Alpha-2-Heremans-Schmid-glycoprotein (AHSG) a potential biomarker associated with prognosis of chromophobe renal cell carcinoma: The PROPOLIS study

Xin Li1 | Cuiyan Ma2,3,4

1Department of Rehabilitation Medicine, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, China
2Department of Computer Science and Technology, BNRist, RIIT, Institute of Internet Industry, Tsinghua University, Beijing, China
3Integrative Medicine Center, School of Life Sciences, Beijing University of Chinese Medicine, Beijing, China
4Tsinghua-Peking Center for Life Sciences, School of Life Sciences, Tsinghua-Peking Center for Life Sciences, Tsinghua University, Beijing, China

Correspondence
Cuiyan Ma, Integrative Medicine Center, School of Life Sciences, Beijing University of Chinese Medicine, No. 11, Bei San Huan Dong Lu, Chaoyang District, Beijing, China
Email: mcymcy402@163.com

Funding information
National Natural Science Foundation of China, Grant/Award Number: NSFC91646202; China Postdoctoral Science Foundation, Grant/Award Number: 2019M650635

Abstract

Background and Aims: Chromophobe renal cell carcinoma (chRCC) is the third common pathological subtype in renal cancers. However, the underlying mechanisms of specific genetic characteristics of chRCC are currently unclear. In this study, protein expression profiles, gene ontology (GO), and survival plots were provided by integrated bioinformatics analysis to investigate key genes associated with the mechanism of tumorigenesis and prognosis of chRCC.

Methods: The chRCC data set of gene expression profiles and clinical data were obtained from the gdc-client (https://portal.gdc.cancer.gov) deposited on The Cancer Genome Atlas (TCGA) data portal. Differentially expressed genes (DEGs) in chRCC, compared with normal samples, were analyzed by R packages “DESeq2,” “edgeR,” and “limma.” Heat maps, volcano plots, and principal component analysis (PCA) were performed for integrated analyses. GUniGO, mutant analysis, and survival plots were performed by R packages. A protein–protein interaction (PPI) network was generated and analyzed by R packages, online String software, and Cytoscape software. Survival analysis and gene expressing comparison in tumor and normal samples were used to detect the core genes of chRCC. Furthermore, the top interacting proteins were reanalyzed.

Results: A total of 306 upregulated genes and 678 downregulated genes were identified by a Venn diagram. Ten hub genes were extracted from PPI network. Furthermore, Alpha-2-Heremans-Schmid-glycoprotein (AHSG), one of 10 hub genes, was found to be associated with chRCC, and had a big difference in expression between survival and dead events. AHSG could predict potential prognostic and may be a diagnostic biomarker in chRCC.

Conclusion: This study illustrated that AHSG may be a potential therapeutic target and prognostic genetic marker for chRCC.

KEYWORDS
bioinformatics analysis, biomarker, chRCC, inflammation, renal cancer

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. Health Science Reports published by Wiley Periodicals LLC.
1 | INTRODUCTION

Kidney cancer is a common cancer both for males and females, with an incidence of about 3.7% of new cancer cases.\(^1\) Renal cell carcinoma (RCC), the most prevalent kind, among kidney cancers, is approximately 85%. RCC is a heterogeneous disease consisting of clear cell renal cell carcinoma (ccRCC) and non-clear cell renal cell carcinoma (nccRCC)\(^2\)–\(^4\). The third most common pathological subtype of renal malignancies is chromophobe renal cell carcinoma (chRCC), and the majority of its patients are elderly. Patients with chRCC often have a better prognosis and satisfactory treatment outcomes than those with other types of RCC.\(^5\) At present, the treatment of RCC has shifted from nonspecific immunotherapy to targeted therapy, and further to new immunotherapeutic agents.\(^6\)–\(^8\) However, the effectiveness of targeted therapy for chRCC is not well defined, and immunotherapeutic agents are still in clinical trials. Many questions remain regarding the effectiveness of biomarker testing and the choice of treatment options for patients. The overall prognosis remains poor, especially for patients with high-stage diseases, and tumor metastasis and death occur frequently.\(^9\),\(^10\) Factors associated with these malignant biological behaviors include, among others, large tumor volume, late pathological stage, coagulative necrosis, sarcomatoid differentiation, and other factors.\(^11\),\(^12\)

Unlike ccRCC and papillary RCC, the pathological stage of chRCC is not related to the prognosis. There are many opinions on the classification of chRCC. In 2010, the Paner system provided a three-stage classification of renal suspicious cell carcinoma without counting the heterotypic cells contained in those renal chromophobe cell carcinomas, and mainly considered the density of cancer nuclei, the depth of nuclear staining, and the state of cell dedifferentiation or sarcomatoid differentiation.\(^11\),\(^13\)

Alpha-2-Heremans-Schmid-glycoprotein (AHSG) has been found to be a circulating plasma protein and was associated with inflammatory conditions. AHSG gene polymorphisms showed an association with aortic calcification in general hemodialysis patients, and Fetuin-A and IL-6 played a dominant role in the development of aortic calcification.\(^14\) On the other hand, researchers found that red cell distribution width (RDW) values were a novel inflammatory marker in routine hemogram in thyroid cancer patients, implying an increased inflammatory burden in cancer patients.\(^15\)

In the absence of definitive research, studying chRCC from a genetic perspective will help to understand the pathogenesis and progression of the disease and play an important role in finding safer and more effective treatments.

2 | METHODS

2.1 | Ethical requirements

This study was not linked to any research on human participants or animals and therefore did not have any ethical compliance. Figure 1 shows the study workflow.

2.2 | Gene expression data

The clinical data and the gene expression data for the chRCC were downloaded from The Cancer Genome Atlas (TCGA) by the gdc-client (https://portal.gdc.cancer.gov). We selected kidney chromophobe (KICH) data from TCGA website to study the chRCC (https://portal.gdc.cancer.gov). The downloaded genomic data were in the format of

---

**FIGURE 1** Workflow of this study to identify marker genes in chRCC. CC, cellular component; chRCC, chromphobe renal cell carcinoma; GO, gene ontology; KM plot, Kaplan-Meier plot; PCA, principal component analysis; TCGA, The Cancer Genome Atlas.
HTSeq counts, and the data contained 89 clinical samples, consisting of 65 tumor samples and 24 normal samples.

2.3 | Differentially expressed genes (DEGs) data

RNA sequence data were analyzed by the R packages "DESeq2," "edgeR," and "limma." We used Venn diagram to analyze the shared upregulated and downregulated genes through VennDiagram package. In the present study, statistically significant DEGs were defined by the values of adjusted $p < 0.05$ and $|\text{logFC}| > 2$ as cutoff criteria.

The DEG profiles from the "DESeq2," "edgeR," and "limma" analyses are shown in the individual volcano and individual heat maps, respectively. The shared upregulated and downregulated genes analyzed by "DESeq2," "edgeR," and "limma" were visualized by a principal components analysis (PCA) plot and a heatmap.

2.4 | DEGs enrichment analysis

Gene ontology (GO) analysis is a commonly method for annotating genes and gene products by biological pathways (BP), molecular functions (MF), and cell components (CC). To obtain more protein functions to explore chRCC biological information, DEGs were integrated by GO analysis through the cluster Profiler R package to identify biological properties. Key results were obtained for CC with a cutoff value of <0.05.

2.5 | PPI network construction

The Search Tool for the Retrieval of Interacting Genes (STRING) database (https://string-db.org/) is a useful tool for the prediction and the analysis of protein–protein interactions (PPI). We retrieved the interacting genes through PPIs database and found the hub genes by the website software of String (http://www.string-db.org/) and cytohubba from cytoscape (http://www.cytoscape.org/), which could be used for the function enrichment analysis and the interaction network analysis.

2.6 | Statistical analysis

We show the mutational landscape of important genes used to discover chRCC tumor samples by the maftools R package. All $p$-values were statistically significant at an adjusted $p < 0.05$. All data processing was performed using R (64-bit, version 4.0.3) software.

2.7 | Overall survival (OS) Kaplan–Meier (KM) estimate

OS was calculated from the diagnosis date to the end of follow-up, or from the patient's death. The follow-up time varied between patients, with a maximum follow-up time of 5132 days. The survival curves were performed by the R packages "survival" and "survminer." For this analysis, only the samples related to the tumor stage were considered.

2.8 | Gene expression comparison in vital status

The expression level of the hub genes was compared by the groups of live and dead events of 65 cancer samples by using the ggstatsplot R package. To investigate the probable regulation mechanisms of Alpha-2-HS-glycoproteinas in chRCC, the tool of cytoscape, GeneMANIA, was used to annotate the biological processes of AHSG.

3 | RESULTS

3.1 | DEGs identification

We filtered a total of 65 tumor samples and 24 normal samples from the expression profile database, TCGA-KICH data in this study. We compared the cancer samples with normal ones to get the DEGs and Volcano plots and heatmaps, respectively by three analyzed methods of "DESeq2," "edgeR," and "limma" (Figure 2). Then we provided the principal component analysis between cancer samples and normal ones by PCA plots (Figure 3A). We confirmed the shared DEGs consisting of 306 upregulated and 678 downregulated genes through a Venn diagram analyzed by "DESeq2," "edgeR," and "limma." A heatmap was performed to analyze the shared DEGs (Figure 3B-D).

3.2 | GO analysis

The common DEGs were divided into the upregulated and downregulated groups. R 4.0.3 software was used to perform the CC analysis for the whole of 984 DEGs. The CC analysis showed that the DEGs particularly came from the top 20 terms, enriched in apical plasma membrane, apical part of cell, brush border, brush border membrane, cluster of actin-based cell projections, intrinsic component of synaptic membrane, integral cell projection membrane, component of synaptic membrane, integral component of postsynaptic membrane, postsynaptic membrane, integral component of postsynaptic membrane, postsynaptic membrane, intrinsic component of presynaptic membrane, presynapse, integral component of presynaptic membrane, axon terminus, presynaptic membrane, basolateral plasma membrane, collagen-containing extracellular matrix, and transmembrane transporter complex (Figure 4). The common genes of the top 5 pathway from CC were displayed in Figure 5.

3.3 | Survival analysis

To identify the significantly survival curves, the R packages "survival" and "survminer" were performed. According to the
analysis, we found that the survival significantly and progressively worsened among chRCC patients in the higher stage. The patients in the first three stages had a slower progress, while the patients in the last stage had a significantly faster progress and had higher mortality (Figure 6).

3.4 | Mutation analysis

DEGs mutation landscape was performed to further analyze chRCC. 47 of 65 cancer samples had at least one mutation. Among them, tumor protein 53 (TP53) had the highest mutation frequency, phosphatase tensin homolog (PTEN), and zonadhesin (ZAN) followed TP53, and one sample among them had the most mutations (603 gene mutations). The mutation landscape was showed by the waterfall plot with the top 30 critical molecules, including six variant types: missense_mutation, multi-hit, splice_site, frame_shift_ins nonsense_mutation, and frame_shift_del. And the missense mutation was the significantly focused type among the six types (Figure 7A). Figure 7B further analyzed the variants of the samples from six different aspects.

3.5 | PPI network and the hub genes analysis

Using STRING tool and Cytoscape software, we built a PPI network to analyze the chRCC DEGs. Cytohubba in Cytoscape software was used to select the top hub genes from PPI network (Figure 8A). The hub genes were fibrinogen alpha chain (FGA, MIM: 134820), fibrinogen beta chain (FGB, MIM: 134830), alpha 2-HS glycoprotein (AHSG, MIM: 138680), albumin (ALB, MIM: 103600), alpha-1-microglobulin/bikunin precursor (AMBP, MIM: 176870), serpin family A member 1 (SERPINA1, MIM: 107400), apolipoprotein H (APOH, MIM: 138700), inter-alpha-trypsin inhibitor heavy chain 2 (ITIH2, MIM: 146640), histidine-rich glycoprotein (HRG, MIM: 142640), and transthyretin (TTR, MIM: 176300), which mostly were serum proteins and plasma proteins, mainly functioning in coagulation, protease inhibitors, transport or inflammatory regulation.25–33

To decide the significant biomarkers, the KM survival curve was constructed for the top 10 hubs. Only the expression of AHSG was found to be associated with chRCC. According to the further analysis, only the events with the high expression of AHSG gene had a significantly worse survival contrasted to normal samples (adjusted p < 0.05) (Figure 8).
There was a significantly difference for AHSG expression between the dead chRCC patients (n = 9) and the survival chRCC patients (n = 56) (Figure 9A). To investigate the regulating mechanisms of AHSG gene in chRCC, we annotated the biological processes of AHSG by GeneMANIA. The result showed that the top 20 proteins interacting with AHSG (Figure 9B) mainly functioned in tumor suppressor, signal transduce, cancer development.26,27,29,30,33

| DISCUSSION |

ccRCC is a malignant tumor originating from the collecting epithelial cells of the renal tubules, which has no obvious clinical symptoms in the early stage and is clinically uncommon. However, some patients still have a poor prognosis and patients with the advanced disease require adjuvant therapy. Compared with other cancers, chRCC remains less studied.1,40

The bioinformatics approach to studying the genes associated with various stages of chRCC and identifying key regulatory genes in disease development is important for further understanding the disease mechanism, improving disease treatment measures, and developing corresponding targeted drugs.

Our study was based on the KICH data from the publicly available database, TCGA database without a large sample size. Serum AHSG protein levels have been found to reflect tumor burden and inflammatory conditions.41 The findings of some studies indicated a modest linear association between the Fetuin-A (AHSG...
protein) concentration and the risk of colorectal cancer, showing that Fetuin-A was significantly different in the diagnosis of prostate cancer. Increased inflammatory burden also has been reported in cancers. AHSG gene is associated with inflammatory conditions. From this perspective, studying AHSG in RCC is also reasonable. Larger-scale prospective studies are needed to determine the specificity and sensibility of AHSG to the disease.

In our study, to determine possible biomarkers and molecular mechanisms of chRCC, we analyzed TCGA-KICH data using integrated bioinformatics. We collected 65 tumor samples and 24 normal samples for analysis. To obtain more reliable results, DESeq2, edgeR, and limma tools were selected, along with heatmaps and volcano plots (Figure 2). Then we used the shared 306 up and 678 downregulated DEGs among these three methods to perform further analysis (Figure 3C,D).

We employed the clusterProfiler R package to perform Functional enrichment and GO analysis to process DEGs, concluding CC showed that the DEGs mainly originate from the apical plasma membrane, apical part of cell, cell projection membrane, synaptic membrane, and presynapse (top 5 cell constituents) (Figure 4). To further study the DEGs functions, the top 5 pathway common genes (62 genes) of GO CC were analyzed (Figure 5).

Waterfall plots of the top 30 key molecules presented six variant mutation types, among which, missense mutation was the focused variant type among the six types (Figure 6). All patients had less than 50 mutations except one patient who had 603 mutations. TP53 had the most mutations (29%) in all patients (Figure 7), consistent with the previous studies that TP53 was a gene with a highly frequent mutation in human cancers.
Ten hub genes (FGA, FGB, AHSG, ALB, AMBP, SERPINA1, APOH, ITIH2, HRG, and TTR) were screened by cytohubba analysis, and we could see that only AHSG expression was significantly associated with chRCC through the KM survival curve. According to further analysis of the gene expression between the dead chRCC patients (n = 9) and the surviving chRCC patients (n = 56) (Figure 9A), only the AHSG gene survived significantly less in chRCC contrasted to normal samples (p < 0.05) (Figure 8), suggesting that AHSG may be a potential biomarker associated with prognosis of chRCC.26,27,29,30,33,34,35,37,38

AHSG, a2-HS-glycoprotein, belongs to the family of cysteine protease inhibitors and is a kind of serum proteins that is abundant in fetal serum. Adult liver synthesizes about 95% of AHSG protein, which is secreted into the blood, with the normal AHSG protein concentration being about 450–600 μg/ml, contributing to biological functions of AHSG,47 with some similar methods to other proteins by interacting with different kinds of serum calciproteins,48 AHSG can activate the toll-like receptor 4 (TLR4) and induce interleukin-1β to secret from macrophages. Some recent studies also found that AHSG was related to small intestinal neuroendocrine tumors and the AHSG content was associated with differentiating tumor stages.49,50 In this study, AHSG was found to be the only key gene significantly
associated with OS in chRCC patients, which showed that AHSG was closely related to the prognosis of chRCC. Therefore, AHSG may be a new direction for the study of chRCC.

In conclusion, we detected genes associated with the occurrence and development of chRCC by multiple integrated bioinformatics approaches performed on TCGA data and further explored key genes associated with CHCCC. Among them, AHSG may be a potential therapeutic target and prognostic genetic marker for chRCC. However, due to the lack of in-depth studies at the level of chRCC-related genes, further evaluations and experiments are needed to assess these bioinformatics results to elucidate the biological function of this gene in the pathogenesis of chRCC and to provide new clues and directions for the treatment of chRCC.

AUTHOR CONTRIBUTIONS
Cuiyan Ma: Conceptualization; Data curation; Formal analysis; Funding acquisition; Resources; Visualization; Writing–original draft; Writing–review & editing. Xin Li: Conceptualization; Funding acquisition; Resources; Supervision; Writing–review & editing. All authors have read and approved the final version of the manuscript. The corresponding author or manuscript guarantor had full access to all of the data in this study and takes complete responsibility for the integrity of the data and the accuracy of the data analysis.

ACKNOWLEDGMENTS
We acknowledged the biotrainee group to share bioinformatic knowledge. This study was supported by the grants from the National Natural Science Foundation of China (Grant NSFC91646202 to Chunx. Xing) and the Postdoctoral Foundation at Center for Life Sciences and the China Postdoctoral Science Foundation (Grant 2019M650635 to Cuiy. Ma).

CONFLICT OF INTEREST
The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT
Data and materials in this study are available. The clinical data and the gene expression data for the chRCC were downloaded from The Cancer Genome Atlas (TCGA) by the gdc-client (https://portal.gdc.cancer.gov). The downloaded data format is HTSeq-Counts containing 89 clinical samples, consisting of 65 tumor ones and 24 normal ones. The authors confirm that the data supporting the findings of this study are available within the article or TCGA database.

TRANSPARENCY STATEMENT
The lead author Xin Li affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

ORCID
Cuiyan Ma http://orcid.org/0000-0002-1474-9034

REFERENCES
1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. CA Cancer J Clin. 2017;67(1):7-30. doi:10.3322/caac.21387
2. D’Angelo A, Bagby S, Di Pierro G, et al. An overview of the clinical use of cabozantinib in the treatment of advanced non-clear-cell
pathway induced large tibial defect rats. *Biomed Pharmacother*. 2021;138:111480. doi:10.1016/j.biopha.2021.111480

39. Tugba T, Sari I, Solmaz D, et al. Fetuin-A is related to syndesmophytes in patients with ankylosing spondylitis: a case control study. *Clinics (Sao Paulo)*. 2014;69(10):688-693.

40. Bai HS, Wang D, Wen L, Shou JZ, Li CL, Xing NZ. Diagnosis and treatment of 126 cases of chromophobe renal cell carcinoma. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao*. 2021;43(2):247-252. doi:10.3881/j.issn.1000-503X.13032

41. Shi F, Wu H, Qu K, et al. Identification of serum proteins AHSG, FGA and APOA-I as diagnostic biomarkers for gastric cancer. *Clin Proteomics*. 2018;15:18. doi:10.1186/s12014-018-9194-0

42. Nimptsch K, Aleksandrova K, Boeing H, et al. Plasma Fetuin-A concentration, genetic variation in the AHSG gene and risk of colorectal cancer. *Int J Cancer*. 2015;137(4):911-920. doi:10.1002/ijc.29448

43. Guzel O, Kosem A, Aslan Y, et al. The role of Pentraxin-3, Fetuin-A and Sirtuin-7 in the diagnosis of prostate cancer. *Urol J*. 2021;19(3):196-201. doi:10.22037/uj.v19i3.6626

44. Aktas G, Sit M, Karagoz I, et al. Could red cell distribution width be a marker of thyroid cancer. *J Coll Physicians Surg Pak*. 2017;27(9):556-558.

45. Muzasti RA, Suhardjono D, Purwanto B, Juwita Sembiring R. Is there an association between alpha-2-Heremans-Schmid glycoprotein (AHSG) gene Thr256Ser polymorphisms with aortic calcification in regular hemodialysis patients in Medan, Indonesia. *Med Glas (Zenica)*. 2020;17(1):46-53. doi:10.17392/1072-20

46. Zhao Y, Li Y, Sheng J, et al. PS3-R273H mutation enhances colorectal cancer stemness through regulating specific lncRNAs. *J Exp Clin Cancer Res*. 2019;38(1):379. doi:10.1186/s13046-019-1375-9

47. Dabrowska AM, Tarach JS, Wojtysiak-Duma B, Duma D. Fetuin-A (AHSG) and its usefulness in clinical practice. Review of the literature. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*. 2015;159(3):352-359. doi:10.5507/bp.2015.018

48. Heo JI, Yoon DW, Yu JH, et al. Melatonin improves insulin resistance and hepatic steatosis through attenuation of alpha-2-HS-glycoprotein. *J Pineal Res*. 2018;65(2):e12493. doi:10.1111/jpi.12493

49. Chang TT, Ho CH. Plasma proteome atlas for differentiating tumor stage and post-surgical prognosis of hepatocellular carcinoma and cholangiocarcinoma. *PLoS One*. 2020;15(8):e0238251. doi:10.1371/journal.pone.0238251

50. Wang Q, Yu C. Expression profiling of small intestinal neuroendocrine tumors identified pathways and gene networks linked to tumorigenesis and metastasis. *Biosci Rep*. 2020;40(6):BSR20193860. doi:10.1042/BSR20193860

How to cite this article: Ma C, Li X. Alpha-2-Heremans-Schmid-glycoprotein (AHSG) a potential biomarker associated with prognosis of chromophobe renal cell carcinoma: the PROPOLIS study. *Health Sci Rep*. 2022;5:e878. doi:10.1002/hsr2.878