

RESEARCH

Open Access



# Differential networking meta-analysis of gastric cancer across Asian and American racial groups

Wentao Dai<sup>1,4</sup>, Quanxue Li<sup>1,2</sup>, Bing-Ya Liu<sup>3</sup>, Yi-Xue Li<sup>1,2,4,5\*</sup> and Yuan-Yuan Li<sup>1,2,4,5\*</sup>

From The 11th International Conference on Systems Biology (ISB 2017)  
Shenzhen, China. 18-21 August 2017

## Abstract

**Background:** Gastric Carcinoma is one of the most lethal cancer around the world, and is also the most common cancers in Eastern Asia. A lot of differentially expressed genes have been detected as being associated with Gastric Carcinoma (GC) progression, however, little is known about the underlying dysfunctional regulation mechanisms. To address this problem, we previously developed a differential networking approach that is characterized by involving differential coexpression analysis (DCEA), stage-specific gene regulatory network (GRN) modelling and differential regulation networking (DRN) analysis.

**Result:** In order to implement differential networking meta-analysis, we developed a novel framework which integrated the following steps. Considering the complexity and diversity of gastric carcinogenesis, we first collected three datasets (GSE54129, GSE24375 and TCGA-STAD) for Chinese, Korean and American, and aimed to investigate the common dysregulation mechanisms of gastric carcinogenesis across racial groups. Then, we constructed conditional GRNs for gastric cancer corresponding to normal and carcinoma, and prioritized differentially regulated genes (DRGs) and gene links (DRLs) from three datasets separately by using our previously developed differential networking method. Based on our integrated differential regulation information from three datasets and prior knowledge (e.g., transcription factor (TF)-target regulatory relationships and known signaling pathways), we eventually generated testable hypotheses on the regulation mechanisms of two genes, XBP1 and GIF, out of 16 common cross-racial DRGs in gastric carcinogenesis.

**Conclusion:** The current cross-racial integrative study from the viewpoint of differential regulation networking provided useful clues for understanding the common dysfunctional regulation mechanisms of gastric cancer progression and discovering new universal drug targets or biomarkers for gastric cancer.

**Keywords:** Gastric carcinoma (GC), Differential networking meta-analysis, Cross-racial, Conditional gene regulatory networks (GRN), Dysfunctional regulation mechanisms

\* Correspondence: [yxli@scbit.org](mailto:yxli@scbit.org); [yyl@scbit.org](mailto:yyl@scbit.org)

<sup>1</sup>Shanghai Center for Bioinformation Technology, 1278 Keyuan Road, Shanghai 201203, People's Republic of China

Full list of author information is available at the end of the article



## Background

Gastric Carcinoma (GC) is one of the most common and lethal tumor around the world, which is characterized by high heterogeneity, easy metastasis, and poor prognosis [1, 2]. The morbidity and mortality of gastric carcinoma in Eastern Asia are much higher than the world average level [3]. During the last decade, quite a series of high-throughput profiling, including genetic variation study [4–7], genome-wide association study (GWAS) [7], gene expression analysis [8–10], epigenetic variation study [11], and integrative genomic analysis [12–15], have greatly help to understand the biology of GC, and identified quite a lot of GC-associated genes. It was noticed that the above studies always collected GC samples from a certain racial group such as Chinese [7, 12, 16], Korean [8], and American [14, 17], respectively, while the common dysregulation mechanisms of gastric carcinogenesis across racial groups has been paid little attention due to lack of integration research based on cross-racial GC datasets.

It has been widely accepted that cancer results from the dysregulation of multiple fundamental cell processes including proliferation, differentiation, migration, apoptosis, and so on [18], which could be captured by gene regulatory network (GRN) modelling, a widely used approach to explore the pathogenesis of complex diseases from the systemic aspect [19–21]. In recent years, a novel theme “differential networking” was put forward and a number of methods have been developed to identify the regulators, the relationships, and even the sub-networks that differ between phenotypes [20–23]. With the rapid accumulation of transcriptomic data, differential network analysis is helpful to survey the dynamics of gene regulation, which is crucial to the understanding of pathophysiological processes [24–26]. In our previous studies, we designed and implemented a differential co-expression analysis (DCEA) approach called DCGL to recognize differential co-expression genes (DCGs) and links (DCLs) in a link-based quantitative way [27–29]. Based on this methodology, we further developed a differential regulation networking (DRN) framework [30, 31], which built conditional gene regulatory network (GRN) or combinatorial GRN (cGRN) and then prioritized differentially regulated genes (DRGs) and links (DRLs). Our DRN strategy proves to substantially reduce the computational burden and leads to insightful comments on selecting subject related genes and their differential regulation mechanisms underlying phenotypic changes.

Based on the previous methodologies, the current study aimed at investigating the common dysregulation mechanisms of gastric carcinogenesis across Chinese, Korean and American. To this end, we constructed a novel integrative analysis framework from the viewpoint

of differential regulation, which integrated a variety of modules with differential regulation networking (DRN) analysis and integrative analysis as the core steps. First of all, the conditional gene regulation networks (GRN) were built from three Gastric Carcinoma datasets (GSE54129, GSE24375 and TCGA-STAD) separately, and differentially regulated genes (DRGs) and differentially regulated links (DRLs) were prioritized then. It was found that known cancer genes and drug targets are significantly ranked higher, and most of top-10 DRGs from the three datasets have been reported to be GC related (~60%), or cancer related (~90%); meanwhile, there is a lack of consistency among the three top DRG lists. By integrating DRGs and DRLs from three datasets to the prior regulation knowledge, it was found that the 16 common DRGs across racial groups are mainly located in the transcription factor complex in nucleus and their functions were enriched in transcriptional regulation of RNA polymerase II, transcriptional activator activity and transcription factor binding. We therefore proposed two common cross-racial DRGs (GIF, XBP1) and their related regulation relationships which might play crucial roles in the dysregulation mechanisms of gastric carcinogenesis. This integration analysis of GC across racial groups provided useful clues for understanding common dysfunctional regulation mechanisms of gastric carcinogenesis and discovering new universal drug targets or biomarkers for gastric cancer, and also indicated the complexity and diversity of gastric carcinogenesis as well.

## Methods

### Gene expression datasets

The Affymetrix GeneChip Operating System (GCOS) was used to measure expression level of 111 Chinese gastric carcinoma samples and 21 Chinese normal mucosa samples (GSE54129). The raw expression datasets were normalized by robust RMA method and log<sub>2</sub> transformed. The evaluation of GSE54129 raw data in terms of expression level distribution, density distribution and correlations of samples were shown in Additional file 1: Figure S1. After mapping probe sets to Gene Symbols based on their platform annotations, 20,307 unique genes were obtained.

We also downloaded the mRNA expression dataset of stomach adenocarcinoma (STAD) from The Cancer Genome Atlas (TCGA) Data Portal (<https://cancergenome.nih.gov/>), which contains sequenced 29 matched American tumor-normal pairs with Illumina HiSeq platform. After discarding genes with more than 20% missing values, we got 19,211 RPKM normalized and log<sub>2</sub> transformed unique genes.

The normalized gene expression profile of Korean gastric carcinoma GSE24375 [7] was downloaded from

Gene Expression Omnibus (GEO) and all measurements were  $\log_2$  transformed. The dataset includes eight patient-matched gastric normal mucosa, adenoma and carcinoma samples and two additional carcinoma samples. Probe sets with more than 20% missing values were discarded, while probe sets with less missing values were filled up with KNN method. After probe sets filtering, 18,468 probe sets were mapped to Gene Symbols based on their platform annotations and 12,658 unique genes were obtained. The distributions of normalized and  $\log_2$  transformed expression levels of genes in three datasets are presented in Additional file 1: Figure S2.

#### Enrichment analysis: Function, pathway, cancer genes and drug targets

The Database for Annotation, Visualization and Integrated Discovery (DAVID) [32] was used to identify over-represented KEGG pathways and GO terms based on the hypergeometric distribution with  $p$ -values  $< 0.05$  were considered statistically significant.

A total of 486 cancer genes and 2093 drug targets were downloaded from Cancer Gene Census (<http://cancer.sanger.ac.uk/cancergenome/projects/census/>) and Drug-Bank (<http://www.drugbank.ca/>), respectively, which were used to validate the differential regulation analysis on the three gene expression datasets.

#### Differential networking meta-analysis framework

In order to implement differential networking meta-analysis, we developed a novel framework which integrated a variety of modules as follows (outlined in Additional file 1: Figure S3).

Differentially co-expressed genes (DCGs) and links (DCLs) were identified with our previously developed differential coexpression analysis (DCEA) methods [27, 28]. DCGs with  $p$ -values less than 0.05 were selected by DCP method, and DCLs were picked out by DCE method with LFC model in DCGL package [28].

The conditional gene regulatory networks (GRN) were constructed for the three preprocessed datasets respectively based on DCGs and DCLs by using the conditional GRN modelling approach developed in our previous work [30]. First, we applied DCGL package to the expression dataset (GSE54129, GSE24375 and TCGA-STAD, respectively) to extract differentially coexpressed genes (DCGs) and differentially coexpressed gene links (DCLs) between normal and cancer. The DCGs and the gene pairs in the DCLs which involved at least one DCG between normal and carcinoma were selected as core seeds for the construction of conditional GRNs. Then we constructed the conditional GRNs based on forward predicted TF-target regulatory relationships and the core seed genes by using stepwise linear regression according to our previous method [30]. In this way, we built three

normal and cancer GRN pairs corresponding to the three expression datasets (GSE54129, GSE24375 and TCGA-STAD) separately.

The differential regulated genes (DRGs) and links (DRLs) in conditional GRNs were ranked by our previously developed quantitatively methods [30], DR measure and modified LFC model, respectively. The power of the above two methods in cancer genes and drug targets prioritization have been validated strictly in our previous work. In order to further test the power of prioritization of DRGs and DRLs methods in the three GRN pairs from 3 GC gene expression datasets, we carried out permutation tests by randomly perturbing the DRGs lists for 5000 times for each gene expression dataset as similar as the method described previously [30, 33].

The integrative analysis was carried out based on both genes and functions. In order to prioritize common DRGs across gastric cancer datasets, we combined DRG lists from the three GC datasets to select cross-racial DRGs which listed in the top  $N$  ( $< 100$ ) DRGs commonly in at least two racial groups. To globally understand the common functional relevance of differential regulation across all the three GC datasets, we performed GO/KEGG functional enrichment analysis on both top 10 DRGs out of every single dataset and common cross-racial DRGs by using DAVID 6.8.

## Results

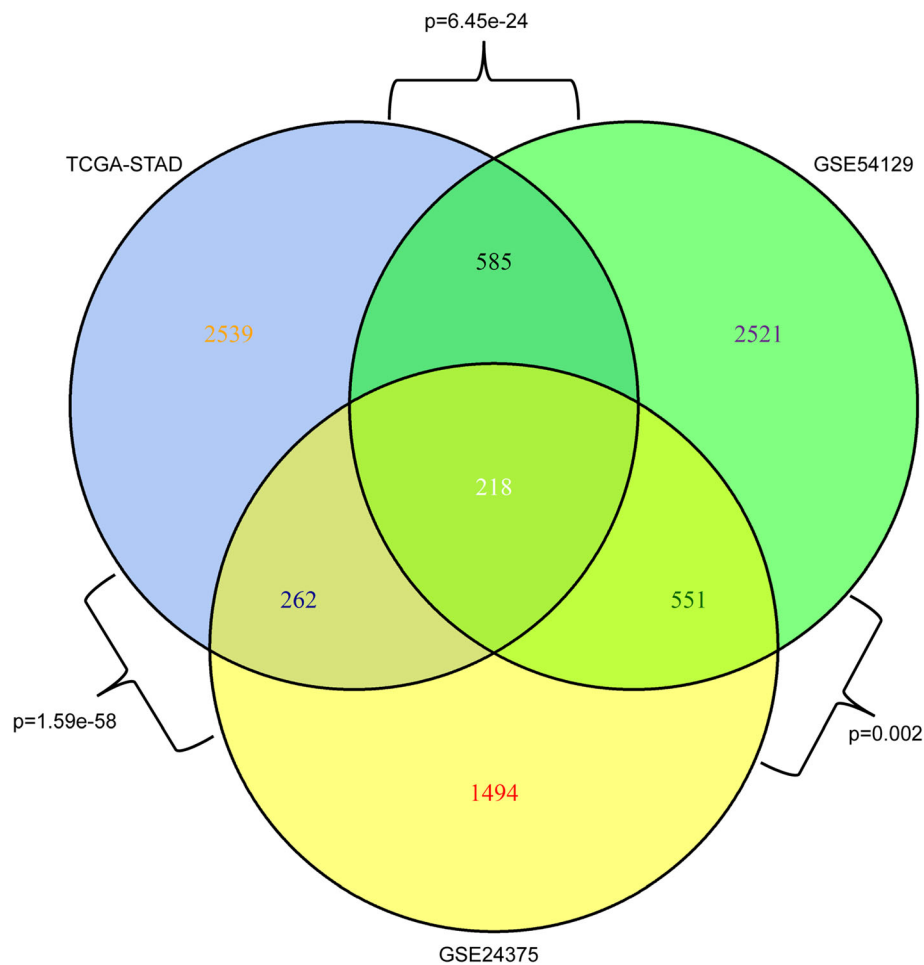
### Screening for differentially co-expressed genes

According to DCGL method [28], the genes with the  $p$ -value of dC less than 0.05 were taken as DCGs between normal and carcinoma. A total of 3875 DCGs were selected from GSE54129 dataset; 3604 DCGs from TCGA-STAD dataset; and 2524 DCGs from GSE24375 dataset.

The intersection and enrichment significance among three DCG sets (Fig. 1) shows that the DCGs between every two datasets are significantly overlapped by Kappa test. This suggests a basically coherent profile of molecular interactions underlying gastric carcinogenesis across racial groups included in this study, Chinese, Korean, and American.

### Construction of GRNs and identification of DRGs and DRLs

Based on the (normal and carcinoma) expression data of the selected DCG sets, we built paired conditional (normal and carcinoma) GRNs (Fig. 2) respectively corresponding to Chinese (GSE54129), Korean (GSE24375) and American (TCGA-STAD) by using the method described in the section of Materials and methods. The statistics of the three pairs of GRNs are listed in Table 1. It was found that the paired GRNs, i.e., normal and carcinoma GRNs for a certain dataset, always share the same regulators and most of the target genes, however,



**Fig. 1** The intersection and enrichment significance among three DCG sets selected from GSE54129 (Chinese, yellow), GSE24375 (Korean, green) and TCGA-STAD (American, cyan) gastric cancer datasets

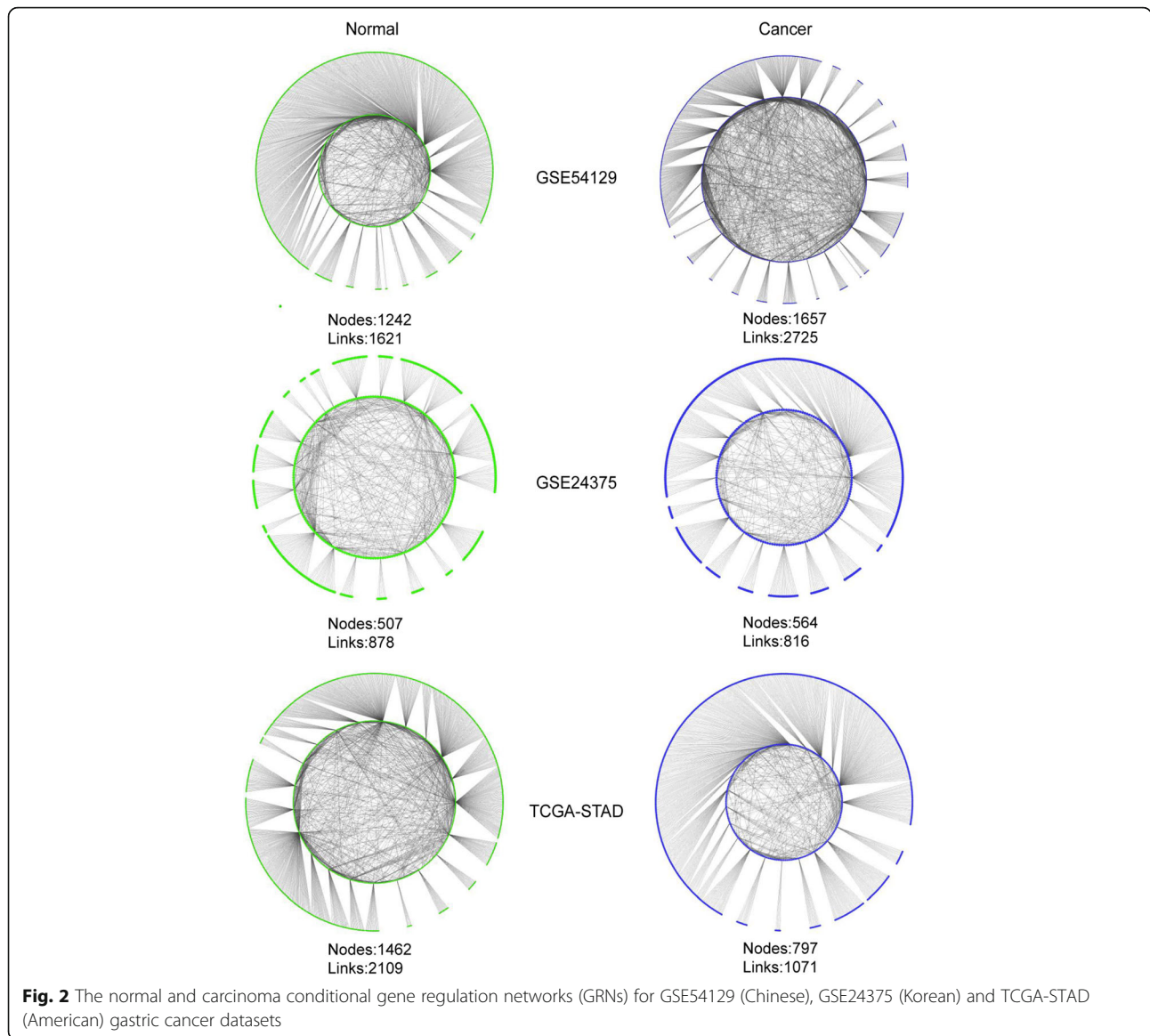
they share only a few of links, indicating that our analysis method did efficiently narrow down the search space and build stage-specific networks enriching subject relevant regulation relationships as expected [30, 31]. We then checked the global topological change between normal and cancer GRNs across the three datasets. According to the node number, the network size for Chinese (GSE54129) and Korean (GSE24375) expanded from normal to cancer, while that for American (TCGA-STAD) shrank. According to the average degree of nodes, the network complexity for Chinese (GSE54129) increased from normal to cancer, while that for Korean (GSE24375) and American (TCGA-STAD) decreased. Considering the inconsistency of the sample size of the three datasets, the above observation need to be further evaluated based on data from larger patient populations.

As shown in Table 2, known cancer genes and drug targets were enriched in all the conditional GRNs by Fisher's Exact Test, demonstrating that our conditional

GRNs have the potential to highlight crucial cancer-related regulation relationships, thus proving the rationality of the three GRN pairs.

After constructing conditional GRNs, a key issue is to quantitatively analyze the dynamic changes of gene regulation during phenotypic changes, i.e., from normal to carcinoma in the current work. Two methods were used to measure the differential regulation of a specific gene or gene link between two conditional GRNs as described in the section of Materials and methods, based on which differentially regulated genes (DRGs) and differentially regulated gene pairs or links (DRLs) were prioritized.

Permutation test showed that gastric cancer genes were significantly ranked higher in the DRG lists from three GC expression datasets, with  $p$ -values of  $1.36e-05$  (GSE24375),  $6.54e-53$  (GSE54129) and  $0.00793$  (TCGA-STAD). The rank of drug targets in these DRG lists presented a similar trend, though not significant in datasets GSE24375 and TCGA-STAD. This indicates that DR



ranking is appropriate to prioritize disease related genes in the current conditional GRNs.

The top 10 DRGs identified from three GC expression datasets were listed in Table 3. Among the top 10 DRGs for Chinese (GSE54129), six genes have been reported to be gastric cancer (GC) related (REG4 [34], ANXA13 [35], C7 [36], ASS1 [37], MSMB [38], CREB1 [39]) and the rest four were directly regulated by known gastric

cancer genes in our GRNs, in which two genes are also cancer related (BPTF [40], CLCA1 [41]). Among the top 10 DRGs identified for Korean (GSE24375), six genes are GC related (ESRRG [42], LIMS1 [43, 44], GATA3 [45], GATA6 [46], SOX9 [47], POU2F1 [48]) and the other four are cancer related (IRF2 [49], RGS3 [50], MRPL36 [51], FOSB [52]). Among the top 10 DRGs identified from TCGA-STAD dataset, six are GC related

**Table 1** Statistics of conditional GRNs from three datasets

GRNs	GSE54129			TCGA-STAD			GSE24375		
	Links (#)	TFs (#)	Targets (#)	Links (#)	TFs (#)	Targets (#)	Links (#)	TFs (#)	Targets (#)
Normal	1621	30	1242	2109	21	1462	878	22	507
Carcinoma	2725	30	1657	1071	21	797	816	22	564
shared by normal and carcinoma GRNs	723	30	972	760	21	718	303	22	331

**Table 2** The enrichment of cancer genes and drug targets in conditional GRNs

GRNs	GSE54129		TCGA		GSE24375	
	Cancer Genes	Drug Targets	Cancer Genes	Drug Targets	Cancer Genes	Drug Targets
Normal	8.40E-10	2.84E-13	0.00012	6.776e-08	0.0034	0.039
Carcinoma	2.48E-12	2.21E-10	2.08E-6	4.841e-07	0.0015	0.0002

Enrichment significance ( $p$ -value) was calculated by Fisher Exact Test

(E2F1 [50], AK1B10 [51], CEBPA [52], PTK7 [53], RASAL1 [54], GKN [55]), and three are cancer related (DPCR1 [56], GGT6 [57], RAB25 [58]). Additionally, 873 DRLs, 590 DRLs and 319 DRLs were identified for Chinese (GSE54129), American (TCGA-STAD) and Korean (GSE24375), respectively (Additional file 2: Table S1). We noticed that the vast majority of top 10 DRGs are gastric cancer related in all three datasets, while the three gene lists do not overlap at all, suggesting that the most crucial factors during gastric carcinogenesis may vary across racial groups. We therefore carried out the following integrative analysis aiming to identify cross-racial dysregulation mechanisms underlying gastric carcinogenesis.

#### Integrative analysis of cross-racial DRGs and the proposed mechanisms

In order to prioritize common DRGs across gastric cancer datasets, we combined the three DRG lists from the three GC datasets in Table 4. Since there are no common DRGs within the top 50 DRG lists among the three datasets, we first checked the overlap between every two datasets. As shown in Table 4, CEBPA, GATA6, GATA3 and GIF genes appear in the top 30 DRGs commonly in at least two different GC datasets, i.e., two racial groups; similarly, GIF, XBP1, CEBPA, GATA6, DPP4, PTK7 and GATA3 appear in the top 50 DRGs. There are 16 cross-racial DRGs in the top 100 DRGs commonly in at least two racial groups, including GIF, XBP1, CEBPA, GATA6,

DPP4, PTK7, GATA3, FOSB, EPAS1, CCND2, MALL, IRF1, SOX9, VILL, GALNT3 and LGALS4. The number of TOP N ( $\leq 100$ ) overlapping DRGs across racial groups are not significantly different as shown in Table 4. Meanwhile, we also noticed that among a total number of 6399 DCGs for Chinese and Korean, 12% (769) were shared by the two DCG lists; among 7479 DCG for American and Korean, 10.7% (803) were shared; while among 6128 DCG for Chinese and American, 7.8% (480) were shared (Fig. 1). This suggests that Chinese and Korean have similar gastric carcinogenesis, compared with American.

To globally understand the common functional relevance of differential regulation across the three GC datasets, we performed GO/KEGG functional enrichment analysis on the top ranked DRGs (as shown in Tables 3 and 4) by using DAVID 6.8 [32]. First, the TOP 10 DRGs from three GC dataset (GSE54129, GSE24375 and TCGA-STAD) were analyzed separately. However, no GO/KEGG terms were significantly enriched in any of the three gene sets after multiple hypothesis test correction. Secondly, the terms “transcription regulatory region DNA binding” and “RNA polymerase II transcription factor binding” were identified based on seven common TOP-50 DRGs (GIF, XBP1, CEBPA, GATA6, DPP4, PTK7 and GATA3) between at least two GC datasets. Similarly, based on 16 common TOP-100 DRGs, the terms of “transcription from RNA polymerase II promoter”, “positive regulation of transcription from RNA

**Table 3** The top 10 ranked genes from three datasets

TCGA-STAD			GSE54129			GSE24375		
DRGs	DR_value	rank	DRGs	DR_value	rank	DRGs	DR_value	rank
<b>E2F1</b>	18.45799	1	SLC7A9	65.03438	1	<b>LIMS1</b>	3401.341	1
<b>AKR1B10</b>	15.58216	2	<b>REG4</b>	22.26006	2	<i>FOSB</i>	582.8947	2
<b>CEBPA</b>	12.34533	3	<i>BPTF</i>	20.02962	3	<i>MRPL36</i>	261.5739	3
<b>PTK7</b>	11.85102	4	DHRS11	19.11607	4	<b>ESRRG</b>	247.57	4
PLAU	11.5865	5	<b>ANXA13</b>	17.92791	5	<b>GATA3</b>	132.0578	5
<i>DPCR1</i>	11.37501	6	<b>C7</b>	15.36403	6	<b>GATA6</b>	85.31825	6
<i>GGT6</i>	11.14235	7	<b>ASS1</b>	15.10191	7	<b>SOX9</b>	47.91384	7
RAB25	9.645954	8	<b>MSMB</b>	13.03145	8	<b>POU2F1</b>	41.98352	8
<b>RASAL1</b>	9.191246	9	<i>CLCA1</i>	12.91625	9	<i>IRF2</i>	37.68406	9
<b>GKN2</b>	8.602126	10	<b>CREB1</b>	12.84939	10	<i>RGS3</i>	35.60814	10

The genes are sorted by DR value. Genes in bold refer to GC-related genes; genes in italic refer to cancer-related genes

**Table 4** Intersection of top ranked DRGs between every two datasets

	Top10	Top20	Top30	Top40	Top50	Top100
<b>TCGA-STAD vs GSE54129</b>	0	CEBPA	<b>CEBPA</b>	CEBPA	CEBPA	CEBPA/GIF/XBP1
<b>GSE54129 vs GSE24375</b>	0	GATA6	<b>GATA6</b>	GATA6	GATA6/DPP4	GATA6/GIF/FOSB/XBP1/ DPP4/EPAS1/CCND2/MALL
<b>TCGA-STAD vs GSE24375</b>	0	0	<b>GATA3/GIF</b>	GIF/GATA3/PTK7	GIF/GATA3/PTK7/XBP1	GIF/GATA3/PTK7/XBP1/ IRF1/SOX9/MILL/GALNT3/LGALS4
<b>Across 3 datasets</b>	0	0	0	0	0	<b>GIF/XBP1</b>

GSE54129, GSE24375 and TCGA-STAD collected Chinese, Korean and American gastric cancer samples respectively. The overlaps of top ranked differentially regulated genes (DRGs) between any two datasets (Row 2nd, 3rd and 4th) and between all three datasets (Row 5th) are listed

polymerase II promoter”, “transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding”, “sequence-specific DNA binding”, “transcription factor complex” and “nucleoplasm” were recognized. That is, the 16 common DRGs across racial groups are mainly located in the transcription factor complex in nucleus and their functions were enriched in transcriptional regulation of RNA polymerase II, transcriptional activator activity and transcription factor binding. This is consistent with the basic understanding that transcription factors (TFs) play crucial roles in the proliferation and differentiation of cells.

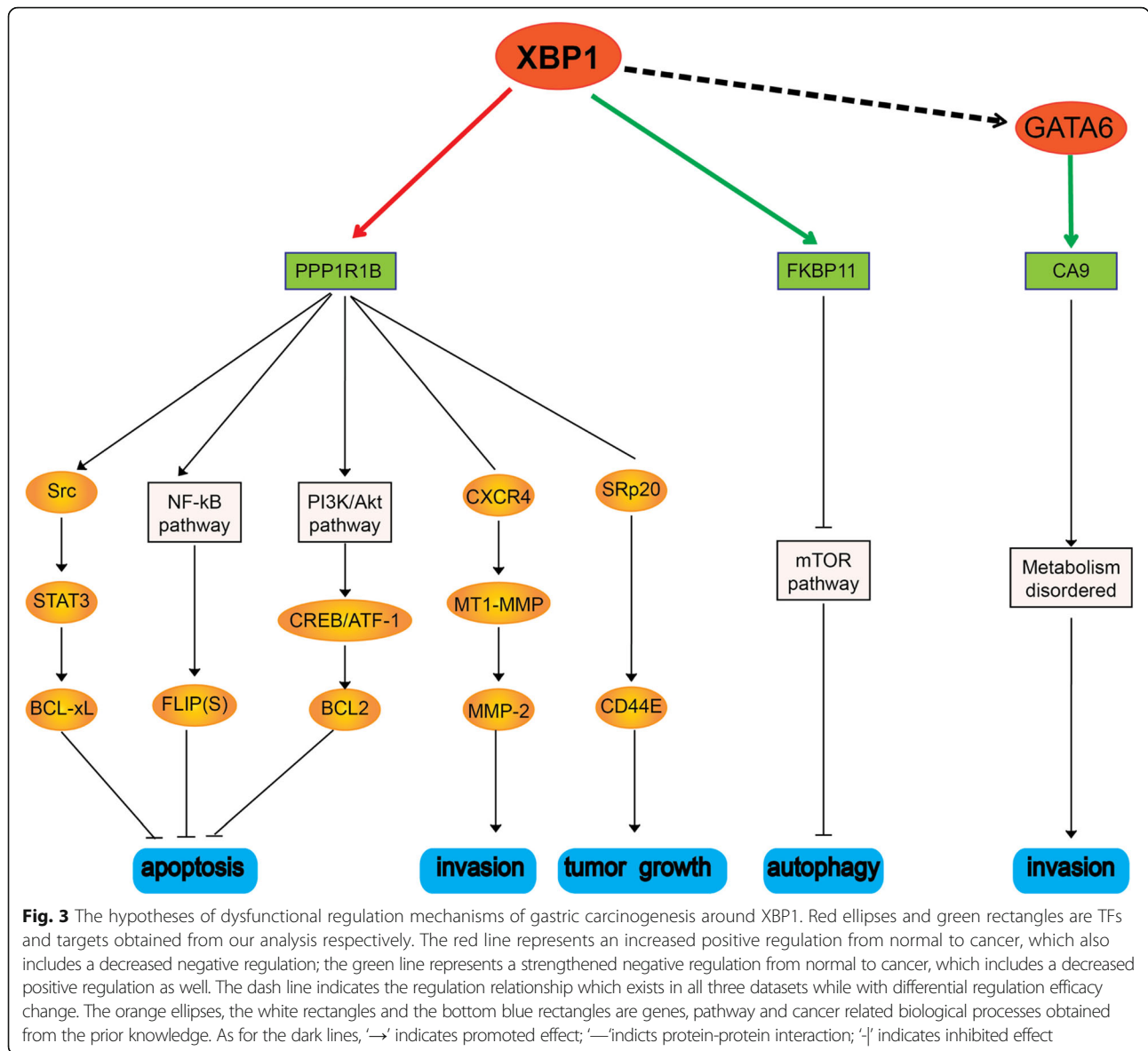
It is interesting that GIF, the overlapping gene in the three top-100 gene lists across Chinese, Korean, and American, which is also a TOP-30 DRG commonly in at least two racial groups, was reported to be a prognosis biomarker of gastric cancer and its decreased expression was reported to be correlated with the progresses of gastritis [53, 54]. The down-expression of GIF in gastric carcinoma samples was indeed observed in all three GC dataset. We proposed GIF might be a common GC related gene across racial groups, which participates in carcinogenesis by differential regulation. We then focused on GIF and its surrounding DRLs to generate hypotheses on the regulation mechanisms of gastric carcinogenesis. In our integrative DRN analysis, GIF were regulated by GATA3 in GSE24375 (− 2.896 vs 0.01; DRL TOP 25) and TCGA-STAD (2.893 vs − 0.01; DRL TOP7) datasets but their regulation efficacy changes were not consistent. However, GIF was regulated by CEBPA in GSE54129 (− 3.412 vs 0.01; DRL TOP7). It was noticed that the DRLs of GIF were all ranked in TOP 30 in every dataset. According to the clinical information of TCGA-STAD, the GIF expression in different patient groups based on Neoplasm Histologic Grade were significantly different with Kruskal Wallis Test *p*-value 0.002 as shown in Additional file 1: Figure S4. We therefore suggest that GIF and its regulation are worthy of further investigation to elucidate its role in gastric carcinogenesis in diverse racial groups.

XBP1 is the only overlapping transcription factor (TF) in the three top-100 gene lists across the three racial groups included in our analysis. Considering the crucial

roles of TF in cancer progression, we put much attention to XBP1 although its DR rank and surrounding DRLs are all out of TOP-30 in every dataset and the mutation frequency of XBP1 is low (0.9%) in 588 gastric cancer samples in ICGC database (<https://icgc.org/>). By integrating our differential regulation analysis results to the prior knowledge, the dysfunctional regulation mechanisms underlying gastric carcinogenesis around XBP1 were proposed in Fig. 3. In all the three GC datasets, the positive regulation of PPP1R1B by XBP1 was increased from normal to cancer, and the negative regulation of FKBP11 by XBP1 was strengthened. The up-regulation of PPP1R1B was reported to inhibit apoptosis through NF-κB/FLIP(S) pathway [55] and promote cell invasion [56] and gastric tumorigenesis [57, 58]. The down-regulation of FKBP11 leads to the inhibition of autophagy through MTOR pathway and induces carcinogenesis [59]. In Chinese and Korean datasets (GSE54129 and GSE24375), the positive regulation of GATA6 by XBP1 was increased from normal to cancer, while it was decreased in American dataset (TCGA-STAD) during carcinogenesis. This might contribute to the differential gastric carcinogenesis between Asian and American. It is interesting that the positive regulation of CA9 by GATA6 was decreased in all three racial groups. This is consistent with the observation that Ca9 expression was frequently lost in gastric cancers in part by methylation [60], while contradictory to another report that the activity of Ca9 contributes to invasion and thus advanced tumor progression in a subset of gastric cancers [60–62].

## Discussion

In this work, we carried out a cross-racial integrative research on gastric cancer in terms of dysfunctional regulation mechanisms by implementing a novel differential networking meta-analysis framework. Differential regulation networking (DRN) analysis aims to identify the regulatory relationships relevant or even causative to phenotypic changes, which is challenging in the field of both computational and experimental biology. Since cancer has a nature of dysregulation mechanisms during carcinogenesis, the DRN analysis is helpful for deciphering differential regulation and differential networking



underlying phenotypic changes in cancer. Even if lots of attention has been paid on traditional differential expression analysis in the past carcinogenesis studies, only in very recent years differential regulation networking analysis has become more and more widely applied [22, 23, 30]. With the transcriptomic data from cancer samples increasingly accumulated in the public domain, it is time to investigate the common dysfunctional regulation mechanisms of carcinogenesis at a broader systematic level. We therefore created a differential networking meta-analysis framework based on our previously developed methodologies, and applied it to three GC gene expression datasets corresponding to Chinese, Korean and American.

First of all, the high rank of cancer genes and drug targets in the three DRG lists (Table 2) as well as the basic

statistics of the three conditional GRN pairs (Table 1) proved the rationality of the whole modelling strategy and the effectiveness of DR measure as expected. Since the DRN analysis module in the framework enables the discovery of novel regulators or regulatory relationships that have not yet been associated to the disease of interest [30, 31, 63–65], we hoped to find out novel dysregulation regulators and their related mechanisms underlying gastric carcinogenesis across different racial groups.

In our framework, DCGs were taken as the seed genes for differential network construction. As shown in Fig. 1, the DCGs between every two datasets are significantly overlapped, indicating a basically coherent profile of molecular interactions across Chinese, Korean, and American; meanwhile, we also noticed a closer overlap



between Chinese and Korean, which was supported by previous reports [66, 67]. It is interesting that from the viewpoint of gene regulatory network (GRN), the three GRN pairs present significant differences from network topologies to DRG identities (Fig. 2). Furthermore, the consistency of highly ranked DRGs across three racial groups is quite limited as shown in Tables 3 and 4. The difference between racial groups could be associated with genetics, diet habits and other environmental factors. These observations strongly support the necessity of integrative studies in terms of differential regulation.

In order to decipher the common carcinogenesis across racial groups, we focus on those common DRGs across datasets. The DRGs commonly in at least two racial groups were presented in Table 4. Four genes (CEBPA, GATA6, GATA3 and GIF) were selected from TOP30 DRG lists, seven genes (GIF, XBP1, CEBPA, GATA6, DPP4, PTK7 and GATA3) were selected from TOP50 DRG lists, and 16 genes (GIF, XBP1, CEBPA, GATA6, DPP4, PTK7, GATA3, FOSB, EPAS1, CCND2, MALL, IRF1, SOX9, VILL, GALNT3 and LGALS4) were selected from TOP100 DRGs. Among the 16 cross-racial DRGs, all are cancer related according to ICGC and three genes (GATA3, SOX9 and CEBPA) have been regarded as cancer driver genes [68–71]. Functional enrichment analysis indicated that the 16 common cross-racial DRGs were enriched in the function of transcriptional regulation. We therefore propose that the highly ranked TFs among the 16 top common DRGs, including XBP1, FOSB, MRPL36, GATA3, GATA6, SOX9, POU2F1, IRF2, PTK7, EPAS1 and IRF1 (Table 4) are worthy of further investigation, especially on their roles of gene transcriptional regulation.

We then narrowed down our attention to GIF and XBP1 out of the 16 DRGs. Considering that GIF has been reported to be a GC prognosis biomarker [54, 72, 73], and GIF's DRLs were all ranked in TOP 30 in every dataset, we suggest that GIF and its regulation are worthy of further investigation to elucidate its role in gastric carcinogenesis. The transcriptional regulation of GIF by GATA3 and CEBPA might be related to the racial specificity of GC. Although XBP1 has not been related to gastric cancer so far, it is the only overlapping transcription factor (TF) in the three top-100 DGR lists across Chinese, Korean and American. By integrating our DRN analysis results to prior knowledge, we proposed a hypothesis of dysfunctional regulation mechanisms for gastric carcinogenesis around XBP1 in Fig. 3. XBP1 and its DRLs seem to play crucial roles during gastric carcinogenesis across three racial groups.

The previous studies on the effects of racial factors on gastric cancer were mainly focused on statistics of clinical phenotypes, such as metastasis and survival [67, 74–76], and a few of literatures also reported genetic polymorphisms

[66, 77]. By means of differential regulation analysis, our cross-racial meta-analysis of gastric cancer demonstrated the complexity and diversity in terms of differential networking, and made an insightful complement to these previous reports. The race specific and common cross-racial DRGs and DRLs are helpful to decipher the phenotypic differences among different racial GC groups from the perspective of transcriptional regulation, and helpful to discover universal or race specific drug targets or biomarkers.

The current work provides a differential networking meta-analysis framework which presents high cohesion and low coupling. This framework is extendable and adaptable to the studies on the regulation mechanisms of other cancer and even other phenotypic changes.

## Conclusions

Our study aimed at investigating the common dysregulation mechanisms of gastric carcinogenesis across racial groups from three Gastric Carcinoma datasets (GSE54129, GSE24375 and TCGA-STAD) for Chinese, Korean and American. We constructed a cross-racial integrative analysis framework from the viewpoint of differential regulation. The cross-racial meta-analysis of gastric cancer demonstrated the complexity and diversity of gastric carcinogenesis, and provided useful clues for understanding the common dysfunctional regulation mechanisms of gastric cancer progression. In addition, the differential networking meta-analysis framework presented here is extendable and adaptable to the studies on the regulation mechanisms of other cancer and even other phenotypic changes.

## Additional files

**Additional file 1:** Supplementary Figure and legends. **Figure S1.** Quality assessment of our data set (GSE54129). **Figure S2.** Expression level distribution of genes in three datasets. **Figure S3.** The differential networking meta-analysis framework. **Figure S4.** The GIF expression in different patient groups based on Neoplasm Histologic Grade. (DOCX 1497 kb)

**Additional file 2: Table S1.** DRN meta-analysis of gastric cancer, <https://www.synapse.org/#!Synapse:syn10154163> (XLSX 98 kb)

## Abbreviations

DCEA: Differential CoExpression Analysis; DCG: Differential co-expression gene; DCL: Differential co-expression gene link; DRG: Differentially regulated gene; DRL: Differentially regulated gene link; DRN: Differential regulation networking; GC: Gastric Carcinoma; GRN: Gene regulatory network; STAD: Stomach adenocarcinoma

## Acknowledgements

Not applicable.

## Funding

This work was supported by the grants from the Shanghai Sailing Program (16YF1408600), the National Natural Science Foundation of China (81672736 and 91529302), Shanghai Municipal Commission of Science and Technology (15ZR1430300 and 14DZ2252000), and the administrative committee of

Shanghai Zhangjiang Hi-Tech Park (2016–08). The cost of publication for this paper was covered by the Shanghai Sailing Program (16YF1408600).

#### Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

#### About this supplement

This article has been published as part of *BMC Systems Biology* Volume 12 Supplement 4, 2018: Selected papers from the 11th International Conference on Systems Biology (ISB 2017). The full contents of the supplement are available online at <https://bmcsystbiol.biomedcentral.com/articles/supplements/volume-12-supplement-4>.

#### Authors' contributions

WD and YYL contributed to the design and conception of the study. WD and QL conducted computational experiments, analyzed and interpreted data, developed the analysis framework and drafted the manuscript. BYL joined in the processing of data materials and helped to improve the manuscript. YYL and YXL conceived of the project and participated in its design, helped to analyze and interpret the data and drafted the manuscript. All authors have read and approved the manuscript for publication.

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### Author details

<sup>1</sup>Shanghai Center for Bioinformatics Technology, 1278 Keyuan Road, Shanghai 201203, People's Republic of China. <sup>2</sup>School of biotechnology, East China University of Science and Technology, Shanghai 200237, China. <sup>3</sup>Shanghai Key Laboratory of Gastric Neoplasms, Shanghai Institute of Digestive Surgery, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, People's Republic of China. <sup>4</sup>Shanghai Engineering Research Center of Pharmaceutical Translation & Shanghai Industrial Technology Institute, 1278 Keyuan Road, Shanghai 201203, People's Republic of China. <sup>5</sup>Key Lab of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China.

Published: 24 April 2018

#### References

- Figueiredo C, Garcia-Gonzalez MA, Machado JC. Molecular pathogenesis of gastric cancer. *Helicobacter*. 2013;18(Suppl 1):28–33.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin*. 2016; 66(1):7–30.
- McGuire S. World cancer report 2014. Geneva, Switzerland: World Health Organization, international agency for research on cancer, WHO press. *Adv Nutr*. 2015;7(2):418–9.
- Buffart TE, Carvalho B, van Grieken NC, van Wieringen WN, Tijssen M, Kranenbarg EM, Verheul HM, Grabsch HI, Ylstra B, van de Velde CJ, et al. Losses of chromosome 5q and 14q are associated with favorable clinical outcome of patients with gastric cancer. *Oncologist*. 2012;17(5):653–62.
- Holbrook JD, Parker JS, Gallagher KT, Halsey WS, Hughes AM, Weigman VJ, Lebowitz PF, Kumar R. Deep sequencing of gastric carcinoma reveals somatic mutations relevant to personalized medicine. *J Transl Med*. 2011;9:119.
- Song HR, Kim HN, Kweon SS, Choi JS, Shim HJ, Cho SH, Chung IJ, Park YK, Kim SH, Choi YD, et al. Genetic variations in the PRKAA1 and ZBTB20 genes and gastric cancer susceptibility in a Korean population. *Mol Carcinog*. 2013;52:E155.
- Wang K, Kan J, Yuen ST, Shi ST, Chu KM, Law S, Chan TL, Kan Z, Chan AS, Tsui WY, et al. Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat Genet*. 2011;43(12):1219–23.
- Kim H, Eun JW, Lee H, Nam SW, Rhee H, Koh KH, Kim H. Gene expression changes in patient-matched gastric normal mucosa, adenomas, and carcinomas. *Exp Mol Pathol*. 2011;90(2):201–9.
- Wu XM, Shao XQ, Meng XX, Zhang XN, Zhu L, Liu SX, Lin J, Xiao HS. Genome-wide analysis of microRNA and mRNA expression signatures in hydroxycamptothecin-resistant gastric cancer cells. *Acta Pharmacol Sin*. 2011;32(2):259–69.
- Cao WJ, Wu HL, He BS, Zhang YS, Zhang ZY. Analysis of long non-coding RNA expression profiles in gastric cancer. *World J Gastroenterol*. 2013;19(23): 3658–64.
- Gopal G, Raja UM, Shirley S, Rajalekshmi KR, Rajkumar T. SOSTDC1 down-regulation of expression involves CpG methylation and is a potential prognostic marker in gastric cancer. *Cancer Genet*. 2013;206(5):174–82.
- Fan B, Dachrut S, Coral H, Yuen ST, Chu KM, Law S, Zhang L, Ji J, Leung SY, Chen X. Integration of DNA copy number alterations and transcriptional expression analysis in human gastric cancer. *PLoS One*. 2012;7(4):e29824.
- Wang K, Yuen ST, Xu J, Lee SP, Yan HH, Shi ST, Siu HC, Deng S, Chu KM, Law S, et al. Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. *Nat Genet*. 2014;46(6):573–82.
- Liu J, McClelland M, Stawiski EW, Gnad F, Mayba O, Haverty PM, Durinck S, Chen YJ, Klijn C, Jhunjhunwala S, et al. Integrated exome and transcriptome sequencing reveals ZAK isoform usage in gastric cancer. *Nat Commun*. 2014;5:3830.
- Bass AJ, Thorsson V, Shmulevich I, et al. Comprehensive molecular characterization of gastric adenocarcinoma[J]. *Nature*. 2014;513(7517):202–9.
- Cui J, Yin Y, Ma Q, et al. Comprehensive characterization of the genomic alterations in human gastric cancer[J]. *Int J Cancer*. 2015;137(1):86–95.
- Cancer Genome Atlas Research N. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature*. 2014;513(7517):202–9.
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646–74.
- Kreeger PK, Lauffenburger DA. Cancer systems biology: a network modeling perspective. *Carcinogenesis*. 2010;31(1):2–8.
- Castro MA, de Santiago I, Campbell TM, Vaughn C, Hickey TE, Ross E, Tilley WD, Markowitz F, Ponder BA, Meyer KB. Regulators of genetic risk of breast cancer identified by integrative network analysis. *Nat Genet*. 2016;48(1):12–21.
- Yu H, Tu K, Wang YJ, Mao JZ, Xie L, Li YY, Li YX. Combinatorial network of transcriptional regulation and microRNA regulation in human cancer. *BMC Syst Biol*. 2012;6:61.
- de la Fuente A. From 'differential expression' to 'differential networking' - identification of dysfunctional regulatory networks in diseases. *Trends Genet*. 2010;26(7):326–33.
- Gill R, Datta S, Datta S. Differential network analysis in human cancer research. *Curr Pharm Des*. 2014;20(1):4–10.
- Zeng L, Yu J, Huang T, Jia H, Dong Q, He F, Yuan W, Qin L, Li Y, Xie L. Differential combinatorial regulatory network analysis related to venous metastasis of hepatocellular carcinoma. *BMC Genomics*. 2012;13(Suppl 8):S14.
- Madhamshettiwar PB, Maetschke SR, Davis MJ, Reverter A, Ragan MA. Gene regulatory network inference: evaluation and application to ovarian cancer allows the prioritization of drug targets. *Genome Med*. 2012;4(5):1–16.
- Yosef N, Shalek AK, Gaublotme JT, Jin H, Lee Y, Awasthi A, Wu C, Karwacz K, Xiao S, Jorgolli M. Dynamic regulatory network controlling TH17 cell differentiation. *Nature*. 2013;496(7446):461–8.
- Liu BH, Yu H, Tu K, Li C, Li YX, Li YY. DCGL: an R package for identifying differentially coexpressed genes and links from gene expression microarray data. *Bioinformatics*. 2010;26(20):2637–8.
- Yang J, Yu H, Liu BH, Zhao Z, Liu L, Ma LX, Li YX, Li YY. DCGL v2.0: an R package for unveiling differential regulation from differential co-expression. *PLoS One*. 2013;8(11):e79729.
- Yu H, Liu BH, Ye ZQ, Li C, Li YX, Li YY. Link-based quantitative methods to identify differentially coexpressed genes and gene pairs. *BMC Bioinformatics*. 2011;12:315.
- Cao MS, Liu BY, Dai WT, Zhou WX, Li YX, Li YY. Differential network analysis reveals dysfunctional regulatory networks in gastric carcinogenesis. *Am J Cancer Res*. 2015;5(9):2605–25.

31. Li Q, Li J, Dai W, Li Y-X, Li Y-Y. Differential regulation analysis reveals dysfunctional regulatory mechanism involving transcription factors and microRNAs in gastric carcinogenesis. *Artif Intell Med.* 2017;77:12–22.
32. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44–57.
33. Yang J, Wu SJ, Dai WT, Li YX, Li YY. The human disease network in terms of dysfunctional regulatory mechanisms. *Biol Direct.* 2015;10:60.
34. Ying LS, Yu JL, Lu XX, Ling ZQ. Enhanced RegIV expression predicts the intrinsic 5-fluorouracil (5-FU) resistance in advanced gastric cancer. *Dig Dis Sci.* 2013;58(2):414–22.
35. van Duin M, van Marion R, Vissers KJ, Hop WC, Dinjens WN, Tilanus HW, Siersema PD, van Dekken H. High-resolution array comparative genomic hybridization of chromosome 8q: evaluation of putative progression markers for gastroesophageal junction adenocarcinomas. *Cytogenet Genome Res.* 2007;118(2–4):130–7.
36. Tsuge H, Tsurumura T, Utsunomiya H, Kise D, Kuzuhara T, Watanabe T, Fujiki H, Suganuma M. Structural basis for the Helicobacter pylori-carcinogenic TNF- $\alpha$ -inducing protein. *Biochem Biophys Res Commun.* 2009;388(2):193–8.
37. Delage B, Fennell DA, Nicholson L, McNeish I, Lemoine NR, Crook T, Szlosarek PW. Arginine deprivation and argininosuccinate synthetase expression in the treatment of cancer. *Int J Cancer.* 2010;126(12):2762–72.
38. Ohnuma S, Miura K, Horii A, Fujibuchi W, Kaneko N, Gotoh O, Nagasaki H, Mizoi T, Tsukamoto N, Kobayashi T, et al. Cancer-associated splicing variants of the CDCA1 and MSMB genes expressed in cancer cell lines and surgically resected gastric cancer tissues. *Surgery.* 2009;145(1):57–68.
39. Chen L, Yang Q, Kong WQ, Liu T, Liu M, Li X, Tang H. MicroRNA-181b targets cAMP responsive element binding protein 1 in gastric adenocarcinomas. *IUBMB Life.* 2012;64(7):628–35.
40. Buganim Y, Goldstein I, Lipson D, Milyavsky M, Polak-Charcon S, Mardoukh C, Solomon H, Kalo E, Madar S, Brosh R, et al. A novel translocation breakpoint within the BPTF gene is associated with a pre-malignant phenotype. *PLoS One.* 2010;5(3):e9657.
41. Bustin SA, Li SR, Dorudi S. Expression of the Ca<sup>2+</sup>-activated chloride channel genes CLCA1 and CLCA2 is downregulated in human colorectal cancer. *DNA Cell Biol.* 2001;20(6):331–8.
42. Yin Y, Zhuo W, Zhao Y, Chen S, Li J, Wang L, Zhou T, Si JM. Converting a microarray signature into a diagnostic test: a trial of custom 74 gene array for clarification and prediction the prognosis of gastric cancer. *PLoS One.* 2013;8(12):e81561.
43. Kim SK, Jang HR, Kim JH, Noh SM, Song KS, Kim MR, Kim SY, Yeom YI, Kim NS, Yoo HS, et al. The epigenetic silencing of LIMS2 in gastric cancer and its inhibitory effect on cell migration. *Biochem Biophys Res Commun.* 2006;349(3):1032–40.
44. Zhu ZL, Yan BY, Zhang Y, Yang YH, Wang ZM, Zhang HZ, Wang MW, Zhang XH, Sun XF. PINCH expression and its clinicopathological significance in gastric adenocarcinoma. *Dis Markers.* 2012;33(4):171–8.
45. Keshari RP, Wang W, Zhang Y, Wang DD, Li YF, Yuan SQ, Qiu HB, Huang CY, Chen YM, Xia JC, et al. Decreased expression of the GATA3 gene is associated with poor prognosis in primary gastric adenocarcinoma. *PLoS One.* 2014;9(2):e87195.
46. Sulahian R, Casey F, Shen J, et al. An integrative analysis reveals functional targets of GATA6 transcriptional regulation in gastric cancer. *J. Oncogene.* 2014;33(49):5637–48.
47. Choi YJ, Song JH, Yoon JH, Choi WS, Nam SW, Lee JY, Park WS. Aberrant expression of SOX9 is associated with gastrosone 1 inactivation in gastric cancers. *Gastric Cancer.* 2014;17(2):247–54.
48. Qian J, Kong X, Deng N, et al. OCT1 is a determinant of synbindin-related ERK signalling with independent prognostic significance in gastric cancer. *J. Gut.* 2015;64(1):37–48.
49. Sakai T, Mashima H, Yamada Y, Goto T, Sato W, Dohmen T, Kamada K, Yoshioka M, Uchinami H, Yamamoto Y, et al. The roles of interferon regulatory factors 1 and 2 in the progression of human pancreatic cancer. *Pancreas.* 2014;43(6):909–16.
50. Tatenhorst L, Senner V, Puttmann S, Paulus W. Regulators of G-protein signaling 3 and 4 (RGS3, RGS4) are associated with glioma cell motility. *J Neuropathol Exp Neurol.* 2004;63(3):210–22.
51. Piao L, Li Y, Kim SJ, Byun HS, Huang SM, Hwang SK, Yang KJ, Park KA, Won M, Hong J, et al. Association of LETM1 and MRPL36 contributes to the regulation of mitochondrial ATP production and necrotic cell death. *Cancer Res.* 2009;69(8):3397–404.
52. Kataoka F, Tsuda H, Arai T, Nishimura S, Tanaka H, Nomura H, Chiyoda T, Hirasawa A, Akahane T, Nishio H, et al. EGRI and FOSB gene expressions in cancer stroma are independent prognostic indicators for epithelial ovarian cancer receiving standard therapy. *Genes Chromosomes Cancer.* 2012;51(3):300–12.
53. Wu W, Juan WC, Liang CR, Yeoh KG, So J, Chung MC. S100A9, GIF and AAT as potential combinatorial biomarkers in gastric cancer diagnosis and prognosis. *Proteomics Clin Appl.* 2012;6(3–4):152–62.
54. Lozano-Pope I, Sharma A, Matthias M, Doran KS, Obonyo M. Effect of myeloid differentiation primary response gene 88 on expression profiles of genes during the development and progression of Helicobacter-induced gastric cancer. *BMC Cancer.* 2017;17(1):133.
55. Belkhiri A, Zhu S, Chen Z, Soutto M, El-Rifai W. Resistance to TRAIL is mediated by DARPP-32 in gastric cancer. *Clin Cancer Res.* 2012;18(14):3889–900.
56. Zhu S, Hong J, Tripathi MK, Sehdev V, Belkhiri A, El-Rifai W. Regulation of CXCR4-mediated invasion by DARPP-32 in gastric cancer cells. *Mol Cancer Res.* 2013;11(11):86–94.
57. Zhu S, Chen Z, Katsha A, Hong J, Belkhiri A, El-Rifai W. Regulation of CD44E by DARPP-32-dependent activation of SRp20 splicing factor in gastric tumorigenesis. *Oncogene.* 2016;35(14):1847–56.
58. Belkhiri A, Dar AA, Zaika A, Kelley M, El-Rifai W. T-Darpp promotes cancer cell survival by up-regulation of Bcl2 through Akt-dependent mechanism. *Cancer Res.* 2008;68(2):395–403.
59. Wei FZ, Cao Z, Wang X, Wang H, Cai MY, Li T, Hattori N, Wang D, Du Y, Song B, et al. Epigenetic regulation of autophagy by the methyltransferase EZH2 through an mTOR-dependent pathway. *Autophagy.* 2015;11(12):2309–22.
60. Chen J, Rocken C, Hoffmann J, Kruger S, Lendeckel U, Rocco A, Pastorekova S, Malfertheiner P, Ebert MP. Expression of carbonic anhydrase 9 at the invasion front of gastric cancers. *Gut.* 2005;54(7):920–7.
61. Kato Y, Yashiro M, Noda S, Kashiwagi S, Matsuoaka J, Fuyuhiko Y, Doi Y, Hirakawa K. Expression of a hypoxia-associated protein, carbonic anhydrase-9, correlates with malignant phenotypes of gastric carcinoma. *Digestion.* 2010;82(4):246–51.
62. Giatromanolaki A, Koukourakis MI, Sivridis E, Pastorek J, Wykoff CC, Gatter KC, Harris AL. Expression of hypoxia-inducible carbonic anhydrase-9 relates to angiogenic pathways and independently to poor outcome in non-small cell lung cancer. *Cancer Res.* 2001;61(21):7992–8.
63. Madhamsheetiwar PB, Maetschke SR, Davis MJ, Reverter A, Ragan MA. Gene regulatory network inference: evaluation and application to ovarian cancer allows the prioritization of drug targets. *Genome medicine.* 2012;4(5):41.
64. Ma C, Xin M, Feldmann KA, Wang X. Machine learning-based differential network analysis: a study of stress-responsive transcriptomes in Arabidopsis. *Plant Cell.* 2014;26(2):520–37.
65. Araki R, Seno S, Takenaka Y, Matsuda H. An estimation method for a cellular-state-specific gene regulatory network along tree-structured gene expression profiles. *Gene.* 2013;518(1):17–25.
66. Loh M, Koh KX, Yeo BH, Song CM, Chia KS, Zhu F, Yeoh KG, Hill J, Iacopetta B, Soong R. Meta-analysis of genetic polymorphisms and gastric cancer risk: variability in associations according to race. *Eur J Cancer.* 2009;45(14):2562–8.
67. Yao JC, Schnirer II, Reddy S, Chiang S, Najam A, Yu C, Giacco G, Hess K, Rashid A, Xie K. Effects of sex and racial/ethnic group on the pattern of gastric cancer localization. *Gastric Cancer.* 2002;5(4):0208–12.
68. Yachida S, Wood LD, Suzuki M, Takai E, Totoki Y, Kato M, Luchini C, Arai Y, Nakamura H, Hama N. Genomic sequencing identifies ELF3 as a driver of ampullary carcinoma. *Cancer Cell.* 2016;29(2):229–40.
69. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* 2012;487(7407):330–7.
70. Gleeson FC, Kerr SE, Kipp BR, Voss JS, Minot DM, Tu ZJ, Henry MR, Graham RP, Vasmatazis G, Chevillat JC. Targeted next generation sequencing of endoscopic ultrasound acquired cytology from ampullary and pancreatic adenocarcinoma has the potential to aid patient stratification for optimal therapy selection. *Oncotarget.* 2016;7(34):54526.
71. Rubio-Perez C, Tamborero D, Schroeder MP, Antolin AA, Deu-Pons J, Perez-Llana C, Mestres J, Gonzalez-Perez A, Lopez-Bigas N. In silico prescription of anticancer drugs to cohorts of 28 tumor types reveals targeting opportunities. *Cancer Cell.* 2015;27(3):382–96.
72. Wu W, Juan WC, Liang CR, Yeoh KG, So J, Chung M. S100A9, GIF and AAT as potential combinatorial biomarkers in gastric cancer diagnosis and prognosis. *Proteomics Clin Appl.* 2012;6(3–4):152–62.
73. Gordon MM, Brada N, Remacha A, Badell I, Ed R, Baiget M, Santer R, Quadros EV, Rothenberg SP, Alpers DH. A genetic polymorphism in the

coding region of the gastric intrinsic factor gene (GIF) is associated with congenital intrinsic factor deficiency. *Hum Mutat.* 2004;23(1):85–91.

74. Ikoma N, Blum M, Chiang Y-J, Estrella JS, Roy-Chowdhuri S, Fournier K, Mansfield P, Ajani J, Badgwell BD. Race is a risk for lymph node metastasis in patients with gastric cancer. *Ann Surg Oncol.* 2017;24(4):960–5.
75. Howard JH, Hiles JM, Leung AM, Stern SL, Bilchik AJ. Race influences stage-specific survival in gastric cancer. *Am Surg.* 2015;81(3):259–67.
76. Yang D, Hendifar A, Lenz C, Togawa K, Lenz F, Lurje G, Pohl A, Winder T, Ning Y, Groshen S. Survival of metastatic gastric cancer: significance of age, sex and race/ethnicity. *J Gastrointest Oncol.* 2011;2(2):77–84.
77. Ward E, Jemal A, Cokkinides V, Singh GK, Cardinez C, Ghafoor A, Thun M. Cancer disparities by race/ethnicity and socioeconomic status. *CA Cancer J Clin.* 2004;54(2):78–93.

Submit your next manuscript to BioMed Central  
and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

