

RESEARCH

Open Access

# Genetic variants and risk of gastric cancer: a pathway analysis of a genome-wide association study

Ju-Han Lee<sup>†</sup>, Younghye Kim<sup>†</sup>, Jung-Woo Choi and Young-Sik Kim<sup>\*</sup>

## Abstract

This study aimed to discover candidate single nucleotide polymorphisms (SNPs) for hypothesizing significant biological pathways of gastric cancer (GC). We performed an Identify Candidate Causal SNPs and Pathways (ICSNPPathway) analysis using a GC genome-wide association study (GWAS) dataset, including 472,342 SNPs in 2,240 GC cases and 3,302 controls of Asian ethnicity. By integrating linkage disequilibrium analysis, functional SNP annotation, and pathway-based analysis, seven candidate SNPs, four genes and 12 pathways were selected. The ICSNPPathway analysis produced 4 hypothetical mechanisms of GC: (1) rs4745 and rs12904 → *EFNA1* → ephrin receptor binding; (2) rs1801019 → *UMPS* → drug and pyrimidine metabolism; (3) rs364897 → *GBA* → cyanoamino acid metabolism; and (4) rs11187870, rs2274223, and rs3765524 → *PLCE1* → lipid biosynthetic process, regulation of cell growth, and cation homeostasis. This pathway analysis using GWAS dataset suggests that the 4 hypothetical biological mechanisms might contribute to GC susceptibility.

**Keywords:** Genome-wide association study; Pathway-based analysis; Gastric cancer

## Introduction

Despite a decline in its incidence, gastric cancer (GC) is still the second most common cause of cancer-related death worldwide (Hohenberger and Gretschel 2003). Furthermore, GC remains one of the most prevalent high-mortality cancers in Northeast Asia (Hohenberger and Gretschel 2003). *Helicobacter pylori* infection is the strongest risk factor for GC (Polk and Peek 2010), but only a small proportion of infected individuals develop malignancy. Thus, genetic factors such as polymorphisms in GC-related genes, in addition to dietary factors and environmental factors, substantially contribute to GC susceptibility (Milne et al. 2009).

Genome-wide association studies (GWAS) have proved successful in identifying associations between specific genes and complex diseases (Manolio 2010), and opened a new phase in researching the genetic causes of disease. Furthermore, GWAS datasets are increasingly being used

to recognize the biological pathways underlying complex diseases (Ramanan et al. 2012), because the functional pathway analysis using genomic datasets has high statistical power to detect the biological mechanisms of disease causation (Ramanan et al. 2012).

Recently, (Zhang et al. 2011a) developed the pathway analysis tool called Identify Candidate Causal SNPs and Pathways (ICSNPPathway) analysis. This method highlights the candidate SNPs and their corresponding candidate pathways from GWAS data by integrating linkage disequilibrium (LD) analysis, functional SNP annotation, and pathway-based analysis (PBA) (Zhang et al. 2011a). The ICSNPPathway analysis provides candidate SNPs and their corresponding candidate pathways using GWAS data, thereby making it easier to link variants to biological mechanisms.

We conducted ICSNPPathway analysis using a GC GWAS dataset available online to identify candidate SNPs and promising biological mechanisms that contribute to GC susceptibility.

\* Correspondence: apysk@korea.ac.kr

<sup>†</sup>Equal contributors

Department of Pathology, Korea University Ansan Hospital, 123, Jeokgeum-Ro, Danwon-Gu, Ansan-Si, Gyeonggi-Do 425-707, Republic of Korea

## Methods

### GWAS dataset

The GC GWAS dataset is publicly available from the NCBI dbGap (<http://www.ncbi.nlm.nih.gov/gap>). The dataset includes genotypes of 472,342 SNPs on Illumina 660 W Quad chip from 2,240 GC cases and 3,302 controls of Chinese ethnicity (Abnet et al. 2010; Li et al. 2013). Study participants were drawn from the Shanxi Upper Gastrointestinal Cancer Genetics Project and the Linxian Nutrition Intervention Trial, which included a total of 1,625 GC cases and 2,100 controls. Six hundred and fifteen GC cases and 1,202 controls from the Shanghai Men's Health Study, the Shanghai Women's Health Study, and the Singapore Chinese Health Study were also included in the database. Controls were matched for age ( $\pm 5$  years), sex, and geographical location and they were all cancer-free at the time of enrollment (Abnet et al. 2010; Li et al. 2013). The dataset was filtered to prevent genotyping errors. The SNPs were excluded if they showed a call rate lower than 90% in cases or controls or significant deviation from Hardy-Weinberg equilibrium in the controls ( $P < 10^{-4}$ ). Finally, 470,698 SNPs were left for downstream pathway analysis.

### ICSNPPathway analysis

We conducted ICSNPPathway analysis using the GC GWAS dataset in two-stages (Zhang et al. 2011a). First, candidate causal SNPs were pre-selected by LD analysis and the most significant functional SNPs were annotated. Next, biological mechanisms for the pre-selected candidate causal SNPs were found using PBA. A full list of GWAS SNP  $P$ -values was used for the ICSNPPathway analysis. The ICSNPPathway analysis is based on LD analysis and the discovery of functional SNPs using improved-gene set enrichment analysis (*i*-GSEA). The ICSNPPathway searches for SNPs in LD with the most significant SNPs in a GWAS dataset to identify more possible candidate causal SNPs based on the extended dataset, such as HapMap data. We selected the optional parameters for LD: Chinese Han in Beijing as the HapMap population; a cut-off for LD measurement of  $r^2 = 0.8$ ; and a maximum distance to search LD neighborhoods of 200 kb. The ICSNPPathway pre-selects candidate causal SNPs based on functional SNPs, which are defined as SNPs that may alter protein, gene expression or the role of protein in the context of the pathway. Functional SNPs include deleterious and non-deleterious, non-synonymous SNPs; SNPs leading to the loss or gain of a stop codon; SNPs resulting in a frame shift; SNPs located at essential splice sites; and SNPs in regulatory regions. The ICSNPPathway server detects pathway-associated traits in the full list of GWAS SNP  $P$ -values using *i*-GSEA.

The term "most significant SNPs" denotes SNPs with a  $P$ -value below a certain threshold. The  $P$ -value threshold

to extract the most significant SNPs can be specified using the GWAS SNP  $P$ -values. The ICSNPPathway analysis presents the most significant pathways from the original GWAS when a  $P$ -value threshold less than  $1 \times 10^{-4}$  is chosen. Two parameters were set for the analysis. The first parameter was 'within gene only' meaning that only the  $P$ -values of SNPs located within genes were utilized in the PBA algorithm. The second was a false discovery rate (FDR) cut-off of 0.05 for multiple testing corrections. The FDR, which is defined as the expected proportion of false positives among all significant tests, allows researchers to identify a set of positive candidates. We used cut-offs of 5 minimum and 100 maximum in order to avoid overly narrow or overly broad functional categories.

The ICSNPPathway analysis used four pathway databases including the Kyoto Encyclopedia of Genes and Genomes (<http://www.genome.jp/kegg/pathway.html>) (Kanehisa et al. 2010), BioCarta (<http://www.biocarta.com/genes/index.asp>), Gene Ontology biological process (<http://www.geneontology.org>) (Ashburner et al. 2000), and Gene Ontology molecular function (<http://www.broadinstitute.org/gsea/msigdb/index.jsp>). When a candidate SNP was not present on a particular genotyping array, proxy SNPs in LD for that candidate SNP were identified, based on observed LD patterns in HapMap. Therefore, SNP annotation and proxy search (SNAP) was performed, which is a tool for the identification and annotation of proxy SNPs using HapMap.

## Results

Using GWAS SNP  $P$ -values as inputs, the ICSNPPathway analysis identified seven candidate causal SNPs, four genes, and 12 candidate causal pathways ([http://ICSNPPathway.psych.ac.cn/getResult.do?tag=4904B29B74A51307E8CFD85B4466A802\\_1374721534621](http://ICSNPPathway.psych.ac.cn/getResult.do?tag=4904B29B74A51307E8CFD85B4466A802_1374721534621)) (Tables 1 and 2) (Figure 1). SNPs rs4745 and rs12904, which were not represented in the original GWAS, are in LD with rs4460629 ( $r^2 = 1.0$ ) ( $-\log_{10}(P) = 6.472$ ). SNPs rs1801019, rs364897 and rs11187870, which were not represented in the original GWAS, were in LD with rs4234221, rs4460629 and rs3781264, respectively ( $r^2 = 0.824, 0.924, 0.857$ ) ( $-\log_{10}(P) = 4.087, 6.472, 10.405$ , respectively). SNPs rs2274223 and rs3765524, which were represented in the original GWAS ( $-\log_{10}(P) = 8.633$  and  $8.556$ , respectively), were not in LD with any SNP.

The seven candidate causal SNPs and 12 candidate causal pathways provided four hypothetical biological mechanisms for GC: (1) rs4745 (non-synonymous coding) and rs12904 (regulatory region) to *EFNA1* gene to ephrin receptor binding; (2) rs1801019 (non-synonymous coding) to *Uridine Monophosphate Synthetase (UMPS)* gene to drug and pyrimidine metabolism; (3) rs364897 (non-synonymous coding) to *Glucocerebrosidase (GBA)* gene to cyanoamino acid and starch/sucrose metabolism; and (4)

**Table 1 Candidate causal SNPs and pathways**

Candidate SNP	Functional class	Gene	Candidate causal pathway*	$-\log_{10}(P)\dagger$	In LD with	$r^2$	D'	$-\log_{10}(P)\ddagger$
rs4745	Nonsynonymous coding	<i>EFNA1</i>	1	-	rs4460629	1.0	1.0	6.472
rs12904	Regulatory region	<i>EFNA1</i>	1	-	rs4460629	1.0	1.0	6.472
rs1801019	Nonsynonymous coding	<i>UMPS</i>	2,3	-	rs4234221	0.824	1.0	4.087
rs364897	Nonsynonymous coding	<i>GBA</i>	4,11	-	rs4460629	0.924	1.0	6.472
rs11187870	Regulatory region	<i>PLCE1</i>	5,6,7,8,9,10,12	-	rs3781264	0.857	1.0	10.405
rs2274223	Nonsynonymous coding	<i>PLCE1</i>	5,6,7,8,9,10,12	8.633	rs2274223	-	-	8.633
rs3765524	Nonsynonymous coding	<i>PLCE1</i>	5,6,7,8,9,10,12	8.556	rs3765524	-	-	8.556

SNP; single nucleotide polymorphism, LD; linkage disequilibrium.

\*numbers indicate the indexes of pathway ranked significance (false discovery rate).

† $-\log_{10}(P)$  values of candidate causal SNPs in the original genome wide association studies (GWAS), - denotes that this SNP is not represented in the original GWAS.

‡ $-\log_{10}(P)$  values of SNPs in LD with candidate causal SNPs in the original GWAS.

rs11187870 (regulatory region), rs2274223 (non-synonymous coding) and rs3765524 (non-synonymous coding) to *Phospholipase C epsilon 1 (PLCE1)* gene to growth, kidney development, cellular cation homeostasis, urogenital system development, lipid biosynthetic process, regulation of cell growth, and cation homeostasis (Tables 1 and 2).

## Discussion

The results of previous GWAS studies have suggested that the rs227423 at 10q23 (Abnet et al. 2010), rs1336107 at 5q13, rs9841504 at 3q13 (Shi et al. 2011), and rs2976392 and rs2294008 at 8q24 (Sakamoto et al. 2008) SNPs are associated with GC. Although individual GWAS has been successful in finding new susceptibility genes for various complex diseases, none of the GWAS datasets was analyzed to their full potential (Elbers et al. 2009). Individual GWAS data has focused on SNPs with high statistical significance, whereas the other many SNPs have received

little attention. Thus, pathway analysis of genomic data with gene set enrichment approach could highlight significant SNPs that have otherwise been hidden during gene- or SNP- based analysis (Ramanan et al. 2012; Elbers et al. 2009).

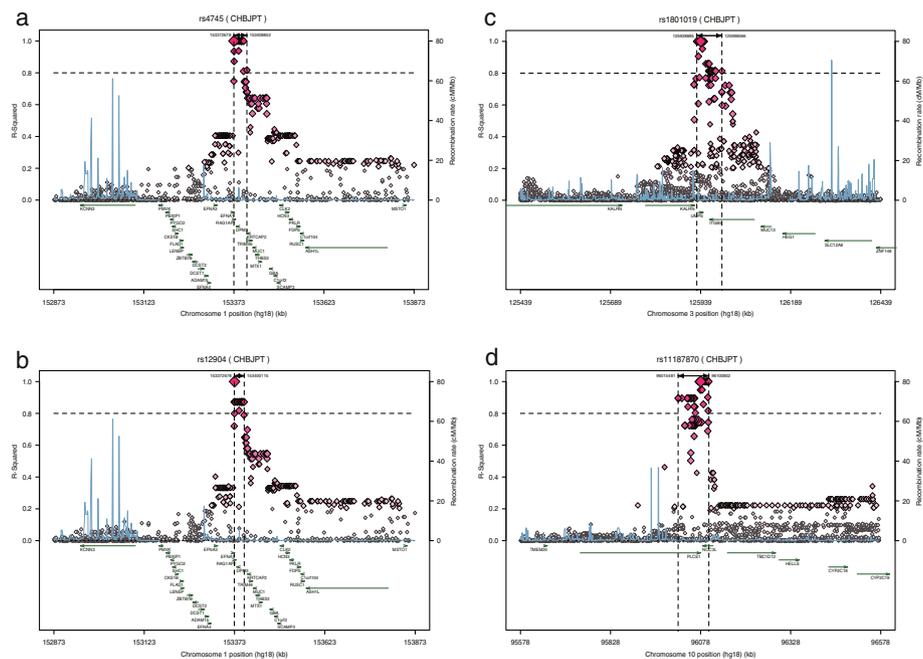
In this study, we identified seven candidate causal SNPs, four genes, and 12 candidate pathways by ICSN-Pathway analysis. The candidate SNPs and pathways provided four hypothetical biological mechanisms. In this genome-wide pathway analysis, the most significant GC-associated pathway was that of ephrin receptor binding.

The *EFNA1* (ephrinA1), located within chromosomal region 1q21-q22, is a glycosylphosphatidylinositol (GPI) linkage ligand with a 205-amino acid chain that preferentially binds to the receptor tyrosine kinase EphA2 at sites, where cell-cell contact occurs (Wykosky and Debinski 2008). Some studies have shown an association between GC and *EFNA1*. (Li et al. 2012) reported that

**Table 2 Candidate causal pathways**

Index	Candidate causal pathway	Description	Nominal P	FDR
1	Ephrin receptor binding	Interacting selectively and non-covalently with an ephrin receptor	<0.001	0.004
2	Hsa00983	Drug metabolism – other enzymes	<0.001	0.032
3	Hsa00240	Pyrimidine metabolism	0.004	0.032
4	Hsa00460	Cyanoamino acid metabolism	0.007	0.033
5	Growth	The increase in size or mass of an entire organism, a part of an organism or a cell	0.006	0.036
6	Kidney development	The process whose specific outcome is the progression of the kidney over time, from its formation to the mature structure	0.003	0.038
7	Cellular cation homeostasis	The regulation of the levels, transport, and metabolism of cations within a cell or between a cell and its external environment	0.009	0.039
8	Urogenital system development	The process whose specific outcome is the progression of the urogenital system over time, from its formation to the mature structure	<0.001	0.040
9	Lipid biosynthetic process	The chemical reactions and pathways resulting in the formation of lipids	0.006	0.040
10	Regulation of cell growth	Any process that modulates the frequency, rate or extent of cell growth	0.002	0.042
11	Hsa00500	Starch and sucrose metabolism	0.003	0.043
12	Cation homeostasis	The regulation of the levels, transport, and metabolism of cations	0.009	0.044

FDR; false discovery rate.



**Figure 1** Regional LD plots of rs4745 (*EFNA1*) (a), rs12904 (*EFNA1*) (b), rs1801019 (*UMPS*) (c), and rs11187870 (*PLCE1*) (d) SNPs. SNPs are plotted along with proxies (based on HapMap data on Chinese Han in Beijing) as a function of genomic location and are annotated by recombination rate across the locus (light-blue line). On the y-axis, pairwise  $r^2$  values are provided for each proxy SNP using color codes.

rs12904, located in the *EFNA1* gene, is significantly associated with GC risk. (Nakamura et al. 2005) also reported that *EFNA1* was overexpressed in 57% of GC tissue samples. Our pathway analysis suggests that the *EFNA1* gene and ephrin receptor binding pathway may play an important role in GC susceptibility.

The *UMPS* is a fundamental enzyme in pyrimidine synthesis (Gusella et al. 2011). A small-scale study in 23 GC patients indicated that *UMPS* polymorphisms are not related to cancer risk in Caucasian GC patients (Gusella et al. 2011). However, further studies are needed to clarify the effects of the candidate *UMPS* gene and *UMPS*-associated pathways on the development of GC.

The *GBA*, located in 1q21-22, encodes a glucocerebrosidase, which catalyzes the hydrolysis of glucocerebroside, a membrane glycolipid, to ceramide and glucose (Velayati et al. 2010). The association of *GBA* gene mutation with Gaucher's disease, Parkinson disease, or Lewy body disorder has been reported (Velayati et al. 2010). However, to the best of our knowledge, candidate *GBA* and its pathways have not been investigated in GC.

The *PLCE1* gene on chromosome 10q23 encodes an enzyme that catalyzes the hydrolysis of phosphatidylinositol-4,5-bisphosphate, generating the secondary messengers inositol 1,4,5-triphosphate and diacylglycerol, which participate in cell growth and differentiation (Bunney et al. 2009). Several studies have reported that the rs2274223 polymorphism in *PLCE1* is a risk factor for

GC in the Chinese Han population (Abnet et al. 2010; Zhang et al. 2011b; Wang et al. 2012). In addition, (Luo et al. 2011) suggested that GC patients with this SNP have a survival advantage. However, no significant association was observed between rs2274223 polymorphism and GC in a Caucasian population (Palmer et al. 2012; Kupcinskis et al. 2014).

Previous studies of GC GWAS have suggested that prostate stem cell antigen (*PSCA*) and *MUC1* SNPs are associated with GC risk (Kupcinskis et al. 2014; Rizzato et al. 2013). However, the pathway analysis in our study could not identify these SNPs. This discordance may be caused by several factors, such as ethnic diversity, differences in sample size or type, and variable GWAS array chips.

Several web servers for pathway analysis of GWAS have been offered: *i*-GSEA4GWAS (<http://gsea4gwas.psy.ch.ac.cn>), VEGAS (<https://vegas2.qimrberghofer.edu.au>) and DAVID (<http://david.abcc.ncifcrf.gov>). We selected the ICSNPathway because it is an updated version of *i*-GSEA4GWAS and has advantages in exploring candidate causal SNPs, genes, and disease associations (Zhang et al. 2011a).

The present ICSNPathway analysis has some limitations. First, incomplete annotation of the human genome is an important shortcoming of the pathway-based approach. In addition, imperfect knowledge of their genetic basis in complex diseases may decrease the ability of ICSNPathway

analysis to explore true causal SNPs and pathways. Subsequent replication studies are required to confirm the candidate SNPs, genes, and associated pathways (Jia et al. 2011). However, replication in independent datasets is beyond the scope of this study. Pathway analyses using GWAS datasets can be a useful tool in discovering novel genes that are associated with disease susceptibility (Ramanan et al. 2012; Zhang et al. 2011a).

It is rather arbitrary to determine the appropriate threshold for significant SNPs (Ramanan et al. 2012; Lambert et al. 2010). The cut-off levels for significant SNPs have ranged from  $P < 0.05$  to  $P < 5 \times 10^{-8}$  (Ramanan et al. 2012). (Lambert et al. 2010) reported that pathway analyses were performed using different levels of cut-off values ( $<0.01$ ,  $<0.001$ , or  $<0.0001$ ). However, there was little difference in finding significant pathways according to these cut-off values. We adopted  $P$  value of  $1 \times 10^{-4}$ , considering the number of significant SNPs which could be entered for the pathway analysis.

In conclusion, we carried out ICSNPPathway analysis using the GC GWAS data to evaluate genetic associations with GC at the SNP and pathway levels, considering that a pathway-based approach improves the results of individual SNP analyses of GWAS datasets. We identified seven candidate causal SNPs and four genes, and developed four hypotheses that possibly contribute to GC susceptibility. Further studies are needed to confirm and explore genetic variations of the hypothetical pathways underlying GC susceptibility.

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

JHL and YK collected data and wrote the manuscript draft. JWC collected data and discussed the study results. YSK designed the study and edited the manuscript. All authors read and approved the final manuscript.

#### Acknowledgements

The authors gratefully acknowledge investigators for sharing their valuable GWAS data. This study was supported by grants from Korea University (grant number: K1421351) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (grant number: 2013 R1A1A1058146).

Received: 28 November 2014 Accepted: 28 April 2015

Published online: 06 May 2015

#### References

- Abnet CC, Freedman ND, Hu N, Wang Z, Yu K, Shu XO, Yuan JM, Zheng W, Dawsey SM, Dong LM, Lee MP, Ding T, Qiao YL, Gao YT, Koh WP, Xiang YB, Tang ZZ, Fan JH, Wang C, Wheeler W, Gail MH, Yeager M, Yuenger J, Hutchinson A, Jacobs KB, Giffen CA, Burdett L, Fraumeni JF Jr, Tucker MA, Chow WH et al (2010) A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. *Nat Genet* 42(9):764–767, doi:10.1038/ng.649
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25(1):25–29, doi:10.1038/75556
- Bunney TD, Baxendale RW, Katan M (2009) Regulatory links between PLC enzymes and Ras superfamily GTPases: signalling via PLCepsilon. *Adv Enzyme Regul* 49(1):54–58, doi:10.1016/j.advenzreg.2009.01.004
- Elbers CC, van Eijk KR, Franke L, Mulder F, van der Schouw YT, Wijmenga C, Onland-Moret NC (2009) Using genome-wide pathway analysis to unravel the etiology of complex diseases. *Genet Epidemiol* 33(5):419–431, doi:10.1002/gepi.20395
- Gusella M, Bertolaso L, Bolzonella C, Pasini F, Padrini R (2011) Frequency of uridine monophosphate synthase Gly(213)Ala polymorphism in Caucasian gastrointestinal cancer patients and healthy subjects, investigated by means of new, rapid genotyping assays. *Genet Test Mol Biomarkers* 15(10):691–695, doi:10.1089/gtmb.2011.0021
- Hohenberger P, Gretschel S (2003) Gastric cancer. *Lancet* 362(9380):305–315, doi:S014067360313975X
- Jia P, Wang L, Meltzer HY, Zhao Z (2011) Pathway-based analysis of GWAS datasets: effective but caution required. *Int J Neuropsychopharmacol* 14(4):567–572, doi:10.1017/S1461145710001446
- Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M (2010) KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res* 38(Database issue):D355–D360, doi:10.1093/nar/gkp896
- Kupcinkas J, Wex T, Link A, Bartuseviciute R, Dedelaite M, Kevalaite G, Leja M, Skieceviciene J, Kiudelis G, Jonaitis L, Kupcinkas L, Malfertheiner P (2014) PSCA and MUC1 gene polymorphisms are linked with gastric cancer and pre-malignant gastric conditions. *Anticancer Res* 34(12):7167–7175
- Lambert JC, Grenier-Boley B, Chouraki V, Heath S, Zelenika D, Fievet N, Hannequin D, Pasquier F, Hanon O, Brice A, Epelbaum J, Berr C, Dartigues JF, Tzourio C, Campion D, Lathrop M, Amouyel P (2010) Implication of the immune system in Alzheimer's disease: evidence from genome-wide pathway analysis. *J Alzheimers Dis* 20(4):1107–1118, doi:10.3233/JAD-2010-100018
- Li Y, Nie Y, Cao J, Tu S, Lin Y, Du Y (2012) G-A variant in miR-200c binding site of EFNA1 alters susceptibility to gastric cancer. *Mol Carcinog* 53(3):219–229, doi:10.1002/mc.21966
- Li WQ, Hu N, Hyland PL, Gao Y, Wang ZM, Yu K, Su H, Wang CY, Wang LM, Chanock SJ, Burdett L, Ding T, Qiao YL, Fan JH, Wang Y, Xu Y, Shi JX, Gu F, Wheeler W, Xiong XQ, Giffen C, Tucker MA, Dawsey SM, Freedman ND, Abnet CC, Goldstein AM, Taylor PR (2013) Genetic variants in DNA repair pathway genes and risk of esophageal squamous cell carcinoma and gastric adenocarcinoma in a Chinese population. *Carcinogenesis* 34(7):1536–1542, doi:10.1093/carcin/bgt094
- Luo D, Gao Y, Wang S, Wang M, Wu D, Wang W, Xu M, Zhou J, Gong W, Tan Y, Zhang Z (2011) Genetic variation in PLCE1 is associated with gastric cancer survival in a Chinese population. *J Gastroenterol* 46(11):1260–1266, doi:10.1007/s00535-011-0445-3
- Manolio TA (2010) Genomewide association studies and assessment of the risk of disease. *N Engl J Med* 363(2):166–176, doi:10.1056/NEJMr0905980
- Milne AN, Carneiro F, O'Morain C, Offerhaus GJ (2009) Nature meets nurture: molecular genetics of gastric cancer. *Hum Genet* 126(5):615–628, doi:10.1007/s00439-009-0722-x
- Nakamura R, Kataoka H, Sato N, Kanamori M, Ihara M, Igarashi H, Ravshanov S, Wang YJ, Li ZY, Shimamura T, Kobayashi T, Konno H, Shinmura K, Tanaka M, Sugimura H (2005) EPHA2/EFNA1 expression in human gastric cancer. *Cancer Sci* 96(1):42–47, doi:10.1111/j.1349-7006.2005.00007.x
- Palmer AJ, Lochhead P, Hold GL, Rabkin CS, Chow WH, Lissowska J, Vaughan TL, Berry S, Gammon M, Risch H, El-Omar EM (2012) Genetic variation in C20orf54, PLCE1 and MUC1 and the risk of upper gastrointestinal cancers in Caucasian populations. *Eur J Cancer Prev* 21(6):541–544, doi:10.1097/CEJ.0b013e3283529b79
- Polk DB, Peek RM Jr (2010) *Helicobacter pylori*: gastric cancer and beyond. *Nat Rev Cancer* 10(6):403–414, doi:10.1038/nrc2857
- Ramanan VK, Shen L, Moore JH, Saykin AJ (2012) Pathway analysis of genomic data: concepts, methods, and prospects for future development. *Trends Genet* 28(7):323–332, doi:10.1016/j.tig.2012.03.004
- Rizzato C, Kato I, Plummer M, Muñoz N, Canzian F (2013) Genetic variation in PSCA and risk of gastric advanced preneoplastic lesions and cancer in relation to *Helicobacter pylori* infection. *PLoS One* 8(9):e73100, doi:10.1371/journal.pone.0073100
- Sakamoto H, Yoshimura K, Saeki N, Katai H, Shimoda T, Matsuno Y, Saito D, Sugimura H, Tanioka F, Kato S, Matsukura N, Matsuda N, Nakamura T, Hyodo I, Nishina T, Yasui W, Hirose H, Hayashi M, Toshiro E, Ohnami S, Sekine A, Sato Y, Totsuka H, Ando M, Takemura R, Takahashi Y, Ohdaira M, Aoki K, Honmyo I, Chiku S et al (2008) Genetic variation in PSCA is associated with susceptibility to diffuse-type gastric cancer. *Nat Genet* 40(6):730–740, doi:10.1038/ng.152

- Shi Y, Hu Z, Wu C, Dai J, Li H, Dong J, Wang M, Miao X, Zhou Y, Lu F, Zhang H, Hu L, Jiang Y, Li Z, Chu M, Ma H, Chen J, Jin G, Tan W, Wu T, Zhang Z, Lin D, Shen H (2011) A genome-wide association study identifies new susceptibility loci for non-cardia gastric cancer at 3q13.31 and 5p13.1. *Nat Genet* 43(12):1215–1218, doi:10.1038/ng.978
- Velayati A, Yu WH, Sidransky E (2010) The role of glucocerebrosidase mutations in Parkinson disease and Lewy body disorders. *Curr Neurol Neurosci Rep* 10(3):190–198, doi:10.1007/s11910-010-0102-x
- Wang M, Zhang R, He J, Qiu L, Li J, Wang Y, Sun M, Yang Y, Wang J, Yang J, Qian J, Jin L, Ma H, Wei Q, Zhou X (2012) Potentially functional variants of PLCE1 identified by GWASs contribute to gastric adenocarcinoma susceptibility in an eastern Chinese population. *PLoS One* 7(3):e31932, doi:10.1371/journal.pone.0031932
- Wykosky J, Debinski W (2008) The EphA2 receptor and ephrinA1 ligand in solid tumors: function and therapeutic targeting. *Mol Cancer Res* 6(12):1795–1806, doi:10.1158/1541-7786.MCR-08-0244
- Zhang K, Chang S, Cui S, Guo L, Zhang L, Wang J (2011a) ICSNPathway: identify candidate causal SNPs and pathways from genome-wide association study by one analytical framework. *Nucleic Acids Res* 39(Web Server issue):W437–W443, doi:10.1093/nar/gkr391
- Zhang H, Jin G, Li H, Ren C, Ding Y, Zhang Q, Deng B, Wang J, Hu Z, Xu Y, Shen H (2011b) Genetic variants at 1q22 and 10q23 reproducibly associated with gastric cancer susceptibility in a Chinese population. *Carcinogenesis* 32(6):848–852, doi:10.1093/carcin/bgr051

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)

---