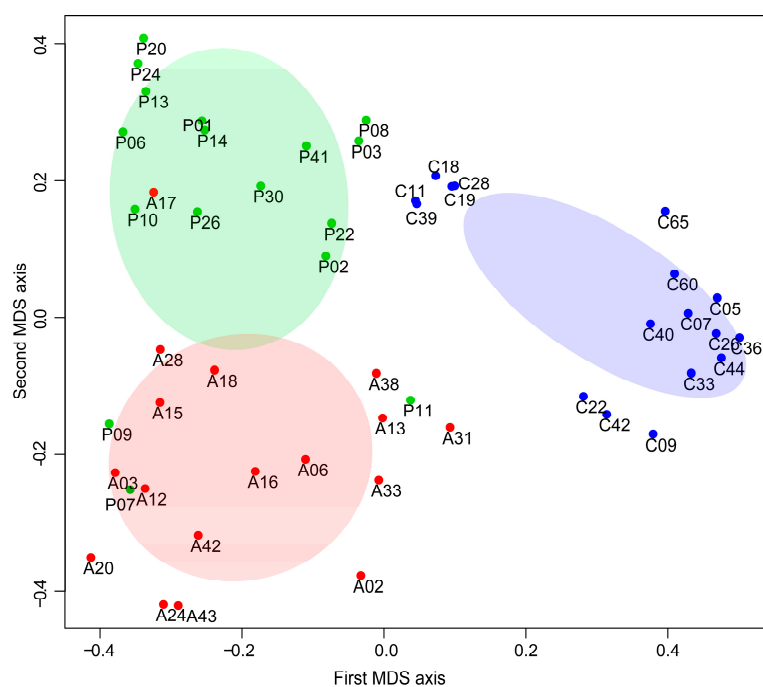


Figure 2. MDS diagram of RF classification, obtained by using the proximity between each pair of samples. A, C and P are abbreviations of AIHp, HCs and PBCp, respectively.

Seven proteins, selected by the Boruta algorithm with high MDG scores, showed also significant differences by Mann–Whitney tests. However, other components, prevalently those with low mean decreased Gini scores, did not show significant changes and, on

the other hand, components with significant changes were not selected by the Boruta algorithm. This apparent contrast is due to the essential nature of RF classification and in general of methods based on decision trees. Indeed, RF is able to find multiple 'split' points of the same variable, a method completely different from that adopted by classical univariate tests that consider the whole distribution of data. This allows two or more groups to be discriminated by a certain variable even when the mean (or average rank) of the variable is the same in the different groups. On the other hand, the use of multiple split points is not very suitable for normal diagnostic purposes, which require values gradually and coherently related to the severity of the disease. Because of this fact, although the classification produced by RF is of considerable interest in several respects, the proteins that collectively contribute to the classification of AIHp and PBCp cannot be tout court considered as candidate markers of the diseases.

The relative abundance of 23 proteins, including the 14 proteins with significant changes based on Mann–Whitney tests and the 17 proteins selected by Boruta algorithm for RF classification (eight proteins were in common) are graphically shown in the heatmap of Figure 3. From the heatmap it is interesting to observe the higher values of Hornerin (Q86YZ3; last row) in the HCs group with respect to both AIHp and PBCp groups, in agreement with the results of the Mann–Whitney tests.

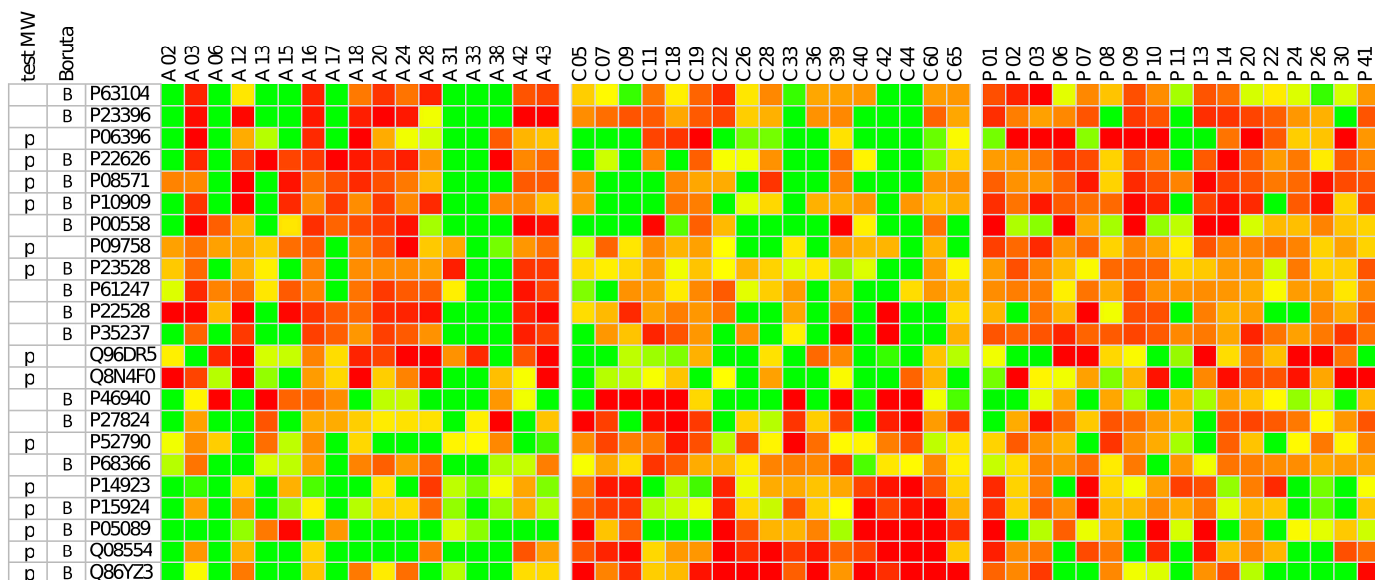


Figure 3. Heatmap of the abundance of proteins that showed significant changes based on M–W tests (denoted by the P code in the first column) and the proteins selected by Boruta algorithm for classification of AIH, PBC and HC groups (denoted by the B code in the second column). A, C and P are abbreviations of AIHp, HCs and PBCp, respectively. Eight proteins were in common in the two lists. To facilitate comparisons, data were row-wise standardized, that is each protein has mean = 0 and standard deviation = 1. Colors range from bright green (mean − 1 SD) to bright red (mean + 1 SD). The last row represents Hornerin protein, the only one that was significantly decreased both in AIHp and PBCp with respect to HCs.

To assess the diagnostic value of Hornerin as a potential salivary biomarker of autoimmune liver diseases, Hornerin data were evaluated by a ROC curve, by setting HCs versus AIHp and PBCp grouped together (Figure 4). The area under the ROC curve (AUC) was 0.939, and the best cut-off resulted in a sensitivity and specificity of 82% and 74%, respectively.

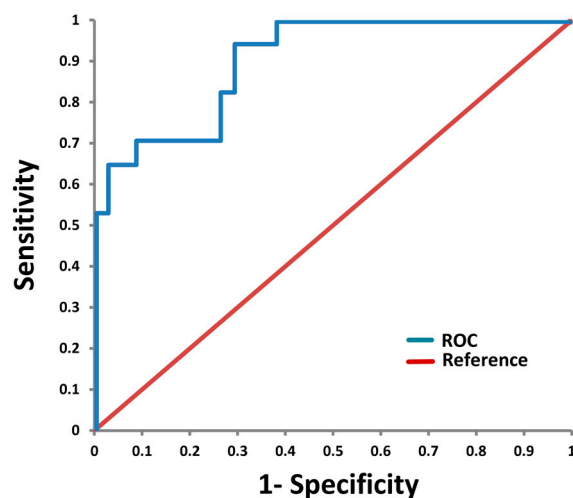


Figure 4. Receiver Operating Characteristic (ROC) curve for Hornerin (Q86YZ3) abundance in HCs versus both forms of AIHp and PBCp, grouped together with an AUC = 0.939.

2.5. Protein-Protein Interaction Network, Topological Analysis and Pathway Enrichment

The topological features of the protein-protein interaction network (PPI_n), such as node degree and betweenness centrality (BC) distributions, were examined using the Network Analyzer tool of Cytoscape. In scale-free networks, often observed in biological systems, most of the nodes (proteins) have few connections to the others, while a small number of nodes (hubs) are extensively connected to numerous others within the network. The nodes with high degrees were identified as hub proteins, which represent important connections within the network structure. On the other hand, nodes with high BC values were identified as bottleneck proteins because of their central role in facilitating communication and information flow between different parts of the network.

2.5.1. Proteins Commonly Found among AIHp, PBCp and HCs

The extended PPI_n obtained from the 467 proteins found among AIHp, PBCp and HCs, generated by the STRING database, is shown in Figure 5A.

The PPI_n was composed by only one giant network containing 449 nodes, interacting with 2293 edges and having a PPI enrichment p -value $< 1.0 \times 10^{-16}$. Among the 449 nodes, 47 nodes were selected for their higher BC value and 47 nodes for their larger degree. The 25 nodes having the highest BC (min. 0.008; max. 0.132) and degree (min. 58; max 177) values, with respect to the full PPI_n, were extracted from the extended network, and constitute the backbone network (Figure 5B). In the backbone network Glyceraldehyde-3-Phosphate Dehydrogenase (*GAPDH*) and albumin (*ALB*) were hub proteins with the largest degree and highest BC values. The other proteins of the backbone network were Cofilin (*CFL1*), T-Complex 1 protein (*TCP1*), Eukaryotic Translation Elongation Factor 2 (*EEF2*), Cell Division Cycle 42 (*CDC42*), Ras Homolog Family Member A (*RHOA*), Matrix Metalloproteinase 9 (*MMP9*), Integrin Subunit Alpha M (*ITGAM*), ATP Synthase F1 Subunit Alpha (*ATP5F1A*), ATP Synthase, F1 Subunit Beta (*ATP5B*), Valosin-Containing Protein (*VCP*), Pyruvate Kinase M1/2 (*PKM*), Heat Shock Protein Family A (*Hsp70*) Member 8 (*HSPA8*), Heat Shock Protein Family A (*Hsp70*) Member 5 (*HSPA5*), Heat Shock Protein Family D (*Hsp60*) Member 1 (*HSPD1*), Heat Shock Protein Family A (*Hsp90*) Member 1 (*HSP90AA1*), Heat Shock Protein Family B (*Hsp90*) Member 1 (*HSP90AA1*), Catalase (*CAT*), Calreticulin (*CALR*), Prohibitin 1 (*PHB1*), Enolase 1 (*ENO1*), Prolyl 4-hydroxylase Subunit Beta (*P4HB*) and Fibronectin 1 (*FN1*).

Finally, the 25 proteins included in the backbone network underwent GO biological process annotation and Reactome pathway analysis (Table 4).

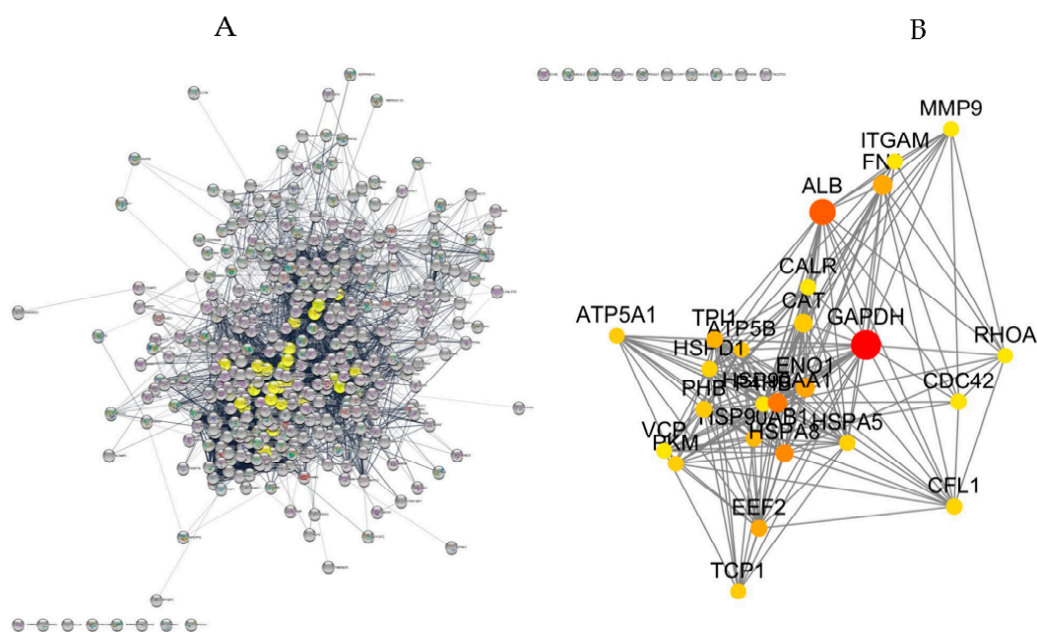


Figure 5. (A) Extended PPI network composed by 449 nodes interacting each other with 2293 edges and a PPI enrichment p -value $< 1.0 \times 10^{-16}$. Yellow nodes denote proteins with high degrees and betweenness centrality values thus representing the backbone network. (B) Topology of the backbone network, extracted from the extended network, containing proteins with higher BC and degree values. The size of the node corresponds to their BC values and colors range from yellow to red based on degree values. Protein names are reported in the text.

Table 4. Top five most significantly enriched Gene Ontology (GO) Biological Process and Reactome Pathway enrichment terms for the proteins in the backbone network. Protein names are reported in the text.

Term ID	Term	Term Size	Enriched Terms	FDR *	Associated Genes
Gene Ontology Biological Process					
GO:0071310	Cellular response to organic substance	2369	21	5.99×10^{-12}	GAPDH CAT ATP5B EEF2 TCP1 PKM CALR HSPA5 P4HB HSP90AA1 FN1 VCP HSP90AB1 MMP9 HSPD1 CDC42 RHOA HSPA8 CFL1 ITGAM PHB
GO:0006457	Protein folding	213	9	6.53×10^{-9}	TCP1 CALR HSPA5 P4HB HSP90AA1 VCP HSP90AB1 HSPD1 HSPA8
GO:0046034	ATP metabolic process	204	8	1.12×10^{-7}	GAPDH TPII ENO1 ATP5B PKM VCP ATP5A1 HSPA8
GO:0043312	Neutrophil degranulation	484	10	1.42×10^{-7}	CAT EEF2 PKM HSP90AA1 VCP HSP90AB1 MMP9 RHOA HSPA8 ITGAM
GO:0051702	Interaction with symbiont	93	6	7.16×10^{-7}	GAPDH FN1 HSPD1 HSPA8 CFL1 PHB
Reactome pathways					
HSA-168256	Immune System	1956	17	1.13×10^{-8}	CAT EEF2 TCP1 PKM CALR HSPA5 P4HB HSP90AA1 FN1 VCP HSP90AB1 MMP9 CDC42 RHOA HSPA8 CFL1 ITGAM

Table 4. Cont.

Term ID	Term	Term Size	Enriched Terms	FDR *	Associated Genes
HSA-3371556	Cellular response to heat stress	89	5	4.23×10^{-05}	<i>HSPA5</i> <i>HSP90AA1</i> <i>VCP</i> <i>HSP90AB1</i> <i>HSPA8</i>
HSA-5336415	Uptake and function of diphtheria toxin	6	3	4.76×10^{-05}	<i>EEF2</i> <i>HSP90AA1</i> <i>HSP90AB1</i>
HSA-6785807	Interleukin-4 and Interleukin-13 signaling	107	5	7.35×10^{-05}	<i>HSP90AA1</i> <i>FN1</i> <i>MMP9</i> <i>HSPA8</i> <i>ITGAM</i>
HSA-9020591	Interleukin-12 signaling	46	4	1.10×10^{-04}	<i>TCP1</i> <i>P4HB</i> <i>CDC42</i> <i>CFL1</i>

* Corrected with Benjamini–Hochberg, Term size: the total number of proteins in the GO and Reactome pathways; Enriched terms: number of terms pertaining to the backbone network; Associated genes: proteins associated to the specific term.

2.5.2. Proteins with Varied Levels among AIHp, PBCp and HCs

The PPI network built by using the 23 proteins obtained by merging those showing significant varied levels among AIHp, PBCp and HCs (14 proteins) based on Mann–Whitney tests, with the proteins selected by the Boruta algorithm (17 proteins), is shown in Figure 6.

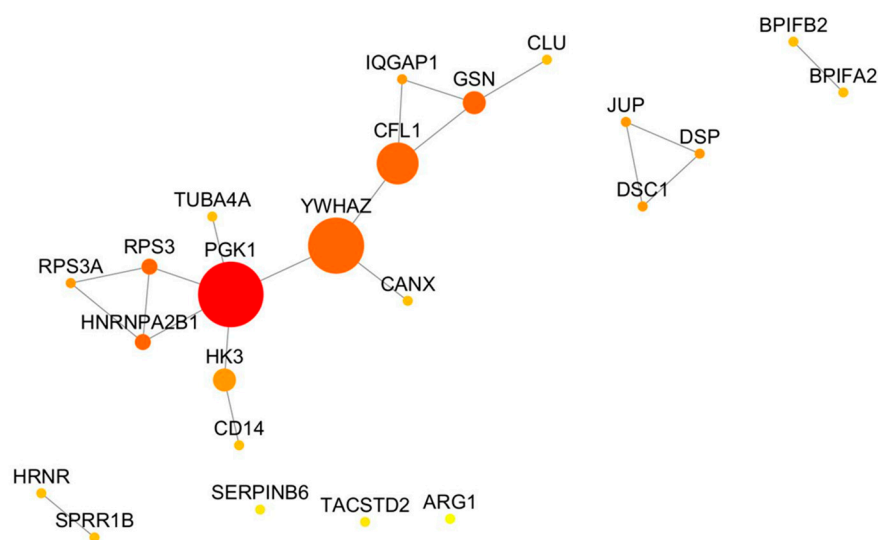


Figure 6. Topology of the network containing varied proteins among AIHp, PBCp and HCs. The size of the node corresponds to their BC values and colors range from yellow to red based on degree values. BC: Betweenness centrality.

The PPI network was composed of one main network containing 13 nodes, interacting with each other through 15 edges, and three minor networks (topological features of the nodes are reported in supplemental material Table S2). The protein with the highest BC and largest degree is Phosphoglycerate Kinase 1 (*PGK1*) followed by 14-3-3 protein zeta/delta (*YWHAZ*) and Cofilin-1 (*CFL1*). One minor network comprised three components of desmosomes, namely Junction plakoglobin (*JUP*), Desmoplakin (*DSP*) and Desmocollin-1 (*DSC1*); the second comprised two antimicrobial peptides, namely BPI fold-containing family A member 2 (*BPIFA2*) and BPI fold-containing family B member 2 (*BPIFB2*) and the last comprised two cytoskeleton components, namely Hornerin (*HRNR*) and Small Proline Rich Protein 1B (*SPRR1B*).

Finally, also these proteins underwent GO biological process annotation and Reactome pathway analysis as reported in Table 5.

Table 5. Gene Ontology (GO) Biological process and Reactome Pathway enrichment terms for the proteins varied among AIHp, PBCp and HCs.

Term ID	Term	Term Size	Enriched Terms	FDR	Associated Genes
Gene Ontology Biological Process					
GO:0045055	Regulated exocytosis	697	12	5.56×10^{-08}	<i>TUBA4A</i> <i>DSC1</i> <i>IQGAP1</i> <i>HK3</i> <i>CD14</i> <i>CLU</i> <i>ARG1</i> <i>HRNR</i> <i>GSN</i> <i>DSP</i> <i>JUP</i> <i>SERPINB6</i>
GO:0044419	Interspecies interaction between organisms	1899	12	1.60×10^{-04}	<i>BPIFB2</i> <i>CANX</i> <i>BPIFA2</i> <i>IQGAP1</i> <i>RPS3</i> <i>CD14</i> <i>CLU</i> <i>RPS3A</i> <i>HNRNPA2B1</i> <i>ARG1</i> <i>GSN</i> <i>CFL1</i>
GO:0060429	Epithelium development	1109	9	0.0011	<i>DSC1</i> <i>IQGAP1</i> <i>SPRR1B</i> <i>HRNR</i> <i>TACSTD2</i> <i>PGK1</i> <i>DSP</i> <i>JUP</i> <i>CFL1</i>
GO:0070268	Cornification	113	4	0.0036	<i>DSC1</i> <i>SPRR1B</i> <i>DSP</i> <i>JUP</i>
GO:2001235	Positive regulation of apoptotic signaling pathway	180	4	0.0162	<i>RPS3</i> <i>CLU</i> <i>GSN</i> <i>YWHAZ</i>
GO:0071345	Cellular response to cytokine stimulus	1013	7	0.0289	<i>CANX</i> <i>RPS3</i> <i>HNRNPA2B1</i> <i>ARG1</i> <i>GSN</i> <i>YWHAZ</i> <i>CFL1</i>
GO:0030155	Regulation of cell adhesion	712	6	0.0328	<i>IQGAP1</i> <i>RPS3</i> <i>ARG1</i> <i>TACSTD2</i> <i>GSN</i> <i>JUP</i>
Reactome pathways					
HSA-168256	Immune System	1956	18	4.76×10^{-11}	<i>BPIFB2</i> <i>CANX</i> <i>TUBA4A</i> <i>BPIFA2</i> <i>DSC1</i> <i>IQGAP1</i> <i>HK3</i> <i>CD14</i> <i>CLU</i> <i>HNRNPA2B1</i> <i>ARG1</i> <i>HRNR</i> <i>GSN</i> <i>DSP</i> <i>JUP</i> <i>YWHAZ</i> <i>CFL1</i> <i>SERPINB6</i>
HSA-6809371	Formation of the cornified envelope	127	4	0.0083	<i>DSC1</i> <i>SPRR1B</i> <i>DSP</i> <i>JUP</i>
HSA-109581	Apoptosis	173	4	0.0159	<i>CD14</i> <i>GSN</i> <i>DSP</i> <i>YWHAZ</i>
HSA-447115	Interleukin-12 family signaling	56	3	0.0159	<i>CANX</i> <i>HNRNPA2B1</i> <i>CFL1</i>
HSA-6803157	Antimicrobial peptides	87	3	0.0338	<i>BPIFB2</i> <i>BPIFA2</i> <i>CLU</i>
HSA-76002	Platelet activation, signaling and aggregation	260	4	0.0461	<i>TUBA4A</i> <i>CLU</i> <i>YWHAZ</i> <i>CFL1</i>

FDR corrected with Benjamini–Hochberg. Term size: the total number of proteins in the GO and reactome pathways; Enriched terms: number of terms pertaining to the backbone network; Associated genes: proteins associated to the specific term.

3. Discussion

The aim of this study was to provide a salivary molecular signature of autoimmune liver diseases. A robust statistical analysis of quantitative salivary proteomic data from AIHp, PBCp and controls provided indication of the proteins statistically varied in levels among groups, considered as potential biomarkers of the disease, and proteins able to classify a subject with good accuracy based on AIH or PBC occurrence. Additionally, proteomic data have been used to build a PPI network and a functional enrichment analysis has been performed to explore the biological significance of salivary proteins and obtain insights regarding the contribution of these proteins to the pathogenesis of autoimmune liver diseases.

3.1. Topology-Based Functional Enrichment Analysis

In the backbone network, Glyceraldehyde-3-Phosphate Dehydrogenase (*GAPDH*) and albumin (*ALB*) were either hub and bottleneck proteins with the largest degree and BC values. It is reported that proteins that are hubs as well as bottlenecks will likely be evolutionarily conserved [29] and are involved in multiple pathways [30].

In the context of chronic liver disease, both quantitative and functional changes have been evidenced in either albumin or *GAPDH* [31]. It has been demonstrated that in patients with liver cirrhosis, albumin undergoes both several reversible and irreversible posttranscriptional changes that alter its properties [32]. Importantly, albumin represented a good prognostic factor, being a significant predictor of death in patients with liver cirrhosis [33,34]. Ideally, topological network analysis identifies proteins susceptible to be potential biomarker of pathologies [35]; however, our proteomic analysis of saliva from AIH and PBC patients did not evidence varied levels of these two important proteins. Indeed, among the 25 proteins constituting the backbone network, and thus showing higher degree and BC values, only Cofilin-1 (degree 68, BC 0.016) was found to be increased in PBC patients with respect to HCs. As reported below, Cofilin-1 is a multifunctional protein involved in many biological processes and functional pathways also in the liver.

The topology-based functional enrichment analysis performed on the backbone's proteins revealed a set of central nodes mainly associated with the immune system; among these Interleukin-4, Interleukin-13 and Interleukin-12 signaling displayed higher *p*-values. Hepatocytes, cholangiocytes, putative hepatobiliary progenitor cells and fibroblasts express functional interleukin-4 and interleukin-13 receptors, and studies conducted on transgenic mice with interleukin-13 signaling genetically disrupted in hepatocytes, cholangiocytes or fibroblasts revealed key roles for interleukin-13 in fibrosis, steatosis, cholestasis and ductular reaction [36]. PBC has a genetic association with interleukin-12 signaling [37], so that modulation of this pathway at an early stage of disease has been proposed as a therapeutic model in the treatment of this pathology [38]. Regarding proteins with varied levels among AIHp, PBCp and HCs, the highest central nodes were Phosphoglycerate Kinase 1 (*PGK1*) followed by 14-3-3 protein zeta/delta (*YWHAZ*) and Cofilin-1 (*CFL1*). The topology-based functional enrichment analysis performed on these salivary proteins highlighted an enrichment of terms with multifaceted biological functions mostly related to the immune system, but also with a strong involvement in the liver fibrosis process and with antimicrobial activity.

3.2. Dysregulated Proteins in AIH and PBC Are Mainly Involved in Liver Fibrosis

Liver fibrosis is an abnormal wound repair response caused by an assortment of chronic liver damages, which is characterized by over-deposition of diffuse extracellular matrix and anomalous hyperplasia of connective tissue [39]. Our analysis performed on the salivary proteome of AIH and PBC evidenced a set of dysregulated proteins with respect to HCs potentially involved in the development and/or progression of this pathological condition. These findings were further confirmed by the topology-based biological and pathways enrichment analysis that evidenced an enrichment of terms related not only to the immune system, but also to epithelium development, regulation of cell adhesion and formation of the cornified envelope.

The concomitant reduced level of Junction plakoglobin, Desmoplakin and Desmocollin-1 in AIH patients is intriguing since all three proteins are major components of desmosomes, cell structures specialized in cell-to-cell adhesion. The reduction in desmosomal protein level in AIHp, evidenced for the first time in this study, could be related to their liver fibrotic state; in fact, 70% of the AIHp enrolled in the study presented the most severe liver fibrotic stage III and IV while the majority of PBCp presented a mild liver fibrotic stage I and II. The exact localization of cell junctions in hepatic epithelial cells has not been determined yet [40], but Zhou et al. have shown that, in the absence of plakoglobin, bile duct ligation resulted in a more severe disease outcome with enhanced liver fibrosis [41]. The correlation between the levels of desmosomal proteins and the fibrotic stage

has been described in mice lacking E-cadherin in the liver which developed periportal inflammation via an impaired intrahepatic biliary network, as well as periductal fibrosis, which resembles primary sclerosing cholangitis [42]. These findings point out that an intact intrahepatic biliary network with normal bile secretion depends on functional E-cadherin. Moreover, Dubash et al. described how the loss of desmosomal proteins in transgenic mice activated a signaling pathway in cardiomyocytes that up-regulated multiple inflammatory and extracellular matrix proteins known to promote tissue fibrosis [43].

Arginase-1 levels were found decreased in AIHp with respect to HCs and important for AIH classification in RF analysis. This protein is a manganese-containing enzyme that catalyzes the final step in the urea cycle. In the liver arginase-1 is mainly expressed in the periportal hepatocytes, but not in the bile ducts, endothelial and Kupffer cells [44]. Serum arginase level is significantly associated with oxidative stress since it is indirectly involved in nitric oxide (NO) regulation by the consumption of L-arginine, which is a common substrate for NO synthase (NOS). The balance between the consumption of L-arginine by arginase-1 (leading to the production of proline for collagen and polyamine production which are essential for cell growth and matrix modelling) and NOS (for NO production) determines the outcome of wound repair, with arginase-1 controlling the healing process and NOS regulating the anti-microbial activity. In a mouse model of liver fibrosis, hepatic stellate cells activation is accompanied by a switch in arginine catabolism resulting from downregulation of NOS and upregulation of arginase-1; therefore, inhibition of arginase-1 has been proposed as an anti-fibrotic target for the treatment of liver fibrosis [45]. Our results, which evidenced reduced arginase-1 levels in AIHp, are apparently in contrast with these findings, however it can be highlighted that most of the patients enrolled for this study were undergoing corticosteroids therapy whose effect on arginase-1 expression remains to be elucidated [46].

Gelsolin and Cofilin-1, both showing increased levels in PBCp with respect to HCs, are actin-binding proteins, key regulators of actin filament assembly and disassembly. In particular, Cofilin-1 is, among the proteins found with varied levels in AIHp and PBCp, the only one included in the backbone network probably thanks to its prominent role in actin dynamics and modulation essential for cell survival. This prominent role in the topological analysis is not surprising since liver injuries create a microenvironment that alters actin dynamics in the hepatic stellate cells which are responsible for type I collagen expression, the major extracellular matrix protein in various types of fibrotic diseases [47]. Hereafter, Cofilin-1 levels were related to the stage of liver fibrosis playing a key role in the progression of liver fibrosis toward hepatocellular carcinoma [48]. Moreover, Cofilin-1 levels have been found to be increased in HBV-associated hepatocellular carcinoma and its levels were correlated with the severity of this liver disease [49].

Early experiments on mice lacking gelsolin evidenced a reduction in dermal fibroblasts migration, establishing that gelsolin is required for rapid motile responses in cell types involved in stress responses such as hemostasis, inflammation and wound healing [50].

Similarly, clusterin, a Golgi extracellular chaperone implicated in cholestatic and fibrotic processes, was also found to be augmented in PBC patients with respect to HCs. It is expressed by hepatocytes and secreted into the bile, possibly acting as a chaperone to protect either bile duct epithelia from damage by toxic bile constituents or biliary proteins from misfolding [51]. Clusterin was found to be upregulated in thioacetamide-induced and bile duct ligation mouse models of liver fibrosis; the upregulation of clusterin attenuated hepatic fibrosis by inhibiting the hepatic stellate cells' activation and Smad3 signaling pathways responsible for the production of extracellular matrix proteins such as type I collagen [52]. On the contrary, circulating clusterin levels were significantly reduced in biliary atresia, which is a rare cholestatic liver disease of neonates characterized by obstruction of the biliary system; in this case clusterin levels were reduced, especially in patients with worse outcomes including jaundice and severe liver fibrosis [53]. Our results, in agreement to these findings, suggested that upregulation of clusterin may act as a defense mechanism to prevent hepatic fibrosis.

Higher levels of Tumor-associated calcium signal transducer 2 (Trop-2) were found in PBC patients with respect to HCs. It has been reported that Trop-2 levels in hepatitis C patients were inversely correlated to AST and ALT values [54], but our results cannot confirm this correlation since PBC patients enrolled for the study presented almost normal values of transaminases. Trop-2 plays a role in tumor progression by actively interacting with several key molecular signaling pathways traditionally associated with cancer development and progression. Conversely to most type of tumor, where Trop-2 was found upregulated [55], in liver cancer this protein has been found downregulated [56]. The different expression pattern of Trop-2 in liver cancer and autoimmune liver diseases is intriguing and deserves further elucidation.

The same controversial results in the expression pattern between liver cancer and autoimmune liver diseases has been highlighted for hornerin; this protein was found to be elevated via proteomic analysis in hepatocellular carcinoma [57] while our proteomic data revealed hornerin as the only protein with levels reduced in both AIH and PBC patients with respect to HCs as well as the protein that, in RF analysis, best discriminated these pathologies from HCs. This contradiction is quite interesting and could be related to different pathophysiological pathways between cancer and AIH and PBC. For instance, hornerin was found to be downregulated in atopic dermatitis, contributing to the epidermal barrier defect observed in this skin disease [58]. Indeed, being a component of the liver matrisome [59], it can be speculated that reduced levels of hornerin in both AIH and PBC patients may be correlated to the anomalous hyperplasia of connective tissue associated with liver fibrosis. The role of hornerin as a potential biomarker for AIH/PBC classification has been further validated by the ROC curve of HCs versus AIHp and PBCp grouped together, which provided an AUC of 0.93 with an accuracy of 78%. However, it may be highlighted that a better classification accuracy of HCs of 100% with respect to AIHp and PBCp grouped together was obtained by RF using the whole panel of 17 proteins selected by the Boruta algorithm.

Our study on the salivary proteome of AIH and PBC revealed also varied levels of proteins with prevalent antimicrobial function, and marginally involved in the liver fibrosis process. Indeed, several studies have shown that metabolites of oral microbiota, through the gut-oral axis, can enter the bloodstream and contribute to the occurrence and progression of many liver diseases [60,61].

BPI fold-containing family A member 2 (*BPIFA2*) and BPI fold-containing family B member 2 (*BPIFB2*) proteins are characterized by the presence of the bactericidal/permeability-increasing protein fold (BPI fold) [62].

Interestingly, despite their similar antimicrobial function, increased levels of *BPIFA2* and *BPIFB2* were observed in AIH and PBC patients, respectively, with respect to HCs. This difference could be related to an oral dysbiosis, which is often observed in such patients and involves the alteration of different bacterial species in the oral microbiota of AIHp [23] and PBCp [63] and towards which *BPIFA2* and *BPIFB2* can show different inhibitory activity. Indeed, a recent study on *BPIFA2* knockout mice highlighted its prominent role in the solubilization of ingested bacterial lipopolysaccharides [64]. From another point of view, the increased level of *BPIFA2* in AIHp only could be related to other biological functions of this multifaceted protein. In fact, besides its prominent role in the local antibacterial response, *BPIFA2* exerted a role in the prevention of early fibrosis progression and the development of chronic kidney disease [65]. A similar protective role in inhibiting liver fibrosis can be speculated for the AIHp since the patients enrolled for this study showed a marked liver fibrotic state.

An impaired gut-oral axis in autoimmune hepatitis could be also related to the increased level of Monocyte differentiation antigen CD14 we observed in PBCp with respect to HCs. CD14 might be an important factor in the pathogenesis of PBC since it is constitutively expressed by human intrahepatic biliary epithelial cells [66]. As a marker of monocyte activation and response to bacterial lipopolysaccharides, CD14 level reflects the host response to products of microbial translocation [67]. Increased levels of CD14 were found in PBC

patients before UDCA therapy [68], but our results showed an increase in CD14 levels also in patients undergoing therapy with UDCA. CD14 protein also plays a key role in liver disease by reducing fibrosis through degradation or clearing up collagen-I deposits [69].

3.3. Machine Learning Analysis of Proteomics Data

RF analysis, one of several up-to-date machine learning methods, evidenced a panel of proteins/peptides present in saliva able to correctly classify the AIHp group with respect to PBCp with 94% accuracy, the PBCp group with respect to the AIHp group with 88% accuracy and PBCp/AIHp grouped together with respect to HCs with 100% accuracy. This result confirms our previous one obtained by applying the same statistical approach to a mass-spectrometry dataset obtained after top-down proteomic analysis of the acid-soluble fraction of saliva from AIHp and PBCp. Moreover, this approach confirms the feasibility of the salivary proteome to discriminate groups of subjects based on physiological or pathological condition not only confined to the oral cavity [28].

3.4. Study Limitation

The low number of subjects involved in the study may not be entirely representative of the considered population; nevertheless, the statistical analysis performed by our group provides a good classification of subjects based on AIH or PBC occurrence. A larger population will be useful to further validate the present findings. Most of the proteins found with altered levels among AIHp, PBCp and HCs and/or selected by Boruta algorithm for patients' classification have multiple isoforms often characterized by different, sometimes even opposite, functional activities, and the bottom-up approach exploited in this study does not allow for characterization of isoforms/proteoforms.

Moreover, the observational nature of the study did not allow a demonstration of the causal effect of the varied proteins on AILDs, which can be considered potential biomarkers rather than causal mediators of these pathologies.

4. Materials and Methods

4.1. Ethical Statement

This is a cross-sectional study performed in 2021 on AIHp and PBCp recruited from the liver unit of University Hospital of Cagliari, Sardinia, Italy. Patients and healthy controls signed the informed written consent that agreed with the latest stipulations established by the Declaration of Helsinki. The Committee of the "Azienda Ospedaliero-Universitaria di Cagliari", Cagliari, Italy, approved the study on 21 July 2021 (reference number PG/2021/11303).

4.2. Study Subjects and Clinical Studies

Patients were diagnosed based on the criteria reviewed by the International Autoimmune Hepatitis Study Group (IAIHG) in 1999 [70] and by the EASL clinical practice guidelines [5]. The study included patients showing, at the time of saliva sampling, almost normal values of ALT and AST; based on these inclusion criteria, only patients that were under pharmacological therapy for at least three years were selected. Only one AIHp was without therapy but was included in this study because of his normal values of transaminases. Patients affected by overlap syndrome, chronic hepatitis induced by HBV or HCV, drug or alcohol abuse, fatty liver disease, primary sclerosing cholangitis and any major oral disease (periodontitis, caries) were excluded.

The HCs included age- and sex-matched healthy volunteers recruited from the local population. Controls were excluded if they were relatives of the patients and had a history of liver diseases, immunological disorders and major oral diseases. Most of the controls were patients' caregivers recruited in the hospital during follow-up and/or medical and research personnel involved in the study.

4.3. Sample Collection, Treatment and Acid-Insoluble Proteins Solubilization

Unstimulated whole saliva (WS) (from 0.2 to 1 mL) was collected with a soft plastic aspirator at the basis of the tongue from 9.00 am to 1.00 pm. in fasting conditions using a standard protocol optimized to preserve salivary proteins from proteolytic degradation. After collection, samples were immediately mixed with an equal volume of 0.2% (*v/v*) 2,2,2-trifluoroacetic acid (TFA) and centrifuged at $14,000\times g$ for 10 min at 4 °C. The insoluble fraction (pellet) was separated and stored at -80° until the solubilization.

Three different solutions were tested for pellet solubilization: (i) Solution 1, 2% SDS, 0.4 M sodium chloride (NaCl); (ii) Solution 2, 0.1 M sodium hydroxide (NaOH); (iii) Solution 3, 2% SDS, 0.5 mM DTT, 30 mM Tris-HCl pH 6.8. To improve solubilization, the three solutions were submitted to three cycles of 1 min sonication/1 min vortex followed by 5 min centrifugation $14,000\times g$, 4 °C. The resulting pellet was submitted to a further solubilization step under the same conditions. The insoluble material was discarded, and the soluble fraction submitted to TPC determination in duplicate by Pierce™ BCA Protein Assay kit (Thermo Fisher Scientific, Waltham, MA, USA), according to the provided kit instructions.

4.4. SDS-PAGE, Bands Excision and Enzymatic Digestion

All chemicals and reagents were purchased from Bio-Rad (Hercules, CA, USA). Ten μ g of proteins were loaded on SDS-PAGE under reducing conditions. Samples were previously treated with Laemmli Buffer [71] for 5 min at 100 °C. The electrophoretic separation, carried out at 180 Volt constant for 30 min, was performed using 4–15% Mini-PROTEAN® TGX™ Precast Protein Gels and Tris/Glycine Running Buffer (0.025 M Tris, 0.192 M Glycine, 0.1% SDS, pH 8.3). Molecular weights were determined by loading Precision Plus Dual Color Protein Standard. After the electrophoretic separation, gels were stained with Bio-Safe™ Coomassie Stain following the provided instructions. After destaining, each lane of the gel was divided into three slices corresponding to the following molecular weights: 250–75 kDa (A), 75–25 kDa (B) and < 25 kDa (C) (Figure S1), and slices, manually excised, cut into small pieces and submitted to in-gel digestion. Trypsin (Trypsin Singles, Proteomics Grade—Sigma-Aldrich/Merck, Darmstadt, Germany) was added to gel samples following the provided instructions in enzyme/proteins ratio of 1/80 (*w/w*) and incubated overnight at 37 °C. Extracted tryptic peptides were lyophilized and then solubilized in 0.1% formic acid (FA) for nano-RP-HPLC-high resolution ESI-MS and MS/MS analysis.

4.5. Nano-RP-HPLC-High Resolution ESI-MS/MS Analysis

All chemicals and reagents were purchased from Sigma-Aldrich/Merck. One hundred fifty-three samples corresponding to the tryptic peptides prepared by in-gel digestion of A, B and C slices obtained in triplicate from the acid-insoluble fraction of 17 saliva samples were analyzed, from September 2021 to November 2022, with an Ultimate 3000 Nano System HPLC (Dionex-Thermo Fisher Scientific) coupled with a LTQ Orbitrap Elite (Thermo Fisher Scientific). The Easy Spray reverse-phase nano column (250 mm \times 75 μ m inner diameter I.D., Thermo Fisher Scientific) was a C18 with 2 μ m beads and elution of peptides was achieved with aqueous solvent A (0.1% FA) and aqueous solvent B (0.1% FA, 80% ACN *v/v*) in 100 min at a flow rate of 0.3 μ L/min with the following gradient: 0–3 min at 4%B, 3–70 min 4–50% B, 70–90 min 50–80%B, 90–92 min 80–90%B, 92–100 min 90%B. The mass spectrometer was operating at 1.7 kV in the data-dependent acquisition mode, with the capillary temperature set at 275 °C and S-Lens RF level 68.4%. Full MS experiments were performed in positive ion mode from 350 to 1600 *m/z* with resolution 120,000 (at 400 *m/z*). The 10 most intense ions were subjected to CID fragmentation setting 35% of normalized collision energy for 10 ms, isolation width of 2 *m/z* and activation *q* of 0.25.

4.6. Protein Identification and Quantitation

Protein characterization was performed by Proteome Discoverer (PD) software (version 2.2, Thermo Fisher Scientific) with the SEQUEST HT cluster search engine (University

of Washington, licensed to Thermo Electron Corporation, San Jose, CA, USA) against the UniProtKB Homo sapiens database (188,453 entries, release 2019_03). MS spectra obtained from the three gel portions of the same sample were analyzed by PD as fractions of that sample. Database search parameters were as follows: carbamidomethylation of cysteine as fixed modification, oxidation of methionine, serine/threonine phosphorylation, N-terminal pyroglutamic residue and N-terminal acetylation as dynamic modifications and allowance for up to two missed tryptic cleavages. The peptide mass tolerance was set to 10 ppm and fragment ion mass tolerance was 0.6 Da. Peptides were filtered for high confidence and a minimum length of 6 amino acids; settings of FDR were 0.01 (strict) and 0.05 (relaxed).

Proteins identified were then filtered for high FDR confidence and for a minimum number of 2 unique peptides, excluding keratins. Grouping and quantification were set specifying categorical factor related to the condition AIHp, PBCp and HC. Proteins identified have been subjected to PD Label-Free Quantification and protein abundances have been determined based on area of unique peptide precursor ions. The mass spectrometry proteomics data have been deposited with the ProteomeXchange Consortium via the PRIDE [72] partner repository with the dataset identifier PXD039847.

4.7. Data Analysis

Only proteins found with at least 30% of distribution among all the three comparison groups (AIHp, PBCp and HCs) were selected for statistical analysis. Protein abundances were first automatically normalized against the total amount of tryptic peptides by PD software, then transformed to \log_2 and submitted to quantile normalization. Values under the detection limit were replaced with $\log_2(1000)$, a value below the minimum of the entire dataset. The choice of this conventional value was not critical for statistical analyses, as non-parametric tests and RF do not consider the continuous data distribution but only the rank order of values.

4.8. Protein-Protein Interaction Network, Topological Analysis and Pathway Enrichment

Protein-protein interaction (PPI) analysis was performed on 467 proteins found with at least 30% of distribution among all the three comparison groups (AIHp, PBCp and HCs) and on the 23 proteins obtained by merging those showing significant varied levels among AIHp, PBCp and HCs (14 proteins) based on Mann-Whitney tests, with the proteins selected by Boruta algorithm (17 proteins). Analysis was performed by STRING v.11.5 [73] (latest access on July 2023) with a default medium confidence of 0.4 and FDR stringency at 5%. Active interaction sources were based on “experiments”, “co-expression” and “co-occurrence”. Topological parameters of the PPI network, such as degree and betweenness centrality (BC) were calculated from Cytoscape v.3.10 [74]. The stringApp for Cytoscape was used to retrieve PPI networks from the STRING database [75] and topological analyses of the networks were performed using the Network Analyzer tool [76]. Two important metrics—degree and betweenness—were utilized to evaluate the importance of nodes in a network [77]. Hub proteins were identified by their very high degree of connectivity. Proteins with high betweenness centrality, namely bottlenecks, are key connectors in the PPI network, controlling the flow of information within a network [78].

For accessing the key nodes in the PPI, network members were first ranked by their degrees and BC afterwards the top-scoring proteins corresponding to 10% of the total number of nodes were selected. These topologically central proteins comprised the backbone of the PPI network. Selected proteins were further submitted to functional enrichment analysis based on reactome pathway distribution and GO biological process by stringApp, after removing redundant terms with default cutoff of 0.5.

4.9. Statistical Analysis

Various goodness-of-fit tests (Kolmogorov-Smirnov, Shapiro-Wilk, Lilliefors, etc.) showed a considerable deviation from normality of protein abundances. Thus, comparisons among the three groups were performed by Mann-Whitney (MW) and Kruskal-Wallis

(KW) non-parametric tests. Significant p -values of multiple tests were adjusted by the Benjamini–Hochberg procedure [79] to keep the cumulative false discovery ratio (FDR) among all tests less than 10%. Classification of subjects was obtained using Random Forest (RF) analysis and optimized by selecting a subset of relevant proteins identified by the Boruta method [80] and by tuning the two main RF parameters (total number of trees and number of features randomly sampled for each split point). RF classification was validated by the ‘out-of-bag’ samples. In detail, this method consists in using only about two-thirds of the samples for each decision tree. The classification obtained with these samples is then tested using the remaining one-third of the samples (hence the term ‘out-of-bag error’). This procedure is repeated for each of the planned number of trees (1000 in our analysis), each time randomly selecting the samples for classification and those for validation. The overall accuracy is ultimately assessed as the average of the ‘out-of-bag’ errors. The relative importance of each protein for classification was evaluated by the decrease in the Gini purity index and decrease in classification accuracy observed after temporarily excluding that protein from the analysis. Dimensionally reduced diagrams of the RF classification were obtained by multidimensional scaling (MDS) using the RF proximity between each two subjects. Proximity was calculated as the normalized frequency of trees containing the two subjects in the same end node. Statistical analyses were made using R (R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014. <http://www.R-project.org/>; latest access on 1 December 2022)

5. Conclusions

The bottom-up proteomic approach applied in this study, associated with a robust statistical analysis, allowed us to highlight a set of potential salivary biomarkers of AIH and PBC. The topology-based functional enrichment analysis performed on these potential salivary biomarkers highlighted an enrichment of proteins with multifaceted biological functions mostly related to the immune system, but also with strong involvement in the liver fibrosis process and with antimicrobial activity. Cofilin-1 was the protein with the highest centrality values in the backbone PPI network among those with varied levels, while Hornerin was the only protein showing reduced levels in both AIHp and PBCp compared to HCs and also able to better discriminate the patients of both groups from HCs in the RF analysis. RF analysis confirmed the feasibility of the salivary proteome to discriminate groups of subjects based on AIH or PBC occurrence as previously suggested by our group.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms241512207/s1>.

Author Contributions: Conceptualization and writing, B.M., A.O., T.C.; R.V., G.D.; mass spectrometry experiments, G.G.; mass spectrometry data analysis, G.G., C.C.; statistical analysis, G.D.; clinical assessment, L.C. and S.O. All authors have contributed to review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by University of Cagliari, FIR 2020.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Ethics Committee of Azienda Ospedaliero-Universitaria di Cagliari, Cagliari; ITALY (protocol code PG/2021/11303, date of approval 21 July 2021).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [72] partner repository with the dataset identifier PXD039847.

Acknowledgments: We acknowledge the CeSAR (Centro Servizi d’Ateneo per la Ricerca) of the University of Cagliari, Italy for the HR-MS/MS experiments performed with the HPLC-LTQ Orbitrap Elite. We acknowledge the Sardinian Regional Government for the financial support of Giulia

Guadalupi PhD scholarship (P.O.R. Sardegna F.S.E.—Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2014–2020—Axis III Education and training, Thematic goal 10, Investment Priority 10ii), Specific goal 10.5.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Heneghan, M.A.; Yeoman, A.D.; Verma, S.; Smith, A.D.; Longhi, M.S. Autoimmune hepatitis. *Lancet* **2013**, *382*, 1433–1444. [[CrossRef](#)] [[PubMed](#)]
2. Manns, M.P.; Czaja, A.J.; Gorham, J.D.; Krawitt, E.L.; Mieli-Vergani, G.; Vergani, D.; Vierling, J.M. American Association for the Study of Liver Diseases. Diagnosis and management of autoimmune hepatitis. *Hepatology* **2010**, *51*, 2193–2213. [[CrossRef](#)] [[PubMed](#)]
3. Mack, C.L.; Adams, D.; Assis, D.N.; Kerkar, N.; Manns, M.P.; Mayo, M.J.; Vierling, J.M.; Alsawas, M.; Murad, M.H.; Czaja, A.J. Diagnosis and Management of Autoimmune Hepatitis in Adults and Children: 2019 Practice Guidance and Guidelines From the American Association for the Study of Liver Diseases. *Hepatology* **2020**, *72*, 671–722. [[CrossRef](#)] [[PubMed](#)]
4. Manns, M.P.; Lohse, A.W.; Vergani, D. Autoimmune hepatitis—Update 2015. *J. Hepatol.* **2015**, *62*, S100–S111. [[CrossRef](#)]
5. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. *J. Hepatol.* **2017**, *67*, 145–172. [[CrossRef](#)] [[PubMed](#)]
6. Lleo, A.; Leung, P.S.C.; Hirschfield, G.M.; Gershwin, E.M. The Pathogenesis of Primary Biliary Cholangitis: A Comprehensive Review. *Semin. Liver Dis.* **2020**, *40*, 34–48. [[CrossRef](#)]
7. Chazouillères, O.; Wendum, D.; Serfaty, L.; Montembault, S.; Rosmorduc, O.; Poupon, R. Primary biliary cirrhosis-autoimmune hepatitis overlap syndrome: Clinical features and response to therapy. *Hepatology* **1998**, *28*, 296–301. [[CrossRef](#)]
8. Czaja, A.J. Diagnosis and management of the overlap syndromes of autoimmune hepatitis. *Can. J. Gastroenterol.* **2013**, *27*, 417–423. [[CrossRef](#)]
9. Schulz, L.; Sebode, M.; Weidemann, S.A.; Lohse, A.W. Variant syndromes of primary biliary cholangitis. *Best Pract. Res. Clin. Gastroenterol.* **2018**, *34*, 55–61. [[CrossRef](#)]
10. Schiulthiß, C.; Steinmann, S.; Willscher, E.; Paschold, L.; Lohse, A.W.; Binder, M. Immune signatures in variant syndromes of primary biliary cholangitis and autoimmune hepatitis. *Hepatol. Commun.* **2023**, *7*, e0123. [[CrossRef](#)]
11. Floreani, A.; Franceschet, I.; Cazzagon, N.; Spinazzè, A.; Buja, A.; Furlan, P.; Baldo, V.; Gershwin, M.E. Extrahepatic autoimmune conditions associated with primary biliary cirrhosis. *Clin. Rev. Allergy Immunol.* **2015**, *48*, 192–197. [[CrossRef](#)] [[PubMed](#)]
12. Wong, G.W.; Yeong, T.; Lawrence, D.; Yeoman, A.D.; Verma, S.; Heneghan, M.A. Concurrent extrahepatic autoimmunity in autoimmune hepatitis: Implications for diagnosis, clinical course and long-term outcomes. *Liver Int.* **2017**, *37*, 449–457. [[CrossRef](#)] [[PubMed](#)]
13. Vierling, J.M. Diagnosis and treatment of autoimmune hepatitis. *Curr. Gastroenterol. Rep.* **2012**, *14*, 25–36. [[CrossRef](#)] [[PubMed](#)]
14. Liu, H.; Norman, G.L.; Shums, Z.; Worman, H.J.; Krawitt, E.L.; Bizzaro, N.; Vergani, D.; Bogdanos, D.P.; Dalekos, G.N.; Milkiewicz, P.; et al. PBC screen: An IgG/IgA dual isotype ELISA detecting multiple mitochondrial and nuclear autoantibodies specific for primary biliary cirrhosis. *J. Autoimmun.* **2010**, *35*, 436–442. [[CrossRef](#)]
15. Granito, A.; Muratori, L.; Muratori, P.; Pappas, G.; Guidi, M.; Cassani, F.; Volta, U.; Ferri, A.; Lenzi, M.; Bianchi, F.B. Antibodies to filamentous actin (F-actin) in type 1 autoimmune hepatitis. *J. Clin. Pathol.* **2006**, *59*, 280–284. [[CrossRef](#)]
16. Gerosa, C.; Fanos, V.; Fanni, D.; Van Eyken, P.; Ambu, R.; Nemolato, S.; Floris, G.; Iacovidou, N.; Gibo, Y.; Nurchi, A.M.; et al. Liver Biopsy Interpretation in the Differential Diagnosis of Autoimmune Liver Disease in Children. *J. Pediatr. Neonat. Individual. Med.* **2013**, *2*, e020229. [[CrossRef](#)]
17. Tan, D.; Goodman, Z.D. Liver biopsy in primary biliary cholangitis: Indications and interpretation. *Clin. Liver Dis.* **2018**, *22*, 579–588. [[CrossRef](#)]
18. Contini, C.; Serrao, S.; Manconi, B.; Olianias, A.; Iavarone, F.; Bizzarro, A.; Masullo, C.; Castagnola, M.; Messina, I.; Diaz, G.; et al. Salivary Proteomics Reveals Significant Changes in Relation to Alzheimer’s Disease and Aging. *J. Alzheimers Dis.* **2022**, *89*, 605–622. [[CrossRef](#)]
19. Sanna, M.; Firinu, D.; Manconi, P.E.; Pisanu, M.; Murgia, G.; Piras, V.; Castagnola, M.; Messina, I.; Del Giacco, S.; Cabras, T. The salivary proteome profile in patients affected by SAPHO syndrome characterized by a top-down RP-HPLC-ESI-MS platform. *Mol. Biosyst.* **2015**, *11*, 1552–1562. [[CrossRef](#)]
20. Cabras, T.; Sanna, M.; Manconi, B.; Fanni, D.; Demelia, L.; Sorbello, O.; Iavarone, F.; Castagnola, M.; Faa, G.; Messina, I. Proteomic investigation of whole saliva in Wilson’s disease. *J. Proteom.* **2015**, *128*, 154–163. [[CrossRef](#)]
21. Manconi, B.; Liori, B.; Cabras, T.; Vincenzoni, F.; Iavarone, F.; Loreface, L.; Cocco, E.; Castagnola, M.; Messina, I.; Olianias, A. Top-down proteomic profiling of human saliva in multiple sclerosis patients. *J. Proteom.* **2018**, *187*, 212–222. [[CrossRef](#)] [[PubMed](#)]
22. Boroumand, M.; Olianias, A.; Cabras, T.; Manconi, B.; Fanni, D.; Faa, G.; Desiderio, C.; Messina, I.; Castagnola, M. Saliva, a bodily fluid with recognized and potential diagnostic applications. *J. Sep. Sci.* **2021**, *44*, 3677–3690. [[CrossRef](#)] [[PubMed](#)]

23. Rao, B.; Lou, J.; Lu, H.; Liang, H.; Li, J.; Zhou, H.; Fan, Y.; Zhang, H.; Sun, Y.; Zou, Y.; et al. Oral Microbiome Characteristics in Patients with Autoimmune Hepatitis. *Front. Cell Infect. Microbiol.* **2021**, *11*, 656674. [[CrossRef](#)] [[PubMed](#)]
24. Messana, I.; Cabras, T.; Pisano, E.; Sanna, M.T.; Olianias, A.; Manconi, B.; Pellegrini, M.; Paludetti, G.; Scarano, E.; Fiorita, A.; et al. Trafficking and postsecretory events responsible for the formation of secreted human salivary peptides. *Mol. Cell Proteom.* **2008**, *7*, 911–926. [[CrossRef](#)]
25. Ghallab, N.A. Diagnostic potential and future directions of biomarkers in gingival crevicular fluid and saliva of periodontal diseases: Review of the current evidence. *Arch. Oral Biol.* **2018**, *87*, 115–124. [[CrossRef](#)]
26. Iavarone, F.; Olianias, A.; Patini, R.; Gallenzi, P.; Di Tonno, L.; Desiderio, C.; Cabras, T.; Manconi, B.; Vincenzoni, F.; Cordaro, M.; et al. Top down proteomic analysis of gingival crevicular fluid in deciduous, exfoliating and permanent teeth in children. *J. Proteom.* **2020**, *226*, 103890. [[CrossRef](#)] [[PubMed](#)]
27. Grassl, N.; Kulak, N.A.; Pichler, G.; Geyer, P.E.; Jung, J.; Schubert, S.; Sinitcyn, P.; Cox, J.; Mann, M. Ultra-deep and quantitative saliva proteome reveals dynamics of the oral microbiome. *Genome Med.* **2016**, *8*, 44. [[CrossRef](#)] [[PubMed](#)]
28. Olianias, A.; Guadalupi, G.; Cabras, T.; Contini, C.; Serrao, S.; Iavarone, F.; Castagnola, M.; Messana, I.; Onali, S.; Chessa, L.; et al. Top-Down Proteomics Detection of Potential Salivary Biomarkers for Autoimmune Liver Diseases Classification. *Int. J. Mol. Sci.* **2023**, *24*, 959. [[CrossRef](#)]
29. Joy, M.P.; Brock, A.; Ingber, D.E.; Huang, S. High-betweenness proteins in the yeast protein interaction network. *J. Biomed. Biotechnol.* **2005**, *2005*, 96–103. [[CrossRef](#)]
30. Nithya, C.; Kiran, M.; Nagarajaram, H.A. Dissection of hubs and bottlenecks in a protein-protein interaction network. *Comput. Biol. Chem.* **2023**, *102*, 107802. [[CrossRef](#)]
31. Shibuya, A.; Ikewaki, N. High serum glyceraldehyde-3-phosphate dehydrogenase levels in patients with liver cirrhosis. *Hepatol. Res.* **2002**, *22*, 174–179. [[CrossRef](#)] [[PubMed](#)]
32. Domenicali, M.; Baldassarre, M.; Giannone, F.A.; Naldi, M.; Mastroroberto, M.; Biselli, M.; Laggetta, M.; Patrono, D.; Bertucci, C.; Bernardi, M.; et al. Posttranscriptional changes of serum albumin: Clinical and prognostic significance in hospitalized patients with cirrhosis. *Hepatology* **2014**, *60*, 1851–1860. [[CrossRef](#)] [[PubMed](#)]
33. D'Amico, G.; Garcia-Tsao, G.; Pagliaro, L. Natural history and prognostic indicators of survival in cirrhosis: A systematic review of 118 studies. *J. Hepatol.* **2006**, *44*, 217–231. [[CrossRef](#)] [[PubMed](#)]
34. Carvalho, J.R.; Verdelho Machado, M. New Insights About Albumin and Liver Disease. *Ann. Hepatol.* **2018**, *17*, 547–560. [[CrossRef](#)] [[PubMed](#)]
35. Xu, J.; Li, Y. Discovering disease-genes by topological features in human protein-protein interaction network. *Bioinformatics* **2006**, *22*, 2800–2805. [[CrossRef](#)] [[PubMed](#)]
36. Gieseck, R.L., III; Ramalingam, T.R.; Hart, K.M.; Vannella, K.M.; Cantu, D.A.; Lu, W.Y.; Ferreira-González, S.; Forbes, S.J.; Vallier, L.; Wynn, T.A. Interleukin-13 Activates Distinct Cellular Pathways Leading to Ductular Reaction, Steatosis, and Fibrosis. *Immunity* **2016**, *45*, 145–158. [[CrossRef](#)]
37. Webb, G.J.; Siminovitch, K.A.; Hirschfield, G.M. The immunogenetics of primary biliary cirrhosis: A comprehensive review. *J. Autoimmun.* **2015**, *64*, 42–52. [[CrossRef](#)]
38. Liaskou, E.; Patel, S.R.; Webb, G.; Bagkou Dimakou, D.; Akiror, S.; Krishna, M.; Mells, G.; Jones, D.E.; Bowman, S.J.; Barone, F.; et al. Increased sensitivity of Treg cells from patients with PBC to low dose IL-12 drives their differentiation into IFN- γ secreting cells. *J. Autoimmun.* **2018**, *94*, 143–155. [[CrossRef](#)]
39. Povero, D.; Busletta, C.; Novo, E.; di Bonzo, L.V.; Cannito, S.; Paternostro, C.; Parola, M. Liver fibrosis: A dynamic and potentially reversible process. *Histol. Histopathol.* **2010**, *25*, 1075–1091. [[CrossRef](#)]
40. Vinken, M.; Papeleu, P.; Snykers, S.; De Rop, E.; Henkens, T.; Chipman, J.K.; Rogiers, V.; Vanhaecke, T. Involvement of cell junctions in hepatocyte culture functionality. *Crit. Rev. Toxicol.* **2006**, *36*, 299–318. [[CrossRef](#)]
41. Zhou, L.; Pradhan-Sundd, T.; Poddar, M.; Singh, S.; Kikuchi, A.; Stolz, D.B.; Shou, W.; Li, Z.; Nejak-Bowen, K.N.; Monga, S.P. Mice with Hepatic Loss of the Desmosomal Protein γ -Catenin Are Prone to Cholestatic Injury and Chemical Carcinogenesis. *Am. J. Pathol.* **2015**, *185*, 3274–3289. [[CrossRef](#)]
42. Nakagawa, H.; Hikiba, Y.; Hirata, Y.; Font-Burgada, J.; Sakamoto, K.; Hayakawa, Y.; Taniguchi, K.; Umemura, A.; Kinoshita, H.; Sakitani, K.; et al. Loss of liver E-cadherin induces sclerosing cholangitis and promotes carcinogenesis. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1090–1095. [[CrossRef](#)] [[PubMed](#)]
43. Dubash, A.D.; Kam, C.Y.; Aguado, B.A.; Patel, D.M.; Delmar, M.; Shea, L.D.; Green, K.J. Plakophilin-2 loss promotes TGF- β 1/p38 MAPK-dependent fibrotic gene expression in cardiomyocytes. *J. Cell Biol.* **2016**, *212*, 425–438. [[CrossRef](#)] [[PubMed](#)]
44. Radwan, N.A.; Ahmed, N.S. The diagnostic value of arginase-1 immunostaining in differentiating hepatocellular carcinoma from metastatic carcinoma and cholangiocarcinoma as compared to HepPar-1. *Diagn. Pathol.* **2012**, *7*, 149. [[CrossRef](#)] [[PubMed](#)]
45. Zhang, M.; Wu, Z.; Salas, S.S.; Aguilar, M.M.; Trillos-Almanza, M.C.; Buist-Homan, M.; Moshage, H. Arginase 1 expression is increased during hepatic stellate cell activation and facilitates collagen synthesis. *J. Cell Biochem.* **2023**, *124*, 808–817. [[CrossRef](#)]
46. Ogino, K.; Kubo, M.; Takahashi, H.; Zhang, R.; Zou, Y.; Fujikura, Y. Anti-inflammatory effect of arginase inhibitor and corticosteroid on airway allergic reactions in a *Dermatophogoides farinae*-induced NC/Nga mouse model. *Inflammation* **2013**, *36*, 141–151. [[CrossRef](#)]
47. Hijazi, N.; Rockey, D.C.; Shi, Z. The cellular microenvironment and cytoskeletal actin dynamics in liver fibrogenesis. *Biocell* **2022**, *46*, 2003–2007. [[CrossRef](#)]

48. Xu, Y.; Duan, J.; Ji, W.; Liu, C.; Li, X.; Wu, Q.; Gao, C.; Su, C. A novel matrine derivative, WM130, inhibits activation and movement of human hepatic stellate LX-2 cells by targeting cofilin 1. *Cytotechnology* **2022**, *74*, 613–622. [[CrossRef](#)]
49. Yao, B.; Li, Y.; Chen, T.; Niu, Y.; Wang, Y.; Yang, Y.; Wei, X.; Liu, Q.; Tu, K. Hypoxia-induced cofilin 1 promotes hepatocellular carcinoma progression by regulating the PLD1/AKT pathway. *Clin. Transl. Med.* **2021**, *11*, e366. [[CrossRef](#)]
50. Witke, W.; Sharpe, A.H.; Hartwig, J.H.; Azuma, T.; Stossel, T.P.; Kwiatkowski, D.J. Hemostatic, inflammatory, and fibroblast responses are blunted in mice lacking gelsolin. *Cell* **1995**, *81*, 41–51. [[CrossRef](#)]
51. Aigelsreiter, A.; Janig, E.; Sostaric, J.; Pichler, M.; Unterthor, D.; Halasz, J.; Lackner, C.; Zatloukal, K.; Denk, H. Clusterin expression in cholestasis, hepatocellular carcinoma and liver fibrosis. *Histopathology* **2009**, *54*, 561–570. [[CrossRef](#)]
52. Seo, H.Y.; Lee, S.H.; Lee, J.H.; Kang, Y.N.; Choi, Y.K.; Hwang, J.S.; Park, K.G.; Jang, B.K.; Kim, M.K. Clusterin Attenuates Hepatic Fibrosis by Inhibiting Hepatic Stellate Cell Activation and Downregulating the Smad3 Signaling Pathway. *Cells* **2019**, *8*, 1442. [[CrossRef](#)]
53. Jimenez-Rivera, C.; Jolin-Dahel, K.S.; Fortinsky, K.J.; Gozdyra, P.; Benchimol, E.I. International incidence and outcomes of biliary atresia. *J. Pediatr. Gastroenterol. Nutr.* **2013**, *56*, 344–354. [[CrossRef](#)]
54. Sahin, M.; Yetim, A.; Ates, F. The role of tacstd-2 level in hepatitis C patients (controlled clinical research). *Libyan J. Med.* **2019**, *14*, 1622363. [[CrossRef](#)] [[PubMed](#)]
55. Alberti, S.; Miotti, S.; Stella, M.; Klein, C.E.; Fornaro, M.; Menard, S.; Colnaghi, M.I. Biochemical characterization of Trop-2, a cell surface molecule expressed by human carcinomas: Formal proof that the monoclonal antibodies T16 and MOv-16 recognize Trop-2. *Hybridoma* **1992**, *11*, 539–545. [[CrossRef](#)] [[PubMed](#)]
56. Sekhar, V.; Pollicino, T.; Diaz, G.; Engle, R.E.; Alayli, F.; Melis, M.; Kabat, J.; Tice, A.; Pomerence, A.; Altan-Bonnet, N.; et al. Infection with hepatitis C virus depends on TACSTD2, a regulator of claudin-1 and occludin highly downregulated in hepatocellular carcinoma. *PLoS Pathog.* **2018**, *14*, e1006916. [[CrossRef](#)] [[PubMed](#)]
57. Fu, S.J.; Shen, S.L.; Li, S.Q.; Hua, Y.P.; Hu, W.J.; Guo, B.; Peng, B.G. Hornerin promotes tumor progression and is associated with poor prognosis in hepatocellular carcinoma. *BMC Cancer* **2018**, *18*, 815. [[CrossRef](#)] [[PubMed](#)]
58. Henry, J.; Hsu, C.Y.; Haftek, M.; Nachat, R.; de Koning, H.D.; Gardinal-Galera, I.; Hitomi, K.; Balica, S.; Jean-Decoster, C.; Schmitt, A.M.; et al. Hornerin is a component of the epidermal cornified cell envelopes. *FASEB J.* **2011**, *25*, 1567–1576. [[CrossRef](#)]
59. Arteel, G.E.; Naba, A. The liver matrisome—Looking beyond collagens. *JHEP Rep.* **2020**, *2*, 100115. [[CrossRef](#)]
60. Bajaj, J.S.; Betrapally, N.S.; Hylemon, P.B.; Heuman, D.M.; Daita, K.; White, M.B.; Unser, A.; Thacker, L.R.; Sanyal, A.J.; Kang, D.J.; et al. Salivary microbiota reflects changes in gut microbiota in cirrhosis with hepatic encephalopathy. *Hepatology* **2015**, *62*, 1260–1271. [[CrossRef](#)]
61. Elghannam, M.T.; Hassanien, M.H.; Ameen, Y.A.; Turky, E.A.; Elattar, G.M.; ElRay, A.A.; Eltalkawy, M.D. Oral microbiota and liver diseases. *Clin. Nutr. ESPEN* **2023**, *54*, 68–72. [[CrossRef](#)] [[PubMed](#)]
62. Beamer, L.J.; Carroll, S.F.; Eisenberg, D. The BPI/LBP family of proteins: A structural analysis of conserved regions. *Protein Sci.* **1998**, *7*, 906–914. [[CrossRef](#)]
63. Lv, L.; Jiang, H.; Chen, X.; Wang, Q.; Wang, K.; Ye, J.; Li, Y.; Fang, D.; Lu, Y.; Yang, L.; et al. The Salivary Microbiota of Patients With Primary Biliary Cholangitis Is Distinctive and Pathogenic. *Front. Immunol.* **2021**, *12*, 713647. [[CrossRef](#)] [[PubMed](#)]
64. Nandula, S.R.; Huxford, I.; Wheeler, T.T.; Aparicio, C.; Gorr, S.U. The parotid secretory protein BPIFA2 is a salivary surfactant that affects lipopolysaccharide action. *Exp. Physiol.* **2020**, *105*, 1280–1292. [[CrossRef](#)] [[PubMed](#)]
65. Sallustio, F.; Stasi, A.; Curci, R.; Divella, C.; Picerno, A.; Franzin, R.; De Palma, G.; Rutigliano, M.; Lucarelli, G.; Battaglia, M.; et al. Renal progenitor cells revert LPS-induced endothelial-to-mesenchymal transition by secreting CXCL6, SAA4, and BPIFA2 antiseptic peptides. *FASEB J.* **2019**, *33*, 10753–10766. [[CrossRef](#)] [[PubMed](#)]
66. Yokoyama, T.; Komori, A.; Nakamura, M.; Takii, Y.; Kamihira, T.; Shimoda, S.; Mori, T.; Fujiwara, S.; Koyabu, M.; Taniguchi, K.; et al. Human intrahepatic biliary epithelial cells function in innate immunity by producing IL-6 and IL-8 via the TLR4-NF-kappaB and -MAPK signaling pathways. *Liver Int.* **2006**, *26*, 467–476. [[CrossRef](#)]
67. Bazil, V.; Strominger, J.L. Shedding as a mechanism of down-modulation of CD14 on stimulated human monocytes. *J. Immunol.* **1991**, *147*, 1567–1574. [[CrossRef](#)]
68. Umemura, T.; Sekiguchi, T.; Joshita, S.; Yamazaki, T.; Fujimori, N.; Shibata, S.; Ichikawa, Y.; Komatsu, M.; Matsumoto, A.; Shums, Z.; et al. Association between serum soluble CD14 and IL-8 levels and clinical outcome in primary biliary cholangitis. *Liver Int.* **2017**, *37*, 897–905. [[CrossRef](#)]
69. Karlmark, K.R.; Tacke, F.; Dunay, I.R. Monocytes in health and disease—Minireview. *Eur. J. Microbiol. Immunol.* **2012**, *2*, 97–102. [[CrossRef](#)]
70. Boberg, K.M. Prevalence and epidemiology of autoimmune hepatitis. *Clin. Liver Dis.* **2002**, *6*, 635–647. [[CrossRef](#)]
71. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680–685. [[CrossRef](#)] [[PubMed](#)]
72. Perez-Riverol, Y.; Bai, J.; Bandla, C.; Garcia-Seisdedos, D.; Hewapathirana, S.; Kamatchinathan, S.; Kundu, D.J.; Prakash, A.; Frericks-Zipper, A.; Eisenacher, M.; et al. The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **2022**, *50*, D543–D552. [[CrossRef](#)]
73. Szklarczyk, D.; Kirsch, R.; Koutrouli, M.; Nastou, K.; Mehryary, F.; Hachilif, R.; Gable, A.L.; Fang, T.; Doncheva, N.T.; Pyysalo, S.; et al. The STRING database in 2023: Protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **2023**, *51*, D638–D646. [[CrossRef](#)]

74. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)] [[PubMed](#)]
75. Doncheva, N.T.; Morris, J.H.; Gorodkin, J.; Jensen, L.J. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J. Proteome Res.* **2019**, *18*, 623–632. [[CrossRef](#)] [[PubMed](#)]
76. Assenov, Y.; Ramírez, F.; Schelhorn, S.E.; Lengauer, T.; Albrecht, M. Computing topological parameters of biological networks. *Bioinformatics* **2008**, *24*, 282–284. [[CrossRef](#)] [[PubMed](#)]
77. Chen, C.; Shen, H.; Zhang, L.G.; Liu, J.; Cao, X.G.; Yao, A.L.; Kang, S.S.; Gao, W.X.; Han, H.; Cao, F.H.; et al. Construction and analysis of protein-protein interaction networks based on proteomics data of prostate cancer. *Int. J. Mol. Med.* **2016**, *37*, 1576–1586. [[CrossRef](#)]
78. Yu, H.; Kim, P.M.; Sprecher, E.; Trifonov, V.; Gerstein, M. The importance of bottlenecks in protein networks: Correlation with gene essentiality and expression dynamics. *PLoS Comput. Biol.* **2007**, *3*, e59. [[CrossRef](#)]
79. Benjamini, Y.; and Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Statist. Soc. B* **1995**, *57*, 289–300. [[CrossRef](#)]
80. Kursa, M.B.; Jankowski, A.; Rudnicki, W.R. Boruta—A system for feature selection. *Fundam. Inform.* **2010**, *101*, 271–285. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.