Pathol. Oncol. Res. (2017) 23:361–368
DOI 10.1007/s12253-016-0116-z

ORIGINAL ARTICLE

Combined Analysis of ChIP Sequencing and Gene Expression Dataset in Breast Cancer

Pengfei Liu[1] · **Wenhua Jiang**[2] · **Shiyong Zhou**[1] · **Jun Gao**[3] · **Huilai Zhang**[1]

Received: 25 March 2016 / Accepted: 9 September 2016 / Published online: 21 September 2016

Abstract Breast cancer is a common malignancy in women and contribute largely to the cancer related death. The purpose of this study is to confirm the roles of GATA3 and identify potential biomarkers of breast cancer. Chromatin Immunoprecipitation combined with high-throughput sequencing (ChIP-Seq) (GSM1642515) and gene expression profiles (GSE24249) were downloaded from the Gene Expression Omnibus (GEO) database. Bowtie2 and MACS2 were used for the mapping and peak calling of the ChIP-Seq data respectively. ChIPseeker, a R bioconductor package was adopted for the annotation of the enriched peaks. For the gene expression profiles, we used affy and limma package to do normalization and differential expression analysis. The genes with fold change >2 and adjusted *P*-Value <0.05 were screened out. Besides, BETA (Binding and Expression Target Analysis) was used to do the combined analysis of ChIP-Seq and gene expression profiles. The Database for Annotation, Visualization and Integrated Discovery (DAVID) was used for the functional enrichment analysis of overlapping genes between the target genes and differential expression genes (DEGs). What's more, the protein-protein interaction (PPI) network of the overlapping genes was obtained through the Human Protein Reference Database (HPRD). A total of 46,487 peaks were identified for GATA3 and out of which, 3256 ones were found to located at −3000 ~ 0 bp from the transcription start sites (TSS) of their nearby gene. A total of 236 down- and 343 up-regulated genes were screened out in GATA3 overexpression breast cancer samples compared with those in control. The combined analysis of ChIP-Seq and gene expression dataset showed GATA3 act as a repressor in breast cancer. Besides, 68 overlaps were obtained between the DEGs and genes included in peaks located at −3000 ~ 0 bp from TSS. Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways related to cancer progression and gene regulation were found to be enriched in those overlaps. In the PPI network, *NDRG1*, *JUP* and etc. were found to directly interact with large number of genes, which might indicate their important roles in the progression of breast cancer.

Keywords Breast cancer · Chromatin Immunoprecipitation combined with high-throughput sequencing (ChIP-seq) · Human protein reference database (HPRD) · Protein-protein interaction (PPI)

Pengfei Liu and Wenhua Jiang contributed equally to this study.

This work was supported by the Municipal Science and Technology Commission of Tianjin (No. 15ZLZLZF00440) and the Health Bureau Science and Technology Foundation of Tianjin (No. 2012KZ063 and No.2014KZ102).

✉ Huilai Zhang
snowlpf@163.com

[1] Department of Lymphoma, Sino-US Center of Lymphoma and Leukemia, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin, Tianjin's Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Huan-Hu-Xi Road, Ti-Yuan-Bei, He Xi District, Tianjin 300060, China

[2] Department of Radiotherapy, Second Hospital of Tianjin Medical University, Tianjin 300211, China

[3] Department of Oncology, Hengshui Harrison International Peace Hospital, Hengshui, Hebei 053000, China

Introduction

Breast cancer is one of the malignancy in women and the second cause of cancer related death in United States [1]. In recent years, more and more studies about it were carried out and lots of novel therapeutics besides traditional radiotherapy,

chemotherapy and etc., were proposed, such as immunotherapy, targeted therapy [2–4], but its prognosis remains unimproved. Further studies are still needed for the understanding of its mechanisms, which would be helpful for the development of novel biomarkers that play important roles in its progression.

Lots of factors were found to involved in its initiation, progression and metastasis, including abnormal metabolize of amino acids components, such as tryptophan, DNA damage, variation of specific gene expression and so on [5–7]. In some times, the expression level of genes might change without the mutations of DNA sequences and induce the emergence of abnormal phenotypes finally. Lots of processes were proved to result this phenomenon and the perturbation of binding of protein in DNA is one of the most common mechanisms.

GATA3 is a zinc-binding transcription factor which regulates the differentiation of many tissues, including breast [8]. It has been proved to be a frequently mutated gene in breast cancer and its expression decreased with the progression of breast cancer in mRNA, as well as protein level [9–11]. So it might be helpful to explore its binding profiles in breast cancer for the identification of therapeutic targets. Meanwhile, the rapid development of high-throughput sequencing technology promoted the explosive growth of binding profiles of many factors which obtained via Chromatin Immunoprecipitation combined with high-throughput sequencing (ChIP-Seq), and compared with ChIP-on-ChIP (Chromatin Immunoprecipitation combined with DNA microarray), ChIP-Seq could identify novel binding sites in genome-wide with lower false positive rate.

In this study, through the combined analysis of ChIP-Seq and gene expression datasets, we predicted the functions of GATA3 in breast cancer. Besides, its direct targets were screened out through the combination of binding and expression profiles, and their potential related biological process were obtained, which would be valuable for the understanding and treatment of breast cancer.

Materials and Methods

ChIP-Seq and Gene Expression Datasets

The ChIP-Seq and gene expression datasets were all downloaded from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). In ChIP-Seq dataset (GSM1642515), a ChIP sample with GATA3 antibody and a input sample (whole cell lysates) from MCF-7 cells were included [9]. The DNA sequences were determined through GPL11154 Illumina HiSeq 2000 (*Homo sapiens*) and stored as .fastq files. For gene expression dataset (GSE24249), a total of 6 breast cancer samples which contained 3 GATA3 overexpression (transduced with GATA3) and 3 control samples (transduced with eGFP) were involved [12]. The expression profiles were detected based on GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array.

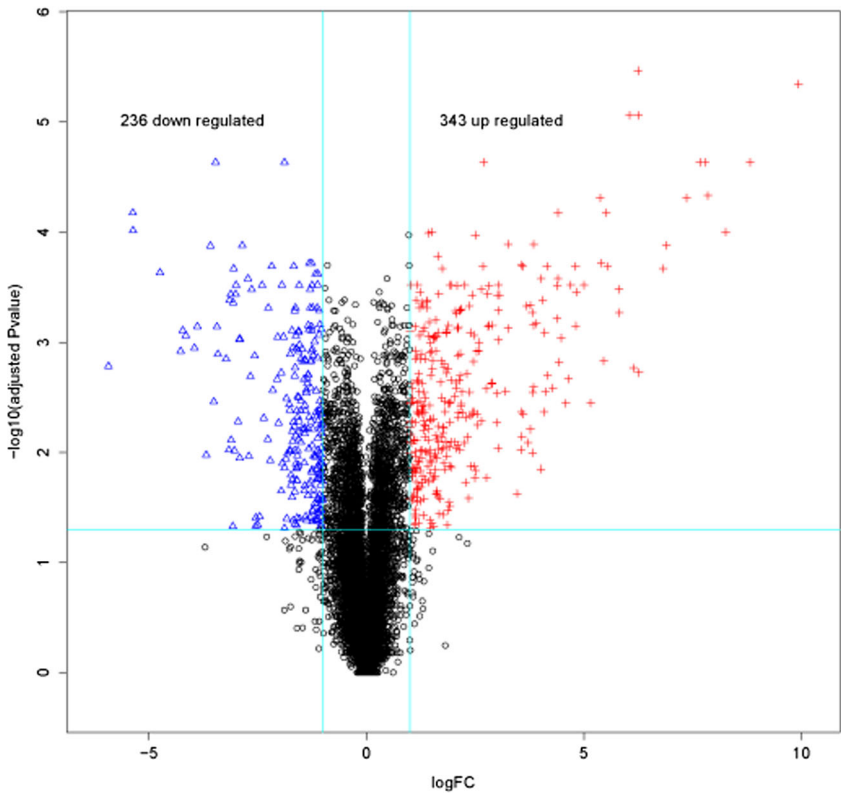

Fig. 1 The distribution of differential expression analysis in GATA3 overexpression breast samples compared with those in normal

Identification of Differential Expression Genes

Data preprocessing based on affy package of *R* were conducted after the raw CEL microarray data were downloaded, which involved background correction and expression level normalization via Robust Multi-array Average (RMA) method [13]. The corresponding gene symbols of the probe sets were obtained through the annotation package and the expression values were summaried for genes corresponding to multi probes. T test followed by Benjamini-Hochberg (BH) correction were applied for the expression matrix, and |fold change| > 2 and adjusted *P*-Value <0.05 were used for the screening of differential expression genes (DEGs).

Preprocessing and Mapping of ChIP-Seq Data

The raw. Fastq files were uploaded to FastaQC, a java-based high-throughput data quality control software, bases quality score less than 20 and reads average quality score less than 25 were filtered out based on a Python code. The remained reads were mapped to the UCSC GRCh37/hg19 genome based on Bowtie2, a fast DNA aligner [14].

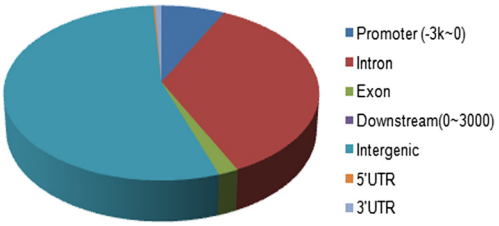

Fig. 3 Distribution of GATA3 enriched peaks across the human genome

Peak Calling and Annotation

To determine the target genes of GATA3, the Model-based Analysis of ChIP-Seq (MACS) version 2 was adopted for the identification of its binding sites [15]. Based on read position on genome and local background normalization, MACS, which contains at least the chromosome name, chromosome start and end positions, could calculate the peak

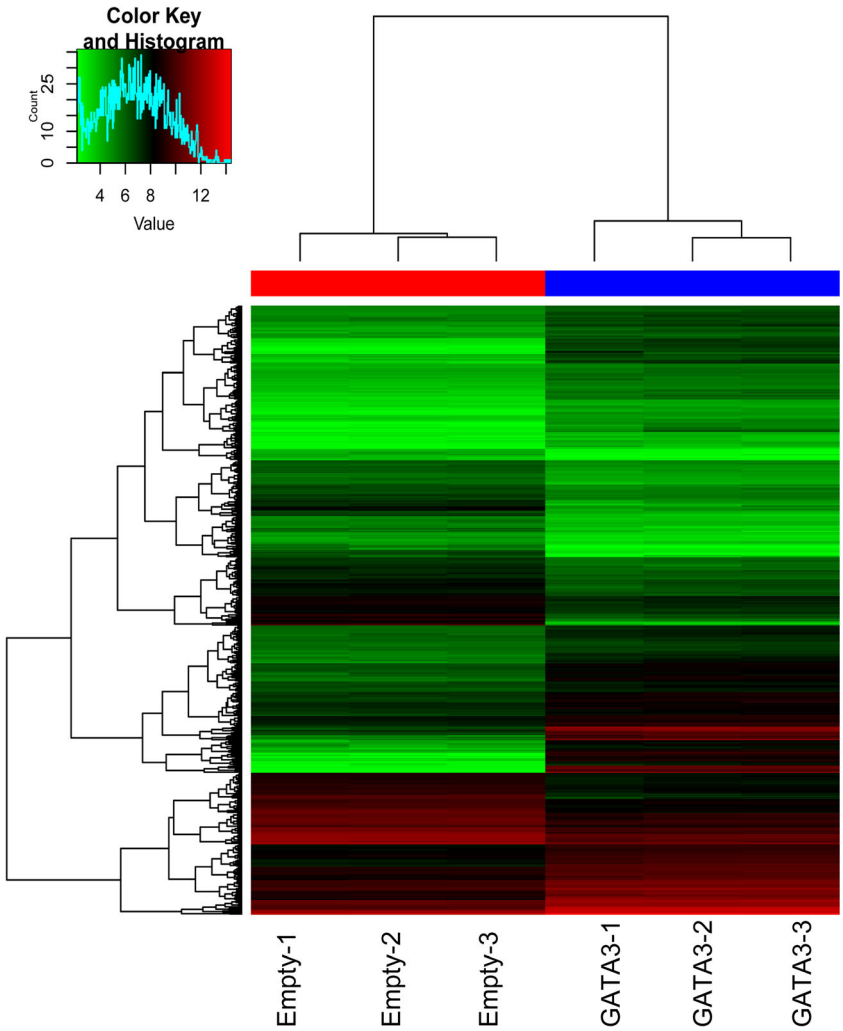

Fig. 2 Supervised clustering of GATA3 overexpression and control samples

enrichment and screen out the potential binding sites, which stored in bed file. Besides, through ChIPseeker, a *R* bioconductor package, which developed by Yu et al. [16], we annotated the peaks with their related genes, distance to the closest transcription start sites (TSS) and etc. To predict the active or repressive function of GATA3, Binding and Expression Target Analysis (BETA) which developed by Wang et al. was used for the combination analysis of gene expression and GATA3 binding sites [17].

Functional Enrichment Analysis

Overlapping genes of DEGs and targets of GATA3 which located in −3000 ~ 0 bp centered the (TSS) were screened out and their significantly enriched Gene Ontology (GO) terms and KEGG pathways were obtained based on the Database for Annotation, Visualization and Integrated Discovery (DAVID, http://david.abcc.ncifcrf.gov/) with the criteria of *P*-Value <0.05 [18].

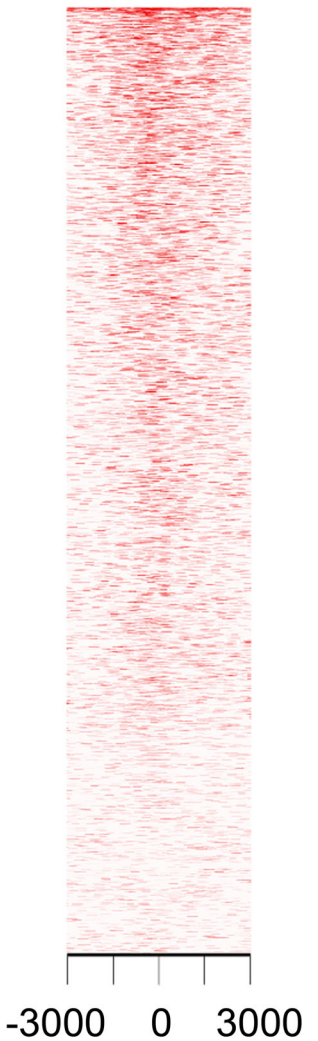

Fig. 4 Density profile of GATA3 ChIP-Seq enriched peak surrounding the TSS of their nearby genes

Protein-Protein Interaction Network

To explore the interactions among those overlapping genes, as well as with the other genes in genome, we downloaded the entire protein-protein pairs from the Human Protein Reference Database (HPRD, http://www.hprd.org) [19] and retrieved the pairs involved the overlaps. The protein-protein interaction (PPI) network was constructed via Cytoscape [20].

Results

Differential Expression Genes

Compared with the control samples, a total of 236 down-regulated and 343 up-regulated genes were identified (shown in Fig. 1). What's more, supervised clustering of GATA3 over-expression and control samples was shown in Fig. 2.

GATA3 Binding Profiles

A total of 44,687 enriched peaks were obtained which contained 3256 ones located in −3000 ~ 0 bp centered the TSS of the corresponding genes. The distribution of all the peaks across the genome was shown in Fig. 3. Besides, the

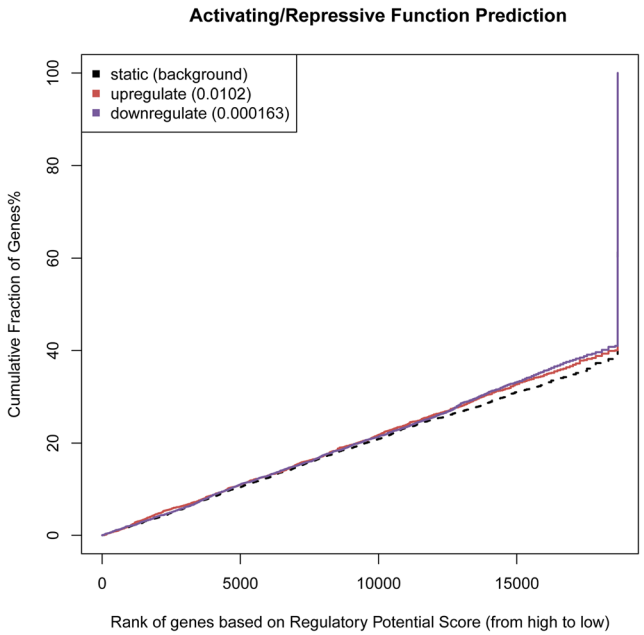

Fig. 5 Function prediction of GATA3 based on the combined analysis of ChIP-Seq and gene expression dataset. The dashed line indicated the non-differential expression genes as the background. *P* values represent the significance of difference in the UP or DOWN groups compared with the NON group by the Kolmogorov-Smirnov test

Table 1 The top 10 enriched gene ontology (GO) terms of overlapping genes according to *P*-Value

Category	GO Name	Gene Number	PValue
CC	Extracellular region	101	9.45×10^{-8}
CC	Extracellular region part	56	2.36×10^{-6}
BP	Positive regulation of transcription, DNA-dependent	33	1.06×10^{-5}
BP	Positive regulation of transcription from RNA polymerase II promoter	28	1.25×10^{-5}
BP	Positive regulation of RNA metabolic process	33	1.28×10^{-5}
MF	Copper ion binding	11	2.79×10^{-5}
BP	Positive regulation of macromolecule biosynthetic process	39	4.10×10^{-5}
BP	Positive regulation of cellular biosynthetic process	40	5.06×10^{-5}
BP	Positive regulation of biosynthetic process	40	6.91×10^{-5}
BP	Female pregnancy	13	8.66×10^{-5}

CC Cellular Component, *BP* Biological Process, *MF* Molecular Function

density profiles of GATA3 surrounding −3000 ~ 3000 of all its targets was obtained and visualized by ngsplot (Fig. 4). The integrated analysis of gene expression and binding profiles indicated that GATA3 might act as a repressor in breast cancer (Fig. 5).

Enriched GO Terms and KEGG Pathways

A total of 2904 genes were involved in the 3256 peaks located in −3000 ~ 0 bp centered the TSS, and 68 overlaps were identified between the 579 DEGs. Biological processes related to cell development, regulation of gene expression and so on were found to be enriched in those overlaps. The top 10 enriched GO terms according to *P*-Value were listed in Table 1. What's more, 3 KEGG pathways including TGF-beta signaling pathway, metabolism of xenobiotics by cytochrome P450 and bladder cancer were also enriched in those overlapping genes (Table 2).

PPI Network

Through HPRD, the PPI network which contained 387 nodes and 385 pairs was constructed (Fig. 6). *NDRG1* directly interact with 61 genes in the network, which down-regulated in the GATA3 overexpression samples compared with the control ones. Besides, another public dataset (GSE5460) also indicated the negative correlation between the expression of GATA3 and *NDRG1* (shown in Fig. 7), which might prove its carcinogenicity.

Discussion

Regulation of gene expression, including protein, such as transcription factor, non-coding RNA, and etc. was proved to involved in many types of diseases, including cancers [21–23]. Many regulators were found to play important roles in the progression of breast cancer and identified as valuable therapeutic targets. But the mechanisms underlying the regulation of gene expression in breast cancer still remains unclear and further studies are still needed.

GATA3 encodes a protein belongs to the GATA family of transcription factors which greatly contribute to T-cell development and endothelial cell biology. GATA3 acts as a suppressor in breast cancer and it expression level decreased with the progression of breast cancer [10, 24]. In the study of Si et al. [9], GATA3 was shown to interact with ZEB2 and G9 A to regulate some co-targets and a reciprocal feedback loop exists between GATA3 and ZEB2, and its dysfunction might contribute to the metastasis of breast cancer. But, no study had conducted integrated analysis of GATA3 binding profiles and expression data in GATA3 abnormal breast cancers.

In this study, combination analysis of ChIP-Seq and gene microarray indicated GATA3 mainly act as a repressor in the gene expression which consistent with many other studies [25, 26]. Besides, among the 68 overlapping genes, 28 ones were found to down-regulated in the GATA-overexpression samples compared with those in control, which might indicate carcinogenicity of those genes. The enriched GO terms and KEGG pathways of the 68 overlapping genes, such as positive

Table 2 The enriched kyoto encyclopedia of genes and genomes (KEGG) pathways of overlapping genes

Pathway Name	Gene Number	Pvalue	Genes
TGF-beta signaling pathway	8	0.0202	INHBA, NOG, LTBP1, ID2, FST, SMAD1, BMP5, PITX2
Metabolism of xenobiotics by cytochrome P450	6	0.0413	GSTM1, AKR1C3, GSTM2, GSTM4, CYP1B1, CYP1A1
Bladder cancer	5	0.0435	RPS6KA5, VEGFC, TP53, CDH1, MMP1

regulation of transcription, DNA-dependent, positive regulation of cellular biosynthetic process, TGF-beta signaling pathway and etc. further proved their roles in cancer and possibility of regulation by GATA3.

In the PPI network, *NDRG1* directly interacted with 61 out of all the 387 genes which might indicate its hub roles in the network. *NDRG1* is a member of the N-myc down-regulated gene family and its expression is a prognostic indictor for several types of cancer [27, 28]. What's more, the decreased expression level in GATA3 overexpression samples and negative correlation with the expression level of GATA3 in the public dataset all indicated its potential role in breast cancer. In

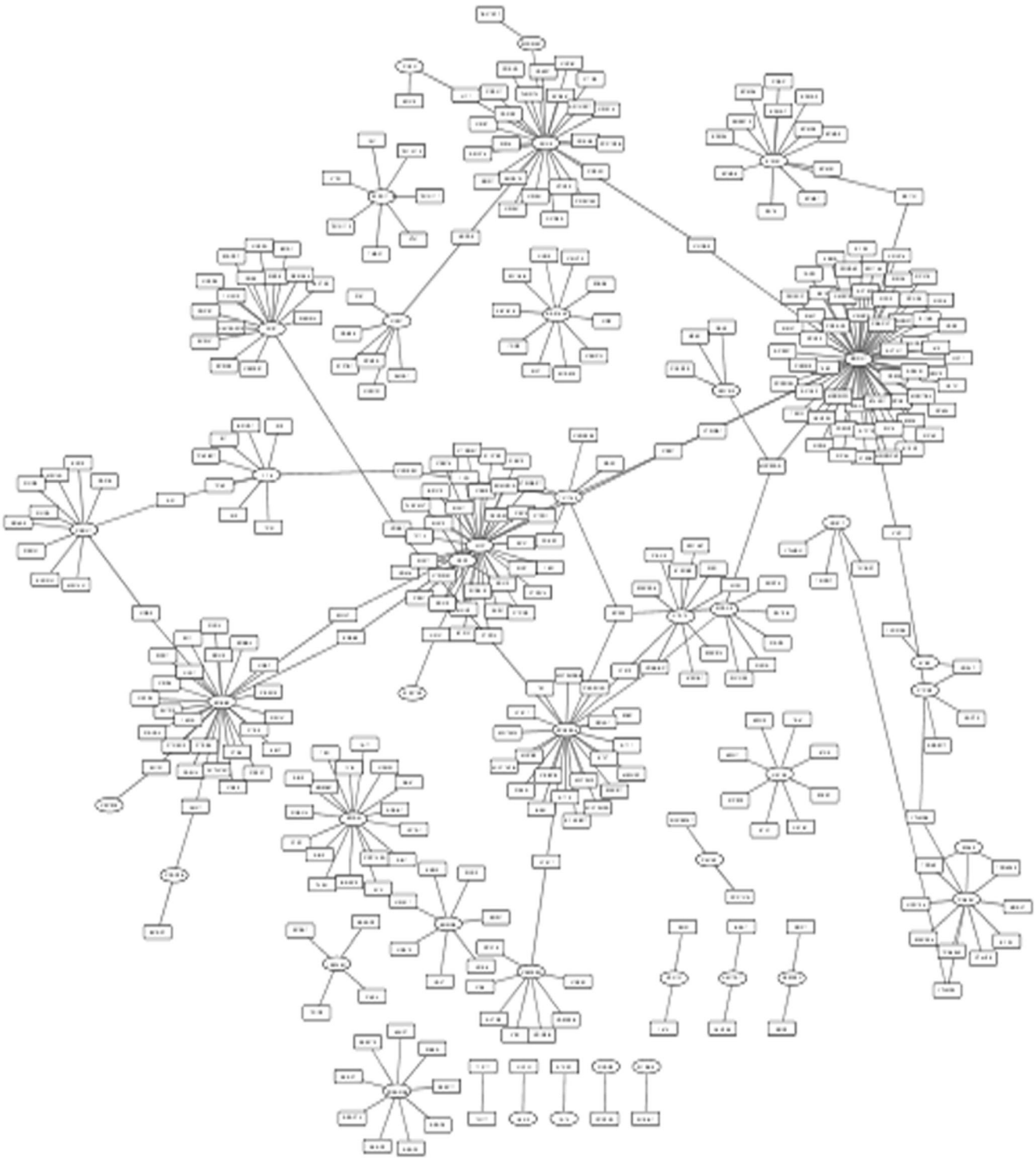

Fig. 6 The PPI network of overlapping genes. The ellipses represent differential expression genes and squares represent non-differential expression genes

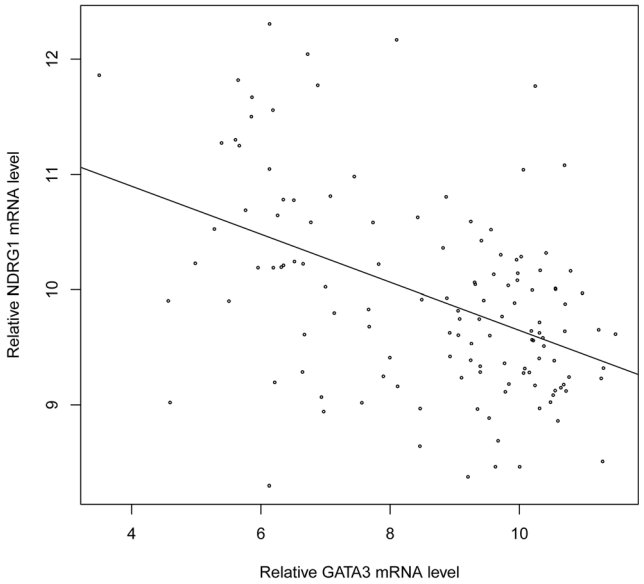

Fig. 7 Correlation analysis of the expression level of GATA3 and NDRG1 based on the public dataset

the study of Song et al., *NDRG1* and GATA3 were found to up- and down-regulated respectively in cervical cancer [29], which also indicated the tumor suppressed and carcinogenic effect of GATA3 and *NDRG1*. *JUP* directly interacted with 28 genes in the PPI network, which is next below *NDRG1* and consistent with Wang's study, which showed *JUP* was a hub gene in colorectal cancer through microarray technology [30]. So genes with high number of direct interactions in the PPI network might be potential biomarkers in the development of breast cancer.

In conclusion, the combined analysis of ChIP-Seq and gene expression profiles indicated GATA3 was a repressor in breast cancer and its variation was associated with the abnormal expression of many cancer related genes. Besides, some novel biomarkers were identified, but further experiments are still needed to confirm their functions.

Acknowledgments This work was supported by the Municipal Science and Technology Commission of Tianjin (No. 15ZLZLZF00440) and the Health Bureau Science and Technology Foundation of Tianjin (No. 2012KZ063 and No. 2014KZ102).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Gradishar WJ, Anderson BO, Balassanian R, Blair SL, Burstein HJ, Cyr A, Elias AD, Farrar WB, Forero A, Giordano SH, Goetz M, Goldstein LJ, Hudis CA, Isakoff SJ, Marcom PK, Mayer IA, McCormick B, Moran M, Patel SA, Pierce LJ, Reed EC, Salerno KE, Schwartzberg LS, Smith KL, Smith ML, Soliman H, Somlo G, Telli M, Ward JH, Shead DA, Kumar R (2016) Invasive breast cancer version 1.2016, NCCN clinical practice guidelines in oncology. Journal of the National Comprehensive Cancer Network. JNCCN 14(3):324–354
2. Pusztai L, Ladanyi A, Szekely B, Dank M (2016) Immunotherapy opportunities in breast cancer. Magyar onkologia 60(1):34–40
3. Liu D (2016) Tumor necrosis factor-alpha, a regulator and therapeutic agent on breast cancer. Curr Pharm Biotechnol 17(6): 486–494
4. Maximiano S, Magalhaes P, Guerreiro MP, Morgado M (2016) Trastuzumab in the Treatment of Breast Cancer. BioDrugs: clinical immunotherapeutics, biopharmaceuticals and gene therapy doi:10.1007/s40259–016–0162-9
5. Kober KM, Dunn L, Mastick J, Cooper B, Langford D, Melisko M, Venook A, Chen LM, Wright F, Hammer M, Schmidt BL, Levine J, Miaskowski C, Aouizerat BE (2016) Gene Expression Profiling of Evening Fatigue in Women Undergoing Chemotherapy for Breast Cancer. Biological research for nursing doi:10.1177/1099800416629209
6. Cao ZG, Qin XB, Liu FF, Zhou LL (2015) Tryptophan-induced pathogenesis of breast cancer. Afr Health Sci 15(3):982–985. doi:10.4314/ahs.v15i3.36
7. Park E (2016) Data on cell cycle in breast cancer cell line, MDA-MB-231 with ferulic acid treatment. Data in brief 7:107–110. doi:10.1016/j.dib.2016.02.001
8. Asch-Kendrick R, Cimino-Mathews A (2016) The role of GATA3 in breast carcinomas: a review. Hum Pathol 48:37–47. doi:10.1016/j.humpath.2015.09.035
9. Si W, Huang W, Zheng Y, Yang Y, Liu X, Shan L, Zhou X, Wang Y, Su D, Gao J, Yan R, Han X, Li W, He L, Shi L, Xuan C, Liang J, Sun L, Wang Y, Shang Y (2015) Dysfunction of the Reciprocal Feedback Loop between GATA3- and ZEB2-Nucleated Repression Programs Contributes to Breast Cancer Metastasis. Cancer Cell 27(6):822–836. doi:10.1016/j.ccell.2015.04.011
10. McCleskey BC, Penedo TL, Zhang K, Hameed O, Siegal GP, Wei S (2015) GATA3 expression in advanced breast cancer: prognostic value and organ-specific relapse. Am J Clin Pathol 144(5):756–763. doi:10.1309/AJCP5MMR1FJVVTPK
11. Dieci MV, Smutna V, Scott V, Yin G, Xu R, Vielh P, Mathieu MC, Vicier C, Laporte M, Drusch F, Guarneri V, Conte P, Delaloge S, Lacroix L, Fromigue O, Andre F, Lefebvre C (2016) Whole exome sequencing of rare aggressive breast cancer histologies. Breast Cancer Res Treat 156(1):21–32. doi:10.1007/s10549–016-3718-y
12. Chu IM, Michalowski AM, Hoenerhoff M, Szauter KM, Luger D, Sato M, Flanders K, Oshima A, Csiszar K, Green JE (2012) GATA3 inhibits lysyl oxidase-mediated metastases of human basal triple-negative breast cancer cells. Oncogene 31(16):2017–2027. doi:10.1038/onc.2011.382
13. Gautier L, Cope L, Bolstad BM, Irizarry RA (2004) Affy–analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20(3):307–315. doi:10.1093/bioinformatics/btg405
14. Langmead B (2010) Aligning short sequencing reads with Bowtie. Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al] Chapter 11:Unit 11 17. doi:10.1002/0471250953.bi1107s32
15. Feng J, Liu T, Qin B, Zhang Y, Liu XS (2012) Identifying ChIP-seq enrichment using MACS. Nat Protoc 7(9):1728–1740. doi:10.1038/nprot.2012.101
16. Yu G, Wang LG, He QY (2015) ChIPseeker: an R/bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics 31(14):2382–2383. doi:10.1093/bioinformatics/btv145
17. Wang S, Sun H, Ma J, Zang C, Wang C, Wang J, Tang Q, Meyer CA, Zhang Y, Liu XS (2013) Target analysis by integration of transcriptome and ChIP-seq data with BETA. Nat Protoc 8(12): 2502–2515. doi:10.1038/nprot.2013.150

18. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 4(5):P3
19. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, Telikicherla D, Raju R, Shafreen B, Venugopal A, Balakrishnan L, Marimuthu A, Banerjee S, Somanathan DS, Sebastian A, Rani S, Ray S, Harrys Kishore CJ, Kanth S, Ahmed M, Kashyap MK, Mohmood R, Ramachandra YL, Krishna V, Rahiman BA, Mohan S, Ranganathan P, Ramabadran S, Chaerkady R, Pandey A (2009) Human Protein Reference Database–2009 update. Nucleic Acids Res 37(Database issue): D767–D772. doi:10.1093/nar/gkn892
20. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13(11):2498–2504. doi:10.1101/gr.1239303
21. Lin Y, Zhao J, Hu X, Wang L, Liang L, Chen W (2016) Transcription factor CCAAT/enhancer binding protein alpha up-regulates microRNA let-7a-1 in lung cancer cells by direct binding. Cancer Cell Int 16:17. doi:10.1186/s12935–016–0294-5
22. Li C, Deng L, Zhi Q, Meng Q, Qian A, Sang H, Li X, Xia J (2016) MicroRNA-183 Functions As an Oncogene by Regulating PDCD4 in Gastric Cancer. Anti Cancer Agents Med Chem 16(4):447–455
23. Mitxelena J, Apraiz A, Vallejo-Rodriguez J, Malumbres M, Zubiaga AM (2016) E2F7 regulates transcription and maturation of multiple microRNAs to restrain cell proliferation. Nucleic Acids Res. doi:10.1093/nar/gkw146
24. Takaku M, Grimm SA, Wade PA (2015) GATA3 in Breast Cancer: Tumor Suppressor or Oncogene? Gene Expr 16(4):163–168. doi:10.3727/105221615X14399878166113
25. Tkocz D, Crawford NT, Buckley NE, Berry FB, Kennedy RD, Gorski JJ, Harkin DP, Mullan PB (2012) BRCA1 and GATA3 corepress FOXC1 to inhibit the pathogenesis of basal-like breast cancers. Oncogene 31(32):3667–3678. doi:10.1038/onc.2011.531
26. Xiong Y, Castro E, Yagi R, Zhu J, Lesourne R, Love PE, Feigenbaum L, Bosselut R (2013) Thpok-independent repression of Runx3 by Gata3 during CD4+ T-cell differentiation in the thymus. Eur J Immunol 43(4):918–928. doi:10.1002/eji.201242944
27. Liu W, Yue F, Zheng M, Merlot A, Bae DH, Huang M, Lane D, Jansson P, Lui GY, Richardson V, Sahni S, Kalinowski D, Kovacevic Z, Richardson DR (2015) The proto-oncogene c-Src and its downstream signaling pathways are inhibited by the metastasis suppressor, NDRG1. Oncotarget 6(11): 8851–8874. doi:10.18632/oncotarget.3316
28. Salis O, Bedir A, Kilinc V, Alacam H, Gulten S, Okuyucu A (2014) The anticancer effects of desferrioxamine on human breast adenocarcinoma and hepatocellular carcinoma cells. Cancer biomarkers: section A of Disease markers 14(6):419–426. doi:10.3233/CBM-140422
29. Song JY, Lee JK, Lee NW, Jung HH, Kim SH, Lee KW (2008) Microarray analysis of normal cervix, carcinoma in situ, and invasive cervical cancer: identification of candidate genes in pathogenesis of invasion in cervical cancer. International journal of gynecological cancer: official journal of the international gynecological cancer. Society 18(5):1051–1059. doi:10.1111/j.1525-1438.2007.01164.x
30. Wang Y, Zheng T (2014) Screening of hub genes and pathways in colorectal cancer with microarray technology. Pathol Oncol Res: POR 20(3):611–618. doi:10.1007/s12253–013–9739-5