Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Medicine

journal homepage:www.elsevier.com/locate/apjtm

Document heading doi:

Bioinformatics analysis of breast cancer bone metastasis related gene-CXCR4

Heng-Wei Zhang, Xian-Fu Sun, Ya-Ning He, Jun-Tao Li, Xu-Hui Guo, Hui Liu*

Department of Breast, Henan Cancer Hospital, Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou 450008, China

ARTICLE INFO

Article history:
Received 10 June 2013
Received in revised form 15 July 2013
Accepted 15 August 2013
Available online 20 September 2013

Keywords:
Breast cancer bone metastasis
CXCR4
Bioinformatics analysis

ABSTRACT

Objective: To analyze breast cancer bone metastasis related gene-CXCR4. **Methods:** This research screened breast cancer bone metastasis related genes by high-flux gene chip. **Results:** It was found that the expressions of 396 genes were different including 165 up-regulations and 231 down-regulations. The expression of chemokine receptor CXCR4 was obviously up-regulated in the tissue with breast cancer bone metastasis. Compared with the tissue without bone metastasis, there was significant difference, which indicated that CXCR4 played a vital role in breast cancer bone metastasis. **Conclusions:** The bioinformatics analysis of CXCR4 can provide a certain basis for the occurrence and diagnosis of breast cancer bone metastasis, target gene therapy and evaluation of prognosis.

1. Introduction

Breast cancer is one of the most common malignancies for women, the incidence of which increases year by year in recent years, but the age of onset keeps dropping. Breast cancer often has distant metastases in bone, liver, lung and others, among which the incidence of bone metastasis is the highest. Breast cancer bone metastasis is a complex process involving multiple genes, factors and steps. The research on the occurrence and molecular mechanism of breast cancer bone metastasis remains a hot spot at present. Since Muller *et al*[1] firstly reported that chemokine receptor was associated with tumor metastasis, and more and more chemokine receptors have been found to be closely related to many processes of tumor development in the subsequent studies. The chemokines are cytokines with low molecular weight (8-10 kD) secreted by different kinds of cells. Their receptors are a kind of G-protein-coupled receptors containing 7 transmembrane domains which can express in different types of cells. The chemokines and their receptors play an important role in the tumor target migration process in bone marrow and other specific organs and tissues containing hematopoietic cells. The interaction of the chemokine in bone matrix and the receptor in breast cancer cell promotes the specific bone metastasis of breast cancer[2]. It has been found that CXCR4 is expressed in at least 23 kinds of tumors till now[3]. Numerous studies demonstrated that CXCR4 is closely associated with tumorigenesis and is highly expressed in human breast cancer cell, malignant breast tumor and distant metastasis[4]. This study analyzed the molecular structure and functional characteristics of CXCR4-the chemokine receptor related to breast cancer metastasis by bioinformatics method, providing theoretical basis for

*Corresponding authors: Hui Liu, Department of Breast, Henan Cancer Hospital, Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou 450008, China
Tel: +86-13803866726
E-mail: lhliuhui1996@163.com
Foundation project: This research is supported by the Project of International Science and Technology Cooperation of Henan Province (094300510014).

the targeted therapy and prognosis of breast cancer bone metastasis.

2. Materials and methods

2.1. General data

A total of 120 breast cancer patients were selected, who visited the Out-patient Department or hospitalized in our hospital from March, 2009 to December, 2011, and all received surgical treatment and were finally diagnosed by pathological examination. The patients were divided into experimental group with76 cases of bone metastasis and control group 44 cases without bone metastasis according to the detection results of the single photon emission computed tomography, CT and others. The group without bone metastasis did not have metastasis of other organs and lymph node through CT, B-ultrasonography and other detections. The age of the bone metastasis group was 28-59 years with the average of 44.5 years. The age of the group without bone metastasis was 29-58 years old with the average of 43.5 years.

2.2. Methods

2.2.1. RNA extraction and probe preparation

The total RNA was extracted by Trizol one step method. The A260/A280 values of the RNA of two kinds of tissues were detected by the ultraviolet spectrophotometer. The agarose gel electrophoresis was performed to identify the quality of RNA. The cDNA probe with fluorescent labeling was synthesized and purified. Cy3 and Cy5 were used to label the control group and the experimental group, respectively. After ethanol precipitation, the probes were dissolved into the 20 μL 5×SSC+0.2% SDS hybridization solution.

2.2.2. Chip hybridization

The chips were pre-heated in the prehybridization buffer at 55 ℃. The labeled probes (Cy3/Cy5) were equally mixed. Then 7.5 μL 4× hybridization solutions, 15.0 μL formamide and some sterile water were added to a total volume of 30.0 μL, with 5 min water bath at 95 ℃ and 16 h incubation in the incubator at 42 ℃ after blended and centrifuged. The chip was placed in the staining jar containing eluant (0.1% SSC) at 55 ℃ in the rocking bed. After slowly weaved for 20 min at room temperature, the chip was repeatedly washed once, and then was air-dried at room temperature.

2.2.3. Detection and analysis

The chip results were scanned by Agilent Microarray Scanner. The data were read by Feature Extraction Software 10.7. Finally, Gene Spring Software11.0 was used to perform normalized processing, and the used algorithm was Lowess.

2.2.4. Bioinformatics analysis of differentially expressed genes

The open reading frames (ORFs) of CXCR4 genes were found out by the online tool NCBI ORF Finder. The structure and function of human CXCR4 gene were analyzed by the online analysis softwares such as ProtParam, ProtScal, PSORTII and SOPMA.

3. Results

3.1. Gene chip

The signal points adjacent to slope rate 1.0 revealed no expression difference. The signal points whose slope rates were more than 2 and 0.5 were differentially expressed genes. The fluorescent labeled signal values were distributed in both sides of the symmetry axis, indicating that there were up-regulated genes and down-regulated genes (Figure 1, 2). Compared with the control group, there were 396 differentially expressed genes with 165 up-regulated genes and 231 down-regulated genes.

Figure 1. Stacking chart of double-color fluorescent label for breast cancer bone metastasis and breast cancer without bone metastasis. Green: down-regulated; red: up-regulated; yellow: similar intensity.

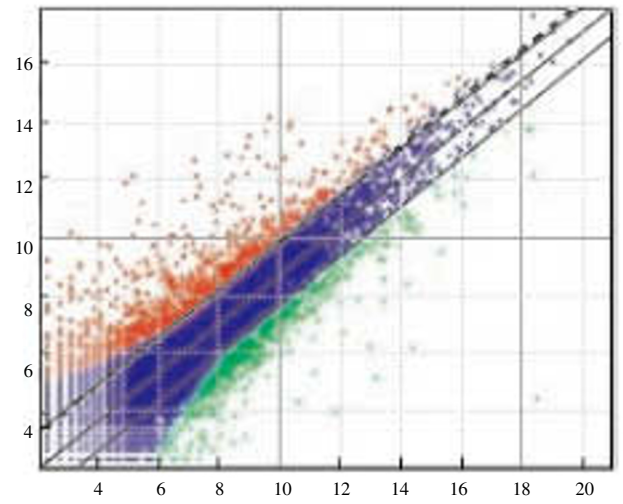


Figure 2. Scatter plot of gene chip hybridization signals.

3.2. Bioinformatics analysis of CXCR4

3.2.1. ORF analysis

ORF is the normal nucleotide sequence of structure gene. The ORF from start codon to stop codon can encode a complete polypeptide chain, and there is no stop codon that can interrupt the translation in ORF. The identification of ORF is the precondition for proving that a new DNA sequence is the partial or full coding gene of a specific protein. The analysis results (Figure 3) demonstrated that the +2 ORF of CXCR4 was the longest with the length of 1 071 bp, encoding 356 amino acids.

Figure 3. ORF analysis of human CXCR4 gene sequence.

3.2.2. Physicochemical property analysis of coding product

The protein physicochemical properties include relative molecular weight, theoretical isoelectric point/atomic composition, instability coefficient, fat coefficient, general mean hydrophobicity and other physicochemical parameters.

Physicochemical property of the coding product of CXCR4 gene was predicted by the ExPASy online tool ProtParam. The amino acids composition of CXCR4 gene is shown in Figure 4. CXCR4 gene encoded 365 amino acids with most of Ile and Leu, accounting for 10.1% and 11.8%, respectively. The total residue number (Arg+Lys) of positive charge was 31, while the total residue number (Asp+Glu) of negative charge was 26, with the theoretical molecular weight of 40.22 kD and the theoretical isoelectric point of 8.61.

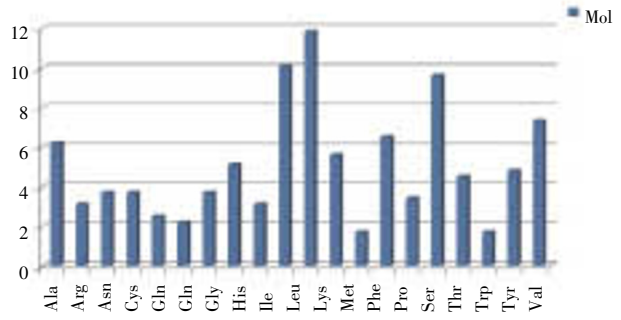


Figure 4. Amino acids composition of CDS sequence of CXCR4 gene.

3.2.3. Hydrophobic / hydrophilic analysis of coding product

The structure of protein determines its physiological function, and the hydrophobicity/hydrophilicity of amino acids can directly influence the structure and function of protein. The ExPASy ProtScal program was used to analyze the hydrophobicity of the encoding product of human CXCR4 gene. The results are shown in Figure 5, and the x-axis and y-axis were amino acid sequence and amino acid scale. The hydrophobicity was stronger as the value was larger. Figure 5 shows that the whole polypeptide chain expressed as hydrophobicity, therefore the CXCR4 was a water-insoluble protein.

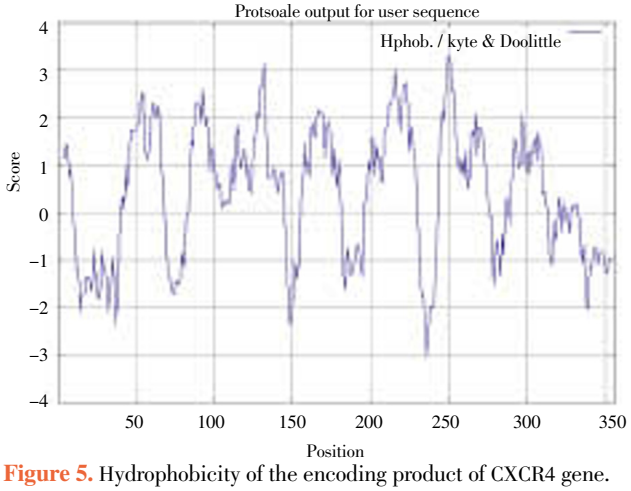


Figure 5. Hydrophobicity of the encoding product of CXCR4 gene.

3.2.4. Signal peptide prediction of the gene coding product

Signal peptide is a short peptide sequence composed by 3-60 amino acids. It is the identification signal that can transport the newly-synthesized protein to the destination, thus it can also be called target signal. SignalP 3.0 Server tool was utilized to predict the signal peptide of human CXCR4. As shown in Figure 6, the C score was the largest, S was steep, and Y value peaks between 68-70 amino acids.

It can be basically determined that signal peptide existed in the cleavage site.

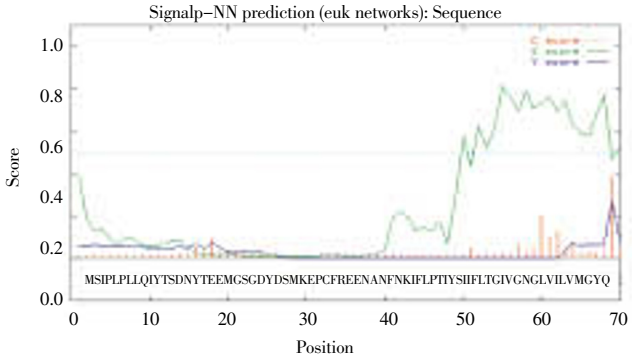


Figure 6. Signal peptide prediction of amino acids sequence of the CXCR4 coding products.

3.2.5. Subcellular localization

The online prediction tool PSORTII was used to analyze the CXCR4 subcellular localization. The analytic results (Table 1) indicated that this protein was probably located in the cellular membrane with the reliability of 56.5%.

Table 1

Prediction of subcellular localization of the CXCR4 encoding product.

Subcellular localization	Possibility (%)
Cellular membrane	56.5
Endocytoplasmic reticulum	21.7
Vacuole	8.7
Golgi apparatus	4.3
Cellular nucleus	4.3
Mitochondria	4.3

3.2.6. Structure prediction

SOPMA online analysis software was used to predict the secondary structure of CXCR4 encoding product (Figure 7). There were 173 α-helixes (48.60%), 61 extended strands (17.13%), 10 β-turns (2.81%) and 112 random coils (31.46%) in the CXCR4 structure.

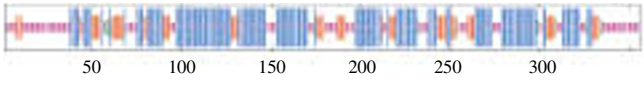


Figure 7. Secondary structure prediction of CXCR4 encoding product by SOPMA.

The analytic result of QuickPhyre online software indicated that the CXCR4 secondary structure had the greatest similarity with d1u19a protein. The sequence alignment between these two kinds of proteins is shown in Figure 8.

The three-dimensional structure analyzed by Phyre2 online software is shown in Figure 9.

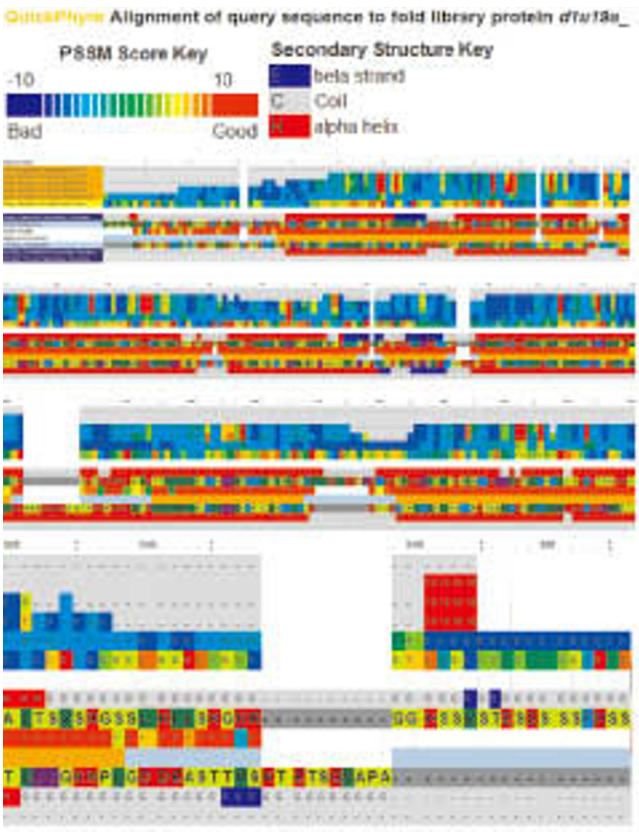


Figure 8. Sequence alignment between CXCR4 protein and d1u19a protein.

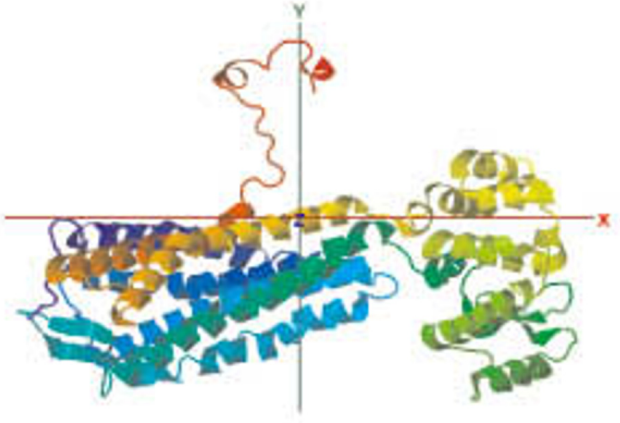
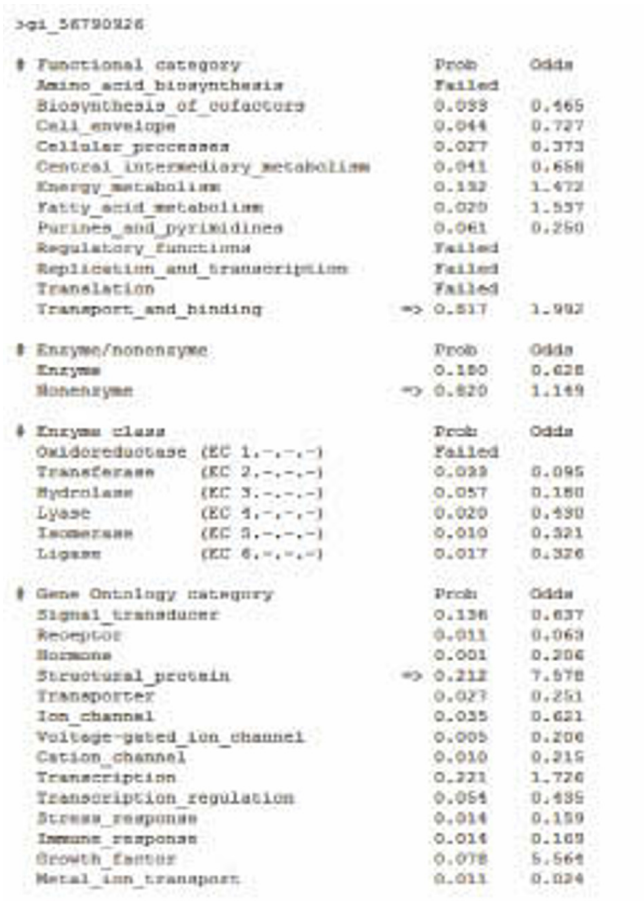


Figure 9. Three-dimensional structure of CXCR4.

3.2.7. CXCR4 protein function analyzed by Protfun online software

Developed by the biological sequence analysis center of Technical University of Denmark, Protfun is a kind of software used to predict and analyze the structural function and classification of protein[5]. The analytic results of Protfun (Figure 10) indicated that CXCR4 protein belonged to non-enzyme. Functional category displayed that CXCR4 had the function of transfer and combination (Prob=0.817,

Odds=1.992). GO category indicated that CXCR4 was a kind of structure protein.



```
>gi_56790926

# Functional category                    Prob     Odds
  Amino_acid_biosynthesis                Failed
  Biosynthesis_of_cofactors              0.033    0.465
  Cell_envelope                          0.044    0.727
  Cellular_processes                     0.027    0.373
  Central_intermediary_metabolism        0.041    0.658
  Energy_metabolism                      0.132    1.472
  Fatty_acid_metabolism                  0.020    1.557
  Purines_and_pyrimidines                0.061    0.250
  Regulatory_functions                   Failed
  Replication_and_transcription          Failed
  Translation                            Failed
  Transport_and_binding               => 0.517    1.992

# Enzyme/nonenzyme                       Prob     Odds
  Enzyme                                 0.180    0.628
  Nonenzyme                           => 0.820    1.169

# Enzyme class                           Prob     Odds
  Oxidoreductase (EC 1.-.-.-)            Failed
  Transferase    (EC 2.-.-.-)            0.033    0.095
  Hydrolase      (EC 3.-.-.-)            0.057    0.180
  Lyase          (EC 4.-.-.-)            0.020    0.490
  Isomerase      (EC 5.-.-.-)            0.010    0.321
  Ligase         (EC 6.-.-.-)            0.017    0.326

# Gene Ontology category                 Prob     Odds
  Signal_transducer                      0.136    0.637
  Receptor                               0.011    0.063
  Hormone                                0.001    0.206
  Structural_protein                  => 0.212    7.578
  Transporter                            0.027    0.251
  Ion_channel                            0.035    0.621
  Voltage-gated_ion_channel              0.005    0.206
  Cation_channel                         0.010    0.215
  Transcription                          0.221    1.726
  Transcription_regulation               0.054    0.435
  Stress_response                        0.014    0.159
  Immune_response                        0.014    0.169
  Growth_factor                          0.078    5.564
  Metal_ion_transport                    0.011    0.024
```

Figure 10. Prediction result of CXCR4 protein function.

3.2.8. Analysis of the structure functional domain

The 7 transmembrane domains in CXCR4 structure functional domain were analyzed by SMART software (Figure 11).

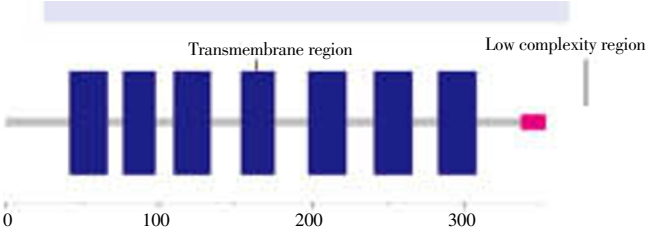


Figure 11. Functional domain distribution of CXCR4 protein structure.

3.2.9. Analysis of phosphorylation sites

Phosphorylation and dephosphorylation play important roles in many eukaryotic cells, including metabolism, cell division, signal transduction and others. The NetPhos2.0 Server online software analytical tool was used to analyze the CXCR4 phosphorylation sites. The results (Figure 12) indicated that there were 15 serines located in the 27, 75, 85, 182, 231, 233, 323, 328, 329, 342, 345, 348, 350, 351 and 352 site, 2 threonines located in the 283 and 346 site and 4 tyrosines located in the 16, 25, 161 and 188 site, which may be phosphorylated.

Figure 12. Phosphorylation sites predicted by NetPhos2.0 Server.

3.2.10. Phylogenetic tree of CXCR4 encoding product

The CXCR4 nucleotide and amino acids sequences were performed Blast homology search in NCBI database. The species with the homology higher than 96% were collected. The pair and multiple sequence alignments were performed using ClustalW, and the UPGMA phylogenetic tree was built by Phylogeny. Figure 13 indicated that the evolutionary distance between human CXCR4 and chimpanzee and Hylobates hoolock was the shortest with the closest genetic relationship, which was consistent with the zootaxy. The homology represented the genetic relationship among species and also reflected the importance of the structural stability of CXCR4 encoding product in different species on the function of organism.

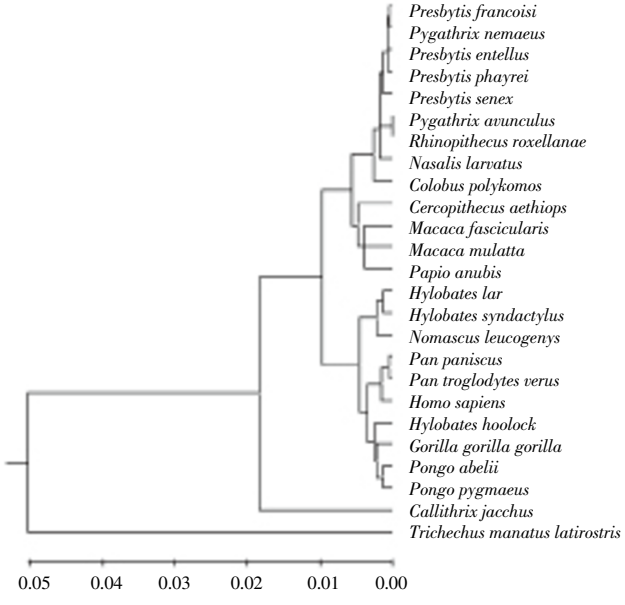


Figure 13. Phylogenetic tree of the CXCR4 encoding products of different species.

4. Discussion

The comparison of breast cancer bone metastasis tissue and breast cancer tissue without bone metastasis by gene

chip indicated that compared with the control group, the expressions of 396 genes had difference with 165 up-regulations and 231 down-regulations in the experimental group. We cannot perform in-depth study one by one due to the numerous genes. This experiment and previous reports[6,7] demonstrated that CXCR4 is not expressed in normal mammary gland, and its expression amount in breast cancer bone metastasis tissue is obviously higher than the breast cancer tissue without bone metastasis. Therefore, we performed bioinformatics analysis of the CXCR4 gene in order to provide a certain basis for the study on the mechanism of breast cancer bone metastasis.

CXCR4 is a kind of protein containing 356 amino acids which is encoded by the gene separated from the cDNA of human monocyte by IL-8R gene probe. It is a kind of highly conserved G-protein-coupled transmembrane receptor[8,9]. It can cross the cellular membrane and transfer the signal in external environment to interior, having control action on the process related to tissue development such as cell growth, hormone secretion and light perception. Generally speaking, CXCR4 can help to activate immune system and stimulate cell motivation. When the signal activating receptor cannot be correctly modulated, CXCR4 would accelerate the aggravation and proliferation of cancer. CXCR4 is a chemokine receptor in GPCR gene family, and exists widely in the central nervous system and immunocytes. After combined with ligand stromal cell-derived factor-1 (SDF-1), CXCR4 can cause the migration of immunocytes and development of nervous system. CXCR4 is also the receptor of HIV virus entering T cell, and its ligand SDF-1 can block HIV infection. CXCR4 can down-regulate many signal pathways. CXCR4 binding SDF-1 can activate the signal pathway mediated by G protein and down-regulate ras and PI3 kinase. CXCR4 can activate AP-1 and chemokines to regulate gene transcription.

Chemokine receptor CXCR4 is the specific receptor of chemokine stromal cell-derived factor-1(CXCL12), playing an important role in the formation and metastasis of human breast cancer[10]. The analysis of differentially expressed genes in this study indicated that the expression of CXCR4 was obviously up-regulated, which is consistent with the previous reports[11-13] that the expression level of CXCR4 obviously increased in many malignant tumor tissues including breast cancer. *In vitro* experiments CXCR4 binding its corresponding ligand SDF-1 can induce the metastasis of breast cancer cell, showing dosage effect relationship. Using anti-CXCR4 antibody or RNA interference technique to inhibit the expression of CXCR4 gene can inhibit the *in vitro* metastasis of breast cancer cells[14]. Some studies indicated that CXCR4 is the only receptor of chemokine SDF-1, which is expressed in many malignant tumor cells and leukemic cells[15-17] and participates in the metastasis of tumor[18]. An SL *et al*[19] studied the CXCR4 expression in human breast cancer and precancerous lesion by immunohistochemistry, analyzed the relationship between the other clinicopathological parameters related to breast cancer, and concluded that CXCR4 may be the early molecular event of breast cancer, its expression in invasive breast cancer was related to the clinicopathological parameters of breast cancer advancement, and CXCR4 can be regarded as the diagnostic indicator of breast cancer and could be the new target for breast cancer treatment. Cabioglu *et al*[20] performed a study to determine whether CXCR4, CCR7, ER, PR and HER2-neu in metastatic breast cancer were expressed in some organ-specific metastasis, and they found that CXCR4 can promote primary breast cancer to have bone metastasis. There are more and more studies on CXCR4 at present, and the targeted therapy to CXCR4 would be a new hot spot on gene therapy of tumors.

Conflict of interest statement

We declare that we have no conflict of interest.

References

[1] Muller A, Homey B, Soto H, Ge N, Catron D, Buchanan ME, et al. Involvement of chemokine receptors in breast cancer metastasis. *Nature* 2001; **410**(6824): 50-56.

[2] Kakinuma T, Hwang ST. Chemokines, chemokine receptors, and cancer metastasis. *J Leukocyte Biol* 2006; **79**: 4639-4651.

[3] Balkwill F. Cancer and the chemokine network. *Nat Rev Cancer* 2004; **4**(7): 540-550.

[4] Shim H, Lau SK, Devi S, Yoon Y, Cho HT, Liang Z. Lower expression of CXCR4 in lymph node metastases than in primary breast cancers: potential regulation by ligand-dependent degradation and HIF-1alpha. *Biochem Biophys Res Commun* 2006; **346**(1): 252-258.

[5] Jensen LJ, Gupta R, Staerfeldt HH, Brunak S. Prediction of human protein function according to Gene Ontology categories. *Bioinformatics* 2003; **19**(5): 635-642.

[6] Cabioglu N, Sahin A, Morandi P, Meric-Bernstam F, Islam R, Lin HY, et al. Chemokine receptor CXCR4 expression in breast cancer as a potential predictive marker of isolated tumor cells in bone marrow. *Clin Exp Metastasis* 2005; **22**(1): 39-46.

[7] Cronin PA, Wang JH, Redmond HP. Hypoxia increase the metastatic ability of breast cancer cells via upregulation of CXCR4. *BMC Cancer* 2010; **10**: 225.

[8] Loetscher M, Geiser T, O'Reilly T, Zwahlen R, BaggioliniM, Moser B. Cloning of a human seven-transmembrane domain receptor,

LESTR, that is highly expressed in leukocytes. *J Biol Chem* 1994; **269**(1): 232-237.

[9] Otsuka S, Bebb G. CXCR4 signaling pathway. *Bio Carta J Thorac Oncol* 2008; **3**: 1379-1383.

[10]Liu FF, Wei J, Lang RG. Significance of expression of stromal cell derived factor 1 and CXCR4 in invasive breast cancer. *Zhonghua Bing Li Xue Za Zhi* 2008; **37**(8): 529-535.

[11]Helbig G, Christopherson KW, Bhat-Nakshatri P, Kumar S, Kishimoto H, Miller KD, et al. NF-kappaB promotes breast cancer cell migration and metastasis by inducing the expression of the chemokine receptor CXCR4. *J Biol Chem* 2003; **278**(24): 21631-21638.

[12]Geminder H, Sagi-Assif O, Goldberg L, Meshel T, Rechavi G, Witz IP, et al. A possible role for CXCR4 and its ligand, the CXC chemokine stromal cell-derived factor-1, in the development of bone marrow metastases in neuroblastoma. *J Immunol* 2001; **167**(8): 4747-4757.

[13]Ma Q, Jones D, Springer TA. The chemokine receptor CXCR4 is required for the retention of B lineage and granulocytic precursors within the bone marrow microenvironment. *Immunity* 1999; **10**(4): 463-471.

[14]Chen Y, Stamatoyannopoulos G, Song CZ. Down-regulation of CXCR4 by inducible small interfering RNA inhibits breast cancer cell invasion *in vitro*. *Cancer Res* 2003; **63**(16): 4801-4804.

[15]Shibuta K, Mori M, Shimoda K, Inoue H, Mitra P, Barnard GF, et al. Regional expression of CXCL12/CXCR4 in liver and hepatocellular carcinoma and cell-cycle variation during *in vitro* differentiation. *Jpn J Cancer Res* 2002; **93**(7): 789-797.

[16]Sehgal A, Keener C, Boynton AL, Warrick J, Murphy GP. CXCR4, a chemokine receptor, is overexpressed in and required for proliferation of glioblastoma tumor cells. *J Surg Oncol* 1998; **69**(2): 99-104.

[17]Mohle R, Bantz, Malcolm AS. The chemokine receptor CXCR4 is expressed on $CD34^+$ hematopoietic progenitors and leukemic cells and mediates transendoelial migration induced by stromalcell-derived factor-1. *Blood* 1998; **91**(16): 4523-4530.

[18]Wang JM, Deng X, Gong W, Su S. Chemokine and their role in tumor growth and metastasis. *Immunol Methods* 1998; **220**(1): 17-21.

[19]An SL, Yu Y, Liu J. Expression and significance of CXCR4 receptor in human breast cancer. *Chin J Surgery Integrated Traditional & Western Med* 2009; **15**(4): 360-363.

[20]Cabioglu N, Sahin AA, Morandi P, Meric-Bernstam F, Islam R, Lin HY, et al. Chemokine receptors in advanced breast cancer: differential expression in metastatic disease sites with diagnostic and therapeutic implications. *Ann Oncol* 2009; **20**: 113-1019.