Published in final edited form as:
Carcinogenesis. 2007 October ; 28(10): 2149–2153.

Candidate markers for the detection of hepatocellular carcinoma in low-molecular weight fraction of serum

Radoslav Goldman*,
Georgetown University, Lombardi Comprehensive Cancer Center, 3970 Reservoir Road Northwest, Washington, DC 20057, USA

Habtom W. Ressom,
Georgetown University, Lombardi Comprehensive Cancer Center, 3970 Reservoir Road Northwest, Washington, DC 20057, USA

Mohamed Abdel-Hamid,
Minia University and Viral Hepatitis Research Laboratory, NHTMRI, Cairo, Egypt

Lenka Goldman,
Georgetown University, Lombardi Comprehensive Cancer Center, 3970 Reservoir Road Northwest, Washington, DC 20057, USA

Antai Wang,
Georgetown University, Lombardi Comprehensive Cancer Center, 3970 Reservoir Road Northwest, Washington, DC 20057, USA

Rency S. Varghese,
Georgetown University, Lombardi Comprehensive Cancer Center, 3970 Reservoir Road Northwest, Washington, DC 20057, USA

Yanming An,
Georgetown University, Lombardi Comprehensive Cancer Center, 3970 Reservoir Road Northwest, Washington, DC 20057, USA

Christopher A. Loffredo,
Georgetown University, Lombardi Comprehensive Cancer Center, 3970 Reservoir Road Northwest, Washington, DC 20057, USA

Steven K. Drake,
Clinical Chemistry Service, Department of Laboratory Medicine, National Institutes of Health, Bethesda, MD, USA

Sohair A. Eissa,
National Cancer Institute, Cairo, Egypt

Iman Gouda,
National Cancer Institute, Cairo, Egypt

Sameera Ezzat, and
Menoufiya University, Shibin El Kom, Egypt

Francoise Seillier Moiseiwitsch
Georgetown University, Lombardi Comprehensive Cancer Center, 3970 Reservoir Road Northwest, Washington, DC 20057, USA

*To whom correspondence should be addressed. Tel: +1 202 687 9868; Fax: +1 202 687 1988; Email: rg26@georgetown.edu.

Conflict of Interest Statement: None declared.

Abstract

Hepatocellular carcinoma (HCC) represents an important public health problem in Egypt where up to 90% of HCC cases are attributable to hepatitis C viral (HCV) infection. Serum alpha-fetoprotein is elevated in only ~60% of HCC patients. The development of effective markers for the detection of HCC could have an impact on cancer mortality and significant public health implications worldwide. The objective of our study was to assess six candidate markers for detection of HCC identified by mass spectrometric analysis of enriched serum. The study examined 78 HCC cases and 72 age- and gender-matched cancer-free controls recruited from the Egyptian population. Matrix-assisted laser desorption–ionization time-of-flight mass spectrometric analysis of enriched low-molecular weight fraction of serum was used for identification of the candidate markers. Our analyses show that all six candidate markers are associated with HCC after adjustment for important covariates including HCV and hepatitis B viral infections. The marker candidates are independently predictive of HCC with areas under the receiver operating characteristic (AuROC) curve ranging from 63–93%. A combination of the six markers improves prediction accuracy to 100% sensitivity, 91% specificity and 98% AuROC curve in an independent test set of 50 patients. Two of the candidate markers were identified by sequencing as fragments of complement C3 and C4. In conclusion, a set of six peptides distinguished with high prediction accuracy HCC from controls in an Egyptian population with a high rate of chronic HCV infection. Further evaluation of these marker candidates for the diagnosis of HCC is needed.

Introduction

Hepatocellular carcinoma (HCC) is a significant worldwide health problem with as many as 500 000 new cases diagnosed each year (1). There is considerable geographical variation in the incidence of HCC (2). In Egypt, HCC is third among cancers in men with >8000 new cases predicted by 2012 (3-5). The HCC epidemic in Egypt is associated with hepatitis C viral (HCV) infection; Egypt has the highest prevalence of HCV in the world with ~13.8% of the population infected and seven million with chronic HCV liver disease (6). Up to 90% of HCC cases in the Egyptian population were attributed to HCV (5,7). In the USA, the increasing incidence of HCC has been associated with HCV infection (8,9). Studies of HCV progression to HCC are expected to provide new insights on the management of this increasing problem and therefore are of great public health interest (10).

The natural progression of HCV infection to hepatitis, cirrhosis and HCC is slow. Chronic hepatitis develops in ~80% of those infected with HCV. Over the course of ≥20 years, 10–30% of HCV carriers develop cirrhosis; patients with cirrhosis have an annual risk of 1–2% for developing HCC (11). The prognosis of patients with HCC remains extremely poor. The currently available systemic therapies demonstrate poor to modest response rates and have not been shown to improve survival in patients with HCC (12). Complete surgical resection and liver transplant are at present the only curative treatment options (13). However, the majority of patients present with advanced unresectable disease not amenable to definitive local therapies (14,15). The slow development and late detection of HCC suggest that the identification of biomarkers of disease progression and early detection represents attractive strategies for potential improvement of the outcome of HCC patients.

Current diagnosis of HCC relies on clinical information, liver imaging and measurement of serum alpha-fetoprotein (AFP). The reported sensitivity (41–65%) and specificity (80–94%) of AFP are not sufficient for early diagnosis, and so additional markers are needed (16,17). The development of effective markers for the diagnosis of HCC could have an impact on HCC-related cancer mortality and significant public health implications worldwide. This is an active area of research with several groups reporting new marker candidates within the last few years (18-21).

The characterization of peptides in serum is a promising strategy for biomarker discovery (22-24). We developed a method for identification of peptides in the enriched low-molecular weight (LMW) fraction of serum based on matrix-assisted laser desorption–ionization time-of-flight (MALDI–TOF) mass spectrometry (MS) (25). In this study, we describe the application of MALDI–TOF MS to the detection of HCC in a cohort of cases and controls recruited from the Egyptian population (5). Our study identified a set of six discriminatory peptide peaks in the sera of HCC patients and controls that could potentially be used for diagnosis of HCC (26). In this paper, we present an analysis of the performance of six candidate markers in an Egyptian study population.

Materials and methods

Materials

Red-top vacutainer blood collection tubes (BD 366430) were obtained from Becton Dickinson (Franklin Lakes, NJ). C8 magnetic beads, α-cyano-4-hy-droxycinaminic acid and MALDI 600 μm AnchorChip were purchased from Bruker Daltonics (Billerica, MA). Microcon 50 kDa ultrafiltration membranes were purchased from Millipore (Bedford, MA). Other chemicals and solvents were purchased from Sigma–Aldrich (St Louis, MO); solvents were of high-performance liquid chromatography grade.

Study population and sample collection

HCC cases and controls were enrolled in collaboration with the National Cancer Institute of Cairo University, Egypt, from 2000 to 2002, as described previously (5). Briefly, adults with newly diagnosed HCC aged 17 and older without a previous history of cancer were eligible for the study. Diagnosis of HCC was confirmed by pathology, cytology, imaging (computer tomography (CT) ultrasound) and serum AFP. Controls were recruited from the orthopedic department of Kasr El Aini Faculty of Medicine, Cairo University. All participants signed informed consent, provided a blood sample and answered a questionnaire with demographic information, personal habits, medical history, occupational history and agricultural activities. The study protocol was approved by the Institutional Review Committee and conformed to the ethical guidelines of the 1975 Helsinki Declaration.

Blood samples were collected by a trained phlebotomist each day ~10 am and processed within a few hours according to a standard protocol. Aliquots of sera were frozen at −80°C immediately after collection until analysis; all mass spectrometric measurements were performed on twice-thawed sera. Each patient's hepatitis B viral (HBV) and C viral infection status was assessed by enzyme immunoassay for anti-HCV, anti-HBC and HBV surface antigen and by polymerase chain reaction for HCV RNA (5,27). Serum samples of 150 patients from the parent study, consisting of 78 cases and 72 controls matched on age and gender, were analyzed by MALDI–TOF MS as described previously (25,26). Characteristics of this population are summarized in Table I which shows, as expected, increased markers of viral infections (HCV RNA, anti-HCV and anti-HBV) in cancer cases (5,27).

MALDI–TOF MS analysis

We utilized an enrichment procedure to analyze native peptides in the LMW fraction of serum (0.9–5 kDa) using MALDI–TOF MS as described previously (25,28). Briefly, serum samples (15 μl) were desalted on C8 magnetic beads and ultrafiltered in 25% acetonitrile on 50 kDa Microcon membranes. Ultra-filtrates were spotted on AnchorChip MALDI target with α-cyano-4-hydroxycinaminic acid matrix (3.3 mg/ml in 50% acetonitrile). Samples were analyzed on an Ultraflex MALDI–TOF/TOF MS (Bruker Daltonics). Ionization was achieved by irradiation with a nitrogen laser (λ = 337 nm) operating at 20 Hz. An average of 50 shots at each of 20 positions was collected for a total of 1000 shots per spot; the positions were taken

in an automated spiral pattern radiating out from the center of the AnchorChip spot. AnchorChip plate locations were calibrated prior to each run. Positive ions were accelerated at 19 kV with 80 ns of pulsed ion extraction delay. Each spectrum was recorded in linear positive mode and was externally calibrated using a standard mixture of peptides. Mass spectra were acquired using the Flex Control and Flex Analysis software (Bruker Daltonics) and raw data were exported as text files for further analysis. To sequence peptides, an exploratory scan from 800 to 5000 Da was performed in the reflectron mode to assign a mass window (~0.5% mass width) for fragmentation and peptide sequencing in the 'LIFT' MS/MS mode. Additional peptide sequencing was carried out on a 4800 MALDI–TOF/ TOF™ Analyzer (Applied Biosystems, CA). Analysis was carried out in positive ion mode in both reflector and MS/MS acquisitions with laser repetition rate at 200 Hz. In both reflector and MS/MS mode, the instrument default calibration was used. The positive ion reflector spectra were obtained for the peptide mix in 800–4000 mass range. In MS/MS mode, 2 kV collision energy (with Collision induced dissociation gas ON) was used to fragment the peptides. The database searches for peptide identification were performed using MASCOT Distiller 2.1 (Matrix Science, London, UK). The MS/MS spectra obtained from MALDI–TOF/TOF were searched against the SwissProt human database. No enzyme was considered in these searches and both MS and MS/MS tolerance was 0.3 Da.

Data processing and analysis

Raw spectra are available at http://microarray.georgetown.edu/web/files/carcinogenesis.htm. Analyses were carried out with the Flex Analysis, ClinProTools (Bruker Daltonics), MATLAB (MathWorks, Natick, MA) and SAS (SAS, Cary, NC) software packages. Spectra were processed as described previously (26,29). Briefly, the dimension of each spectrum was reduced from ~136 000 m/z values to 23 846 bins (100 p.p.m. step). Baseline-corrected spectra were normalized by dividing each spectrum by its total ion current; spectra were scaled to an intensity of 100 assigned to the highest peak in the training dataset. For determination of peaks, the dataset was randomly divided into 100 training spectra (50 HCC and 50 control) and 50 blinded testing spectra (28 HCC and 22 control). Peaks were identified as a change from positive to negative slope and nearby peaks within 300 p.p.m. mass were coalesced into a single window to account for drift in m/z location. This procedure identified 264 peak-containing windows in the training spectra; the maximum intensity in each window was used as the variable of interest. A particle swarm optimization (PSO)–support vector machine (SVM) algorithm was applied to the training dataset to select six mass windows for classification of HCC cases and controls. The PSO–SVM algorithm combines two machine learning methods, PSO and SVM, as described previously (26,29). PSO starts with N randomly selected particles and searches for the optimal particle iteratively. Each particle is an m-dimensional vector and represents a candidate solution. An SVM classifier is built for each candidate solution (particle) to evaluate its performance through the cross-validation method. The PSO algorithm guides the selection of potential biomarkers (mass windows) that lead to best prediction accuracy in distinguishing between two groups. The algorithm uses the most-fit particles to contribute to the next generation of N candidate particles. Thus, on the average, each successive population of candidate particles fits better than its predecessor. This process continues until the performance of the SVM classifier converges.

The testing spectra were scaled based on the parameters used for scaling the training spectra and peaks that fall within the selected six mass windows were quantified. Sensitivity and specificity of these six marker candidates were evaluated on the testing dataset. Logistic regression models were used to determine association of the marker candidates and covariates including HCV and HBV viral infections (independent variables) with HCC status (dependent variable).

Results

Comparison of MALDI–TOF mass spectra of 78 HCC patients and 72 controls matched on age and gender showed marked differences in intensity of several peptide peaks (Figure 1). Serum samples were enriched as described previously (25). Consistency of the differences is visible in the overlay of spectra in Figure 1; differences in the relative ion intensities of average spectra in the mass range of 0.9–5 kDa are presented in supplementary Figure 1 (available at *Carcinogenesis* Online). We selected 6 of 264 peaks for classification of HCC using PSO–SVM computational methods described previously (26,29). Four of the peaks selected for classification of HCC are pointed out in Figure 1 which expands the 0.9–1.9 kDa region; two additional markers were selected in the 2.5 and the 4.1 kDa region, respectively. Intensities of the six marker candidates (MALDI A–F) are summarized in Table II; the values represent baseline-corrected and -normalized intensities (29). The results show that three candidate markers are higher in cases (MALDI B, MALDI C and MALDI D); the remaining three are higher in controls. Overlay of all 150 spectra for two of the peptide peaks in Figure 2 shows the magnitude of the differences. MALDI A is consistently higher in controls ($n = 72$; black traces); MALDI B is higher in cases ($n = 72$; green traces). The smallest difference was observed for MALDI E (Table II). MALDI E was selected based on criteria of prediction accuracy in combination with the other marker candidates even though the difference in mean abundance for some other peptides was greater (see below).

Table III summarizes results of fitting univariate logistic regression models for each marker candidate and important covariate adjusted for matching variables age and gender. The analysis shows that, besides the six marker candidates, several covariates are significantly associated with HCC status. These include HBV and HBC viral infections, residency (urban versus rural) and date of sample collection (see below). In multivariate logistic regression models, we considered each individual marker candidate (MALDI A–F) together with the relevant covariates (HCV RNA, anti-HBC, residency, date of collection, age and gender). As shown in Table IV, the association of marker candidates remains significant after adjustment for all the important covariates. We did not include anti-HCV in the regression models because it is correlated with HCV RNA (correlation coefficient = 0.823). The six selected markers are weakly correlated with each other; the highest correlation coefficient of 0.392 was observed between MALDI C and MALDI D (supplementary Table 1 is available at *Carcinogenesis* Online). This suggests that the markers are independently predictive of HCC and that their combination should have increased prediction accuracy.

The multivariate analysis showed that HCV RNA and date of collection remain significantly associated with the markers. To examine whether MALDI A–F are associated with HCV infection, we fitted univariate logistic regressions with HCV RNA status as the dependent variable and each of the marker candidates in the control population ($n = 72$). The analysis shows that none of the six selected markers is significantly associated with the presence of viral RNA (supplementary Table 2 is available at *Carcinogenesis* Online). We also showed that the HCV presence is not associated with MALDI A–F in cases. This suggests that the observed association of the peptides with HCC is not driven by the viral infection but rather by the presence of the tumor. A potential association of the markers with liver cirrhosis that accompanies most of the HCC cases needs to be addressed in future studies.

Recruitment of cases for the study began prior to the recruitment of controls; controls were subsequently matched to the recruited cases on age and gender. To make sure that the observed association of MALDI A–F is not biased by the time of storage of samples at −80°C, we analyzed the association of MALDI A–F with HCC status in a subset of samples collected simultaneously between January and April of 2002 (supplementary Table 3 is available at *Carcinogenesis* Online). The subset of cases for this analysis is relatively small (22 HCC and

72 controls) but all the marker candidates except for MALDI E remain significantly associated with HCC. This shows that for at least five of the six markers the time of storage does not affect the association with HCC status.

The long-term goal of our study is to identify markers of HCC that would improve the sensitivity and specificity of AFP (16,17). We determined sensitivity and specificity of MALDI A–F on a blinded independent test set of 50 samples (28 HCC cases and 22 controls). The prediction accuracy of individual marker candidates is summarized in Table V. Sensitivity of the markers ranges from 50–96% and specificity ranges from 36–91%; these values are comparable with the reported performance of AFP (20). Correlation of any of the six peak intensities in 78 HCC cases with AFP was not significant with the correlation coefficients ranging from −0.132 to 0.158; the AFP concentration in controls was not determined. The prediction accuracy in combination with AFP should be evaluated in the future. We were most interested, however, in the performance of a combination of the markers because we expected that a combination of markers would better address the heterogeneity of the disease process. Our analysis shows that the combination of the six selected markers achieved 100% sensitivity and 91% specificity in classification of HCC.

It is important to note that our comparison group represents the general population in Egypt; 33% of the controls carry HCV antibodies and 22% tested positive for HCV RNA. This group of controls is not an arbitrary healthy comparison group and reflects the magnitude of the HCV epidemic in Egypt. Prediction accuracy of the marker candidates is further demonstrated by the receiver operating characteristic curves presented in supplementary Figure 3 (available at *Carcinogenesis* Online). Increased area under the curve for a combined classifier (97% compared with 63–93% for the individual markers as listed in Table V) suggests that a combination of markers is more effective in predicting HCC status than individual markers. Further studies including a cirrhosis comparison group will be needed to evaluate the potential of these diagnostic marker candidates for early detection.

Preliminary sequencing results show that the markers are peptides. Sequencing of MALDI A identified with high probability (MASCOT ion score 99, expectation value 3.5e-6), a fragment of complement C3 with sequence SSKITHRIHWESASLL. Sequencing of MALDI D identified with high probability (MASCOT ions score 131, expect 5.9e-10), a fragment of complement C4a with sequence NGFKSH-ALQLNNRQI (supplementary Figure 2 is available at *Carcinogenesis* Online). Sequencing of the remaining peptides is under way.

Discussion

AFP is the only marker used currently in the clinic for detection of HCC. Although AFP improves detection of HCC, a significant number of HCC patients present without elevated AFP, and therefore additional markers are needed to increase the sensitivity and specificity of detection (16,30). This study describes an initial validation of six marker candidates selected by MALDI–TOF analysis of enriched LMW fraction of serum.

A recent study identified native peptides associated with cancers of the prostate, breast and bladder by MALDI–TOF/TOF analysis of serum desalted on C8 magnetic beads (22). It was suggested that tumor-associated proteolytic activity is responsible for generation of the diagnostic peptides. Analysis of native peptides begins to provide interesting biological insights and potentially new disease markers (23,31,32). Our study used a newly optimized enrichment procedure for a MALDI–TOF peptidomic analysis of LMW fraction of serum in patients with HCC (25).

We selected 6 of 264 peptide peaks for classification of HCC using previously described computational methods (26). We compared intensities of the six marker candidates in serum

samples of 150 patients and found marked differences between HCC cases and controls as demonstrated in Table II and Figures 1-2. The differences are associated with the presence of HCC; the association with HCC remains significant after adjustment for all important covariates (Table IV). We took the following steps to limit the number of false discoveries (33,34). The cases and controls are a representative sample of the Egyptian population (5) with a substantial proportion of controls carrying HCV infection; it is not a convenience sample. A standardized sample collection and processing protocol was used to minimize variability in freeze–thaw cycles and other factors suggested to affect peptide abundance (35). Our analytical methods were optimized to limit variability of the measurements to a mean Coefficient of variation (CV) of ~10% in analysis of 15 replicates of a serum standard (25). We focused our analysis on peaks (as opposed to all observed mass points) and introduced rigorous guidelines for biomarker selection (26,29). The performance of the markers was determined on an independent blinded dataset that was not used for selection of the candidate markers. The identity of two peptides was verified by sequencing.

It is important to note that the selected markers are not the ones with greatest differences between cases and controls based on statistical tests. Each of the six marker candidates is significantly associated with HCC with three of them having higher intensities in cancer patients (Tables II and III) but our computational methods selected markers based on prediction accuracy (29). We expected that a combination of the markers would better classify samples and that interactions may be important in their performance. The results show that the observed correlation coefficients between markers are all smaller than 0.392 which suggests that they are independently predictive of HCC. The individual marker candidates have good prediction accuracy, in general, comparable with AFP (16). MALDI A has highest sensitivity and MALDI B highest specificity; MALDI E has the lowest prediction accuracy (Table V). We have observed higher sensitivity (100%) and specificity (91%) for the combination of six markers. The prediction accuracy of the six most significantly different peak intensities is lower (96% sensitivity and 82% specificity) which suggests that it is not an optimal criterion for selection of marker combinations.

The presented results provide evidence that our MALDI–TOF analysis of LMW serum fraction identified novel candidate markers of HCC. We expect that these peptides and protein fragments are associated with modified proteolytic activity in the HCC patients; at present, this hypothesis remains a speculation. Our current effort focuses on the sequencing of the candidates and on the analysis of their biological origin. Preliminary sequencing results identified MALDI A and MALDI D as fragments of complement C3 and C4a, respectively (supplementary Figure 2 is available at *Carcinogenesis* Online). Functional implications of these fragments of proteins involved in innate immune response in the serum of patients need further study. Future expansion of the study and comparison with a group of cirrhotic patients is needed to determine if the candidate markers are correlated with cirrhosis and have potential clinical utility for the detection of HCC. The sequenced peptides can be quantified by an appropriate analytical method, for example an isotope dilution mass spectrometric assay.

In summary, we present evidence that six peptide peaks present in LMW fraction of serum are good candidate markers of HCC. The observation should be repeated in an independent set of samples and expanded outside of the Egyptian population. Sequencing of the marker candidates is essential to provide further biological insights and means for accurate quantification.

Supplementary material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We want to acknowledge Dr Alaa Ismail and Dr Hussain Khaled for invaluable help with design and execution of the studies in Egypt.

Funding An Associate Membership from National Cancer Institute's Early Detection Research Network, U.S. Army Medical Research and Material Command, Prostate Cancer Research Program (W81XWH-04-1-0294); National Cancer Institute (R03 CA119288 and R01 CA115625-01A2 to R.G., R03 CA119313 H.W.R. and RO1CA85888 to C.L.).

References

1. Montalto G, et al. Epidemiology, risk factors, and natural history of hepatocellular carcinoma. Ann N Y Acad Sci 2003;963:13–20. [PubMed: 12095924]
2. Parkin DM. The global health burden of infection-associated cancers in the year 2002. Int J Cancer 2006;118:3030–3044. [PubMed: 16404738]
3. Mizokami M, et al. Tracing the evolution of hepatitis C virus in the United States, Japan, and Egypt by using the molecular clock. Clin Gastroenterol Hepatol 2005;3:S82–S85. [PubMed: 16234067]
4. Deuffic-Burban S, et al. Expected increase in hepatitis C-related mortality in Egypt due to pre-2000 infections. J Hepatol 2006;44:455–461. [PubMed: 16310281]
5. Ezzat S, et al. Associations of pesticides, HCV, HBV, and hepatocellular carcinoma in Egypt. Int J Hyg Environ Health 2005;208:329–339. [PubMed: 16217918]
6. Perz JF, et al. The coming wave of HCV-related liver disease: dilemmas and challenges. J Hepatol 2006;44:441–443. [PubMed: 16426700]
7. Hassan MM, et al. The role of hepatitis C in hepatocellular carcinoma: a case control study among Egyptian patients. J Clin Gastroenterol 2001;33:123–126. [PubMed: 11468438]
8. El-Serag HB, et al. Rising incidence of hepatocellular carcinoma in the United States. N Engl J Med 1999;340:745–750. [PubMed: 10072408]
9. Wong JB, et al. Estimating future hepatitis C morbidity, mortality, and costs in the United States. Am J Public Health 2000;90:1562–1569. [PubMed: 11029989]
10. Di Bisceglie AM, et al. Hepatitis C-related hepatocellular carcinoma in the United States: influence of ethnic status. Am J Gastroenterol 2003;98:2060–2063. [PubMed: 14499788]
11. Ikeda K, et al. Disease progression and hepatocellular carcinogenesis in patients with chronic viral hepatitis: a prospective observation of 2215 patients. J Hepatol 1998;28:930–938. [PubMed: 9672166]
12. Llovet JM. Updated treatment approach to hepatocellular carcinoma. J Gastroenterol 2005;40:225–235. [PubMed: 15830281]
13. Schwartz JM, et al. Treatment of hepatocellular carcinoma. Curr Treat Options Gastroenterol 2003;6:465–472. [PubMed: 14585235]
14. Steel LF, et al. A proteomic approach for the discovery of early detection markers of hepatocellular carcinoma. Dis Markers 2001;17:179–189. [PubMed: 11790885]
15. Lopez LJ, et al. Hepatocellular carcinoma. Curr Opin Gastroenterol 2004;20:248–253. [PubMed: 15703649]
16. Marrero JA. Screening tests for hepatocellular carcinoma. Clin Liver Dis 2005;9:235–251. [PubMed: 15831271]
17. Gupta S, et al. Test characteristics of alpha-fetoprotein for detecting hepatocellular carcinoma in patients with hepatitis C. A systematic review and critical analysis. Ann Intern Med 2003;139:46–50. [PubMed: 12834318]
18. Filmus J, et al. Glypican-3 and alphafetoprotein as diagnostic tests for hepatocellular carcinoma. Mol Diagn 2004;8:207–212. [PubMed: 15887976]
19. Block TM, et al. Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans. Proc Natl Acad Sci USA 2005;102:779–784. [PubMed: 15642945]
20. Chignard N, et al. Cleavage of endoplasmic reticulum proteins in hepatocellular carcinoma: detection of generated fragments in patient sera. Gastroenterology 2006;130:2010–2022. [PubMed: 16762624]

21. Lee IN, et al. Identification of complement C3a as a candidate biomarker in human chronic hepatitis C and HCV-related hepatocellular carcinoma using a proteomics approach. Proteomics 2006;6:2865–2873. [PubMed: 16586433]
22. Villanueva J, et al. Differential exoprotease activities confer tumor-specific serum peptidome patterns. J Clin Invest 2006;116:271–284. [PubMed: 16395409]
23. Tammen H, et al. Peptidomic analysis of human blood specimens: comparison between plasma specimens and serum by differential peptide display. Proteomics 2005;5:3414–3422. [PubMed: 16038021]
24. Hortin GL, et al. Proteomics: a new diagnostic frontier. Clin Chem 2006;52:1218–1222. [PubMed: 16675505]
25. Orvisky E, et al. Enrichment of low molecular weight fraction of serum for mass spectrometric analysis of peptides associated with hepatocellular carcinoma. Proteomics 2006;6:2895–2902. [PubMed: 16586431]
26. Ressom, H., et al. Analysis of MALDI-TOF serum profiles for bio-marker selection and sample classification; Proceedings of the IEEE Symposium on Computational Inteligence in Bioinformatics and Computational Biology; La Jolla, CA. 2005. p. 378-384.
27. Abdel-Hamid M, et al. Optimization, assessment, and proposed use of a direct nested reverse transcription-polymerase chain reaction protocol for the detection of hepatitis C virus. J Hum Virol 1997;1:58–65. [PubMed: 10195232]
28. Tirumalai RS, et al. Characterization of the low molecular weight human serum proteome. Mol Cell Proteomics 2003;2:1096–1103. [PubMed: 12917320]
29. Ressom H, et al. Analysis of mass spectral serum profiles for bio-marker selection. Bioinformatics 2005;21:4039–4045. [PubMed: 16159919]
30. Zhang BH, et al. Randomized controlled trial of screening for hepatocellular carcinoma. J Cancer Res Clin Oncol 2004;130:417–422. [PubMed: 15042359]
31. Schulz-Knappe P, et al. The peptidomics concept. Comb Chem High Throughput Screen 2005;8:697–704. [PubMed: 16464157]
32. Hortin GL. The MALDI-TOF mass spectrometric view of the plasma proteome and peptidome. Clin Chem 2006;52:1223–1237. [PubMed: 16644871]
33. Ransohoff DF. Bias as a threat to the validity of cancer molecular-marker research. Nat Rev Cancer 2005;5:142–149. [PubMed: 15685197]
34. Ransohoff DF. Rules of evidence for cancer molecular-marker discovery and validation. Nat Rev Cancer 2004;4:309–314. [PubMed: 15057290]
35. Villanueva J, et al. Correcting common errors in identifying cancer-specific peptide signatures. J Proteome Res 2005;4:1060–1072. [PubMed: 16083255]

Abbreviations

AFP
alpha-fetoprotein

HBV
hepatitis B virus

HCC
hepatocellular carcinoma

HCV
hepatitis C virus

LMW
low-molecular weight

MALDI–TOF
matrix-assisted laser desorption–ionization time-of-flight

MS
: mass spectrometry

PSO
: particle swarm optimization

SVM
: support vector machine

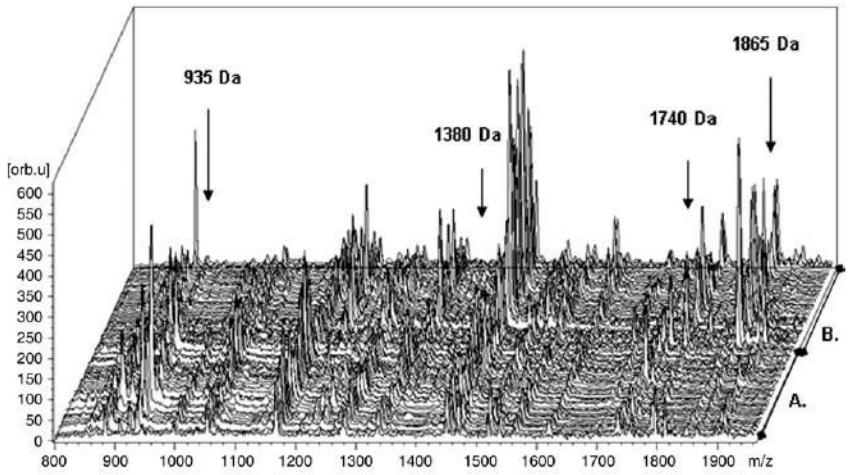

Fig. 1.

Overlay of MALDI–TOF spectra in the mass region 0.9–1.9 kDa; Arrows indicate marker candidates selected in this mass region. (**A**) Cases (*n* = 78); (**B**) controls (*n* = 72).

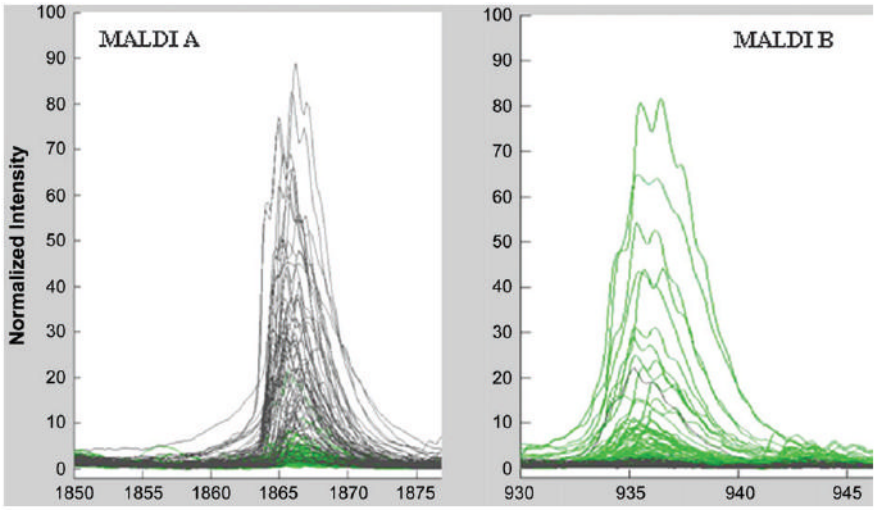

Fig. 2.

Overlay of HCC (green traces, $n = 78$) and controls (black traces, $n = 72$) for two selected biomarker candidates; x-axis indicates m/z values.

Table I

IDemographic variables and viral infections

	Cases (n = 78)	**Controls (*n* = 72)**	***P* value**
Mean age (SD)	54 (9.1)	52 (12.0)	0.3197
Male gender	57 (73%)	50 (69%)	0.6231
Smokers	43 (55%)	39 (54%)	0.8373
HCV RNA+	62 (80%)	16 (22%)	<0.0001
Anti-HCV+	69 (88%)	24 (33%)	<0.0001
Anti-HBV+	60 (77%)	40 (56%)	0.0050
HbsAg+	5 (6%)	1 (1%)	0.2291

Presence of HCV virus (HCV RNA), HCV antibodies (anti-HCV), HBV surface antigen (HBsAg) and HBV antibodies (anti-HBV) was tested as described in Materials and methods; SD = standard deviation.

Table II

Descriptive statistics

Marker	m/z window (Da)	Cases (*n* = 78)			Controls (*n* = 72)		
		Mean	Median	SD	Mean	Median	SD
MALDI A	1863–1871	3.7	2.9	3.0	25.9	21.4	21.1
MALDI B	934–938	9.6	4.5	14.5	1.3	0.9	2.5
MALDI C	2529–2536	2.0	1.6	1.3	1.2	1.1	0.5
MALDI D	1737–1744	5.2	3.8	4.5	2.6	2.1	2.0
MALDI E	1379–1381	1.1	1.1	0.5	1.6	1.3	0.9
MALDI F	4086–4098	1.2	0.7	1.2	3.1	3.0	1.9

Normalized peak intensities of marker candidates in cases and controls; SD = standard deviation.

Table III

Association of candidate markers and covariates with HCC

	P value	OR	95% CI	
Anti-HCV	<0.0001	15.3	6.55	35.87
HCV RNA	<0.0001	9.6	4.47	20.70
Anti-HBV	0.0062	2.7	1.32	5.38
HBsAg	0.1535	4.9	0.55	42.66
Residency	0.0017	3.0	1.50	5.82
Smoking	0.8370	1.1	0.56	2.04
Date of collection	<0.0001	10.9	4.60	26.02
MALDI A	<0.0001	0.7	0.65	0.83
MALDI B	<0.0001	2.1	1.52	2.78
MALDI C	<0.0001	3.3	1.87	5.89
MALDI D	<0.0001	1.6	1.26	1.93
MALDI E	<0.0001	0.3	0.16	0.55
MALDI F	<0.0001	0.4	0.32	0.58

Univariate logistic regression adjusted for age and gender (n = 150). OR, odds ratio; 95% CI, 95% confidence interval; HBsAg, HBV surface antigen.

Table IV

Association of candidate markers with HCC

	***P* value**	**OR**	**95% CI**	
MALDI A	<0.0001	0.70	0.59	0.82
MALDI B	<0.0001	1.76	1.33	2.33
MALDI C	0.0045	2.98	1.40	6.35
MALDI D	0.0011	1.39	1.14	1.69
MALDI E	0.0044	0.33	0.16	0.71
MALDI F	<0.0001	0.45	0.32	0.63

Multivariate logistic regression model controlled for HCV RNA, anti-HBC, residency, date of collection, age and gender (n = 150). OR, odds ratio; 95% CI, 95% confidence interval.

Table V

Prediction accuracy of individual markers

	Sensitivity (%)	Specificity (%)	AuROC (%)
MALDI A	96	73	93
MALDI B	82	91	91
MALDI C	50	91	70
MALDI D	64	82	77
MALDI E	71	36	63
MALDI F	82	59	83
All combined	100	91	98

Sensitivity, specificity and AuROC curve derived from an independent test set of 28 cases and 22 controls.