

ORIGINAL ARTICLE OPEN ACCESS

Integration of Gastric Cancer RNA-Seq Datasets Along With PPI Network Suggests That Nonhub Nodes Have the Potential to Become Biomarkers

Akram Siavoshi¹ | Mehran Piran² | Ali Sharifi-Zarchi³ | Fatemeh Ataellahi⁴

¹Department of Alborz Health Technology Development Center, Alborz University of Medical Sciences, Alborz, Iran | ²Department of Medical Biotechnology, Drug Design and Bioinformatics Unit, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran | ³Department of Computer Engineering, Sharif University of Technology, Tehran, Iran | ⁴Department of Biology, College of Sciences, Shiraz University, Shiraz, Iran

Correspondence: Akram Siavoshi (siavoshi.ss@gmail.com) | Ali Sharifi-Zarchi (asharifi@sharif.ir)

Received: 21 February 2024 | **Revised:** 22 December 2024 | **Accepted:** 15 January 2025

Funding: The authors received no specific funding for this work.

Keywords: biomarker discovery | gastric cancer | IGFBP2 | integration of RNA-seq datasets | nonhubs | protein–protein interaction network

ABSTRACT

Background: The breakthrough discovery of novel biomarkers with prognostic and diagnostic value enables timely medical intervention for the survival of patients diagnosed with gastric cancer (GC). Typically, in studies focused on biomarker analysis, highly connected nodes (hubs) within the protein–protein interaction network (PPIN) are proposed as potential biomarkers. However, this study revealed an unexpected finding following the clustering of network nodes. Consequently, it is essential not to overlook weakly connected nodes (nonhubs) when determining suitable biomarkers from PPIN.

Methods and Results: In this study, several potential biomarkers for GC were proposed based on the findings from RNA-sequencing (RNA-Seq) datasets, along with differential gene expression (DGE) analysis, PPINs, and weighted gene co-expression network analysis (WGCNA). Considering the overall survival (OS) analysis and the evaluation of expression levels alongside statistical parameters of the PPIN cluster nodes, it is plausible to suggest that THY1, CDH17, TGIF1, and AEBP1, categorized as nonhub nodes, along with ITGA5, COL1A1, FN1, and MMP2, identified as hub nodes, possess characteristics that render them applicable as biomarkers for the GC. Additionally, insulin-like growth factor (IGF)-binding protein-2 (IGFBP2), classified as a nonhub node, demonstrates a significant negative correlation with both groups within the same cluster. This observation underscores the conflicting findings regarding IGFBP2 in various cancer studies and enhances the potential of this gene to serve as a biomarker.

Conclusion: The findings of the current study not only identified the hubs and nonhubs that may serve as potential biomarkers for GC but also revealed a PPIN cluster that includes both hubs and nonhubs in conjunction with IGFBP2, thereby enhancing the understanding of the complex behavior associated with IGFBP2.

1 | Background

The modern world of biology and concerned researchers believe that prolonged delays in the pathological diagnosis and prognosis of cancers result in missed treatment opportunities and impose substantial financial burdens associated with care.

Exploring novel biomarkers with prognostic and diagnostic value provides timely medical intervention for survival of patients with gastric cancer (GC) [1–3].

Currently, GC is recognized as the fourth leading cause of cancer-related fatalities globally, particularly affecting populations in

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Cancer Reports* published by Wiley Periodicals LLC.

East Asia. A contributing factor to the high mortality rate is that patients with GC often seek medical intervention only when the disease has progressed to an advanced stage with the tumor having metastasized to other parts of the body. Consequently, developing early and noninvasive diagnostic and prognostic biomarkers that can effectively monitor and predict cancer progression at various stages is paramount [4–6].

In numerous studies focused on cancer, RNA-Seq has been extensively employed as a standalone technique for DGE analysis. However, a large number of differentially expressed genes (DEGs) achieve statistical significance, prompting the use of both PPIN and WGCNA as methodologies to filter these genes and associate them with disease phenotypes, cellular structures, and biological pathways. The WGCNA facilitates the identification of co-expressed genes, while PPIN aids in clustering proteins linked to each gene, thereby contributing to identifying genes closely associated with the disease [7–10].

The PPIN serves as a collective representation and association of proteins arising from biochemical processes within the cell. In this kind of network, the nodes symbolize proteins, while the edges denote the interactions between the proteins. The topological parameters of the network are established based on the connections and interactions among the nodes [11, 12].

The primary topological parameters of PPIN, such as degree (K), Betweenness Centrality (BC), and Closeness Centrality (CC), play a crucial role in elucidating the function of proteins (network nodes) within disease-related signaling pathways. In this context, the degree of a node is quantified by the number of edges associated with it. The BC serves as a centrality measure within a graph, focusing on the shortest paths (SPs) and indicating the frequency with which a node appears on these paths between pairs of nodes [13]. This metric helps identify nodes that act as intermediaries connecting different sections of a graph. CC, on the other hand, provides an estimate of the speed at which information can flow from a given node to other nodes. Consequently, proteins exhibiting high values of K and BC are considered to be significantly interconnected with other proteins, forming clusters within the PPIN, and are commonly referred to as hub nodes. Conversely, nodes characterized by low K and BC, situated among hub nodes, are identified as nonhubs [14].

In numerous studies, it has been demonstrated that clustered hub nodes significantly influence the majority of activities essential for cellular survival and the progression of diseases, making them suitable candidates for evaluation through various laboratory and clinical phases before being recognized as biomarkers [15, 16].

Nonhubs are considered to have less impact on the development of biological processes, leading to the lack of the capacity to serve as clinical biomarkers. Despite the limited number of bioinformatics studies exploring the potential of nonhubs as biomarkers, this topic has garnered interest due to their significant impact on pharmacological and biological characteristics. According to these investigations, nonhubs in biological systems influence biological processes primarily through interacting with hub nodes [17–19].

Clearly, a full explanation of the interaction between PPIN nodes (hubs and nonhubs) plays a crucial role in identifying approaches that enhance exploratory processes for evaluating various biomarkers. In contrast to earlier research, this study uniquely investigates the potential of utilizing hubs and nonhubs (The less prominent nodes within the PPI network) as biomarkers in GC [20–22].

2 | Material and Methods

2.1 | Datasets Resources and Integration Pipeline

The workflow of the current study includes 15 steps presented in the pipeline (Figure 1). All the Fastq paired-end of 48 mRNA sequencing profiles from GC were retrieved from the NCBI-SRA using SRA-Toolkit (<https://github.com/ncbi/sra-tools/wiki/02.-Installing-SRA-Toolkit>) (Table 1). In order to increase the accuracy of the integration results, only paired-end GC datasets containing normal and tumor samples were selected.

2.2 | RNA-Seq Data Analysis, Batch Effect Removal, and DGE Analysis

After the quality check of Fastq paired-end reads, more than 30% of the reads by quality below Q20 and adapters (first 12bp of the reads) were trimmed utilizing the final version of FastQC [23] and Trimmomatic [24]. Subsequently, transcript-level quantifications for each GC-RNA-Seq dataset were accurately quantified from cleaned reads according to the annotation file (Gencode.v99) using Salmon [25].

After obtaining all quant files related to each dataset, the tximport, GenomicFeatures, and Reader packages were employed to import the transcript level to the gene-level files [26]. SVA-seq (ComBat function) and DESeq2 packages were run on merged normalized values to detect and remove batch effects, low-expression genes, and other undesirable sources of variation in merged datasets [27, 28]. To characterize significant DEGs and normalized log expression values, Limma with an adjusted p value threshold (Benjamini–Hochberg) of 0.05 and a fold-change limit of $|\log_2FC| \geq 1$ were used [29].

2.3 | WGCN Analysis and PPIN Construction

Following the DGE analysis, WGCN analysis was applied to DEGs to filter and screen the co-expression genes related to GC. All interconnected nodes and their interactions were derived through the igraph package using a specific formula and set of criteria [30]. Subsequently, the edges of the WGC network were then converted into weights assigned to the nodes, resulting in a giant component of the weighted network. Following this, the weighted adjacency matrix of the giant component was converted to a symmetric matrix, which was then used to create a new adjacency matrix employing the topological overlapping measure (TOM) function available in the WGCNA R package [31].

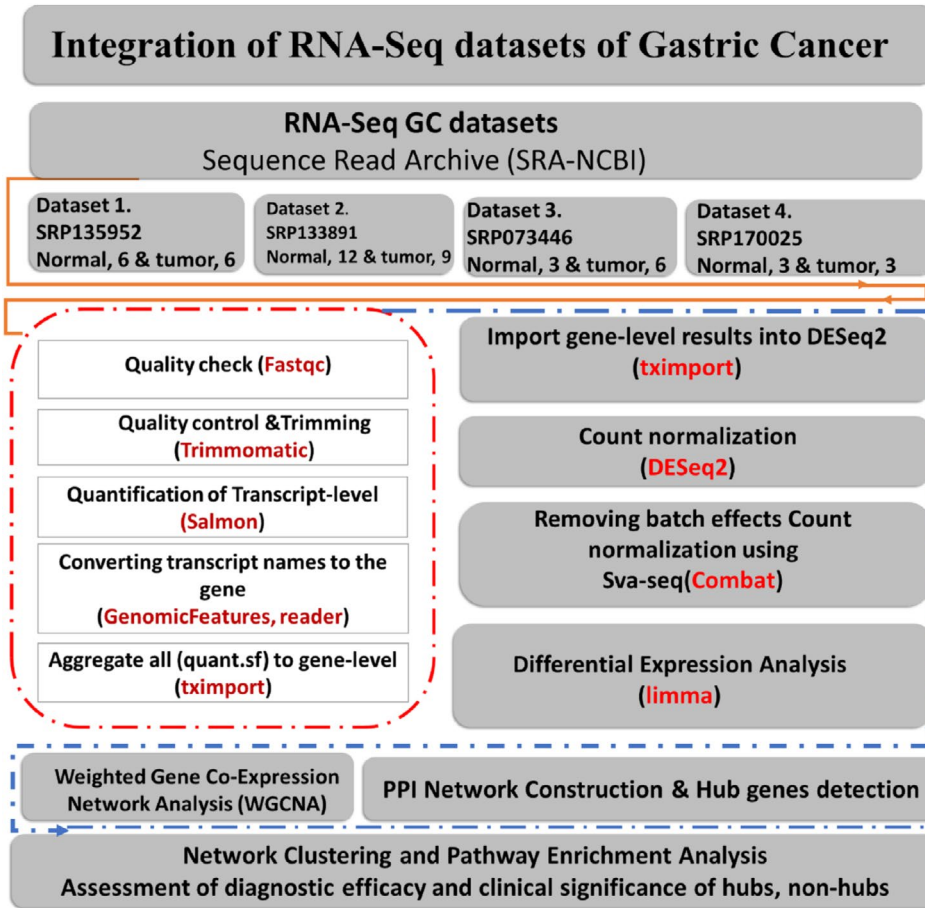


FIGURE 1 | The pipeline of integration of gastric cancer RNA-Seq datasets. The orange solid lines, orange dashed lines, and blue solid lines represent the various steps of the study, specifically the collection of datasets, the analysis of Fastq datasets, and the interpretation of results using Cytoscape and the Enrichment database, respectively.

TABLE 1 | The RNA-seq datasets of gastric cancer employed in the present study.

| R | SRA study | Library layout | Experimental design (tumor/normal) | Organization name |
|---|-----------|----------------|------------------------------------|---|
| 1 | SRP135952 | Paired | 12 (6/6) | East China Normal University |
| 2 | SRP133891 | Paired | 21 (12/9) | Hallym University Medical Center |
| 3 | SRP170025 | Paired | 6 (3/3) | Fudan University Shanghai Cancer Center |
| 4 | SRP073446 | Paired | 9 (3/6) | National Chung Cheng University |

A comprehensive matrix representing all SPs within the weighted network was developed for every pair of nodes, implementing Dijkstra's algorithm [32]. This resulting matrix (D_j) served as the basis for determining the distance value associated with each node in the network. The average SPs from all non-core genes to node j were computed to derive D_j for each node. Then, the average SPs were subtracted from the core genes to reach node j . This value was divided by the average SPs to node j in the entire network. This rating system implies how close each node is to the core nodes [33, 34].

$$D_j = \frac{\frac{\sum_{i \notin c} SP_{ij}}{NC} - \frac{\sum_{i \in c} SP_{ij}}{C}}{\frac{\sum_{i} SP_{ij}}{NC + C}}$$

The STRING database [35] was recruited to build a PPIN based on five sources of evidence. The extracted interactions were imported into the R program to visualize the PPIN.

2.4 | Investigation of the Correlation Between the Expression Level of Hubs, Nonhubs, and IGFBP2

To evaluate the clinical significance of the identified PPIN nodes (both hubs and nonhubs), in conjunction with IGFBP2, Pearson's correlation and the statistical relationship were calculated pairwise among these groups based on their expression levels, employing the corTest package.

2.5 | Comparative Examination of IGFBP2 Expression Levels, Mutation Frequencies, and Patient Survival Outcomes Across Different Types of Cancer

The comparison of the expression levels of IGFBP2 across different types of cancer, highlighting instances of both downregulation and upregulation of this gene, was analyzed using TCGA data through the UALCAN database.

An analysis was conducted to compare patient survival rates at both up and down expression levels of IGFBP2, utilizing various survival metrics such as disease-specific survival (DSS), progression-free survival (PFS), disease-free survival (DFS), and overall survival (OS). This analysis was based on the TCGA dataset obtained from CBioPortal and employed the Survminer R package along with Kaplan–Meier (kME) methods. Additionally, genetic alterations of IGFBP2 were assessed across different types of cancer via Oncoprint [36].

2.6 | Functional and Pathway Enrichment Analysis

Functional pathway enrichment analysis for all the significant PPIN clustered nodes was conducted using g:Profiler (<http://biit.cs.ut.ee/gprofiler/>) and EnrichR (<https://cran.r-project.org/package=enrichR>).

3 | Results

3.1 | RNA-Seq Data Normalization, Batch Effect Removal, and DGE Analysis

Following the integration of the four datasets, we performed clustering on the raw read counts (Figure 2B) as well as on the counts' postbatch effect removal (generated by ComBat) (Figure 2D) by principal component analysis (PCA) separately. The box plot illustrates the distribution of counts across samples for both unnormalized counts (Figure 2A) and batch-corrected counts (Figure 2C). When normalization was performed separately for each batch followed by ComBat application, the optimal outcomes were achieved regarding batch similarity and the creation of a distinct normalized expression matrix suitable for DGE analysis.

After the normalization process and batch effect removal, a unique and normalized matrix with all samples (48 samples) was created and prepared to screen the DEGs between the GC and the control samples through the limma package. A total of 1000 DEGs ($\text{padj} < 0.05$) were obtained, of which 439 genes ($\text{padj} < 0.05$, $\log_2\text{FC} \geq 1$) were significantly upregulated (uDEGs), and 133 genes ($\text{padj} < 0.05$, $\log_2\text{FC} \leq -1$) were significantly downregulated (dDEGs) (Table S1).

3.2 | Discovery of Hubs and Nonhubs From the WGC and Clustered PPI Networks

The SP of the WGCN analysis-based scoring system explained in the material and methods section was used to detect nodes

with shorter distances from the core genes (nodes with a positive D -score), thereby extracting a smaller network from the huge component network. This network consists of 83 nodes containing 19 core nodes and 64 neighbors. The giant component of this network (core network) includes 16 seed nodes and 61 neighboring nodes. The diameter of the WGCN was six genes containing LIFR, LIF, PDGFR, COL1A2, FN1, CEACAM1, and RAB37. The network transitivity was about 50%, the edge density was about 14%, and the mean distance was 2.54 (Table S2).

In general, 83 co-expressed DEGs ($\text{padj} < 0.05$) extracted from WGCN were selected to construct the PPIN. The PPIN nodes with degree greater than 13 and BC greater than 63.94 were considered hubs, including COL1A1, COL1A2, COL3A1, COL5A1, COL4A1, FN1, COL5A2, MMP2, ITGA2, SPARC, THBS1, and ITGA5. The direct neighbors of the hubs with a degree less than three and a BC less than nine, as well as indirect neighbors with two edge distances of the hubs with a degree less than three and a BC equal to zero were considered nonhubs, including NID2, MXRA5, FKBP10, SULF1, THY1, AEBP1, TGM2, CDH17, CDH6, TGIF1, IGFBP2, and PAPP A (Table 2). The clustering of PPIN was performed via a greedy optimization algorithm, resulting in four clusters (Figure 3).

3.3 | Assessment of Diagnostic Efficacy and Clinical Significance of Hubs and Nonhubs

In the current study, the nonhub nodes within each cluster exhibited significant upregulation, and since they have been proposed as potential biomarkers in numerous investigations, we speculated that these nodes may be associated with the hubs and could play a vital role in the progression of GC. To substantiate this hypothesis, we performed Pearson's correlation coefficient (r) test, which confirmed that all nonhubs correlate positively with the hubs. However, a statistically nonsignificant and negative correlation was observed between the expression levels of IGFBP2 and all hubs and nonhubs (Table S3). The results of this correlation analysis were derived consistent with our initial findings (Figure 4).

3.4 | Examining the Expression Levels, the Mutation Percentage, and the Survival Outcomes Associated With IGFBP2 Across Various Cancer Types

Based on correlation analysis, the IGFBP2 gene, identified as a nonhub, exhibited significantly divergent trends compared to both hub and nonhub genes. Consequently, an independent assessment of this gene was conducted, focusing on comparative expression levels across various cancer types, mutation rate analysis, and multiple survival metrics, including OS, DSS, PFS, and DFS, to gain a deeper understanding of its characteristics.

As illustrated in Figure 5A,B, The expression level of IGFBP2 in numerous tumor samples is significantly downregulated, aligning with the findings of our study analysis (Table 2). This evidence further underscores the overall importance of our results.

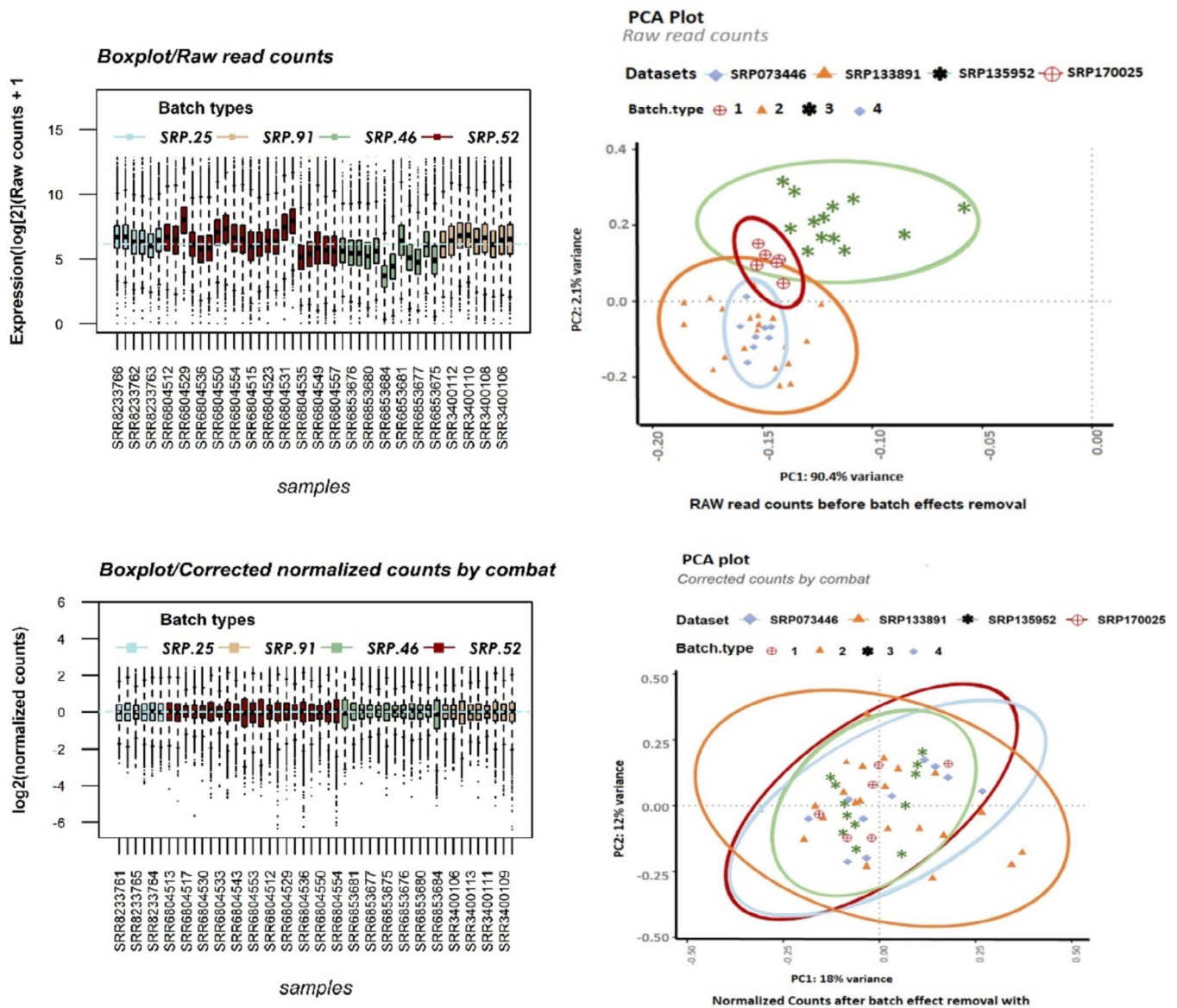


FIGURE 2 | PCA plot and boxplot of the qualified samples based on Salmon's estimate before normalization and batch effect removal. Distribution of unnormalized counts (\log_2 (counts + 1)) per 48 samples (A) Normalized counts by combat (C) are displayed as boxplots and colored by dataset (batches). PCA plots show the first two components (PC1, PC2) of the principal component analysis, with percentages of variance associated with each axis. Colors identify datasets, and shapes identify batches. It can be seen from the PCA plot (B) that the first principal component shows a strong separation between the four batches. After applying ComBat to the counts (D), the differences between the four batches were largely reduced, and the percentage variation along PC1 and PC2 accounts for 17.9% and 12%, respectively, which is very small, and all symbol schemes (batches) remain the same place.

The percentage of mutations in GC samples, including amplification, severe depletion, mRNA upregulation, and various mutations, was assessed using OncoPrint to ascertain whether the conflicting trend of this gene is associated with its mutation rate. In this context, the mutation rate of this gene was initially evaluated across all cancers exhibiting downregulation of this gene, including BLCA (1.6%) and CESC (2.7%), ESCA (1.3%), KICH (0.9%), KIRC (1.2%), SKCM (1%), STAD (1.6%), and PAAD (1.1%) (Figure 5C). Then, the same mutational analysis was performed for cancer types exhibiting upregulation of IGF2BP2, including BRCA (0.5%), GBM (0.3%), HNSC (1.6%), LUAD (0.6%), PRAD (1.1%), THCA (0%), and

UCEC (1.6%) (Figure 5D). In both cases, it was unexpectedly discovered that IGF2BP2 displayed a mutation rate that was nearly identical with a minimal percentage of mutation in different types of cancers.

The findings illustrated in Figure 5G,H,J indicate that the outcome of the OS analysis, considering both the downregulation and upregulation of IGF2BP2, reveals that downregulation of this gene is associated with more prolonged survival in GC patients. Similarly, as indicated in Figure 5E,F,I, the prognostic impact of IGF2BP2 on DSS, DFS, and PFS shows that the downregulation of this gene is associated with prolonged survival.

TABLE 2 | Statistical information of significant nodes (hubs and nonhubs) represented in PPIN based on *K*, BC, CC, and distance.

| | Gene symbol | Features | <i>K</i> | BC | CC | Distance |
|-----------|-------------|----------|----------|--------|--------|----------|
| Cluster A | COL1A2* | UP | 37 | 450.23 | 0.5984 | 1.65 |
| | COL4A1* | UP | 27 | 73.23 | 0.5205 | 1.9 |
| | COL5A2* | UP | 23 | 65.7 | 0.481 | 2.05 |
| | COL3A1* | UP | 32 | 213.01 | 0.539 | 1.83 |
| | COL5A1* | UP | 28 | 100.81 | 0.5241 | 1.88 |
| | COL1A1* | UP | 39 | 436.86 | 0.608 | 1.62 |
| | MMP2* | UP | 20 | 91.19 | 0.5067 | 1.95 |
| | FKBP10 | UP | 2 | 0 | 0.3838 | 2.57 |
| | THY1 | Up | 1 | 0 | 0.3762 | 2.62 |
| | MXRA5 | UP | 3 | 0 | 0.3918 | 2.52 |
| | SPARC* | UP | 18 | 70.08 | 0.4935 | 2 |
| Cluster B | THBS1* | UP | 18 | 65.72 | 0.4935 | 2 |
| | TGIF1 | UP | 1 | 0 | 0.2923 | 3.38 |
| | TGFBI | UP | 4 | 0 | 0.4368 | 2.26 |
| | IGFBP2 | Down | 1 | 0 | 0.3115 | 3.17 |
| | PAPPA | UP | 1 | 0 | 0.3115 | 3.17 |
| Cluster C | FN1* | UP | 24 | 496.28 | 0.5507 | 1.79 |
| | ITGA5* | UP | 18 | 269.01 | 0.5135 | 1.92 |
| | TGM2 | UP | 2 | 0.66 | 0.3878 | 2.55 |
| | NID2 | Up | 3 | 1.39 | 0.3938 | 2.51 |
| | CDH17 | Up | 2 | 0 | 0.2901 | 3.4 |
| Cluster D | CDH6 | Up | 2 | 0 | 0.2901 | 3.4 |
| | SULF1 | UP | 2 | 0 | 0.3838 | 2.57 |
| | AEBP1 | Up | 1 | 0 | 0.3762 | 2.62 |
| | TGM2 | Up | 2 | 0.66 | 0.3878 | 2.55 |

Note: Hub nodes exhibiting the highest values of *K*, BC, and CC, along with the lowest distance, are indicated with asterisks.

The comparative analysis of gene expression alongside survival outcomes may validate the contradictory function of IGFBP2 in GC and other malignancies.

3.5 | Functional and Pathway Enrichment Analysis

A noteworthy finding from the correlation analysis was the negative correlation of IGFBP2 with both hubs and nonhubs. To gain a deeper understanding of the relationship between these two groups and IGFBP2, a pathway analysis was conducted to identify the biological processes and signaling pathways in which IGFBP2 participates, apart from its association with hubs and nonhubs. The outcomes of the signaling pathway analysis indicated that multiple signaling pathways were significantly enriched, demonstrating the co-presence of these groups in relation to IGFBP2, with correlations that were either positive or negative, as detailed in Table S3.

According to the results of both correlation and pathway analysis, as shown in Table 3, The main hubs (MMP2, COL1A1, FN1) and nonhubs (PAPPA, TGM2, TGFBI, IGF1) jointly cooperate in vital biological pathways including regulation of transport and uptake of insulin-like growth factor (IGF), Oncostatin M and TNF-alpha effects on cytokine activity, cell motility, and apoptosis. Meanwhile, IGFBP2 is associated with both nonhubs and hubs, and it is found in the majority of pathways.

4 | Discussion

In most biomarker-based studies, hubs extracted from the PPIN are introduced as biomarkers and used in clinical trials following a series of statistically validated methodologies [1, 37–40]. Surprisingly, the present investigation aimed to identify new biomarkers for GC by retrieving and clustering both hubs and nonhubs from the PPIN using stringent statistical techniques. The analysis revealed that specific subcluster

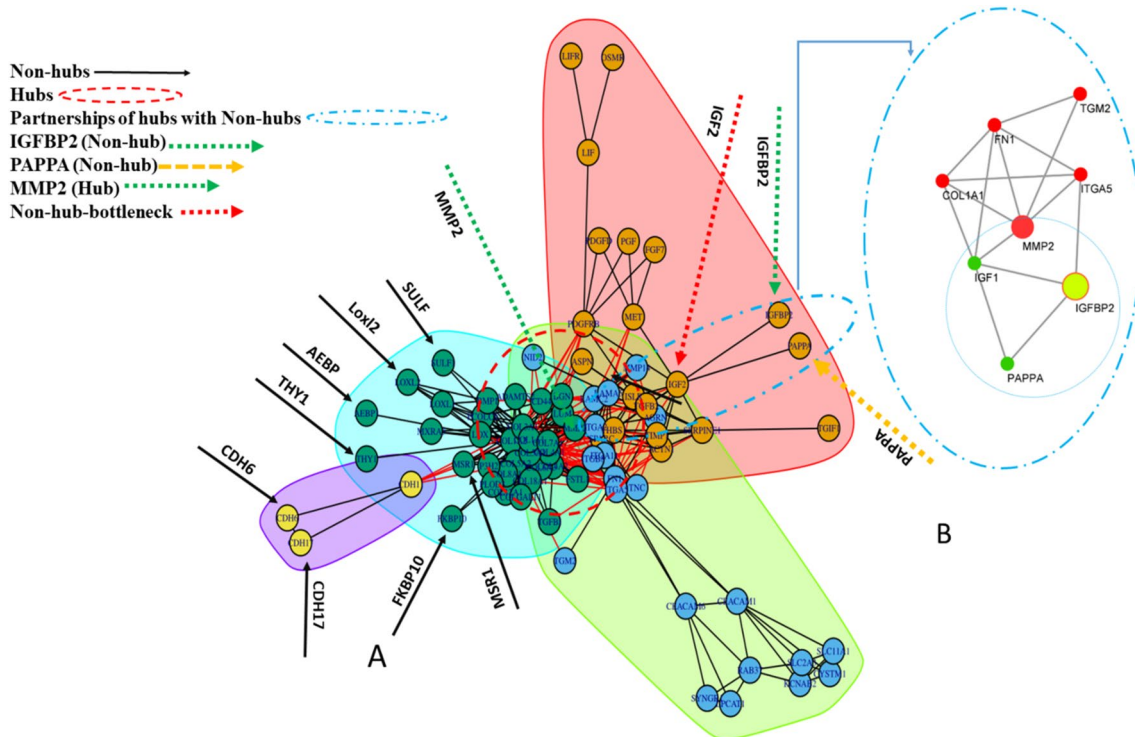


FIGURE 3 | The clustering results of the PPIN. Four clusters are shown in the scheme, including Cluster A (dark green nodes), Cluster B (brown nodes), Cluster C (blue nodes), and Cluster D (yellow nodes). The nonhubs marked by the black arrows and the region by circular dashed red lines were hub nodes. The relationship between IGFBP2 and nonhubs (PAPPA, IGF1) and hubs (MMP2, Col1A1, FN1, TGM2, ITGA5) is shown as circular dashed light blue lines. Cluster A is the largest cluster with 17 genes and hosts the collagen family genes, a key component of the core network.

hubs and nonhubs within each cluster exhibit interconnections regarding their prognostic and diagnostic capabilities in GC.

Given the implementation of previous scientific reports, the significant hub nodes within the PPIN, such as COL1A1, COL1A2, COL3A1, COL5A1, COL4A1, COL5A2, FN1, MMP2, ITGA2, SPARC, THBS1, and ITGA5, play a crucial role in promoting cell proliferation, resistance to chemotherapy, invasion, and migration of GC cells, as well as facilitating distant metastasis. These nodes warrant further statistical and experimental validation in the context of GC research [41–46].

However, considering the significant validation tests related to nonhubs (nodes with lower K , BC), including NID2, MXRA5, FKBP10, SULF1, THY1, AEBP1, TGM2, CDH17, CDH6, TGIF1, IGFBP2, and PAPPA, we hypothesize that these nodes can be associated with hubs and have the potential to be GC biomarkers.

An inconsistent trend of IGFBP2 was observed in both hubs and nonhubs. Therefore, various statistical validations were conducted to assess the specific potential of these two groups in relation to IGFBP2 and to determine whether the aforementioned elements are interconnected and possibly implicated in the GC signaling pathways.

Previous studies showed both types of nodes as validated biomarkers exhibiting upregulation in tumor conditions [47–49]. Overall, several nonhub proteins, including IGFBP2, PAPPA, CDH17, TGIF1, THY1, NID2, TGM2, and AEBP1, have been

identified as oncogenic in specific stages of cancer [49–51]. Nevertheless, several contradictory studies (both in vitro and in vivo) suggested that these proteins may function as tumor suppressors [52–54]. Furthermore, several observations have illustrated the interaction between the identified nonhubs and hubs within our research [55–57].

Pearson's correlation analysis was conducted to validate the significance of relationships between the expression level of hubs and nonhubs. While, all nonhubs showed a strong, significant positive correlation with the hubs, a significant negative correlation was observed between the expression level of IGFBP2 and both hubs and nonhubs (Table S3). Pearson's correlation analysis conducted on TCGA data involving 293 patients with GC corroborated the findings of our study; however, the difference in IGFBP2 levels between the two groups of patients with GC was statistically insignificant. Given that IGFBP2 has not demonstrated a consistent trend in prior research or our current study, it was subjected to further examination through comparative analyses of various validation studies. The mutation frequency of this gene across multiple cancer types (1512 samples) revealed minimal variation in downregulated states (BLCA; 1.6%; KIRC; 1.2%; KICH; 0.9%; ESCA; 1.3%) and upregulated states (BRCA; 0.5%; GBM; 0.3%; PRAD; 1.1%; THCA; 0; LUAD; 0.6%). Consequently, it can be suggested that the fluctuating behavior of this gene in different cancers is not influenced by its mutation rate. As illustrated in Figure 5G,J, the low expression of this gene across assorted cancers, whether upregulated or downregulated, correlates with significantly longer survival. The alterations in the expression levels of this gene indicate an increase in various cancers such as BRCA, GBM,

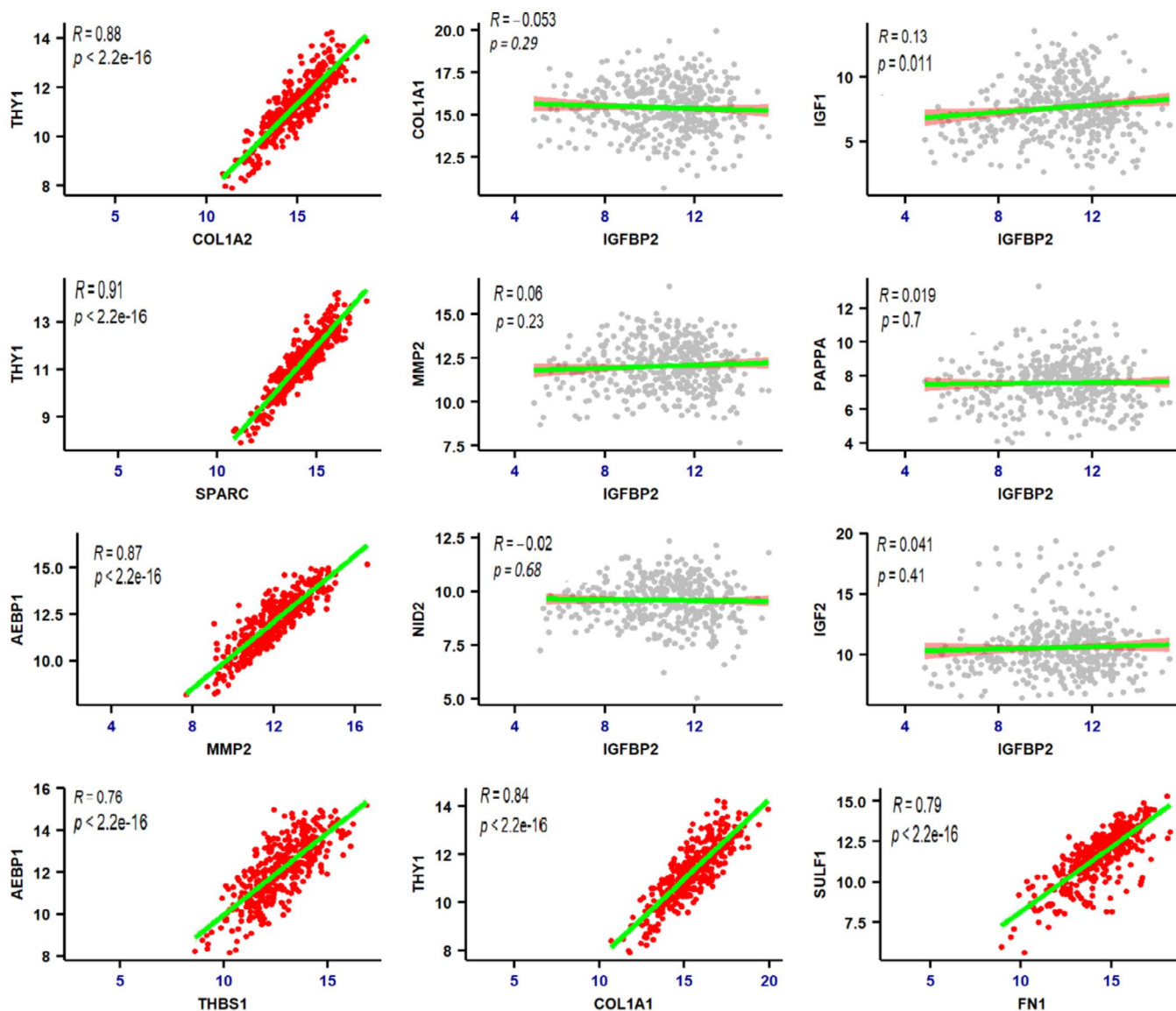


FIGURE 4 | Pearson's correlation between expression levels based on the GC-TCGA dataset. The correlation coefficients (Pearson) between the expression level of hubs, nonhubs, and IGFBP2 in GC samples (Nature TCGA, Pan-Cancer Atlas, 293 patients). As shown, there is a statistically strong positive correlation between the expression level of hubs (MMP2, FN1, COL1A1, SPARC) and nonhubs (THY1, SULF1, AEBP1, NID2) (all plots with regression line colored green and dot (samples) colored red). The distribution of Pearson's correlation coefficients between the expression levels of hubs and nonhubs with IGFBP2 is shown as plots with gray dot plots. It is clear that the correlation between IGFBP2 and IGF1 was statistically positive, and contrary to the correlation analysis in our study, there is no significant relationship between hubs and nonhubs with IGFBP2.

HNSC, LUAD, and UCEC, while a decrease is observed in other types, including BLCA, KIRC, KICH, and ESCA (Figure 5A,B). IGFBP2 is a member of a family comprising six insulin-like growth factor binding proteins (IGFBPs). Current research indicates a varied expression pattern of this gene across different cancer types. For instance, studies have shown an upregulation of IGFBP2 in breast and cervical cancers [58–60], while downregulation has been noted in glioma, bladder cancer, and rhabdomyosarcoma (RMS) [61–63]. Consequently, it remains to be determined whether the overexpression or downregulation of this gene should be regarded as a potential biomarker. It raises the question of whether IGFBP2 functions as an oncogene or a tumor suppressor in cases of upregulation versus downregulation.

Considering the contradictory results, the multiple validation tests for IGFBP2, and the lack of a consistent trend, it seems that the best way to determine the exact role of this gene is to evaluate the interaction of this gene with hubs and nonhubs in the PPIN.

A selection of clustered genes suitable for investigation includes nonhub genes, such as PAPP and IGF1, and hub genes, including MMP2, COL1A1, and FN1. Upon examining the interrelations among these genes, it became evident that they play a significant role in the regulation of IGF, as well as in the transport and uptake facilitated by IGFBPs and the regulation of IGFBP. Moreover, the activity of IGF activity is regulated via the IGFBPs pathway.

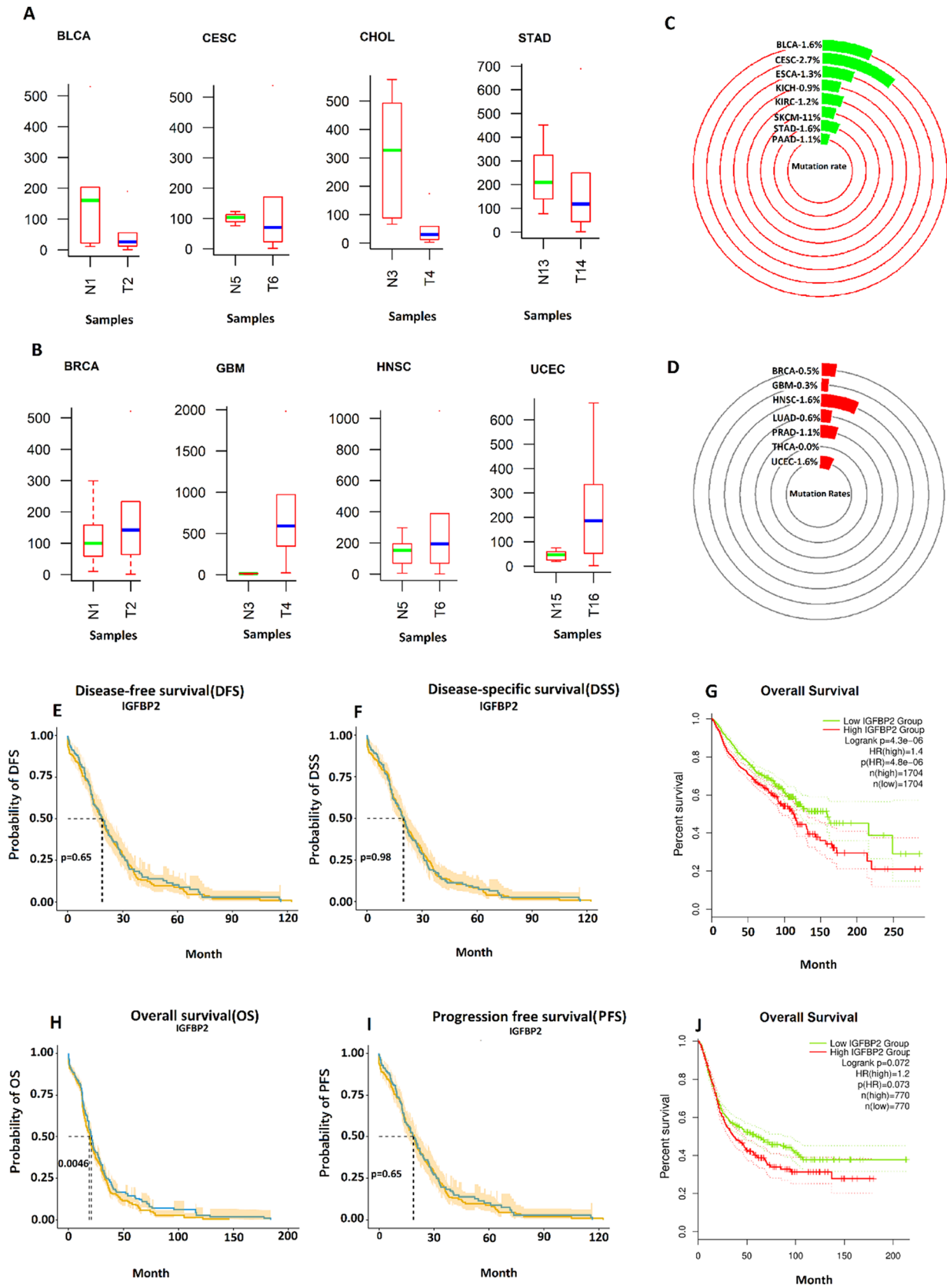


FIGURE 5 | Legend on next page.

Therefore, to clarify the function of this gene alongside its neighboring genes, a closer look at the IGF pathway should be taken. In addition to IGF-I, IGF-II, and IGFBP2, their receptors

(IGF-IRs) are components of the IGF pathway. Less is known about this pathway to date; however, the presence, interaction, and potency of IGFs with this pathway and both PAPPAs and

TABLE 3 | Pathway analysis based on BioPlanet pathway maps.

| Series | Description | Adj-p | Hubs | Nonhubs |
|--------|---|-------------|---|----------------------|
| 1 | Beta-1 integrin cell surface interactions | 2.14E-16 | COL1A1; COL3A1; COL5A1; COL4A1; COL5A2; FN1; ITGA5; THBS1; LAMA5 | TGM2 |
| 2 | Beta-3 integrin cell surface interactions | 4.52E-09 | COL1A1; COL4A1; FN1; THBS1 | THY1, TGFBI |
| 3 | NCAM1 interactions | 2.943e-9 | COL1A1; COL3A1; COL1A2 | AGRN |
| 4 | Syndecan 1 pathway | 3.41E-09 | COL1A1; COL3A1; COL1A2 | |
| 5 | Extracellular matrix organization | 9.41E-14 | COL1A2; COL4A1; COL5A2; COL3A1; COL5A1; COL1A1; MMP2; SPARC; THBS1; FN1; ITGA5; MMP14 | NID2, PCOLCE |
| 6 | TGF-beta regulation of extracellular matrix | 0.0234\3885 | COL1A1; FN1; TGFBI | NID2, AEBP1 |
| 7 | Alpha-9 beta-1 integrin pathway | 0.020489945 | FN1 | TGM2 |
| 8 | Interleukin-4 signaling pathway | 2.89E-04 | COL1A1; COL1A2 | THY1 |
| 9 | Integrin beta-2 pathway | 0.025135122 | TGFBI | THY1 |
| 10 | Alpha-M beta-2 integrin signaling | 0.046128092 | MMP2 | THY1 |
| 11 | Signaling by PDGF | 0.042372297 | PDGFRB; COL1A1; COL3A1; COL1A2; COL4A2; COL5A1; COL4A1; PDGFD; COL5A2; COL6A3; THBS1 | THY1 |
| 12 | Oncostatin M | 0.002014676 | MMP2, COL1A1; FN1 | IGFBP2, TGM2 |
| 13 | Regulation of insulin-like growth factor (IGF) transport and uptake by insulin-like growth factor binding proteins (IGFBPs) | 0.020360665 | MMP2, FN1 | IGFBP2, PAPP A |
| 14 | Diabetes pathways | 0.025436156 | MMP2 | IGFBP2, PAPP A |
| 15 | Insulin-like growth factor (IGF) activity regulation by insulin-like growth factor binding proteins (IGFBPs) | 9.64E-05 | MMP2 | IGFBP2, PAPP A, IGF1 |

Note: Significantly enriched pathway terms for hubs, nonhubs, and IGFBP2 are shown. The results of the presence of two nodes with IGFBP2 in the biological signaling pathways are consistent with the results of cluster analysis (Figure 3). As shown by the cluster analysis, IGFBP2 is clustered with PAPP A, MMP2, IGF1, TGM2, FN1, and COL1A1.

MMP2 could shed light on the complex role of this pathway in cancers [64–67]. PAPP A (nonhub) and MMP2 (hub) have been introduced as cancer biomarkers.

Since high levels of IGFBPs can inhibit IGF-stimulated functions by blocking binding to its receptors, the proteolysis of IGFBP2 is a pivotal mechanism in regulating and enhancing the carcinogenic effects of IGF-signal transmission. A number

of studies reported that PAPP A and MMP2 act as IGFBP2 proteinases, leading to cleavage of the IGF1/IGFBP2 complex. The breakdown of this system leads to the release of IGF1 from the complex, and this process (accumulation of IGF1) can facilitate cell migration toward the tumor [68, 69].

Given the role of IGFBP2 in the IGF system, it can be assumed that upregulation of this gene with blockade of the IGF pathway

FIGURE 5 | Prognostic value of IGFBP2: IGFBP2 mRNA expression was up- and downregulated in two states (normal samples [N], tumor samples [T]) in different cancer types (A, B). The y-axis in (A, B) represents the expression level of IGFBP2 in different cancer types and the median of boxplots colored by solid red and blue lines. The genetic alterations of IGFBP2 in different cancer types (C, D) and overall survival (OS) with 95% confidence intervals (CIs) (dotted lines) of patients with GC (G, J) were shown. The prognostic impact of IGFBP2 on OS (H), DSS (F), PFS (I), and DFS (E) was evaluated. The y-axis in (E, F) and (H, I) plots represents probabilities of DFS, DSS, OS, and PFS, respectively. In (E, F) and (H, I) plots, the green solid line represents the patient group with downregulated expression of IGFBP2, and the solid red line represents the patient group with upregulated expression of IGFBP2. OS analysis of GC patients with downregulation of IGFBP2 shows a significantly prolonged overall survival ($p=0.0046$) in patients with GC.

reduces the effect of tumorigenesis. On the other hand, the up-regulation of PAPP A and MMP2 contributes to an increase in the accumulation of IGFBP2 in the cell through proteolysis of the IGF1/IGFBP2 complex, which in turn is involved in the stimulation of tumorigenesis by the IGF pathway. However, to justify the downregulation of this gene in our study and the TCG data, it can be proposed that the IGF pathway for the survival of its topicality also induces the expression of the hypoxia-inducible factor 1 (HIF-1), which can significantly reduce the expression of IGFBP2 [70–74].

From all the above interpretations, it can be concluded that the activity of IGFBP2 is dependent on the IGF pathway. Hence, up-regulating the gene by blocking the IGF complex favors tumor suppression and is also a sign of activity accelerating proteolysis of two proteinase genes (PAPP A and MMP2). On the other hand, the downregulation of this gene may also be a biomarker of cancer recurrence since, in the absence of this gene, the IGF pathway will play a crucial role in tumorigenesis.

Given the statistically significant validation of nonhubs and hubs, it can be speculated that nonhubs and hubs have the potential to become GC biomarkers for the following reasons:

In light of the compelling evidence regarding IGFBP2 from both previous and current studies, we propose that IGFBP2, previously regarded as a dark horse, is influenced by the expression changes of nonhub genes (PAPP A, IGF1) and hub genes (COL1A1, FN1, MMP2) in GC. Significant downregulation and upregulation of this nonhub gene play a crucial role in tumorigenesis and tumor suppression. Alongside its neighboring genes within biological networks and signaling pathways, IGFBP2 plays a vital role in establishing a comprehensive signature throughout various stages of GC [75–79].

5 | Conclusion

In summary, when presenting all the cumulative results of the study and considering the association of the IGFBP2 gene with hub and nonhub, it is essential to recognize that the designation of a gene as a biomarker is significantly influenced by its role within biological systems and gene networks. Specifically, nonhub genes within the PPIN exhibit lower *K* and BC values. However, it can be assumed that these elements appear to exert their effect on disease progression or suppression across the hubs. Therefore, additional computational, predictive modeling, and clinical research are required to determine which group, either hubs or nonhubs, exerts the most significant influence on the others and to assess the potential of nonhubs to serve as biomarkers. Undoubtedly, the advent of big data and machine learning present opportunities for understanding the interactions between hubs and nonhubs in tracking GC states, potentially leading to innovative solutions. Given that all selected datasets pertain to East Asia, we intend to investigate the relationship of this gene with its neighboring genes in the PPIN across various global populations. Furthermore, incorporating diverse variables such as age, gender, and climate into the DGE analysis enhances the precision of the research.

Author Contributions

Akram Siavoshi: conceptualization, data curation, formal analysis, methodology, investigation, visualization, writing – original draft, writing – review and editing. **Mehran Piran:** data curation, formal analysis, software, methodology, investigation, validation, visualization. **Ali Sharifi-Zarchi:** conceptualization, data curation, methodology, investigation, formal analysis, writing – review and editing. **Fatemeh Ataellahi:** writing – review and editing, methodology.

Acknowledgments

We would like to express our heartfelt appreciation to Dr. Ali Sharifi-Zarchi for his unwavering support during the course of this research. We also extend our gratitude to all the participants who contributed to the successful execution of this study.

Ethics Statement

The authors have nothing to report.

Consent

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that supports the findings of this study are available in the [Supporting Information](#) of this article.

References

1. N. McGarvey, M. Gitlin, E. E. Fadli, and K. Chung, “Increased Healthcare Costs by Later Stage Cancer Diagnosis,” *BMC Health Services Research* 22 (2022): 1155, <https://doi.org/10.1186/s12913-022-08457-6>.
2. K. T. Han, W. Kim, and S. Kim, “Does Delaying Time in Cancer Treatment Affect Mortality? A Retrospective Cohort Study of Korean Lung and Gastric Cancer Patients,” *International Journal of Environmental Research and Public Health* 18, no. 7 (2021): 3462, <https://doi.org/10.3390/ijerph18073462>.
3. T. P. Hanna, W. D. King, S. Thibodeau, et al., “Mortality due to Cancer Treatment Delay: Systematic Review and Meta-Analysis,” *British Medical Journal* 4, no. 371 (2020): m4087, <https://doi.org/10.1136/bmj.m4087>.
4. E. Morgan, M. Arnold, M. C. Camargo, et al., “The Current and Future Incidence and Mortality of Gastric Cancer in 185 Countries, 2020–40: A Population-Based Modelling Study,” *EclinicalMedicine* 21, no. 47 (2022): 101404, <https://doi.org/10.1016/j.eclinm.2022.101404>.
5. T. Matsuoka and M. Yashiro, “Biomarkers of Gastric Cancer: Current Topics and Future Perspective,” *World Journal of Gastroenterology* 24, no. 26 (2018): 2818–2832, <https://doi.org/10.3748/wjg.v24.i26.2818>.
6. C. Röcken, “Predictive Biomarkers in Gastric Cancer,” *Journal of Cancer Research and Clinical Oncology* 149 (2023): 467–481, <https://doi.org/10.1007/s00432-022-04408-0>.
7. M. Hong, S. Tao, L. Zhang, et al., “RNA Sequencing: New Technologies and Applications in Cancer Research,” *Journal of Hematology and Oncology* 13, no. 1 (2020): 166, <https://doi.org/10.1186/s13045-020-01005-x>.

8. Y. Wang, "Identifying Key Stage-Specific Genes and Transcription Factors for Gastric Cancer Based on RNA-Sequencing Data," *Medicine* 96, no. 4 (2017): e5691, <https://doi.org/10.1097/md.0000000000005691>.
9. X. Li, J. Wang, B. Zhao, F. X. Wu, and Y. Pan, "Identification of Protein Complexes From Multi-Relationship Protein Interaction Networks," *Human Genomics* 10, no. S2 (2016): 17, <https://doi.org/10.1186/s40246-016-0069-z>.
10. A. Siavoshi, M. Taghizadeh, E. Dookhe, and M. Piran, "Gene Expression Profiles and Pathway Enrichment Analysis to Identification of Differentially Expressed Gene and Signaling Pathways in Epithelial Ovarian Cancer Based on High-Throughput RNA-Seq Data," *Genomics* 114, no. 1 (2022): 161–170, <https://doi.org/10.1016/j.ygeno.2021.11.031>.
11. R. J. De Las and C. Fontanillo, "Protein-Protein Interactions Essentials: Key Concepts to Building and Analyzing Interactome Networks," *PLoS Computational Biology* 6, no. 6 (2010): e1000807, <https://doi.org/10.1371/journal.pcbi.1000807>.
12. S. Khuri and S. Wuchty, "Essentiality and Centrality in Protein Interaction Networks Revisited," *BMC Bioinformatics* 16, no. 1 (2015): 109, <https://doi.org/10.1186/s12859-015-0536-x>.
13. C. J. Tsai, B. Ma, and R. Nussinov, "Protein-Protein Interaction Networks: How Can a Hub Protein Bind So Many Different Partners?," *Trends in Biochemical Sciences* 34, no. 12 (2009): 594–600, <https://doi.org/10.1016/j.tibs.2009.07.007>.
14. V. Grolmusz, G. Ivan, D. Banky, and B. Szerencsi, "How to Find Non-Hub Important Nodes in Protein Networks?," *Biophysical Journal* 102, no. 3 (2012): 184, <https://doi.org/10.1016/j.bpj.2011.11.1004>.
15. L. Cai, S. Qin, and Z. Pei, "Novel Hub Protein Classification and Interaction Rules in Protein-Protein Interaction Network in *Saccharomyces cerevisiae*," (2011), *Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, <https://doi.org/10.1109/FSKD.2011.6019657>.
16. B. Manna, T. Bhattacharya, B. Kahali, and T. C. Ghoshet, "Evolutionary Constraints on Hub and Non-hub Proteins in Human Protein Interaction Network: Insight From Protein Connectivity and Intrinsic Disorder," *Gene* 434, no. 1–2 (2009): 50–55, <https://doi.org/10.1016/j.gene.2008.12.013>.
17. X. Chang, T. Xu, Y. Li, and K. Wang, "Dynamic Modular Architecture of Protein-Protein Interaction Networks Beyond the Dichotomy of 'Date' and 'Party' Hubs," *Scientific Reports* 3, no. 1 (2013): 1691, <https://doi.org/10.1038/srep01691>.
18. Y. Y. Dong, Y. Jun, Q. N. Liu, and Z. H. Wang, "Exploring Centrality-Lethality Rule From Evolution Constraint on Essential Hub and Non-essential Hub in Yeast Protein-Protein Interaction Network," (2011), *Seventh International Conference on Computational Intelligence and Security*, Sanya, China, 63–65, <https://doi.org/10.1109/cis.2011.22>.
19. X. Q. Lu, J. Q. Zhang, S. X. Zhang, et al., "Identification of Novel Hub Genes Associated With Gastric Cancer Using Integrated Bioinformatics Analysis," *BMC Cancer* 21, no. 1 (2021): 697, <https://doi.org/10.1186/s12885-021-08358-7>.
20. F. Sun, C. Zhang, S. Ai, Z. Liu, and X. Lu, "Identification of Hub Genes in Gastric Cancer by Integrated Bioinformatics Analysis. Translational," *Cancer Research* 10, no. 6 (2021): 2831–2840, <https://doi.org/10.21037/tcr-20-3540>.
21. H. Liu, Y. Qu, H. Zhou, Z. Zheng, J. Zhao, and J. Zhang, "Bioinformatic Analysis of Potential Hub Genes in Gastric Adenocarcinoma," *Science Progress* 104, no. 1 (2021): 368504211004260, <https://doi.org/10.1177/00368504211004260>.
22. T. Ma, F. Liang, and G. Tseng, "Biomarker Detection and Categorization in Ribonucleic Acid Sequencing Meta-Analysis Using Bayesian Hierarchical Models," *Applied Statistics* 66, no. 4 (2017): 847–867, <https://doi.org/10.1111/rssc.12199>.
23. Babraham Bioinformatics, "FastQC: A Quality Control Tool for High Throughput Sequence Data. v. 0.11.5," (2010), <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
24. A. M. Bolger, M. Lohse, and B. Usadel, "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data," *Bioinformatics* 30, no. 15 (2014): 2114–2120, <https://doi.org/10.1093/bioinformatics/btu170>.
25. R. Patro, G. Duggal, I. L. Michael, R. A. Irizarry, and C. Kingsford, "Salmon Provides Fast and Bias-Aware Quantification of Transcript Expression," *Nature Methods* 14, no. 4 (2017): 417–419, <https://doi.org/10.1038/nmeth.4197>.
26. M. S. Lawrence, W. Huber, H. Pagès, et al., "Software for Computing and Annotating Genomic Ranges," *PLoS Computational Biology* 9, no. 8 (2013): e1003118, <https://doi.org/10.1371/journal.pcbi.1003118>.
27. J. T. Leek, W. E. Johnson, H. S. Parker, et al., "The sva Package for Removing Batch Effects and Other Unwanted Variation in High-Throughput Experiments," *Bioinformatics* 28, no. 6 (2019): 882–883, <https://doi.org/10.1093/bioinformatics/bts034>.
28. M. I. Love, W. Huber, and S. Anders, "Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data With DESeq2," *Genome Biology* 15, no. 12 (2014): 550, <https://doi.org/10.1186/s13059-014-0550-8>.
29. M. E. Ritchie, B. Phipson, D. Wu, et al., "Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies," *Nucleic Acids Research* 43, no. 7 (2015): e47, <https://doi.org/10.1093/nar/gkv007>.
30. G. Csardi and T. Nepusz, "The Igraph Software Package for Complex Network Research," *International Journal of Complex Systems* 1695 (2006): 1–9.
31. P. Langfelder and S. Horvath, "WGCNA: An R Package for Weighted Correlation Network Analysis," *BMC Bioinformatics* 9 (2008): 559, <https://doi.org/10.1186/1471-2105-9-559>.
32. M. Sniedovich, "Dijkstra's Algorithm Revisited: The Dynamic Programming Connexion," *Control and Cybernetics* 35, no. 3 (2006): 599–620.
33. S. I. Berger, A. Ma'ayan, and R. Iyengar, "Systems Pharmacology of Arrhythmias," *Science Signaling* 3, no. 118 (2010): ra30, <https://doi.org/10.1126/scisignal.2000723>.
34. M. Piran, N. Sepahi, A. Moattari, A. Rahimi, and A. Ghanbariasad, "Systems Biomedicine of Colorectal Cancer Reveals Potential Targets for CRC Treatment," *Frontiers in Oncology* 24, no. 11 (2021): 597536, <https://doi.org/10.3389/fonc.2021.597536>.
35. A. Franceschini, D. Szklarczyk, D. Frankild, et al., "STRING v9.1: Protein-Protein Interaction Networks, With Increased Coverage and Integration," *Nucleic Acids Research* 41, no. D1 (2013): D808–D815, <https://doi.org/10.1093/nar/gks1094>.
36. E. Cerami, J. Gao, U. Dogrusoz, et al., "The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data," *Cancer Discovery* 2, no. 5 (2012): 401–404, <https://doi.org/10.1158/2159-8290.cd-12-0095>.
37. X. He and J. Zhang, "Why Do Hubs Tend to Be Essential in Protein Networks?," *PLoS Genetics* 2, no. 6 (2006): e88, <https://doi.org/10.1371/journal.pgen.0020088>.
38. M. W. Hahn and A. D. Kern, "Comparative Genomics of Centrality and Essentiality in Three Eukaryotic Protein-Interaction Networks," *Molecular Biology and Evolution* 22 (2005): 803–806.
39. H. Jeong, S. P. Mason, A. L. Barabasi, and Z. N. Oltvai, "Lethality and Centrality in Protein Networks," *Nature* 411 (2001): 41–42.
40. C. Nithya, K. Manjari, and H. A. Nagarajaram, "Comparative Analysis of Pure Hubs and Pure Bottlenecks in Human Protein-Protein Interaction Networks," (2021), *bioRxiv*, <https://doi.org/10.1101/2021.04.06.438602>.

41. F. Wang, Q. Xue, D. Xu, Y. Jiang, C. Tang, and X. Liu, "Identifying the Hub Gene in Gastric Cancer by Bioinformatics Analysis and In Vitro Experiments," *Cell Cycle* 19, no. 11 (2020): 1326–1337, <https://doi.org/10.1080/15384101.2020.1749789>.
42. S. Zhang, X. Xiang, L. Liu, H. Yang, D. Cen, and G. Tang, "Bioinformatics Analysis of Hub Genes and Potential Therapeutic Agents Associated With Gastric Cancer," *Cancer Management and Research* 13 (2021): 8929–8951.
43. C. Shao, R. Wang, D. Kong, Q. Gao, and C. Xu, "Identification of Potential Core Genes in Gastric Cancer Using Bioinformatics Analysis," *Journal of Gastrointestinal Oncology* 12, no. 5 (2021): 2109–2122, <https://doi.org/10.21037/jgo-21-628>.
44. A. Li, Y. Li, Y. Li, M. Zhang, H. Zhang, and F. Chen, "Identification and Validation of Key Genes Associated With Pathogenesis and Prognosis of Gastric Cancer," *PeerJ* 16, no. 11 (2023): e16243, <https://doi.org/10.7717/peerj.16243>.
45. S. Wang, S. Tao, Y. Liu, Y. Shi, and M. Liu, "Identification of Significant Genes Associated With Prognosis of Gastric Cancer by Bioinformatics Analysis," *Journal of the Egyptian National Cancer Institute* 34 (2022): 55, <https://doi.org/10.1186/s43046-022-00157-w>.
46. Z.-S. Zhao, Y.-Y. WangWang, Y.-Q. Chu, Z.-Y. Ye, and H.-Q. TaoTao, "SPARC Is Associated With Gastric Cancer Progression and Poor Survival of Patients," *Clinical Cancer Research* 16, no. 1 (2010): 260–268, <https://doi.org/10.1158/1078-0432.CCR-09-1247>.
47. S. Furlong, R. Melanie, P. Coombs, and J. Ghassemi-Rad, "Thy-1 (CD90) Signaling Preferentially Promotes ROR γ t Expression and a Th17 Response," *Frontiers in Cell and Developmental Biology* 6 (2018): 158, <https://doi.org/10.3389/fcell.2018.00158>.
48. S. Sinha, A. Renganathan, P. B. Nagendra, V. Bhat, B. S. Mathew, and M. R. S. Rao, "AEBP1 Down Regulation Induced Cell Death Pathway Depends on PTEN Status of Glioma Cells," *Scientific Reports* 9, no. 1 (2019): 14577, <https://doi.org/10.1038/s41598-019-51068-1>.
49. D. Aintree-Tacha, J. Tyrrell, and T. Haas, "CDH17 Is a More Sensitive Marker for Gastric Adenocarcinoma Than CK20 and CDX2," *Archives of Pathology and Laboratory Medicine* 141, no. 1 (2017): 144–150, <https://doi.org/10.5858/arpa.2015-0404-oa>.
50. Y. Hu, D. Jin, Y. Zhou, et al., "Multiple Roles of THY1 in Gastric Cancer Based on Data Mining," *Translational Cancer Research* 9, no. 4 (2020): 2748–2757, <https://doi.org/10.21037/tcr.2020.02.51>.
51. R. C. Baxter, "IGF Binding Proteins in Cancer: Mechanistic and Clinical Insights," *Nature Reviews Cancer* 14, no. 5 (2014): 329–341, <https://doi.org/10.1038/nrc3720>.
52. J. Zhang, F. Zhang, J. Fan, and B. Feng, "TGIF1 Knockdown Inhibits the Proliferation and Invasion of Gastric Cancer via AKT Signaling Pathway," *Cancer Management and Research* 13 (2021): 2603–2612, <https://doi.org/10.2147/cmar.s254348>.
53. V. S. Donnenberg, A. D. Donnenberg, L. Zimmerlin, et al., "Localization of CD44 and CD90 Positive Cells to the Invasive Front of Breast Tumors," *Cytometry Part B—Clinical Cytometry* 78, no. 5 (2010): 287–301, <https://doi.org/10.1002/cyto.b.20530>.
54. F. Buccisano, F. Maria Rossi, A. Venditti, et al., "CD90/Thy1 Is Preferentially Expressed on Blast Cells of High Risk Acute Myeloid Leukemias," *British Journal of Hematology* 125, no. 2 (2004): 203–212, <https://doi.org/10.1111/j.1365-2141.2004.04883.x>.
55. Z. H. Yu, Y. M. Wang, Y. Z. Jiang, et al., "NID2 Can Serve as a Potential Prognosis Prediction Biomarker and Promotes the Invasion and Migration of Gastric Cancer," *Pathology Research and Practice* 215, no. 10 (2019): 152553, <https://doi.org/10.1016/j.prp.2019.152553>.
56. Y. Di, D. Chen, Y. Wei, and L. Yan, "Bladder Cancer Stage-Associated Hub Genes Revealed by WGCNA Co-Expression Network Analysis," *Hereditas* 156, no. 1 (2019): 7, <https://doi.org/10.1186/s4106-019-0083-y>.
57. T. Li, C. Zhang, G. Zhao, et al., "IGFBP2 Regulates PD-L1 Expression by Activating the EGFR-STAT3 Signaling Pathway in Malignant Melanoma," *Cancer Letters* 477 (2020): 19–30, <https://doi.org/10.1016/j.canlet.2020.02.036>.
58. Z. S. Yuan, Y. Cao, and Z. Y. Li, "IGFBP2 Induces SPRY1 Expression via NF- κ B Signaling Pathway in Glioblastoma Multiforme (GBM)," *European Review for Medical and Pharmacological Sciences* 21, no. 22 (2017): 5072–5080, https://doi.org/10.26355/eurrev_201711_13821.
59. C. Guo, H. Lu, W. Gao, et al., "Insulin-Like Growth Factor Binding Protein-2 Level Is Increased in Blood of Lung Cancer Patients and Associated With Poor Survival," *PLoS One* 8, no. 9 (2013): e74973, <https://doi.org/10.1371/journal.pone.0074973>.
60. C. M. Perks, E. G. Vernon, A. H. Rosendahl, A. H. Rosendahl, D. Tonge, and J. M. P. Holly, "IGF-II and IGFBP-2 Differentially Regulate PTEN in Human Breast Cancer Cells," *Oncogene* 26, no. 40 (2007): 5966–5972, <https://doi.org/10.1038/sj.onc.1210397>.
61. G. Kaur, S. D. Balasubramaniam, and Y. Lee, "IGFBP-2 in Cervical Cancer Development," *Experimental and Molecular Pathology* 113 (2020): 104362, <https://doi.org/10.1016/j.yexmp.2019.104362>.
62. K. L. McDonald, M. G. O'Sullivan, J. F. Parkinson, et al., "IQGAP1 and IGFBP2: Valuable Biomarkers for Determining Prognosis in Glioma Patients," *Journal of Neuropathology and Experimental Neurology* 66, no. 5 (2007): 405–417, <https://doi.org/10.1097/nen.0b013e31804567d7>.
63. Z. Tang, D. Gillatt, E. Rowe, A. Koupparis, J. M. P. Holly, and C. M. Perks, "IGFBP-2 Acts as a Tumor Suppressor and Plays a Role in Determining Chemosensitivity in Bladder Cancer Cells," *Oncotarget* 10, no. 66 (2019): 7043–7057, <https://doi.org/10.18632/oncotarget.27355>.
64. Z. Kang, Y. Yu, Y. J. Zhu, et al., "Downregulation of IGFBP2 Is Associated With Resistance to IGF1R Therapy in Rhabdomyosarcoma," *Oncogene* 33, no. 50 (2014): 5697–5705, <https://doi.org/10.1038/nc.2013.509>.
65. V. C. Russo, W. J. Azar, S. W. Yau, M. A. Sabin, and G. A. Werther, "IGFBP-2: The Dark Horse in Metabolism and Cancer," *Cytokine and Growth Factor Reviews* 26, no. 3 (2015): 329–346, <https://doi.org/10.1016/j.cytogfr.2014.12.001>.
66. Y. Guo, Y. Bao, D. Guo, and W. Yang, "Pregnancy-Associated Plasma Protein a in Cancer: Expression, Oncogenic Functions and Regulation," *American Journal of Cancer Research* 8 (2018): 955–963.
67. A. Gennari, F. Foca, R. Zamarchi, et al., "Insulin-Like Growth Factor-1 Receptor (IGF-1R) Expression on Circulating Tumor Cells (CTCs) and Metastatic Breast Cancer Outcome: Results From the TransMYME Trial," *Breast Cancer Research and Treatment* 181, no. 1 (2020): 61–68, <https://doi.org/10.1007/s10549-020-05596-4>.
68. H. Wang, H. Wang, W. Shen, et al., "Insulin-Like Growth Factor Binding Protein 2 Enhances Glioblastoma Invasion by Activating Invasion-Enhancing," *Cancer Research* 63, no. 15 (2003): 4315–4321.
69. S. P. Guan, A. T. L. Lam, J. F. Newman, et al., "Matrix Metalloproteinase-1 Facilitates MSC Migration via Cleavage of IGF-2/IGFBP2 Complex," *FEBS Open Bio* 8, no. 1 (2017): 15–26, <https://doi.org/10.1002/2211-5463.12330>.
70. P. B. Varsha, C. Prasanna, and H. S. Atreya, "Insulin-Like Growth Factor System in Cancer: Novel Targeted Therapies," *BioMed Research International* 2015 (2015): 538019, <https://doi.org/10.1155/2015/538019>.
71. E. Schiegnitz, P. W. Kämmerer, H. Schön, et al., "The Matrix Metalloproteinase and Insulin-Like Growth Factor System in Oral Cancer—A Prospective Clinical Study," *Oncotargets and Therapy* 10 (2017): 5099–5105, <https://doi.org/10.2147/ott.s149231>.
72. M. Mancini, M. B. Gariboldi, E. Taiana, et al., "Co-Targeting the IGF System and HIF-1 Inhibits Migration and Invasion by (Triple-Negative) Breast Cancer Cells," *British Journal of Cancer* 110, no. 12 (2014): 2865–2873, <https://doi.org/10.1038/bjc.2014.269>.

73. D. Feldser, F. Agani, N. V. Iyer, B. Pak, G. Ferreira, and G. L. Semenza, "Reciprocal Positive Regulation of Hypoxia-Inducible Factor 1 α and Insulin-Like Growth Factor 2," *Cancer Research* 59, no. 16 (1999): 3915–3918.
74. J. Pereira, T. C. Meyer, and S. E. Docherty, "Bimolecular Interaction of Insulin-Like Growth Factor (IGF) Binding Protein-2 With $\alpha\text{v}\beta\text{3}$ Negatively Modulates IGF-I-Mediated Migration and Tumor Growth 1," *Cancer Research* 64, no. 3 (2004): 977–984, <https://doi.org/10.1158/0008-5472.can-03-3056>.
75. J. Saegusa, S. Yamaji, K. Ieguchi, et al., "The Direct Binding of Insulin-Like Growth Factor-1 (IGF-1) to Integrin $\alpha\text{v}\beta\text{3}$ Is Involved in IGF-1 Signaling," *Journal of Biological Chemistry* 284, no. 36 (2009): 24106–24114, <https://doi.org/10.1074/jbc.m109.013201>.
76. T. Li, M. E. Forbes, G. N. Fuller, J. Li, X. Yang, and W. Zhang, "IGFBP2: Integrative Hub of Developmental and Oncogenic Signaling Network," *Oncogene* 39, no. 11 (2020): 2243–2257, <https://doi.org/10.1038/s41388-020-1154-2>.
77. M. Habeb, S. Embarak, A. Fathy, and M. Zalut, "Non-Invasive IGFBP1, IGFBP2 Biomarkers as Predictors and Outcomes of Usual Interstitial Pneumonia (UIP) Therapeutic Response," *Egyptian Journal of Bronchology* 15 (2021): 6, <https://doi.org/10.1186/s43168-020-00050-x>.
78. B. Zhang, C. Q. Hong, Y. H. Luo, et al., "Prognostic Value of IGFBP2 in Various Cancers: A Systematic Review and Meta-Analysis," *Cancer Medicine* 11, no. 16 (2022): 3035–3047, <https://doi.org/10.1002/cam4.4680>.
79. L.-F. Wei, X.-F. Weng, and X.-C. Huang, "IGFBP2 in Cancer: Pathological Role and Clinical Significance," *Oncology Reports* 45 (2020): 427–438.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.