1 Quantitative Analysis of Tissue Secretome Reveals the 2 Diagnostic and Prognostic Value of Carbonic Anhydrase II in 3 Hepatocellular Carcinoma

4 Xiaohua Xing^{1, 2, 3, 4, a}, Hui Yuan^{2, 5, b}, Hongzhi Liu^{2, c}, Xionghong Tan^{2, d}, Bixing
5 Zhao^{2, e}, Yingchao Wang^{2, f}, Jiahe Ouyang^{2, g}, Minjie Lin^{2, h}, Aimin Huang^{1, 3, 4, *, i},
6 Xiaolong Liu^{1, 2, *, j}

7 ¹ Department of Pathology, School of Basic Medical Sciences of Fujian Medical
8 University, Fuzhou 350004, China

9 ² The United Innovation of Mengchao Hepatobiliary Technology Key Laboratory of
10 Fujian Province, Mengchao Hepatobiliary Hospital of Fujian Medical University,
11 Fuzhou 350025, China

12 ³ Institute of Oncology, Fujian Medical University, Fuzhou 350004, China

13 ⁴ Pathology Diagnostic Center, Fujian Medical University, Fuzhou 350004, China

14 ⁵ Liver Disease Center, The First Affiliated Hospital of Fujian Medical University,
15 Fuzhou 350007, China

16

17 * Corresponding author(s).

18 E-mail: xiaoloong.liu@gmail.com (Liu X), aimin@fjmu.edu.cn (Huang A).

19

20 **Running title:** Xing X et al / Diagnostic and Prognostic value of CA2 in HCC

21 ^aORCID: 0000-0001-8971-8577

22 ^bORCID: 0000-0002-7858-3539

23 ^cORCID: 0000-0001-7102-6855

24 ^dORCID: 0000-0001-9113-4159

25 ^eORCID: 0000-0002-0055-3989

26 ^fORCID: 0000-0003-3590-9233

27 ^gORCID: 0000-0001-9248-8805

28 ^hORCID: 0000-0002-0523-2922

29 ⁱORCID: 0000-0002-3141-0315

30 ^jORCID: 0000-0002-3096-4981

31

32 List the counts of words: 6499 words

33 Tables and figures: 0 table and 6 figures

34 Supplementary figures: 4

35 Supplementary tables: 6

36 **Abstract**

37 Early detection and intervention are key strategies to reduce mortality, increase
38 long-term survival and improve the therapeutic effects of hepatocellular carcinoma
39 (HCC) patients. Herein, the isobaric tag for relative and absolute quantitation (iTRAQ)
40 quantitative proteomic strategy was used to study the secretome in conditioned media
41 from HCC cancerous tissues, surrounding noncancerous and distal noncancerous
42 tissues to identify diagnostic and prognostic biomarkers for HCC. In total, 22 and 49
43 secretory proteins were identified to be dysregulated in the cancerous and surrounding
44 noncancerous tissues compared with the distal noncancerous tissues. Among these
45 proteins, carbonic anhydrase II (CA2) was identified to be significantly upregulated in
46 the secretome of cancerous tissues; correspondingly, the serum concentrations of CA2
47 were remarkably increased in HCC patients than that in normal populations.
48 Interestingly, a significant increase of serum CA2 in recurrent HCC patients after
49 radical resection was also confirmed compared with HCC patients without recurrence,
50 and the serum level of CA2 could act as an independent prognostic factor for time to
51 recurrence (TTR) and overall survival (OS). Regarding the mechanism, the secreted
52 CA2 enhances the migration and invasion of HCC cells by activating the epithelial
53 mesenchymal transition (EMT) pathway. Taken together, this study identified a novel
54 biomarker for HCC diagnosis and prognosis and provides a valuable resource of the
55 HCC secretome for investigating serological biomarkers.

56

57 **KEYWORDS:** Hepatocellular carcinoma (HCC); Tissue secretome; Carbonic
58 anhydrase II (CA2); Early diagnosis/prognosis; Epithelial mesenchymal transition
59 (EMT).

60 Introduction

61 Hepatocellular carcinoma (HCC) has a high incidence and mortality, making it the
 62 fifth most common malignant cancer worldwide [1]. Although surgical strategy was
 63 proven to be the most suitable option for HCC therapy, most HCC patients were not
 64 diagnosed or intervened until the advanced stage, rendering them unsuitable for
 65 surgical treatments [2] or having a poor prognosis after surgical excision. Currently,
 66 the five-year survival rate of HCC patients is less than 20%, while the five-year
 67 recurrence and metastasis rate is greater than 80% [3–6]. So far, alpha fetoprotein
 68 (AFP) and des-gamma-carboxy prothrombin (DCP) are the most widely accepted and
 69 clinically applied biomarkers for HCC diagnosis and monitoring. However, the
 70 sensitivity and specificity of AFP and DCP in the early diagnosis as well as prognosis
 71 evaluation of HCC are still insufficient [7–9]. Therefore, it's urgently required to find
 72 novel biomarkers that are highly specific and sensitive to provide an early diagnosis
 73 and prognosis evaluation for HCC.

74 Proteomic-based technology has become a very useful and powerful analytical
 75 tool for biomarker screening [10–13]. A desirable biomarker for HCC diagnosis or
 76 monitoring should be able to be measured in body fluid samples such as serum and
 77 plasma [14], because these samples are low-cost, easy to collect and process, and are
 78 amenable to repeat sampling whenever it is necessary. Therefore, the serum and
 79 plasma are also the ideal targets for proteomic studies that aim to identify diagnostic
 80 or prognostic biomarkers for HCC [15,16]. However, the complex nature of serum
 81 and plasma, as well as their large dynamic concentration range of different proteins,
 82 significantly hinders the progress of proteomic-based biomarker screening.

83 Secreted proteins play important roles in signal transduction, cellular growth,
 84 proliferation, apoptosis and even in tumorigenesis, development, invasion and
 85 metastasis, and are ideal sources for biomarker screening [17]. Investigating the
 86 secretome of HCC tissues or cells may provide valuable information for relevant
 87 studies. Recently, the application of secretomics in screening diagnostic or prognostic

protein biomarkers in HCC cell lines have been reported by many groups [18–20]. However, these results must still be clinically validated [21]. Therefore, it would be more straightforward and convincing to analyze the secretome of primary tumour tissue cultures to identify the diagnostic or prognostic biomarkers for HCC. For example, Yang et al. have established an *in vitro* tissue culture system for HCC and identified matrix metalloproteinase 1 (MMP1) as a diagnostic biomarker for HCC [22]; however, the influence of hepatitis B virus (HBV) infection was not analyzed in this study.

In the present study, we collected serum-free culture media (CM) from the tissue cultures of tumour tissues, surrounding non-tumoral tissues and distal non-tumoral tissues of HCC patients and analyzed the secretome to identify potential diagnostic and prognostic biomarkers for HCC via an iTRAQ-based quantitative proteomic approach. Meanwhile, the sensitivity, specificity and clinical significance of the identified biomarkers was also carefully validated in a large-scale HCC patient cohort by ELISA assay and targeted proteomics of parallel reaction monitoring (PRM). Furthermore, the corresponding molecular mechanisms of the identified biomarker was also carefully explored.

Results and Discussion

Cells in the *in vitro*-cultured tissues were alive and secretory

To ensure that the cells in the *in vitro*-cultured tissues were alive and that the secretome was not contaminated by intracellular proteins, a series of analyses were performed (Figure 1A). We used haematoxylin and eosin (HE) staining to evaluate the cell morphology changes of tissues cultured for 0 day, 1 day and 2 days. As revealed in Figure 1B, the HE-stained tissue sections showed the corresponding characteristic anatomical details of HCC cancerous, surrounding noncancerous and distal noncancerous tissues in all of the respective cultures. With the extension of incubation time, the cells in the cultured tissues were starved and showed necrosis or apoptosis

115 due to the lack of nutrients. The morphology of tissues cultured for 1 day was still
 116 very similar to that of the fresh tissues (cultured 0 days). By contrast, the number of
 117 cell nuclei in tissues cultured for 2 days significantly decreased due to cell necrosis or
 118 apoptosis during the culture process (Figure 1B). TdT-mediated dUTP nick-end
 119 labeling (TUNEL) staining was further used to evaluate the apoptosis rates of tissues
 120 cultured for 0, 1 and 2 days. As revealed in Figure 1C, the tissues cultured for 2 days
 121 had a significantly higher apoptosis rate compared with tissues cultured for 0 or 1
 122 days. In addition, the proteins extracted from culture supernatants with different
 123 culture times were examined by SDS-PAGE. As shown in Figure 1D, the molecular
 124 mass distribution of the extracted proteins was significantly changed with the increase
 125 of culture time. Furthermore, western blot assays clearly demonstrated the prevention
 126 of contamination by intracellular proteins (Figure 1E). Taken together, these results
 127 suggested that the 1 day culture time was the optimal time point for collecting culture
 128 supernatants.

129 **The secretome characteristics were comprehensive and reasonable**

130 iTRAQ labeling combined with mass spectrometry was applied to investigate the
 131 entirety of secretome changes in primary HCC tissues from patients. The features of
 132 the HCC patients in the current study were listed in Table S1. Total proteins extracted
 133 from the supernatant collected from the culture medium of HCC tissues and their
 134 surrounding and distal noncancerous tissues were analyzed using 2D LC-MS/MS, as
 135 shown in Figure 2. We quantified 2388 proteins in total using Scaffold_4.3.2, of
 136 which 1312 proteins were annotated or predicted as secretory proteins, accounting for
 137 54.9% of the total quantified proteins. This result covered 75.7% of the secretory
 138 proteins previously reported by Yang et. Al [19] (Figure S1A), and the percentage of
 139 secretory proteins was much higher than in the human protein database (23%) (Figure
 140 S1B). Among these proteins, 936 secretory proteins were overlapped in the 5 different
 141 replicates, which accounted for 71.4% of the quantified secretory proteins. The
 142 complete list of identified secretory proteins was shown in Table S2. The detailed
 143 features of identified secretory proteins, including the NM-score, MetazSecKB

characteristics, isoelectric point (pI), molecular weight (MW), hydrophobicity and quantification results were also included in the list. The MS-based proteomic data were deposited in the integrated proteome resources of iProX public repository (iProX Accession: IPX0001425001) at <http://www.iprox.org/page/MSV022.html> [23].

The MW distribution of the secretory proteins ranged from 7396 to 628685 Da, with a primary range of 10~40 kD, indicating smaller molecular weight for the secretory proteins (Figure 3A). The pI values ranged from 3.67 to 12.56 and were mainly in the range of 4.4~7.6, which coincides with the microenvironment of liver tissues (Figure 3B). The hydrophobicity of the proteins ranged from 1.1 to 3.3 and was mainly in the range of 1.6~2.0, which implied the enrichment of membrane or transmembrane proteins (Figure S1C).

We also investigated the GO annotations containing cell components, molecular functions and biological processes of the secretory proteins. Cell component category showed that the secretory proteins were primarily extracellular, which suggests the excellent purity of HCC the tissue secretome (Figure 3C). The biological process category indicated that these secretory proteins were mainly involved in biosynthetic process, signal transduction, and transport process (Figure 3D). The molecular function category indicated that these secretory proteins played major roles in RNA binding, enzyme binding, DNA binding and transmembrane transporter activities (Figure 3E). These results were consistent with those of previously reported secretory proteins from HCC tissues [19].

Various functions of secretory proteins in different HCC-related tissues

For proteomic analysis, HCC tissues, surrounding noncancerous tissues and distal noncancerous tissues were obtained from 10 patients who underwent surgery, and the samples were divided into 3 groups as follows: cancerous tissue group (C group); surrounding noncancerous tissue group (SN group) and distal noncancerous tissue group (DN group). In the current study, the differentially abundant proteins had the

172 same change tendency in all 10 biological replicates and presented a fold change of
 173 approximately ± 1.5 ($\log_2 0.58$) in at least 5 biological replicates with a p value < 0.05 .
 174 Under this standard, there were 22 differentially abundant proteins in C/DN group
 175 (Table S3), and 49 differentially abundant proteins in SN/DN group (Table S4). The
 176 numbers of overlapped differentially abundant proteins among two comparisons was
 177 showed by the Venn diagram in Figure 3F. Among these differentially abundant
 178 proteins, 13 differentially abundant proteins were shared in two groups. As the GO
 179 annotation analysis indicated, these proteins mainly participated in extracellular
 180 matrix organization, extracellular structure organization, and tissue morphogenesis
 181 (Figure S1D), which might be a contributing factors for HCC development. According
 182 to the result, 9 dysregulated proteins were found in C/DN group but not in SN/DN
 183 group. These proteins were the mainly participated in the disorder of primary
 184 metabolism (Figure S1E), which is tightly linked to the development of HCC, and
 185 even invasion, infiltration and metastasis. Meanwhile, 36 dysregulated proteins were
 186 found in SN/DN group but not in C/DN group. Similarly, we found these proteins
 187 were the mainly participated in acute-phase response, acute inflammatory response
 188 and post-transcriptional regulation on gene expression (Figure S1F), suggesting that
 189 the surrounding noncancerous tissues are distinctly different from the distal
 190 noncancerous tissues. Interestingly, these processes were all related to primary
 191 metabolism, suggesting that changes in primary materials might play a crucial role in
 192 the occurrence of HCC.

193 To further study the potential molecular mechanisms of the occurrence and
 194 development of HCC, IPA analysis was applied to investigate the signaling pathways
 195 in which the dysregulated proteins participated. The results showed that the two
 196 comparisons (C/DN group and SN/DN group) indeed had specific signaling pathways,
 197 although they also had common shared signaling pathways. As the IPA investigation
 198 showed, the dysregulated proteins in C/DN group were mostly involved in
 199 ERK/MAPK signaling, while the dysregulated proteins are mainly participated in
 200 PI3K-Akt signaling pathway in SN/DN group. There are 11 proteins participated in

201 ERK/MAPK signaling (2 up-regulated and 9 down-regulated) (Figure 3G). The
202 ERK/MAPK pathway can transduce extracellular signals through intracellular signal
203 transduction cascades to finally control the expression of proteins that regulate
204 tumorigenesis and aggressive behaviours [24,25]. The dysregulation of ERK/MAPK
205 signaling pathway in C/DN group revealed that the secreted factors from HCC
206 cancerous tissue might modulate the tumour micro-environment to exert important
207 roles in tumorigenesis.

208 In SN/DN group, 20 proteins (4 up-regulated and 16 down-regulated) participated
209 in PI3K-Akt signaling (Figure 3H). PI3K-Akt signaling plays a crucial role in the
210 regulation of inflammation and metabolism [26–28]. Hence, alterations of the
211 PI3K-Akt pathway might be closely linked to the occurrence and development of
212 tumours. The dysregulation of PI3K-Akt signaling pathway in the secretory
213 environment of HCC adjacent noncancerous tissues compared with that of distal
214 noncancerous tissues suggested the importance of changing microenvironment in
215 tumorigenesis. Targeting these important effectors in tumour microenvironment might
216 be a promising therapeutic strategy.

217 **CA2 might be a valuable biomarker for HCC diagnosis**

218 According to the IPA analysis, we found that carbonic anhydrase II (CA2) was
219 dysregulated between the two comparison (C/DN group and SN/DN group). CA2, a
220 zinc metal enzyme, carries out the reversible hydration of carbon dioxide, and plays a
221 key role in adjusting the pH of tumour microenvironment [29–34]. It is frequently
222 abnormally expressed in different cancers [35–41]. Here, the CA2 relative intensity of
223 the reporter ions of the 8-plex iTRAQ reagent in the MS/MS spectra were in good
224 agreement with the protein levels. As revealed in Figure S2A, the levels of CA2 were
225 remarkably increased in C/DN group and SN/DN group (114>115>116 and
226 117>118>119). Therefore, CA2 might be a potential interesting biomarker for HCC
227 diagnosis and prognosis prediction. Here, we further verified the clinical significance

228 of CA2 on the diagnosis and prognosis of HCC in two additional serum cohorts via
229 the PRM targeted proteomic method and traditional ELISA assays.

230 The CA2 serum concentration was monitored in 49 HCC patients and 23 healthy
231 volunteers through PRM. There were 9 identified peptides from CA2 in total, 7 of
232 which were overlapped in the five 8-plex iTRAQ experiments, and 1 of which had a
233 mis-cleavage site. Therefore, only 6 peptides from CA2 could be used for the PRM
234 experiment, and the annotated spectra and detailed PSM information for these 6
235 identified CA2 peptides were provided in Table S5 and Figure S3. For the skyline
236 results, the peak contributions of the individual fragment ions from the unique peptide
237 was showed in Figure S2B, and the representative quantification information based on
238 peak areas of the peptides including endogenic and synthetic heavy peptides was
239 displayed in Figure S2C. The CA2 serum concentration in HCC patients was
240 remarkably higher than that in normal populations ($p < 0.01$, 7.01 and 13.72 pg/mL
241 average serum CA2 in healthy volunteers and HCC patients, respectively). The
242 receiving operating character (ROC) curve analysis of CA2 revealed that the area
243 under the curve (AUC) was 0.715 for HCC patients relative to the healthy volunteers
244 (Figure 4A). These results were consistent with the data from the proteomic studies,
245 indicating that CA2 might be a valuable biomarker for HCC diagnosis.

246 **CA2 is a novel prognostic biomarker for HCC**

247 According to our previous study, the serum levels of CA2 were steadily increased
248 from 3 months to 9 months after radical resection in the patients with short-term
249 recurrence compared with the patients without recurrence (Figure S4A). According to
250 these data, we expect to find an appropriate time point after radical resection to
251 predict the recurrence of HCC. The serum samples taken 5 months after surgery were
252 collected blindly. An ELISA assay was used to analyze the association between the
253 CA2 serum concentrations and the prognosis for 159 HCC patients, containing 94
254 relapsed patients and 65 relapse-free patients. As revealed in Figure 4B, serum CA2
255 was remarkably upregulated in HCC patients with recurrence compared with HCC

256 patients without recurrence ($p < 0.01$). According to the ROC curve, the AUC reached
 257 0.708 in the validation cohort, indicating that CA2 might be a HCC prognostic
 258 biomarker. To evaluate the prognostic significance of serum CA2 in HCC, we further
 259 studied the overall survival and recurrence rate of patients with low and high serum
 260 levels of CA2 using the 159 samples. The median concentration (105.3 ng/mL) was
 261 used as the optimal cutoff value to stratify patients into low (≤ 105.3 ng/mL) and
 262 high (> 105.3 ng/mL) CA2 groups. As shown by the Kaplan-Meier analysis, HCC
 263 patients with a higher serum CA2 level had a higher recurrence rate than those with a
 264 low serum CA2 level. Furthermore, the overall survival (OS) of HCC patients who
 265 had a high serum CA2 level was remarkably shorter than those with a low CA2 level
 266 (Figure 4C). Furthermore, we repeated this study using the training cohort of 49 HCC
 267 patients and obtained identical results (Figure S4B). In addition, we found a strong
 268 negative linear correlation between serum CA2 levels and the recurrence time,
 269 indicating that the higher the serum CA2, the shorter the recurrence time (Figure S4C).
 270 These results suggested that the serum CA2 level might be a novel prognostic
 271 biomarker for HCC.

272 **CA2 might be a novel prognostic biomarker for AFP-negative HCC patients**

273 As the gold standard for clinical diagnosis and monitoring in HCC, AFP was
 274 remarkably higher in relapsed HCC patients than that in patients without relapse ($p <$
 275 0.001) (Figure S4D). Although the recurrence rate was substantially lower while the
 276 overall survival was remarkably higher in AFP-negative patients than that in
 277 AFP-positive patients in general (AFP-negative and AFP-positive patients were
 278 divided according to clinical detected AFP serum concentration, 20 ng/mL) (Figure
 279 S4E), some patients with negative serum AFP still experienced rapid recurrence or
 280 metastasis (Figure S4F). Therefore, there is an urgent need to identify novel
 281 biomarkers that can predict the recurrence/metastasis in AFP-negative HCC patients.
 282 Interestingly, among the AFP-negative HCC patients, the serum concentration of CA2
 283 in relapsed patients was remarkably higher than that in relapse-free patients ($p < 0.001$)

(Figure 4D). Furthermore, in the AFP-negative HCC patients, the recurrence rates were also higher in patients with high CA2, suggesting the prognostic values of CA2 in AFP-negative patients (Figure 4E). The overall survival rate did not show significant differences between the high-CA2 and low-CA2 groups (Figure S4G).

Combination of serum CA2 and AFP improves prognostic performance in HCC

According to the above results, we confirmed that CA2 had a prognostic value in AFP-negative patients. We then jointly considered the CA2 and AFP serum levels. As shown in Figure 4F, when CA2 and serum AFP were considered together, the AUC reached to 0.803 for the combination of CA2/AFP (AUC of 0.708 for CA2 and 0.765 for AFP), suggesting a clinical values for the combination of these 2 markers. Next, we further evaluated the prognostic value of CA2/AFP in combination for HCC patients in the validation cohort. As shown in Figure 4G, the cumulative recurrence rate in patients in the low-CA2 and AFP-negative group was remarkably lower than that of patients in the high-CA2 and AFP-positive group. Correspondingly, the overall survival of patients in the low-CA2 and AFP-negative group was substantially higher than that of patients in high-CA2 and AFP-positive group. CA2-positive and AFP-negative or CA2-negative and AFP-positive patients had a median OS and TTR in the validation cohorts. Taken together, these results suggested an improved prognostic value when using the serum CA2/AFP levels in combination for HCC patients.

Serum CA2 was significantly correlated with tumour number and BCLC stage

To further study molecular mechanisms of CA2 on the prognosis of HCC, we further checked the association between serum CA2 and various clinicopathologic features of HCC patients. As revealed in Table S6, the Pearson's chi-square test has indicated that higher serum CA2 level was significantly correlated with more tumour number ($p = 0.026$) and more advanced BCLC stage ($p = 0.042$), but did not associated with other clinicopathologic features. These results suggested that CA2 might be a metastasis/recurrence-related protein in HCC.

312 **Secreted CA2 increases the migration and invasion abilities of HCC cells by** 313 **activating of EMT signaling**

314 To investigate the molecular mechanism of extracellular CA2 in HCC
315 recurrence/metastasis, exogenous recombinant CA2 protein (obtained from Abcam)
316 was used to examine its influences on the migration and invasion ability of HCC cells
317 using a trans-well strategy. As shown in Figure 5A, exogenous recombinant CA2
318 treatment of MHCC97L cells significantly promoted the migration and invasion of
319 cells compared to control cells with the same concentration of BSA treatment ($p <$
320 0.05), suggesting that extracellular CA2 enhances cell migration and invasion in HCC,
321 which could explain the observed clinical data. The epithelial-mesenchymal transition
322 (EMT) is well known to be responsible for tumour metastasis. As shown in Figure 5B,
323 the MHCC97L cells treated with exogenous recombinant CA2 exhibited a
324 spindle-like fibroblastic morphology, while the corresponding control cells were
325 round with a more epithelial morphology. Such morphological changes indicated that
326 the secreted CA2 might be involved in the EMT process. To further investigate the
327 involved molecular mechanisms, the related key markers of EMT were also analyzed.
328 As shown in Figure 5C, the addition of exogenous recombinant CA2 upregulated
329 N-cadherin (a mesenchymal marker) and downregulated E-cadherin (a epithelial
330 maker); meanwhile, the expression levels of vimentin and Zeb1 (EMT-promoting
331 transcription factors) were also significantly upregulated by the addition of exogenous
332 recombinant CA2. Taken together, these results demonstrated that the secreted CA2
333 promoted the EMT transition to further modulate the migration and invasion of HCC
334 cells, in turn affecting HCC metastasis.

335 **Intracellular CA2 might perform opposite functions from its extracellular form**

336 We also studied the intracellular levels of CA2 in HCC patients. In this experiment,
337 28 paired HCC tissues and their corresponding noncancerous tissues were detected by
338 western blot. As revealed in Figure 6A and B, the intracellular CA2 expression was
339 remarkably decreased in HCC tissues when compared with their paired noncancerous

tissues, and the representative images of CA2-immunostained HCC tissues were shown in Figure S5. Meanwhile, the downregulation of intracellular CA2 in tumour tissues was further confirmed by immunostaining using in-house TMAs containing 75 HCC tissues and their paired noncancerous tissues (Figure 6C). The results were exactly opposite from those of the extracellular CA2 in the serum of HCC patients, suggesting that intracellular and extracellular CA2 might carry out opposite functions [42]. However, the involved molecular mechanisms should be further explored.

Conclusion

Herein, we applied the iTRAQ-based quantitative proteomic approach to investigate the secretome of the primary culture of HCC tissues, and identified secreted CA2 as a diagnostic and prognostic biomarker for HCC. In particular, CA2 showed good predicative performance in AFP-negative HCC patients, and the combination of CA2 and AFP improved the sensitivity and specificity of HCC prognosis. Regarding the mechanism, extracellular CA2 might regulate the HCC cell migration and invasion by targeting the EMT signaling pathway to affect HCC patients' prognosis. This secretome investigation enabled us to identify a novel HCC diagnostic and prognostic biomarker. The information from this study constitutes a valuable resources for further HCC investigation and for identifying potential serological biomarkers.

Materials and methods

Patients and follow-up

In total, 293 HCC patients and 23 healthy volunteers were enrolled in the current study. All of the HCC patients underwent surgical procedures at the Mengchao Hepatobiliary Hospital of Fujian Medical University (Fuzhou, China). Computer tomography (CT) scanning, angiography and ultrasonography (US) were used to monitor the absence of intrahepatic recurrence and metastasis in the residual liver. All patients met the enrolment eligibility criteria as follows: (1) The patients were diagnosed with HCC through pathological examination after operation; (2) Serum

367 hepatitis B surface antigen (HBs Ag) was positive and hepatitis B surface antibody
 368 (HBs Ab) and hepatitis C virus (HCV) were negative before surgery; (3) Standard
 369 radical resection was performed: no distal metastasis was found before or during the
 370 operation; intraoperative ultrasound examination revealed no other liver lesions; no
 371 obvious tumour thrombus was found in the hepatic portal vein or primary venous
 372 branch; postoperative pathological examination showed no cancer cells at the cutting
 373 edge; no recurrence/metastasis was found in the ultrasound and CT examination 2
 374 months after the operation; (4) Serum AFP was increased before operation and
 375 decreased to normal level two months after surgical operation; (5) The patient had not
 376 received any other intervention or treatment before surgery.

377 For the proteomic analysis, HCC tissues, surrounding noncancerous tissues and
 378 distal noncancerous tissues were obtained from 10 patients who underwent surgical
 379 operation. Serum from 49 HCC patients before hepatectomy and 23 healthy
 380 volunteers in a health screening program at the Mengchao Hepatobiliary Hospital
 381 were collected for PRM analysis. Another validation serum cohort from 159 HCC
 382 patients after hepatectomy with long-term follow-up that ended with death was used
 383 for the ELISA assay. The collection of serum samples was strictly controlled
 384 following the in-house standard operating procedure, which was established on the
 385 basis of previous studies [43]. In addition, a total of 75 formalin-fixed and
 386 paraffin-embedded HCC tissues and their adjacent noncancerous tissues from patients
 387 who underwent surgical resection were collected to fabricate the tissue microarrays
 388 (TMAs) for immunohistochemistry investigation.

389 This project was approved by the Institution Review Board of Mengchao
 390 Hepatobiliary Hospital of Fujian Medical University. Informed consent was obtained
 391 from each participant before the operation. The use of clinical specimens was
 392 completely in compliance with the “Declaration of Helsinki”.

393 **Tissue culture and quality control *in vitro***

394 Following surgical operation, the primary tissue specimens were immediately

395 transferred into PBS on ice and sent to the laboratory within 30 min. The tissues were
396 then rinsed and cut into 2-3 mm³ pieces, and extensively washed several times at room
397 temperature with PBS to eliminate major blood and serum contaminants.
398 Subsequently, the samples were transferred to 10 cm cell culture dishes and incubated
399 in serum-free DMEM medium supplemented with 1% penicillin (cat No. 15070-063,
400 Gibco, USA) and streptomycin (cat No. 15070-063, Gibco, USA) at 37 °C. For
401 protein extraction, the supernatants were centrifuged at 100 g for 2 min at 4 °C to
402 further remove all remaining cells and debris; the samples were then concentrated
403 with a 3 K cutoff centrifugal filter device (cat No. UFC500396, Millipore, USA) and
404 stored at -80 °C until further analysis.

405 Culture medium was collected every 4 hours in following 72 h period (a total of
406 18 time points). The proteins extracted from every time point were subjected to 10%
407 SDS-PAGE to analyze the molecular mass distribution and protein blotting was
408 performed to determine any contamination by intracellular proteins. At the same time,
409 the cultured tissues at every time point were evaluated by histological observations.
410 The cultured tissues were fixed in 10 % neutral formaldehyde, and paraffin sections
411 were made in the conventionally way. One slide was stained with HE (cat No. D006,
412 Nanjing Jiancheng Bioengineering Institute, China) and observed under an optical
413 microscope (AE2000, Motic, USA); another slide was biotin-labeled and stained by
414 TUNEL to detect cell apoptosis during the culture period. Finally, the slides were
415 observed under a confocal microscope (LSM 780, Carl Zeiss, Germany).

416 **Bottom-up proteomics and data analysis**

417 The proteomic studies and data analyses were modified from our previously reported
418 protocols [44]. Briefly, the thiol groups of the above extracted protein samples from
419 the culture supernatant of 3 groups were alkylated with 50 mM iodoacetamide (cat No.
420 I6125, Sigma, USA) (30 min in dark at room temperature) after reduction with 8 mM
421 DTT (cat No. D0632, Sigma-Aldrich, USA) (55 °C, 1 h). Then, the proteins were then
422 precipitated by ice-cold acetone (5 times volume), and re-dissolved in 100 mM

423 tetraethyl-ammonium bromide (TEAB) (cat No. 90360, Sigma, USA). Subsequently,
 424 100 µg of each proteins sample was digested by trypsin (cat No. V511, Promega,
 425 USA) using filter aided sample preparation (FASP), and the peptides were labeled
 426 with the 8-plex iTRAQ reagent (cat No. 4381663, AB SCIEX, USA) as follows: C
 427 group, SN group and DN group were labeled with 114, 115 and 116, respectively; and
 428 one biological repetition of the above 3 groups was labeled with 117, 118 and 119,
 429 respectively. A, B, C, D and E were defined as the 5 independent iTRAQ 8-plex
 430 labeling repetitions. In addition, the DN group samples mixed in equal amounts were
 431 labelled with 113 and included in every 8-plex labelling reaction as an internal
 432 standard to balance each 8-plex labeling. The labeled peptides were mixed in equal
 433 amounts in every 8-plex labeling and were desalted by Sep-Pak Vac C18 cartridges
 434 (cat No. WAT023590, Waters Corporation, USA). The samples were then dried by the
 435 vacuum centrifuge (cat No. 7310038, LABCONCO, USA) for further use.

436 The peptide mixture was separated using an offline LC system (Acquity UPLC,
 437 Waters Corporation, USA) via high-pH separation. High-pH (pH = 10) separation was
 438 performed in a reverse-phase column (C18, 1.7 µm, 2.1×50 mm) (cat No. 186002350,
 439 Waters Corporation, USA) using a 20 min linear gradient from 5% B to 35% B (A: 20
 440 mM ammonium formate (cat No. 70221, Sigma, USA) in water, B: 20 mM
 441 ammonium formate added in 90% ACN (cat No. A998, Thermo Fisher, USA),
 442 ammonium hydroxide (cat No. 17093, Sigma, USA) was used to adjust the pH).
 443 Finally, a total of 30 fractions were collected and 2 equal-interval fractions were
 444 combined to reduce the MS running time: for example, 1 and 16, 2 and 17, etc. [45].
 445 In total, 15 fractions were dried and subsequently separated on a Nano-LC system
 446 (Nano-Aquity UPLC, Waters Corporation, USA) with a 75 min linear gradient from 2%
 447 D to 40% D (C: 0.1% formic acid in water, D: 0.1% formic acid in ACN), which was
 448 performed on an analytical column (C18, 75 µm*15 cm, 3 µm) (cat No. 164534,
 449 Thermo Fisher Scientific, Germany). Next, the peptides were detected by mass
 450 spectrometry (Q-Exactive, Thermo Fisher Scientific, Germany) with 2.1 kV
 451 electrospray voltage at the mass spectrometer inlet. In addition, 70 K mass resolution

452 was applied in the full-scan of MS spectra processing (m/z 350-1200), and 17.5 K
453 resolution was used in the following 15 sequential MS/MS scans of high energy
454 collisional dissociation (HCD). In all studies, 1 microscan was recorded by using a
455 dynamic exclusion of 30 s.

456 The data processing strategy was also carried out according to our previous
457 publications, with some modification [44]. The LC-MS/MS data acquisition was
458 processed using Proteome Discoverer (version 1.4, Thermo Fisher Scientific,
459 Germany) and searched using the Sequest HT (version 2.5.1, Matrix Science, United
460 Kingdom) search algorithms against the human database (Uniprot, 20,264 entries,
461 released at April 10, 2014). Proteome Discoverer were searched with trypsin for
462 protease digestion and maximally we only allowed 2 missed cleavages. 10 parts per
463 million (ppm) parent ion tolerance and 0.02 Da fragment ion mass tolerance were set
464 according to MS precision. Fixed modifications included the iTRAQ modification of
465 lysine residues and the peptide N-terminus, as well as carbamidomethylation of
466 cysteine, while the variable modifications included iTRAQ labeling of tyrosine and
467 oxidation of methionine. The peptide and protein identification false discovery rate
468 (FDR) was calculated using the percolator algorithm against a decoy database and
469 was set at 1% FDR. For protein quantitation, we only considered peptides that were
470 unique to a certain given protein. The fold change between different samples was
471 calculated by the ratio between iTRAQ reporter ion intensity and MS/MS spectra (m/z
472 113-119). The ratios were derived from criteria as follows: 20 ppm fragment ion
473 tolerance was applied for the most confident centroid peak. In Sequest HT, the
474 quantitative protein ratios were normalized and weighted by the median ratio. In
475 addition, Scaffold (version 4.3.2, Proteome Software Inc., USA) was applied to verify
476 the MS/MS-based peptides and identified proteins. The Scaffold was used to analyze
477 and evaluate the quantitative results of iTRAQ, as well as to combine multiple
478 quantitative experiments according to the designed internal standard in each 8-plex
479 iTRAQ experiment for subsequent analysis. In detail, the raw data was searched by
480 Proteome Discoverer using Sequest HT search algorithms, and the generated file was

481 imported into the Scaffold software. The integrated information of the five 8-plex
482 iTRAQ experiments could then be obtained directly by grouping. In the same way, a
483 peptide FDR < 1% and a protein probability > 99.0% were accepted.

484 Two complementary methods were combined to further characterize the secretory
485 proteins. First, MetazSecKB, a secretome proteome knowledgebase of metazoan, was
486 performed to screen secretory proteins, which is the most direct source by which to
487 characterize the secretory proteins. Second, SecretomeP (version 2.0, DTU Health
488 Tech, Denmark), a sequence-based prediction strategy for mammalian secretory
489 proteins, was used to classify the secretory proteins. The classical secreted proteins
490 could be correctly predicted with an N-terminal signal peptide; non-classical secretory
491 proteins without an N-terminal signal peptide could be correctly predicted as
492 secretory according to an NN-score > 0.6. Those proteins without an N-terminal
493 signal peptide and that also had a low NN-score were not identified as secretory
494 proteins.

495 The GO annotation and signaling pathway investigation of differentially abundant
496 proteins were performed using the free online tool DAVID
497 (<http://david.abcc.ncifcrf.gov/>). The enriched signaling pathways of the dysregulated
498 proteins were analyzed using Ingenuity Pathways Analysis (IPA) (version 7.5,
499 Ingenuity Systems, Inc., USA).

500 **Targeted proteomics and data analysis**

501 Before applying the targeted proteomic method of parallel reaction monitoring (PRM)
502 analysis, the serum samples were immunoaffinity depleted of the 14 most abundant
503 proteins using IgY14 LC20 (cat No. 5188-6557, Agilent, USA). The depletion was
504 performed on an Agilent 1260 HPLC (1260 Infinity, Agilent, USA) system following
505 the manufacturer's protocol. The depleted serum was concentrated and the buffer was
506 exchanged to 100 mM TEAB using Amicon 3 K concentrators (cat No. UFC500396,
507 Millipore, USA). The procedures for serum protein denaturation, reduction, alkylation
508 and digestion were described above.

509 For PRM analysis, the unique peptides for CA2 were synthesized by Fmoc
510 solid-phase synthesis with isotope-labeled on the carboxyl side of the amino acid
511 lysine ($^{13}\text{C}_6$, $^{15}\text{N}_2$) and purified by HPLC with purity > 99% (Anhui Guoping
512 Pharmaceutical Co., Ltd, China). The heavy-labeled peptides were mixed and spiked
513 into the tryptic digests of serum proteins at a concentration of 30 fg/ μL . Detailed
514 information for the unique CA2 peptides were displayed in Table S5. The unique
515 peptide for CA2 should preferably have a narrow, symmetrical chromatographic peak,
516 be 8-25 amino acids in length, ionize efficiently, provide a stable and intense signal
517 without any modification, and not elute at the beginning or end time.

518 The PRM analyses were also performed on a mass spectrometer (Q Exactive Plus,
519 Thermo Fisher Scientific, Germany). LC separation was executed with a trap column
520 (C18, 75 μm *2 cm, 3 μm) (cat No. 164946, Thermo Fisher Scientific, Germany) and
521 an analytical C18 column (C18, 75 μm *15 cm, 3 μm) (cat No. 164534, Thermo Fisher
522 Scientific, Germany) on a nano-LC system (EASY-nLC 1000, Thermo Fisher
523 Scientific, Germany). 2 μL of the tryptic digests and depleted serum samples were
524 injected, and a gradient of 2% D up to 35% D over 30 min was applied. In all
525 experiments, a PRM scan was performed at m/z 200, with a resolution of 70 K, an
526 AGC target of 1×10^6 , a maximum injection time of 200 ms and an isolation window
527 of ± 2 . In addition, a normalized collision energy of 27% was used for ion dissociation,
528 and a fixed first mass at 120 m/z was set. The inclusion list including the m/z and
529 corresponding retention times of precursor peptides of interest were displayed in
530 Table S5.

531 The PRM raw data was analyzed with Skyline software (version 4.2, University of
532 Washington, USA), and a 5-minute window was used in the Skyline software. The
533 sum of peak areas of the 5 most intense product ions was considered for protein
534 quantification. Some ions needed to be excluded: for example, those ions that did not
535 match the retention time of other monitored ions, or that showed interference signals,
536 or that gave intense signals at other retention time.

537 **ELISA assay**

538 The serum levels of CA2 in HCC patients were analyzed using an ELISA assay kit
 539 (cat No. LS-F29508, LifeSpan BioScience Inc., USA) following the manufacturer's
 540 instructions. Briefly, the standard proteins and patient serum samples were diluted
 541 with the sample dilution buffer and then 50µL of the diluted standards proteins and
 542 samples were added to 96-well plates. Next, 100 µL of HRP-conjugate solution was
 543 carefully added to each well and incubated for 1 hour at 37 °C. Then, the plate was
 544 then carefully washed 4 times with PBS, and 100 µL of chromogen solution (1:1
 545 solution A and solution B) was added to each well and further incubated for 15 min at
 546 37 °C. Finally, 50µL of stop solution was added to stop the reaction, and the optical
 547 density at 450 nm was measured by a spectrophotometer (M5e, Molecular Devices,
 548 USA).

549 **Western blot**

550 Western blotting was performed according to a previous publication [44]. Briefly,
 551 proteins were separated on a 12% SDS-PAGE, and the gel was transferred to a
 552 nitrocellulose membrane (cat No. HATF00010, Millipore, USA). The membrane was
 553 then blocked in 5% BSA (cat No. A7906, Sigma-Aldrich, USA) at room temperature
 554 for 2 hours, then further incubated in primary antibody against CA2 (1/1000 dilution,
 555 cat No. ab226987, Abcam, USA) at 4°C overnight. The corresponding secondary
 556 antibody was incubated with the membrane at RT for 1 hour after which the blots
 557 were carefully washed 4 times with TBST buffer and revealed using enhanced
 558 chemiluminescence reagents (cat No. 34080, Thermo Scientific, USA) and visualized
 559 by autoradiography.

560 **Immunohistochemistry (IHC)**

561 Immunohistochemistry (IHC) was carried out on HCC TMAs according to a previous
 562 publication [46]. Briefly, after pre-treating at pH6 and blocking with peroxidase, the
 563 sections were further incubated with primary antibody against CA2 (1/50 dilution, cat

No. ab226987, abcam, USA) for 0.5 hour. The sections were then treated with envision FLEX/HRP reagent for 20 min and then washed and stained by the envision FLEX-DAB chromogen (cat No. DM827, Dako, Denmark) and by Mayer's Hematoxylin (Lille's Modification) Histological Staining Reagents (cat No. ab220365, Abcam, USA) for 3 minutes. They were then washed in distilled water for 5 minutes. All pathological sections underwent double-blind scoring as follows: negative (0), weak (1), strong (2) or very strong (3) by two different pathologists double-blindly. A score of 0-1 indicates low expression, and a score of 2-3 indicates high expression.

Cell migration and invasion assays

The cell migration and invasion assays were carried out according to our previously published protocols [44]. The cell migration ability was investigated by using transwell units with 8 μ m pores (cat No. 3428, Corning Costar, USA), and cell invasion was investigated by using 24-well transwell inserts that had been precoated with matrigel with 8 μ m pores (cat No. 354480, BD Biosciences, USA). A total of 1×10^5 cells from the indicated treatments were cultured in serum-free medium in the upper chamber. To induce cell invasion and migration, DMEM supplemented with 10% FBS was placed in the lower chamber. The cells that adhered to the lower surface were then fixed with paraformaldehyde after 18 h of incubation and then stained by crystal violet (0.1%). The cells that adhered to the bottom surface were carefully counted within 5 different views under a microscope using $20 \times$ magnification.

Statistical analysis

A threshold for the iTRAQ ratio was set to screen secretory proteins whose abundance was substantially changed in the tumour tissue group compared with its corresponding noncancerous tissue group. The proteins were defined as differentially dysregulated if the iTRAQ ratio was higher than 1.5 or lower than 0.67 in at least 5 patients, and they also must had the same direction of alteration in all 10 biological replicates. The iTRAQ ratio was based on a comparison of the reporter ion intensities in the tumour tissue group compared to the corresponding noncancerous tissue group.

SPSS 19.0 was used for statistical analysis. Two-tailed paired Student's t-test was used to compare quantitative data between two groups. Fisher's exact test was applied to analyze the relationships between CA2 and clinical-pathological features. The Kaplan-Meier method was applied to calculate survival curves, while the differences were determined by using a log-rank test. In all analyses, $p < 0.05$ was recognized as statistically significant.

Authors' contributions

XHX, AMH and XLL designed the project, analyzed the data and revised the manuscript; XHX executed the study, performed the tissue culture, bottom-up proteomics and targeted proteomics assays, statistical analysis, and drafted the manuscript; HY and XHT carried out the IHC and transwell assays; HZL carried out the ELISA assay and participated in the statistical analysis; BXZ and YCW participated in the study; JHO performed the western blotting; MJL was involved in the clinical sample collection. All authors have carefully read and approved the submitted manuscript.

Competing interests

The authors have declared no competing interests.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 81702910 and Grant No. 81672376); the Educational Commission of Fujian Province (Grant No. 2018B013); and the Natural Science Foundation of Fujian Province (Grant No. 2017J01159 and Grant No. 2016J01417).

Reference

[1] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in

617 GLOBOCAN 2012. *Int J Cancer* 2015;136:E359–86.

618 [2] Wang FS, Fan JG, Zhang Z, Gao B, Wang HY. The global burden of liver disease:
619 the major impact of China. *Hepatology* 2014;60:2099–108.

620 [3] Zhou XD, Tang ZY, Yang BH, Lin ZY, Ma ZC, Ye SL, et al. Experience of 1000
621 patients who underwent hepatectomy for small hepatocellular carcinoma. *Cancer*
622 2001;91:1479–86.

623 [4] Ng KM, Yan TD, Black D, Chu FC, Morris DL. Prognostic determinants for
624 survival after resection/ablation of a large hepatocellular carcinoma. *HPB (Oxford)*
625 2009;11:311–20.

626 [5] Lee YY, McKinney KQ, Ghosh S, Iannitti DA, Martinie JB, Caballes FR, et al.
627 Subcellular tissue proteomics of hepatocellular carcinoma for molecular signature
628 discovery. *J Proteome Res* 2011;10:5070–83.

629 [6] Iizuka N, Oka M, Yamada-Okabe H, Nishida M, Maeda Y, Mori N, et al.
630 Oligonucleotide microarray for prediction of early intrahepatic recurrence of
631 hepatocellular carcinoma after curative resection. *Lancet* 2003;361:923–9.

632 [7] Sterling RK, Wright EC, Morgan TR, Seeff LB, Hoefs JC, Di Bisceglie AM, et al.
633 Frequency of elevated hepatocellular carcinoma (HCC) biomarkers in patients with
634 advanced hepatitis C. *Am J Gastroenterol* 2012;107:64–74.

635 [8] Liu X, Cheng Y, Sheng W, Lu H, Xu Y, Long Z, et al. Clinicopathologic features
636 and prognostic factors in alpha-fetoprotein-producing gastric cancers: analysis of 104
637 cases. *J Surg Oncol* 2010;102:249–55.

638 [9] El-Bahrawy M. Alpha-fetoprotein-producing non-germ cell tumours of the female
639 genital tract. *Eur J Cancer* 2010;46:1317–22.

640 [10] Qiao B, Wang J, Xie J, Niu Y, Ye S, Wan Q, et al. Detection and identification of
641 peroxiredoxin 3 as a biomarker in hepatocellular carcinoma by a proteomic approach.
642 *Int J Mol Med* 2012;29:832–40.

- 643 [11] Song IS, Kim HK, Jeong SH, Lee SR, Kim N, Rhee BD, et al. Mitochondrial
644 peroxiredoxin III is a potential target for cancer therapy. *Int J Mol Sci*
645 2011;12:7163–85.
- 646 [12] Xing X, Huang Y, Wang S, Chi M, Zeng Y, Chen L, et al. Comparative analysis
647 of primary hepatocellular carcinoma with single and multiple lesions by
648 iTRAQ-based quantitative proteomics. *J Proteomics* 2015;128:262–71.
- 649 [13] Feng JT, Liu YK, Song HY, Dai Z, Qin LX, Almofti MR, et al. Heat-shock
650 protein 27: a potential biomarker for hepatocellular carcinoma identified by serum
651 proteome analysis. *Proteomics* 2005;5:4581–8.
- 652 [14] Veenstra TD, Conrads TP, Hood BL, Avellino AM, Ellenbogen RG, Morrison RS.
653 Biomarkers: mining the biofluid proteome. *Mol Cell Proteomics* 2005;4:409–18.
- 654 [15] Anderson NL, Anderson NG. The human plasma proteome: history, character,
655 and diagnostic prospects. *Mol Cell Proteomics* 2002;1:845–67.
- 656 [16] Hanash SM, Pitteri SJ, Faca VM. Mining the plasma proteome for cancer
657 biomarkers. *Nature* 2008;452:571–9.
- 658 [17] Paltridge JL, Belle L, Khew-Goodall Y. The secretome in cancer progression.
659 *Biochim Biophys Acta* 2013;1834:2233–41.
- 660 [18] Slany A, Haudek-Prinz V, Zwickl H, Stattner S, Grasl-Kraupp B, Gerner C.
661 Myofibroblasts are important contributors to human hepatocellular carcinoma:
662 evidence for tumor promotion by proteome profiling. *Electrophoresis*
663 2013;34:3315–25.
- 664 [19] Yu Y, Pan X, Ding Y, Liu X, Tang H, Shen C, et al. An iTRAQ based quantitative
665 proteomic strategy to explore novel secreted proteins in metastatic hepatocellular
666 carcinoma cell lines. *Analyst* 2013;138:4505–11.
- 667 [20] Cao J, Hu Y, Shen C, Yao J, Wei L, Yang F, et al. Nanozeolite-driven approach
668 for enrichment of secretory proteins in human hepatocellular carcinoma cells.

669 Proteomics 2009;9:4881–8.

670 [21] Xing X, Liang D, Huang Y, Zeng Y, Han X, Liu X, et al. The application of
671 proteomics in different aspects of hepatocellular carcinoma research. J Proteomics
672 2016;145:70–80.

673 [22] Yang L, Rong W, Xiao T, Zhang Y, Xu B, Liu Y, et al. Secretary/releasing
674 proteome-based identification of plasma biomarkers in HBV-associated hepatocellular
675 carcinoma. Sci China Life Sci 2013;56:638–46.

676 [23] Ma J, Chen T, Wu S, Yang C, Bai M, Shu K, et al. iProX: an integrated proteome
677 resource. Nucleic Acids Res 2019;47:D1211–7.

678 [24] Chang L, Karin M. Mammalian MAP kinase signalling cascades. Nature
679 2001;410:37–40.

680 [25] Gaesser JM, Fyffe-Maricich SL. Intracellular signaling pathway regulation of
681 myelination and remyelination in the CNS. Exp Neurol 2016;283:501–11.

682 [26] Gherardi E, Birchmeier W, Birchmeier C, Vande Woude G. Targeting MET in
683 cancer: rationale and progress. Nat Rev Cancer 2012;12:89–103.

684 [27] Veillette A, Grenier K, Brasseur K, Frechette-Frigon G, Leblanc V, Parent S, et al.
685 Regulation of the PI3-K/Akt survival pathway in the rat endometrium. Biol Reprod
686 2013; 79:1–11.

687 [28] Solit DB, Basso AD, Olshen AB, Scher HI, Rosen N. Inhibition of heat shock
688 protein 90 function down-regulates Akt kinase and sensitizes tumors to Taxol. Cancer
689 Res 2003;63:2139–44.

690 [29] Noor SI, Jamali S, Ames S, Langer S, Deitmer JW, Becker HM. A surface proton
691 antenna in carbonic anhydrase II supports lactate transport in cancer cells. Elife
692 2018;7: e35176.

693 [30] Chiang WL, Chu SC, Yang SS, Li MC, Lai JC, Yang SF, et al. The aberrant

- 694 expression of cytosolic carbonic anhydrase and its clinical significance in human
695 non-small cell lung cancer. *Cancer Lett* 2002;188:199–205.
- 696 [31] Mallory JC, Crudden G, Oliva A, Saunders C, Stromberg A, Craven RJ. A novel
697 group of genes regulates susceptibility to antineoplastic drugs in highly tumorigenic
698 breast cancer cells. *Mol Pharmacol* 2005;68:1747–56.
- 699 [32] Haapasalo J, Nordfors K, Jarvela S, Bragge H, Rantala I, Parkkila AK, et al.
700 Carbonic anhydrase II in the endothelium of glial tumors: a potential target for therapy.
701 *Neuro Oncol* 2007;9:308–13.
- 702 [33] Zhou Y, Mokhtari RB, Pan J, Cutz E, Yeger H. Carbonic anhydrase II mediates
703 malignant behavior of pulmonary neuroendocrine tumors. *Am J Respir Cell Mol Biol*
704 2015;52:183–92.
- 705 [34] Parks SK, Pouyssegur J. Targeting pH regulating proteins for cancer
706 therapy-Progress and limitations. *Semin Cancer Biol* 2017;43:66–73.
- 707 [35] Viikila P, Kivela AJ, Mustonen H, Koskensalo S, Waheed A, Sly WS, et al.
708 Carbonic anhydrase enzymes II, VII, IX and XII in colorectal carcinomas. *World J*
709 *Gastroenterol* 2016;22:8168–77.
- 710 [36] Zhou R, Huang W, Yao Y, Wang Y, Li Z, Shao B, et al. CA II, a potential
711 biomarker by proteomic analysis, exerts significant inhibitory effect on the growth of
712 colorectal cancer cells. *Int J Oncol* 2013;43:611–21.
- 713 [37] Liu LC, Xu WT, Wu X, Zhao P, Lv YL, Chen L. Overexpression of carbonic
714 anhydrase II and Ki-67 proteins in prognosis of gastrointestinal stromal tumors. *World*
715 *J Gastroenterol* 2013;19:2473–80.
- 716 [38] Takahashi M, Yang XJ, Sugimura J, Backdahl J, Tretiakova M, Qian CN, et al.
717 Molecular subclassification of kidney tumors and the discovery of new diagnostic
718 markers. *Oncogene* 2003;22:6810–8.
- 719 [39] Hynninen P, Parkkila S, Huhtala H, Pastorekova S, Pastorek J, Waheed A, et al.

720 Carbonic anhydrase isozymes II, IX, and XII in uterine tumors. *APMIS*
721 2012;120:117–29.

722 [40] Liu CM, Lin YM, Yeh KT, Chen MK, Chang JH, Chen CJ, et al. Expression of
723 carbonic anhydrases I/II and the correlation to clinical aspects of oral squamous cell
724 carcinoma analyzed using tissue microarray. *J Oral Pathol Med* 2012;41:533–9.

725 [41] Parkkila S, Rajaniemi H, Parkkila AK, Kivela J, Waheed A, Pastorekova S, et al.
726 Carbonic anhydrase inhibitor suppresses invasion of renal cancer cells in vitro. *Proc*
727 *Natl Acad Sci U S A* 2000;97:2220–4.

728 [42] Zhang C, Wang H, Chen Z, Zhuang L, Xu L, Ning Z, et al. Carbonic anhydrase 2
729 inhibits epithelial-mesenchymal transition and metastasis in hepatocellular carcinoma.
730 *Carcinogenesis* 2018;39:562–70.

731 [43] Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N,
732 et al. Procedures for large-scale metabolic profiling of serum and plasma using gas
733 chromatography and liquid chromatography coupled to mass spectrometry. *Nat Protoc*
734 2011;6:1060–83.

735 [44] Liu H, Wang Y, Xing X, Sun Y, Wei D, Chen G, et al. Comparative proteomics of
736 side population cells derived from human hepatocellular carcinoma cell lines with
737 varying metastatic potentials. *Oncol Lett* 2018;16:335–45.

738 [45] Song C, Ye M, Han G, Jiang X, Wang F, Yu Z, et al.
739 Reversed-phase-reversed-phase liquid chromatography approach with high
740 orthogonality for multidimensional separation of phosphopeptides. *Anal Chem*
741 2010;82:53–6.

742 [46] Huang X, Zeng Y, Xing X, Zeng J, Gao Y, Cai Z, et al. Quantitative proteomics
743 analysis of early recurrence/metastasis of huge hepatocellular carcinoma following
744 radical resection. *Proteome Sci* 2014; 22:1–14.

745

746 **Figure legends**

747 **Figure 1 Strict quality control of the tissue secretome**

748 **A.** The quality control workflow for the cultured tissues and the proteins secreted
 749 proteins in the supernatant. **B.** HE staining of cultured tissues. The three columns
 750 were the staining of cancerous HCC tissues (left column, 200 ×), surrounding
 751 noncancerous tissues (middle column, 200 ×) and distal noncancerous tissues (middle
 752 column, 200×), cultured for 0, 1 and 2 days, respectively. **C.** TUNEL staining of
 753 cultured tissues. We examined the densities of DAPI (blue) and FITC (green) for
 754 every tissues. **D.** The molecular weight distribution of secretory proteins through
 755 SDS-PAGE. **E.** Western blot of secretory proteins.

756 **Figure 2 Experimental workflow**

757 The HCC tissues, surrounding noncancerous tissues and distal noncancerous tissues
 758 were cultured *in vitro*, and the conditioned medium was collected to extract secretory
 759 proteins. Secretory proteins were digested with trypsin, directly labeled using
 760 iTRAQ-8plex and analyzed through 2D LC-MS/MS. The target proteins screened by
 761 bioinformatics were then verified *in vitro* and *in vivo* to find potential biomarkers of
 762 HCC and to investigate the molecular mechanisms of HCC recurrence.

763 **Figure 3 Objective and reasonable features of the HCC tissue secretome**

764 The distribution of **(A)** molecular weight and **(B)** isoelectric point of the identified
 765 secretory proteins. GO analysis of the **(C)** cell component, **(D)** biological processes
 766 and **(E)** molecular functions of the secretory proteins. **(F)** Venn diagrams of secretory
 767 proteins for the two comparisons (C/DN group and SN/DN group). The signaling
 768 pathway networks involved in the **(G)** C/DN group and **(H)** in SN/DN group.

769 **Figure 4 Clinical values of serum CA2 in HCC diagnosis and prognosis**

770 **A.** Distribution and ROC curve of serum CA2 levels in HCC patients and healthy
 771 volunteers from the training cohort through PRM. **B.** Distribution and ROC curve of

772 serum CA2 levels in HCC patients, with and without recurrence, from the validation
773 cohort assessed by ELISA. **C.** TTR and OS were compared between the high- and
774 low-CA2 groups in the validation cohort by Kaplan-Meier analysis. **D.** Distribution of
775 serum CA2 levels in AFP-negative HCC patients. **E.** TTR was compared between the
776 high- and low-CA2 groups in AFP-negative HCC patients by Kaplan-Meier analysis.
777 **F.** AUC of the CA2/AFP combination in HCC patients. **G.** Prognostic values of the
778 CA2/AFP combination in HCC patients.

779 **Figure 5 Extracellular CA2 promotes cell migration and invasion through** 780 **activating EMT**

781 **A.** Representative images and quantification results of cell migration and invasion in
782 MHCC97L cells after adding exogenous recombinant CA2. **B.** Cellular morphology
783 of HCC cells with the addition of exogenous recombinant CA2; Magnification, $\times 200$;
784 Scale bar, 50 μm . **C.** Western blot analysis showed that adding exogenous
785 recombinant CA2 downregulated the expression of E-cadherin, while upregulating the
786 expression of Zeb1 and N-cadherin.

787 **Figure 6 Expression of intracellular CA2 was opposite of the extracellular form**

788 **A.** Intracellular CA2 expression in HCC tissues and their paired surrounding
789 non-tumoral tissues. **B.** Intracellular CA2 expression was decreased in HCC tissues
790 compared with their paired surrounding non-tumoral tissues. **C.** IHC analysis of the
791 intracellular CA2 expression in HCC tissues and their paired surrounding non-tumoral
792 tissues on in-house TMAs.

793

794 **Supplementary material**

795 **Figure S1 Features of the dysregulated secretory proteins**

796 **A.** Comparison of the current dataset in this study with the only published dataset for
797 the HCC tissue secretome. **B.** Comparison of the enrichment percentage of secretory

798 proteins between our study and that in the human protein database. **C.** Hydrophobicity
 799 distribution of the identified secretory proteins. The GO analysis of overlapped
 800 different abundance secretory proteins between the comparison of C/DN group and
 801 SN/DN group involved biological processes, **(D)** different abundance secretory
 802 proteins in the C/DN group involved biological processes, and **(E)** different
 803 abundance secretory proteins in the SN/DN group involved biological processes **(F)**.

804 **Figure S2 Representative MS spectra of CA2 in shotgun and targeted**
 805 **proteomics**

806 **A.** Reporter ion relative intensity of the 8-plex iTRAQ reagents related to CA2 in the
 807 MS/MS spectra. **B.** Peak contributions of the individual fragment ions from the
 808 unique peptide of CA2. **C.** Representative PRM results of CA2 in HCC patients and
 809 healthy volunteers from the training cohort through PRM.

810 **Figure S3 Representative annotated spectra for the 6 identified peptides of**
 811 **CA2.**

812 **Figure S4 Validation of the clinical values of serum CA2 concentration**

813 **A.** The variation tendency of serum CA2 during the follow-up time after radical
 814 resection but before recurrence (right), and the sampling time in every patient (left). **B.**
 815 TTR and OS were compared between the high- and low-CA2 groups by
 816 Kaplan-Meier analysis in the training cohort. **C.** The serum level of CA2 was
 817 negatively correlated with the time to recurrence in HCC. **D.** Distribution of serum
 818 AFP levels in the validation cohort by ELISA, *** $p < 0.001$. **E.** TTR and OS were
 819 compared between the AFP-positive and AFP-negative groups in the validation cohort
 820 by Kaplan-Meier analysis. **F.** Distribution of serum AFP levels in AFP-negative HCC
 821 patients. **G.** OS was compared between the high and low CA2 groups in AFP-negative
 822 HCC patients by Kaplan-Meier analysis.

823 **Figure S5 Representative images of CA2 immunostaining in tumour tissues**

824 **Table S1 Basic information and features of HCC patients enrolled in this study.**

825 **Table S2 The complete list of quantified secretory proteins from this study.**

826 **Table S3 Differential abundant secretory proteins in C group compared to DN**
 827 **group.**

828 **Table S4 Differential abundant secretory proteins in SN group compared to DN**
 829 **group.**

830 **Table S5 Detailed PSM information for the 6 identified peptides of CA2.**

831 **Table S6 Univariate logistic regression analysis between serum CA2 levels and**
 832 **clinicopathological characteristics.**

A **B** **C** **D** **E**

HCC patients

Tissue culture

iTRAQ labeling

5 biological repeats

Fraction 1

Fraction 2

Fraction 30

LC-MS/MS (Q Exactive Plus)

Signal P

MetazSecKB

NN-score

Secretome

ELISA

PRM

Potential biomarkers

Clinical analysis

Molecular mechanisms

Signal pathway analysis

HCC cell Model

A

B

C

D

E