

Research Article

Identifying Gastric Cancer Related Genes Using the Shortest Path Algorithm and Protein-Protein Interaction Network

Yang Jiang,¹ Yang Shu,² Ying Shi,³ Li-Peng Li,¹ Fei Yuan,² and Hui Ren⁴

¹ Colorectal Surgery Department, China-Japan Union Hospital of Jilin University, Changchun 130033, China

² State Key Laboratory of Medical Genomics, Institute of Health Sciences, Chinese Academy of Sciences, Shanghai Jiao Tong University School of Medicine and Shanghai Institutes for Biological Sciences, Shanghai 200025, China

³ Breast and Thyroid Surgery Department, The Second Hospital of Jilin University, Changchun 130041, China

⁴ Colorectal Surgery Department, The Second Hospital of Jilin University, Changchun 130041, China

Correspondence should be addressed to Hui Ren; hren@jlu.edu.cn

Received 29 December 2013; Accepted 3 February 2014; Published 5 March 2014

Academic Editor: Tao Huang

Copyright © 2014 Yang Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Gastric cancer, as one of the leading causes of cancer related deaths worldwide, causes about 800,000 deaths per year. Up to now, the mechanism underlying this disease is still not totally uncovered. Identification of related genes of this disease is an important step which can help to understand the mechanism underlying this disease, thereby designing effective treatments. In this study, some novel gastric cancer related genes were discovered based on the knowledge of known gastric cancer related ones. These genes were searched by applying the shortest path algorithm in protein-protein interaction network. The analysis results suggest that some of them are indeed involved in the biological process of gastric cancer, which indicates that they are the actual gastric cancer related genes with high probability. It is hopeful that the findings in this study may help promote the study of this disease and the methods can provide new insights to study various diseases.

1. Introduction

Gastric carcinogenesis is a multistep process involving genetic and epigenetic alteration of protein-coding protooncogenes and tumor-suppressor genes. Gastric cancer (GC) is the fourth most commonly diagnosed cancer and is estimated to be the second most common cause of cancer related death and causes about 800,000 deaths worldwide per year [1, 2]. Because of the improvement of the dietary structure, the mortality rate shows a declining trend worldwide [3]. However, the incidences of gastric cancer are still remarkable in areas where infection by *Helicobacter pylori* is prevalent [4]. Besides *H. pylori*, smoking and alcohol consumption also increase the risk of developing gastric cancer significantly [5, 6]. Compared with women, men have a higher incidence, while estrogen may protect women against the gastric cancer [7].

In the previous cases, over 90% gastric cancers are adenocarcinomas, which could be divided into two major types

in terms of the histopathology [8]. Intestinal type gastric cancer is often related to environmental factors such as *H. pylori*, while diffuse type gastric cancer is more often associated with genetic abnormalities. Caldas et al. reviewed that the diffuse type gastric cancer tended to occur in female and young individuals [9]. Besides adenocarcinomas, other types of gastric cancers like lymphomas occurred in a very low incidence [10]. Since the gastric cancer leads to high mortality, the early diagnose especially the molecular diagnose is particularly important for the therapy.

So far, numerous genes have been found involved in gastric tumorigenesis. Among the reported gastric cancer related genes, most of them could have also been found in other types of carcinomas. p53, famous for its tumor-suppressing role, has a mutated rate ranging from 0 to 21% in diffuse type GC and 36–43% in intestinal type GC [11]; E-cadherin, which plays a pivotal role in EMT (Epithelial Mesenchymal Transition), is predisposed to mutagenesis in sporadic diffuse type GC (33–50%) [12]; another star gene

harboring high correlation with gastric cancer is RNUX3, which manifests to be a tumor-suppressor gene of GC [13]. Although dozens of genes have been found related to gastric cancer, they are insufficient to elucidate the tumorigenesis of GC unless more relevant genes being uncovered.

It is time-consuming to discover novel gastric cancer related genes by experiment alone, because the search space is very large. Computational approach is an alternative way which can help investigators screen out some related genes. On the other hand, lots of computational approaches have been developed to settle various biological problems, such as drug design [14–19] and analysis of complicated biological network [20–24]. In this study, a computational method was built to discover novel gastric cancer related genes based on some known related ones retrieved from Gastric Cancer Database, UniProtKB, and TSGene Database. After applying the shortest path algorithm in protein-protein interaction network to search the shortest path connecting any pair of known gastric cancer related genes, the candidate genes were found. Further analysis suggests that some of them are related to the formation and development of gastric cancer. We hope that this contribution may give help to uncover the mechanism of this disease, thereby designing effective treatments.

2. Materials and Methods

2.1. Materials. Gastric cancers related genes are collected from the following three datasets: (1) 102 genes are picked up from Gastric Cancer Database (<http://www.gastric-cancer.site40.net/>); (2) 128 reviewed gastric cancer related genes were found in the UniProtKB (Protein Knowledgebase, <http://www.uniprot.org/uniprot/>) by setting the keyword as human gastric cancer oncogene/suppressor gene, where 86 are oncogenes and 42 are suppressor genes; (3) 9 genes were obtained from TSGene Database (Tumor Suppressor Gene Database, <http://bioinfo.mc.vanderbilt.edu/TSGene/>) by searching the human gastric cancer in the Literature Search box. After combining these genes, we obtained 150 gastric cancer related genes, which were available in Supplementary Material I (available online at <http://dx.doi.org/10.1155/2014/371397>).

2.2. Protein-Protein Interaction (PPI) Network. It is known that interactions of proteins are important for the majority of biological functions. Many studies have shown that proteins in one interaction always share similar functions [25–29]. Since gastric cancers related genes may have some common features, it is feasible to discover novel gastric cancers related genes based on known related ones and PPI network. In this study, the PPI network was constructed based on the protein interaction information retrieved from STRING (Search Tool for the Retrieval of Interacting Genes/Proteins, <http://string.embl.de/>) (version 9.0) [30], a well-known database integrating known and predicted protein interactions. In the obtained file, each interaction consists of two protein IDs and a score measuring the likelihood of the interaction's occurrence. For later formulation,

the score of the interaction between proteins p_1 and p_2 was denoted by $I(p_1, p_2)$. To construct the weighted network, proteins in the STRING were taken as nodes and two nodes were adjacent if and only if the score of the interaction between the corresponding proteins was greater than zero. In addition, the score of the interaction was used to label the weight of the corresponding edge as follows:

$$w(v_1, v_2) = 1000 - I(p_1, p_2), \quad (1)$$

where p_i ($i = 1, 2$) was the corresponding protein of node v_i .

2.3. Shortest Path Genes. As described in Section 2.1, 150 gastric cancer related genes were collected, which must have some common features related to gastric cancer. On the other hand, according to Section 2.2, two proteins in one interaction, that is, they are adjacent in the constructed PPI network, always share common features. It can be further deduced that proteins in the shortest path connecting two known gastric cancer related genes may share some common features that the two known gastric cancer related genes have. Therefore, we searched the shortest path between any pair of known gastric cancer related genes by Dijkstra's algorithm, the most famous shortest path algorithm proposed by Dijkstra in 1956 [31].

After collecting the shortest paths connecting any pair of known gastric cancer related genes, we found that some nodes/genes occurred in many paths, while the majority of nodes/genes in PPI network were not in any path. To distinguish these nodes/genes, the betweenness of each node/gene was calculated, which is defined as the number of the shortest paths containing the node/gene as an inner node. Since the concept of betweenness accounts for direct and indirect influences of proteins at distant network [32], it has been employed in the study of various natural and man-made networks [33–38].

It is easy to see that genes with high betweenness may share more features related to gastric cancer than those with low betweenness, while the likelihood of gene with betweenness equal to 0 to be the novel gastric cancer related gene is zero. Accordingly, we picked out genes with betweenness greater than 0 and termed them as the shortest path genes. Since the main purpose of this study is to discover novel gastric cancer related genes, the known gastric cancer related genes were not included in the set of shortest path genes.

2.4. Further Filtering Based on Permutation Test. As described in Section 2.3, some of the shortest path genes can be obtained based on their betweenness. However, the betweenness of some nodes may be strongly influenced by the essential structure of the network. For example, the cut-vertex of the network may always receive high betweenness easier than other vertices. To control this false discovery, a permutation test was conducted to further filter these shortest path genes as follows.

- (i) Randomly select 1,000 gene sets $G_1, G_2, \dots, G_{1000}$ in PPI network with the same size of known gastric cancer related gene set.

- (ii) Calculate the betweenness of each shortest path gene on each gene set G_i ($1 \leq i \leq 1000$).
- (iii) The permutation FDR of the shortest path gene p was computed by

$$\text{FDR}(p) = \frac{\sum_{i=1}^{1000} \delta_i}{1000}, \quad (2)$$

where δ_i was defined to be 1 if the betweenness of p on G_i was greater than that of p on the known gastric cancer related gene set.

It is obvious that smaller permutation FDR of one shortest path gene indicates that it is the actual gastric cancer related gene with high possibility.

2.5. Gene Set Enrichment Analysis. DAVID [39] is a functional annotation tool, which has been widely used to analyze gene lists derived from different biological problems [40–45]. Here, it was also employed for KEGG pathway and GO enrichment analysis of the obtained gene set. The enrichment P value was corrected to control family-wise false discovery rate under certain rate (e.g., ≤ 0.05) with Benjamin multiple testing correction method [46]. During the enrichment analysis, all genes in the human genome were considered. 13 items in the output of DAVID and their meanings are listed in Table 1. For detailed description, please see Huang et al.'s study [39].

3. Results and Discussion

3.1. Candidate Genes. Of the 150 known gastric cancer related genes, the shortest path connecting any pair of them was searched in PPI network constructed in Section 2.2. After counting the betweenness of each gene in PPI network, 466 shortest path genes with betweenness greater than zero were retrieved. These 466 genes and their betweenness can be found in Supplementary Material II. To exclude the false discovery, the permutation test was conducted. The permutation FDRs of 466 shortest path genes were calculated by (2) and also listed in Supplementary Material II. It can be observed that 144 genes were with permutation FDRs no more than 0.1. These genes were considered to have a strong relationship with gastric cancer.

3.2. Results of Gene Set Enrichment Analysis. DAVID, as a functional annotation tool, was employed to analyze the 144 shortest path genes. The analysis results included two categories: GO and KEGG. These results were available in Supplementary Materials III and IV, respectively. The detailed discussion based on these results was as follows.

From Supplementary Material III, 294 GO terms were enriched by the 144 genes. We investigated the first 10 GO terms in the list, which were shown in Figure 1. The "Count" items in the output of DAVID for these 10 GO terms were also shown in Figure 1. Among these 10 GO terms, 5 out of the 10 GO terms are cellular component (CC) GO terms including (1) GO:0005654: nucleoplasm ("count" = 30); (2) GO:0031981: nuclear lumen ("count" = 34); (3) GO:0043233:

TABLE 1: Items in the output of DAVID and their meanings.

Item	Meaning
Category	DAVID category, that is, KEGG or GO
Term	Gene set name
Count	The number of genes associated with this gene set
Percentage	Calculated by "gene associated with this gene set"/"total number of query genes"
P value	Modified Fisher Exact P value
Genes	The list of genes from your query set that are annotated to this gene set
List total	The number of genes in your query list mapped to any gene set in this ontology
Pop hits	The number of genes annotated to this gene set on the background list
Pop total	The number of genes on the background list mapped to any gene set in this ontology
Fold enrichment	The ratio of the proportions on query genes and the background information which are associated with the gene set
Bonferroni	Bonferroni adjusted P value
Benjamini	Benjamini adjusted P value
FDR	FDR adjusted P value

organelle lumen ("count" = 37); (4) GO:0031974: membrane-enclosed lumen ("count" = 37); (5) GO:0005829: cytosol ("count" = 30). As we know, tumorigenesis is a very complicated biological process which means the transform processes could take place everywhere in the cells [47]. In our analysis results, the related proteins distribute both in nuclear and cytosol which is in accordance with the characters of the gastric cancer. The remaining 5 GO terms are biological process (BP) GO terms: (1) GO:0032268: regulation of cellular protein metabolic process ("count" = 20); (2) GO:0009725: response to hormone stimulus ("count" = 17); (3) GO:0031399: regulation of protein modification process ("count" = 15); (4) GO:0009719: response to endogenous stimulus ("count" = 17); (5) GO:0010604: positive regulation of macromolecule metabolic process ("count" = 25). Liu et al. reported that the cancer cells usually harbor abnormal metabolic status [48]. In our results 80% (4/5) BP are relative to the metabolic stress response by means of direct regulation of the metabolic process or indirect regulation by altering the stimulus-related pathways. Besides, protein modification, which is also enriched in our results, plays an important role in the carcinogenesis by altering the pivotal proteins [49]. Although these genes may not be the indispensable factors in gastric cancer, the common points among them would give us the hints about the tumorigenesis of the gastric cancer.

From Supplementary Material IV, 8 KEGG pathways were enriched by 144 genes, which were shown in Figure 2. It can be observed that 6 out of 8 KEGG pathways were with P value less than 0.05, which were investigated as follows. The first pathway was hsa04110: cell cycle pathway ("count" = 10). 10 genes including PCNA, MYC, and CCND1 are enriched

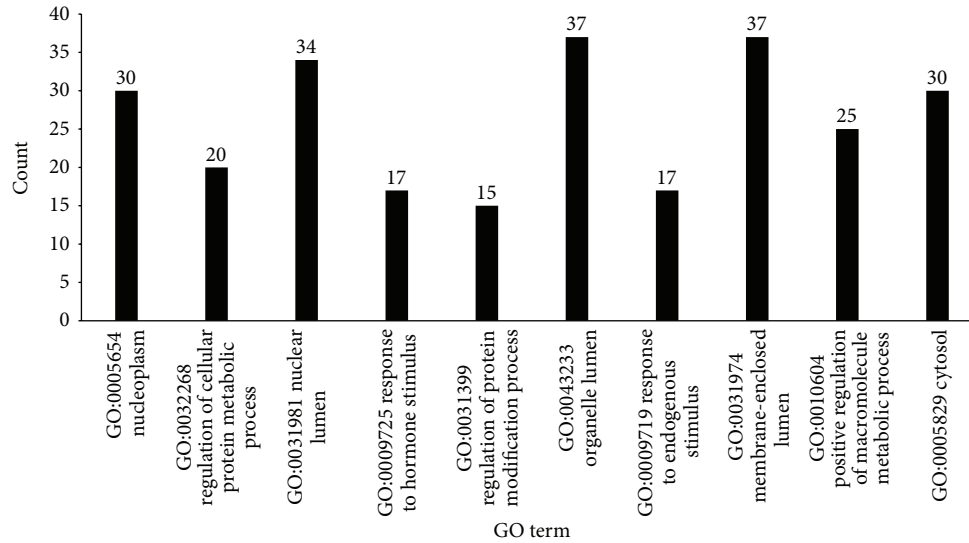


FIGURE 1: The top 10 GO terms enriched by 144 genes. The *x*-axis lists GO's ID and name, while the *y*-axis represents the number of genes that shared the GO term among the 144 genes.

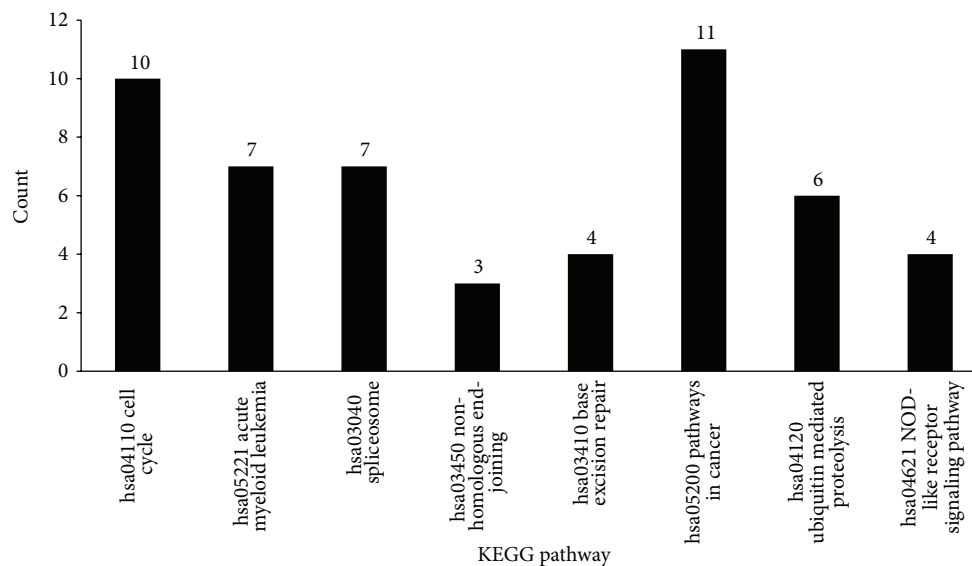


FIGURE 2: The 8 KEGG pathways enriched by 144 genes. The *x*-axis lists pathway's ID and name, while the *y*-axis represents the number of genes that shared the pathway among the 144 genes.

in this pathway. One of the significant characters of gastric cancer is the abnormal activated cell cycle [50]. Among these genes, PCNA is responsible for the DNA synthesis and CCND1 could alter cell cycle by regulating the CDK kinases [51, 52]. Other 2 pathways found in our study are related to the DNA repair which are also very critical for the carcinogenesis. Hsa03450: nonhomologous end joining (NHEJ) ("count" = 3) is a pathway that repairs double-strand breaks in DNA and base excision repair (BER) is a cellular mechanism that repairs damaged DNA throughout the cell cycle [53, 54]. Another intriguing pathway is hsa03040: spliceosome pathway ("count" = 7) which was always abnormal in cancer cells [55]. We speculate that the spliceosome could modify

the expression of the oncogenes or tumor-suppress genes which eventually lead to the tumorigenesis. Finally, we also find the hsa05221: acute myeloid leukemia (AML) pathway ("count" = 7) and cancer related pathways in our list. The results imply that the gastric cancer has the common mechanism as well as other cancers especially the AML. Look has reviewed that RUNX1 is the key factor in the hematopoietic development and highly correlated with AML [56]. However, its homologous protein RUNX3 that shares 70% similarity has been reported playing pivotal role in gastric cancer [57]. The finding unravels that cancer normally has the common molecular mechanism as well as the specific pathway with type-dependent pattern. Although several reported pathways

are included in our study, the novel pathways with gastric cancer would expand our views of mechanisms about the tumorigenesis of gastric cancer. On the other hand, we have observed that some genes in these pathways could play a very important role in the carcinogenesis of gastric cancer.

3.3. Analysis of the Relationship of Some Candidate Genes and Gastric Cancer. As described in Section 3.1, 144 genes were discovered by our method. Some of them may have strong relationship with gastric cancer and were discussed as follows. Table 2 listed these genes and their betweenness and permutation FDRs.

Proliferating Cell Nuclear Antigen (PCNA) (see row 2 of Table 2), also known as cyclin, is an auxiliary protein of DNA polymerase- δ that plays important roles both in DNA synthesis and DNA repair [51, 58]. PCNA could act as a homotrimer and helps increase the processivity of leading strand synthesis during DNA replication [59, 60]. In response to DNA damage, this protein is ubiquitinated and is involved in the RAD6-dependent DNA repair pathway [61, 62]. As we know, DNA repair is the main way to remove the carcinogenic lesions caused by UV or other common mutagens [63]. Pascucci et al. have reviewed that the NER (nucleotide excision repair) was highly correlated with skin cancer and intestinal cancer [64]. Intriguingly, numerous works have considered PCNA labeling rate as the prognostic indicator of gastric cancer because its expression was consistent with malignant potential of gastric cancer [65–67]. Ji et al. have found the abnormal increase of PCNA expression in 58 gastric carcinoma tissues [68]. Similar conclusion was also achieved by Takamura et al. who have performed immunohistochemical study on 164 patients with gastric carcinomas [69]. Although the strong correlation is observed between PCNA and gastric cancer, the detailed mechanism of how PCNA promotes the gastric cancer needs further elucidation.

Besides PCNA, another protein in highly conserved cyclin family was also found in our study. CCND1 (see row 3 of Table 2), with official full name of cyclin D1, was firstly described by Motokura et al. in 1991 [70]. In the following decades, the importance of CCND1 in cell cycle and tumorigenesis was underlined by different labs. Because of the amplification of the 11q13 region where CCND1 locates, CCND1 is frequently overexpressed in human cancers accompanied with abnormalities that are driven by multiple mechanisms including genomic alternations, post-transcriptional regulation and posttranslational protein stabilization [71–73]. On one hand, cyclin D1 could increase CDK activity and consequently result in continuous proliferation which is necessary for tumorigenesis [74, 75]. On the other hand, cyclin D1 may induce the tumorigenesis in certain types of cancers by means of its nuclear receptor-agonistic activity in the CDK-independent way [52, 76].

MYC (see row 4 of Table 2) is a regulator gene that codes for a transcription factor, and it is frequently mutated in many cancers. In Myc-related cancers, Myc is constitutively expressed and leads to the abnormal expression of many genes which may be involved in cell proliferation, differentiation and apoptosis, and these uncontrolled biological

processes finally underlie the cancer. Myc is believed to regulate expression of 15% of all genes [77]. Similar with CCND1, Myc expression could be regulated transcriptionally, posttranscriptionally, or posttranslationally [78]. Chung and Levens have reviewed that the deregulated expression of Myc is sufficient to lead to cellular transformation *in vitro* and tumorigenesis *in vivo* [79]. Besides the transforming role, Myc could also promote chromosomal instability by means of its function as a transcriptional regulator [80]. In the previous reports, Myc overexpression has been described in over 40% of gastric cancer [81]. Among nearly half the gastric cancer, copy number gains are frequently detected along chromosome 8 where Myc locates [82, 83]. As the key factor of tumorigenesis, Myc could provide potential target for therapy for gastric cancer [84].

FOS (see row 5 of Table 2), well known as c-fos, encodes a 62 kDa protein, which forms heterodimer with c-jun and subsequently results in the formation of AP-1 complex. FOS has been found to be overexpressed in a variety of cancers. Bakin and Curran have found that c-fos could change DNA methylation pattern by regulating DNMT1 and thereby cause the downregulation of tumor suppressor genes [85]. In addition, c-fos could lead to the loss of cell polarity and EMT which is critical for the metastatic and invasive growth of cancer cells [86]. Hu et al. also found that c-fos is required for the expression of matrix metalloproteinases that are indispensable for invasive growth of cancer cells [87]. However, some recent studies have unraveled the tumor suppressor activity of c-fos, including prohibition of the cell cycle progression, promotion of cell death, or repressing the anchorage-independent growth [88]. In coincidence with the negative role of c-fos in tumorigenesis, Jin et al. analyzed 625 consecutive gastric cancers; 388 cases (62.1%) showed loss of nuclear c-fos expression [89]. Consistent results were concluded by Zhou et al. in 58 patients with gastric cancer [90]. However, Mazurenko et al. reported that high level of c-fos expression was observed in stomach carcinomas [91]. The discordance may be caused by the different stages of progression in different studies. In conclusion, c-fos is a double-edged sword, which could promote or suppress tumorigenesis of gastric cancer.

RUNX1 (see row 6 of Table 2), better known as AML1, plays a critical role in hematopoietic development [92]. RUNX1 belongs to the RUNX family whose 3 members (RUNX1, RUNX2, and RUNX3) share 70% resemblance. Unlike its familial protein RUNX3 that is a strong candidate as a gastric cancer tumor suppressor. RUNX1 is always considered as a tumor suppressor for acute lymphoblastic leukemia (AML) [56]. Usui et al. have examined mRNA expression of all three RUNX genes in the gastric mucosa, and they found that RUNX1 was coexpressed with RUNX3 in pit cells [93]. Sakakura et al. observed remarkable downregulation of RUNX1 and RUNX3 in 9 gastric cancer cell lines and 56 primary gastric cancer specimens [94]. Although RUNX1 is famous for its involving in AML, more lines of evidence shed light to its anticarcinogenesis activity in other carcinoma including gastric cancer.

Other genes found in our study have also been reported relating with gastric cancer. Specific SNPs (Single

TABLE 2: Important candidate shortest path genes and their betweenness and permutation FDRs.

Ensemble ID of shortest path genes	Gene name	Betweenness	Permutation FDR
ENSP00000368438	PCNA	454	0.083
ENSP00000227507	CCND1	594	0.02
ENSP00000367207	MYC	779	0.01
ENSP00000306245	FOS	318	0.035
ENSP00000300305	RUNX1	224	0.002
ENSP00000262887	XRCC1	152	0.094
ENSP00000352516	DNMT1	169	0.093
ENSP00000379110	CXCL1	107	0.033

Nucleotide Polymorphism) in XRCC1 (X-ray repair cross-complementing 1) (see row 7 of Table 2) are highly associated with gastric cancer [95]. DNMT1 (DNA methyltransferase 1) (see row 8 of Table 2), which is overexpressed in gastric cancer, is associated with increased risks of gastric atrophy with its abnormal polymorphisms [96]. The expression of CXCL1 (chemokine (C-X-C motif) ligand 1) (see row 9 of Table 2) is higher in gastric cancer tissues and endows the cancer cells with more powerful migration and invasion ability [97]. Beyond these genes, more genes associated with gastric tumorigenesis require more evidences for validation or further exploration.

4. Conclusion

Identification of disease genes is one of the most important problems in biomedicine and genomics. For gastric cancer, as one of the leading causes of cancer related deaths worldwide, it is eager to discover its related genes, which can help to uncover its mechanism and design effective treatments. This contribution presented a computational method to identify novel gastric cancer related genes based on known related ones by shortest path algorithm and PPI network. The analysis implies that some genes discovered in this study have direct or indirect relationship with gastric cancer. It is hopeful that this contribution would give a new insight to study this disease and other diseases.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Author's Contribution

Yang Jiang and Yang Shu contributed equally to this work.

Acknowledgments

This paper is supported by Natural Science Fund Projects of Jilin province (201215059), Development of Science and Technology Plan Projects of Jilin province (20100733, 201101074),

SRF for ROCS, SEM (2009-36), Scientific Research Foundation (Jilin Department of Science & Technology, 200705314, 20090175, 20100733), Scientific Research Foundation (Jilin Department of Health, 2010Z068), and SRF for ROCS (Jilin Department of Human Resource & Social Security, 2012-2014).

References

- [1] D. M. Parkin, F. Bray, J. Ferlay, and P. Pisani, "Estimating the world cancer burden: globocan 2000," *International Journal of Cancer*, vol. 94, no. 2, pp. 153-156, 2001.
- [2] J. Ferlay, H.-R. Shin, F. Bray, D. Forman, C. Mathers, and D. M. Parkin, "Estimates of worldwide burden of cancer in 2008: globocan 2008," *International Journal of Cancer*, vol. 127, no. 12, pp. 2893-2917, 2010.
- [3] D. Palli, "Epidemiology of gastric cancer: an evaluation of available evidence," *Journal of Gastroenterology*, vol. 35, supplement 12, pp. 84-89, 2000.
- [4] E. M. El-Omar, M. Carrington, W.-H. Chow et al., "Interleukin-1 polymorphisms associated with increased risk of gastric cancer," *Nature*, vol. 404, no. 6776, pp. 398-402, 2000.
- [5] A. Nomura, J. S. Grove, G. N. Stemmermann, and R. K. Severson, "Cigarette smoking and stomach cancer," *Cancer Research*, vol. 50, article 7084, 1990.
- [6] N. Y. Sung, K. S. Choi, E. C. Park et al., "Smoking, alcohol and gastric cancer risk in Korean men: the National Health Insurance Corporation Study," *British Journal of Cancer*, vol. 97, no. 5, pp. 700-704, 2007.
- [7] E. Chandanos and J. Lagergren, "Oestrogen and the enigmatic male predominance of gastric cancer," *European Journal of Cancer*, vol. 44, no. 16, pp. 2397-2403, 2008.
- [8] P. Lauren, "The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An attempt at a histo-clinical classification," *Acta Pathologica et Microbiologica Scandinavica*, vol. 64, pp. 31-49, 1965.
- [9] C. Caldas, F. Carneiro, H. T. Lynch et al., "Familial gastric cancer: overview and guidelines for management," *Journal of Medical Genetics*, vol. 36, no. 12, pp. 873-880, 1999.
- [10] V. Kumar, A. K. Abbas, N. Fausto, and J. C. Aster, *Robbins & Cotran Pathologic Basis of Disease*, Elsevier Health Sciences, Philadelphia, Pa, USA, 2009.
- [11] C. Maesawa, G. Tamura, Y. Suzuki et al., "The sequential accumulation of genetic alterations characteristic of the colorectal adenoma-carcinoma sequence does not occur between gastric

- adenoma and adenocarcinoma," *The Journal of Pathology*, vol. 176, no. 3, pp. 249–258, 1995.
- [12] K.-F. Becker, M. J. Atkinson, U. Reich et al., "E-cadherin gene mutations provide clues to diffuse type gastric carcinomas," *Cancer Research*, vol. 54, no. 14, pp. 3845–3852, 1994.
 - [13] Q.-L. Li, K. Ito, C. Sakakura et al., "Causal relationship between the loss of *RUNX3* expression and gastric cancer," *Cell*, vol. 109, no. 1, pp. 113–124, 2002.
 - [14] L. Chen, W.-M. Zeng, Y.-D. Cai, K.-Y. Feng, and K.-C. Chou, "Predicting anatomical therapeutic chemical (ATC) classification of drugs by integrating chemical-chemical interactions and similarities," *PLoS ONE*, vol. 7, no. 4, Article ID e35254, 2012.
 - [15] Y. Yamanishi, M. Araki, A. Gutteridge, W. Honda, and M. Kanehisa, "Prediction of drug-target interaction networks from the integration of chemical and genomic spaces," *Bioinformatics*, vol. 24, no. 13, pp. i232–i240, 2008.
 - [16] Y. Yamanishi, M. Kotera, M. Kanehisa, and S. Goto, "Drug-target interaction prediction from chemical, genomic and pharmacological data in an integrated framework," *Bioinformatics*, vol. 26, no. 12, pp. i246–i254, 2010.
 - [17] B. M. Padhy and Y. K. Gupta, "Drug repositioning: re-investigating existing drugs for new therapeutic indications," *Journal of Postgraduate Medicine*, vol. 57, no. 2, pp. 153–160, 2011.
 - [18] L. Chen, J. Lu, X. Luo, and K.-Y. Feng, "Prediction of drug target groups based on chemical-chemical similarities and chemical-chemical/protein connections," *Biochimica et Biophysica Acta*, vol. 1844, no. 1, part B, pp. 207–213, 2014.
 - [19] L. Chen, J. Lu, N. Zhang, T. Huang, and Y. Cai -D, "A hybrid method for prediction and repositioning of drug Anatomical Therapeutic Chemical classes," *Molecular BioSystems*, 2014.
 - [20] H.-W. Ma and A.-P. Zeng, "The connectivity structure, giant strong component and centrality of metabolic networks," *Bioinformatics*, vol. 19, no. 11, pp. 1423–1430, 2003.
 - [21] J. M. Dale, L. Popescu, and P. D. Karp, "Machine learning methods for metabolic pathway prediction," *BMC Bioinformatics*, vol. 11, article 15, 2010.
 - [22] L. Chen, T. Huang, X.-H. Shi, Y.-D. Cai, and K.-C. Chou, "Analysis of protein pathway networks using hybrid properties," *Molecules*, vol. 15, no. 11, pp. 8177–8192, 2010.
 - [23] T. Huang, L. Chen, Y.-D. Cai, and K.-C. Chou, "Classification and analysis of regulatory pathways using graph property, biochemical and physicochemical property, and functional property," *PLoS ONE*, vol. 6, no. 9, Article ID e25297, 2011.
 - [24] L. Chen, W.-M. Zeng, Y.-D. Cai, and T. Huang, "Prediction of metabolic pathway using graph property, chemical functional group and chemical structural set," *Current Bioinformatics*, vol. 8, no. 2, pp. 200–207, 2013.
 - [25] Y. F. Gao, L. Chen, Y. D. Cai, K. Y. Feng, T. Huang, and Y. Jiang, "Predicting metabolic pathways of small molecules and enzymes based on interaction information of chemicals and proteins," *PLoS ONE*, vol. 7, no. 9, Article ID e45944, 2012.
 - [26] L. Hu, T. Huang, X. Shi, W.-C. Lu, Y.-D. Cai, and K.-C. Chou, "Predicting functions of proteins in mouse based on weighted protein-protein interaction network and protein hybrid properties," *PLoS ONE*, vol. 6, no. 1, Article ID e14556, 2011.
 - [27] R. Sharan, I. Ulitsky, and R. Shamir, "Network-based prediction of protein function," *Molecular Systems Biology*, vol. 3, no. 1, article 88, 2007.
 - [28] K.-L. Ng, J.-S. Ciou, and C.-H. Huang, "Prediction of protein functions based on function-function correlation relations," *Computers in Biology and Medicine*, vol. 40, no. 3, pp. 300–305, 2010.
 - [29] P. Bogdanov and A. K. Singh, "Molecular function prediction using neighborhood features," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 7, no. 2, pp. 208–217, 2010.
 - [30] L. J. Jensen, M. Kuhn, M. Stark et al., "STRING 8—a global view on proteins and their functional interactions in 630 organisms," *Nucleic Acids Research*, vol. 37, supplement 1, pp. D412–D416, 2009.
 - [31] T. H. Gormen, C. E. Leiserson, R. L. Rivest, and C. Stein, Eds., *Introduction to Algorithms*, The MIT Press, Cambridge, Mass, USA, 1990.
 - [32] J. B. M. Craven, *Markov Networks for Detecting Overlapping Elements in Sequence Data*, The MIT Press, Cambridge, Mass, USA, 2005.
 - [33] J. Davis and M. Goadrich, "The relationship between precision-recall and ROC curves," in *Proceedings of the 23rd International Conference on Machine Learning (ICML '06)*, pp. 233–240, June 2006.
 - [34] R. Bunescu, R. Ge, R. J. Kate et al., "Comparative experiments on learning information extractors for proteins and their interactions," *Artificial Intelligence in Medicine*, vol. 33, no. 2, pp. 139–155, 2005.
 - [35] D. E. Johnson and G. H. I. Wolfgang, "Predicting human safety: screening and computational approaches," *Drug Discovery Today*, vol. 5, no. 10, pp. 445–454, 2000.
 - [36] B.-Q. Li, B. Niu, L. Chen et al., "Identifying chemicals with potential therapy of HIV based on protein-protein and protein-chemical interaction network," *PLoS ONE*, vol. 8, no. 6, Article ID e65207, 2013.
 - [37] L. Chen, B.-Q. Li, M.-Y. Zheng, J. Zhang, K.-Y. Feng, and Y.-D. Cai, "Prediction of effective drug combinations by chemical interaction, protein interaction and target enrichment of KEGG pathways," *BioMed Research International*, vol. 2013, Article ID 723780, 10 pages, 2013.
 - [38] J. Zhang, M. Jiang, F. Yuan et al., "Identification of age-related macular degeneration related genes by applying shortest path algorithm in protein-protein interaction network," *BioMed Research International*, vol. 2013, Article ID 523415, 8 pages, 2013.
 - [39] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources," *Nature Protocols*, vol. 4, no. 1, pp. 44–57, 2009.
 - [40] M. J. Clark, N. Homer, B. D. O'Connor et al., "U87MG decoded: the genomic sequence of a cytogenetically aberrant human cancer cell line," *PLoS Genetics*, vol. 6, no. 1, Article ID e1000832, 2010.
 - [41] P. Teekakirikul, S. Eminaga, O. Toka et al., "Cardiac fibrosis in mice with hypertrophic cardiomyopathy is mediated by non-myocyte proliferation and requires Tgf- β ," *The Journal of Clinical Investigation*, vol. 120, no. 10, pp. 3520–3529, 2010.
 - [42] C. Chu, K. Qu, F. Zhong, S. Artandi, and H. Chang, "Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions," *Molecular Cell*, vol. 44, no. 4, pp. 667–678, 2011.
 - [43] B. Kalverda, H. Pickersgill, V. V. Shloma, and M. Fornerod, "Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm," *Cell*, vol. 140, no. 3, pp. 360–371, 2010.
 - [44] Y. Mayshar, U. Ben-David, N. Lavon et al., "Identification and classification of chromosomal aberrations in human induced

- pluripotent stem cells," *Cell Stem Cell*, vol. 7, no. 4, pp. 521–531, 2010.
- [45] S. J. Sanders, A. G. Ercan-Sencicek, V. Hus et al., "Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism," *Neuron*, vol. 70, no. 5, pp. 863–885, 2011.
- [46] Y. Benjamini and D. Yekutieli, "The control of the false discovery rate in multiple testing under dependency," *The Annals of Statistics*, vol. 29, no. 4, pp. 919–1188, 2001.
- [47] P. Hohenberger and S. Gretschel, "Gastric cancer," *The Lancet*, vol. 362, no. 9380, pp. 305–315, 2003.
- [48] R. Liu, Z. Li, S. Bai et al., "Mechanism of cancer cell adaptation to metabolic stress: proteomics identification of a novel thyroid hormone-mediated gastric carcinogenic signaling pathway," *Molecular and Cellular Proteomics*, vol. 8, no. 1, pp. 70–85, 2009.
- [49] D. F. Sun, Y. J. Zhang, X. Q. Tian, Y. X. Chen, and J. Y. Fang, "Inhibition of mTOR signalling potentiates the effects of trichostatin A in human gastric cancer cell lines by promoting histone acetylation," *Cell Biology International*, vol. 38, no. 1, pp. 50–63, 2014.
- [50] E. Tahara, "Genetic pathways of two types of gastric cancer," *IARC Scientific Publications*, no. 157, pp. 327–349, 2004.
- [51] J. Essers, A. F. Theil, C. Baldeyron et al., "Nuclear dynamics of PCNA in DNA replication and repair," *Molecular and Cellular Biology*, vol. 25, no. 21, pp. 9350–9359, 2005.
- [52] O. Coqueret, "Linking cyclins to transcriptional control," *Gene*, vol. 299, no. 1–2, pp. 35–55, 2002.
- [53] J. K. Moore and J. E. Haber, "Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*," *Molecular and Cellular Biology*, vol. 16, no. 5, pp. 2164–2173, 1996.
- [54] Y. Liu, R. Prasad, W. A. Beard et al., "Coordination of steps in single-nucleotide base excision repair mediated by apurinic/apyrimidinic endonuclease 1 and DNA polymerase β ," *The Journal of Biological Chemistry*, vol. 282, no. 18, pp. 13532–13541, 2007.
- [55] C. Naro and C. Sette, "Phosphorylation-mediated regulation of alternative splicing in cancer," *International Journal of Cell Biology*, vol. 2013, Article ID 151839, 15 pages, 2013.
- [56] A. T. Look, "Oncogenic transcription factors in the human acute leukemias," *Science*, vol. 278, no. 5340, pp. 1059–1064, 1997.
- [57] Q.-L. Li, K. Ito, C. Sakakura et al., "Causal relationship between the loss of *RUNX3* expression and gastric cancer," *Cell*, vol. 109, no. 1, pp. 113–124, 2002.
- [58] M. K. K. Shivji, M. K. Kenny, and R. D. Wood, "Proliferating cell nuclear antigen is required for DNA excision repair," *Cell*, vol. 69, no. 2, pp. 367–374, 1992.
- [59] G.-L. Moldovan, B. Pfander, and S. Jentsch, "PCNA, the maestro of the replication fork," *Cell*, vol. 129, no. 4, pp. 665–679, 2007.
- [60] G. Prelich, C.-K. Tan, and M. Kostura, "Functional identity of proliferating cell nuclear antigen and a DNA polymerase- δ auxiliary protein," *Nature*, vol. 326, no. 6112, pp. 517–520, 1987.
- [61] C. Hoeghe, B. Pfander, G.-L. Moldovan, G. Pyrowolakis, and S. Jentsch, "RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO," *Nature*, vol. 419, no. 6903, pp. 135–141, 2002.
- [62] M. K. K. Shivji, "Nucleotide excision repair DNA synthesis by DNA polymerase ϵ in the presence of PCNA, RFC, and RPA," *Biochemistry*, vol. 34, no. 15, pp. 5011–5017, 1995.
- [63] J. de Boer and J. H. J. Hoeijmakers, "Nucleotide excision repair and human syndromes," *Carcinogenesis*, vol. 21, no. 3, pp. 453–460, 2000.
- [64] B. Pascucci, M. D'Errico, E. Parlanti, S. Giovannini, and E. Dogliotti, "Role of nucleotide excision repair proteins in oxidative DNA damage repair: an updating," *Biochemistry*, vol. 76, no. 1, pp. 4–15, 2011.
- [65] S. Ohshima, Y. Yonemura, and I. Miyazaki, "Proliferative activity and malignancy in human gastric cancers: significance of the proliferation rate and its clinical application," *Cancer*, vol. 69, no. 2, pp. 314–321, 1992.
- [66] R. G. Kuang, H. X. Wu, G. X. Hao, J. W. Wang, and C. J. Zhou, "Expression and significance of IGF-2, PCNA, MMP-7, and α -actin in gastric carcinoma with Lauren classification," *The Turkish Journal of Gastroenterology*, vol. 24, no. 2, pp. 99–108, 2013.
- [67] S. Jain, M. I. Filipe, P. A. Hall, N. Waseem, D. P. Lane, and D. A. Levison, "Prognostic value of proliferating cell nuclear antigen in gastric carcinoma," *Journal of Clinical Pathology*, vol. 44, no. 8, pp. 655–659, 1991.
- [68] J. Ji, P. Zhao, and B. Huang, "Study of gastric carcinoma and PCNA and c-met gene abnormality," *Wei Sheng Yan Jiu*, vol. 37, no. 4, pp. 479–482, 2008.
- [69] H. Takamura, Y. Yonemura, L. Fonseca et al., "Correlation of DNA ploidy, c-erbB-2 protein tissue status, level of PCNA expression and clinical outcome in gastric carcinomas," *Nihon Geka Gakkai Zasshi*, vol. 96, no. 4, pp. 213–222, 1995.
- [70] T. Motokura, T. Bloom, H. G. Kim et al., "A novel cyclin encoded by a bcl1-linked candidate oncogene," *Nature*, vol. 350, no. 6318, pp. 512–515, 1991.
- [71] W. Jiang, S. M. Kahn, N. Tomita, Y.-J. Zhang, S.-H. Lu, and I. B. Weinstein, "Amplification and expression of the human cyclin D gene in esophageal cancer," *Cancer Research*, vol. 52, no. 10, pp. 2980–2983, 1992.
- [72] M. F. Buckley, K. J. E. Sweeney, J. A. Hamilton et al., "Expression and amplification of cyclin genes in human breast cancer," *Oncogene*, vol. 8, no. 8, pp. 2127–2133, 1993.
- [73] J. K. Kim and J. A. Diehl, "Nuclear cyclin D1: an oncogenic driver in human cancer," *Journal of Cellular Physiology*, vol. 220, no. 2, pp. 292–296, 2009.
- [74] M. Malumbres, I. P. de Castro, M. I. Hernández, M. Jiménez, T. Corral, and A. Pellicer, "Cellular response to oncogenic ras involves induction of the Cdk4 and Cdk6 inhibitor p15(INK4b)," *Molecular and Cellular Biology*, vol. 20, no. 8, pp. 2915–2925, 2000.
- [75] M. Serrano, E. Gomez-Lahoz, R. A. DePinho, D. Beach, and D. Bar-Sagi, "Inhibition of ras-induced proliferation and cellular transformation by p16(INK4)," *Science*, vol. 267, no. 5195, pp. 249–252, 1995.
- [76] R. M. L. Zwijsen, E. Wientjens, R. Klompmaaker, J. van der Sman, R. Bernards, and R. J. A. M. Michalides, "CDK-independent activation of estrogen receptor by cyclin D1," *Cell*, vol. 88, no. 3, pp. 405–415, 1997.
- [77] P. C. Fernandez, S. R. Frank, L. Wang et al., "Genomic targets of the human c-Myc protein," *Genes and Development*, vol. 17, no. 9, pp. 1115–1129, 2003.
- [78] R. C. Sears, "The life cycle of c-Myc: from synthesis to degradation," *Cell Cycle*, vol. 3, no. 9, pp. 1133–1137, 2004.
- [79] H.-J. Chung and D. Levens, "c-myc expression: keep the noise down!," *Molecules and Cells*, vol. 20, no. 2, pp. 157–166, 2005.
- [80] E. V. Prochownik and Y. Li, "The ever expanding role for c-Myc in promoting genomic instability," *Cell Cycle*, vol. 6, no. 9, pp. 1024–1029, 2007.
- [81] A. N. Milne, R. Sitarz, R. Carvalho, F. Carneiro, and G. J. A. Offerhaus, "Early onset gastric cancer: on the road to unraveling

- gastric carcinogenesis," *Current Molecular Medicine*, vol. 7, no. 1, pp. 15–28, 2007.
- [82] C. Sakakura, T. Mori, T. Sakabe et al., "Gains, losses, and amplifications of genomic materials in primary gastric cancers analyzed by comparative genomic hybridization," *Genes Chromosomes Cancer*, vol. 24, no. 4, pp. 299–305, 1999.
- [83] S. Yang, H.-C. Jeung, H. J. Jeong et al., "Identification of genes with correlated patterns of variations in DNA copy number and gene expression level in gastric cancer," *Genomics*, vol. 89, no. 4, pp. 451–459, 2007.
- [84] M. Vita and M. Henriksson, "The Myc oncoprotein as a therapeutic target for human cancer," *Seminars in Cancer Biology*, vol. 16, no. 4, pp. 318–330, 2006.
- [85] A. V. Bakin and T. Curran, "Role of DNA 5-methylcytosine transferase in cell transformation by fos," *Science*, vol. 283, no. 5400, pp. 387–390, 1999.
- [86] I. Fialka, H. Schwarz, E. Reichmann, M. Oft, M. Busslinger, and H. Beug, "The estrogen-dependent c-JunER protein causes a reversible loss of mammary epithelial cell polarity involving a destabilization of adherens junctions," *Journal of Cell Biology*, vol. 132, no. 6, pp. 1115–1132, 1996.
- [87] E. Hu, E. Mueller, S. Oliviero, V. P. Papaioannou, R. Johnson, and B. M. Spiegelman, "Targeted disruption of the c-fos gene demonstrates c-fos-dependent and -independent pathways for gene expression stimulated by growth factors or oncogenes," *The EMBO Journal*, vol. 13, no. 13, pp. 3094–3103, 1994.
- [88] M. Mikula, J. Gotzmann, A. N. M. Fischer et al., "The proto-oncoprotein c-Fos negatively regulates hepatocellular tumorigenesis," *Oncogene*, vol. 22, no. 42, pp. 6725–6738, 2003.
- [89] S. P. Jin, J. H. Kim, M. A. Kim et al., "Prognostic significance of loss of c-fos protein in gastric carcinoma," *Pathology and Oncology Research*, vol. 13, no. 4, pp. 284–289, 2007.
- [90] L. Zhou, J.-S. Zhang, J.-C. Yu et al., "Negative association of c-fos expression as a favorable prognostic indicator in gastric cancer," *Archives of Medical Research*, vol. 41, no. 3, pp. 201–206, 2010.
- [91] N. N. Mazurenko, E. A. Kogan, N. M. Sukhova, and I. B. Zborovskaia, "Synthesis and distribution of oncoproteins in tumor tissue," *Voprosy Meditsinskoj Khimii*, vol. 37, no. 6, pp. 53–59, 1991.
- [92] W.-J. Song, M. G. Sullivan, R. D. Legare et al., "Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia," *Nature Genetics*, vol. 23, no. 2, pp. 166–175, 1999.
- [93] T. Usui, K. Aoyagi, N. Saeki et al., "Expression status of *RUNX1/AML1* in normal gastric epithelium and its mutational analysis in microdissected gastric cancer cells," *International Journal of Oncology*, vol. 29, no. 4, pp. 779–784, 2006.
- [94] C. Sakakura, A. Hagiwara, K. Miyagawa et al., "Frequent down-regulation of the runt domain transcription factors *RUNX1*, *RUNX3* and their cofactor *CBFB* in gastric cancer," *International Journal of Cancer*, vol. 113, no. 2, pp. 221–228, 2005.
- [95] H. Xue, P. Ni, B. Lin, H. Xu, and G. Huang, "X-ray repair cross-complementing group 1 (*XRCC1*) genetic polymorphisms and gastric cancer risk: a HuGe review and meta-analysis," *The American Journal of Epidemiology*, vol. 173, no. 4, pp. 363–375, 2011.
- [96] J. Jiang, Z. Jia, D. Cao et al., "Polymorphisms of the DNA methyltransferase 1 associated with reduced risks of *Helicobacter pylori* infection and increased risks of gastric atrophy," *PLoS ONE*, vol. 7, no. 9, Article ID e46058, 2012.
- [97] W.-L. Cheng, C.-S. Wang, Y.-H. Huang, M.-M. Tsai, Y. Liang, and K.-H. Lin, "Overexpression of CXCL1 and its receptor CXCR2 promote tumor invasion in gastric cancer," *Annals of Oncology*, vol. 22, no. 10, pp. 2267–2276, 2011.