RESEARCH ARTICLE

Integrated Bioinformatics Approach Reveals Crosstalk Between Tumor Stroma and Peripheral Blood Mononuclear Cells in Breast Cancer

Lang He[1,2]*, Dan Wang[2], Na Wei[3], Zheng Guo[1,4]

Abstract

Breast cancer is now the leading cause of cancer death in women worldwide. Cancer progression is driven not only by cancer cell intrinsic alterations and interactions with tumor microenvironment, but also by systemic effects. Integration of multiple profiling data may provide insights into the underlying molecular mechanisms of complex systemic processes. We performed a bioinformatic analysis of two public available microarray datasets for breast tumor stroma and peripheral blood mononuclear cells, featuring integrated transcriptomics data, protein-protein interactions (PPIs) and protein subcellular localization, to identify genes and biological pathways that contribute to dialogue between tumor stroma and the peripheral circulation. Genes of the integrin family as well as CXCR4 proved to be hub nodes of the crosstalk network and may play an important role in response to stroma-derived chemoattractants. This study pointed to potential for development of therapeutic strategies that target systemic signals travelling through the circulation and interdict tumor cell recruitment.

Keywords: Breast cancer - tumor stroma - peripheral blood - molecular crosstalk - integrated bioinformatics

Asian Pac J Cancer Prev, **17 (3)**, 1003-1008

Introduction

Breast cancer is a major cause of morbidity and the leading global cause of cancer death in women. Breast cancer constitutes 16% of all female-related cancers yet resulting in approximately half a million deaths every year (2015a). Clinical management relies on known prognostic factors, such as hormone receptor (HR) and HER2 status, for predicting responsiveness to therapies (Kwast et al., 2014; Song et al., 2015). However, Even when morphological characteristics and ER phenotypes are similar, patients have varying prognosis or chemotherapy response (2015b). Gradually increasing resistance against conventional therapies requires improving strategies to diagnose and treat breast cancer.

Recent insights into the cancer biology has provided evidence that cancer progression is driven not only by a tumor's genetic alterations and interactions within the tumor microenvironment (TME), but also by complex systemic processes (McAllister and Weinberg, 2014a). Tumor-derived factors may mobilize host cells from distant tissues, for instance, host circulating cells, the bone marrow and spleen (Shaked et al., 2014), recruiting various peripheral blood cells from the circulation into the TME, resulting impingement on cancer progression and metastasis (Christopher et al., 2011; Hanahan and Coussens, 2012; Joyce and Fearon, 2015).

In addition to tumor-driven systemic perturbations , the stroma cells of TME be educated and sculpted by tumor cells via paracrine and juxtacrine, also have capacity of promoting tumor progress through systemic effects (Quail and Joyce, 2013). For example, Cancer-associated fibroblasts (CAFs) derived circulating CXCL12 was shown to mobilize the progenitor cells into the circulation, subsequently leading to their recruitment into the TME to promote angiogenesis (Hattori et al., 2001; Orimo et al., 2005). These TME-driven systemic perturbations , which can also influence disease outcome, remain ambiguous.

However, the composition of TME is very complicated, including immune cells, CAF, blood vessels, lymphatic vessels. This cell-type heterogeneity makes it very difficult to study their crosstalk with peripheral blood cells via co-culture. The genome-wide ''omics'' seems more important due to the absence of such an ''experimental testing ground''. Re-analysis of large data sets with advanced analysis tools will offer valuable clues.

Herein, we analyzed two mRNA expression profiling datasets, including peripheral blood mononuclear cells (PBMCs) and tumor stroma samples from breast cancer patients respectively. To provide biologically meaningful

*[1]Bioinformatics Centre, School of Life Science, University of Electronic Science and Technology of China, [2]Department of Biomedicine, Chengdu Medical College, [3]Sichuan Cancer Hospital and Institute, Chengdu, [4]Department of Bioinformatics, School of Basic Medical Sciences, Fujian Medical University, Fuzhou, China *For correspondence: helang79@sohu.com*

results, we propose comprehensive bioinformatics approaches to construct crosstalk network between the breast tumor stroma and the peripheral blood cells.

Materials and Methods

Data sources

Microarray data were downloaded from Gene Expression Omnibus (GEO) (Barrett et al., 2013), Accession number GSE9014, includes 53 cases of tumor stroma laser-capture microdissected from IDC breast cancer cases, and 31 cases of individual-matched normal adjacent stroma. Genes whose expression varied most between tumor tissue and normal stroma for the 31 tissue-matched pairs of this data had been identified to be independent of ER, HER2 and lymph node status, as well as age, grade and tumor size (Finak et al., 2008). Accession number GSE27562 (LaBreche et al., 2011), includes gene expression analysis of PBMCs from 57 women with a diagnosis of breast cancer and 31 healthy women with normal mammograms.

Basic stages of comprehensive bioinformatics approach

I. Data preparation: Obtain transcriptional profiling data from peripheral blood and stromal tissue compartments from samples of breast cancer patients.

II. Crosstalk components analysis: connect network of ligands-receptors-interacting proteins. Identify the key components as follows:

i.Upregulated ligands in tumor stromal cells compared to normal counterparts;

ii.Their corresponding receptors expressed in peripheral blood cells of breast cancer patients;

iii. Intracellular interacting proteins of these receptors expressed in peripheral blood cells of breast cancer patients;

III. Network analysis: Perform network analyses to identify network hubs, differential network features, and differential functional enrichment.

Data Preprocessing and Identification of differentially-expressed genes (DEGs)

The raw data were analyzed using the integrated software package BRB-Array Tools version 4.0 (2015c). BRB-Array Tools performed a series of preprocessing steps including filtering and computing hybridization data of each probe in the probe set. Probes with the following conditions were removed: 1) greater than 20% of expression data values had more than 1.5-fold change in either direction from the median value; 2) more than 50% of the gene expression data were "absent". The Robust Multiarray Average (RMA) method (Irizarry et al., 2003) was used for normalization. The SAM (Statistical Analysis of Microarray) method (Tusher et al., 2001) was used to analyze the transcription profiles and screen for DEGs (False Discovery Rate [FDR]=1%, fold change [FC]≥1.5).

Screening of DEGs for ligands and their cognate receptors

The significantly upregulated ligands of breast cancer stroma and their cognate receptors were sorted based on two databases, the Database of Ligand-Receptor Partners (DLRP) (Graeber and Eisenberg, 2001), a database of protein ligand and protein receptor pairs that are known to interact with each othe, and The Universal Protein Resource (UniProt) (2015d), UniProtKB entry: Cellular component/Subcellular locations/Secreted. The cognate receptors of these ligands were identified by analyzing protein interactions from the two databases simultaneously. To verify interaction relationships the ligand-receptor pairs were evaluated manually using published literature.

Selecting intracellular proteins interacting with receptors

The interaction proteins of the receptors was selected based on the UniProt database and the Search Tool for the Retrieval of Interacting Genes (STRING) database (Franceschini et al., 2013) (combined score >0.4, proteins with subcellular localizations in "extracellular" and "secreted" were excluded). The crosstalk network of ligands-receptors-interacting proteins was integrated and visualized by Cytoscape software (Shannon et al., 2003). The property of the network was evaluated using the network analyzer plugin of Cytoscape.

GO and KEGG pathway enrichment analysis

To uncover further insights into the precise biological function and signaling pathways involved with the crosstalk genes identified in the present study, Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were performed for the genes in the crosstalk network. The online based tool of the database for annotation, visualization and integrated discovery (DAVID) (Huang et al., 2009) was used to perform these analyses with $P<0.05$.

Results

Differential up-regulation ligands in breast cancer stroma samples

A total of 3420 transcripts, corresponding to 2737

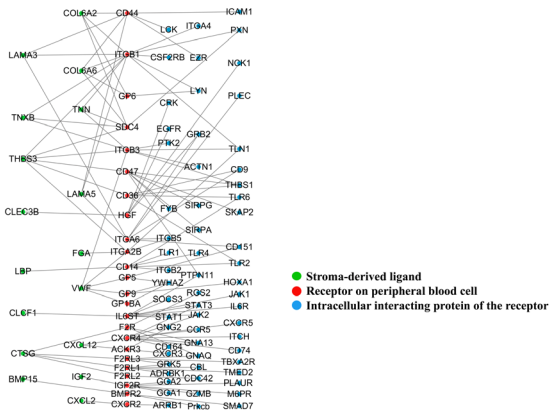

Figure 1. The Network of Molecular Interactions between Breast Tumor Stroma and PBMCs. The green nodes represent upregulated differentially expressed breast tumor stroma-derived ligands, the red represent coresponding PBMC membrane receptors, while the blue indicate intracellular proteins interacting with receptors.Grey lines stand for the interaction between two proteins

DEGs, including 1122 upregulated DEGs and 1615 downregulated DEGs were identified between the breast cancer stroma samples and the controls. Among these upregulated DEGs, 19 were define as ligand according to the two databases, DLRP and UniProt. There were 26 corresponding receptors expressed in PBMCs were retrieved as potential crosstalk components. The complete list of ligand-receptor pairs were show in Table 1.

Crosstalk network of ligands-receptors-interacting proteins

A total of 73 proteins (genes) and 189 interactions with the receptors met the selection criteria. The interactions include direct (physical) and indirect (functional) associations as derived from four sources, including literature reported protein interactions, genome analysis and prediction, high-throughput experiments and co-expression studies (Franceschini et al., 2013). Associated with ligand data, the integrated interaction relationships visualized with Cytoscape revealed a complex interlaced network of crosstalk between tumor stroma and peripheral blood, as shown in Figure 1. This network consisted of 19 tumor stroma cell-derived differentially up-regulated ligands and 26 PBMC membrane receptors interacting with 73 intracellular proteins. The top 4 proteins with the highest degrees (hub nodes) in the PPI network were ITGB1, ITGB3, ITGA2B, CXCR4, with the degrees of 15, 12, 10, 10, respectively. All the four hub nodes were

Table 1. The Selected Ligand-Receptor Pairs

Up regulated ligands in breast tumor stroma	Membrane receptor expressed in PBMCs
CLEC3B	HGF
FGA	ITGA2B
KISS1	KISS1R
LBP	CD14
CLCF1	IL6ST
CXCL2	CXCR2
CX3CL1	CX3CR1
CXCL12	ACKR3, CXCR4
COL6A2	ITGB1, GP6, SDC4, CD44
COL6A6	ITGB1, GP6, SDC4, CD44
LAMA3	ITGA6, ITGB1, CD44
LAMA5	ITGA6, ITGB1, CD44
TNN	ITGB1, ITGB3, SDC4
THBS3	ITGA2B, ITGB1, ITGB3, SDC4, CD36, CD47
TNXB	ITGB1, ITGB3, SDC4
VWF	GP1BA, GP5, GP9, ITGA2B, ITGB3
CTSG	F2R, F2RL1, F2RL2, F2RL3
IGF2	IGF2R
BMP15	BMPR2

Table 2. Top 20 Significantly Enriched GO Biology Processes for Crosstalk Network Nodes

Term	Count	P-Value
GO:0009611~response to wounding	36	1.22E-24
GO:0001775~cell activation	28	6.59E-23
GO:0007155~cell adhesion	34	1.2E-18
GO:0022610~biological adhesion	34	1.26E-18
GO:0007166~cell surface receptor linked signal transduction	51	1.79E-18
GO:0050817~coagulation	16	6.18E-16
GO:0007596~blood coagulation	16	6.18E-16
GO:0007599~hemostasis	16	1.6E-15
GO:0045321~leukocyte activation	20	9.37E-15
GO:0042060~wound healing	18	3.46E-14
GO:0051270~regulation of cell motion	18	4.1E-14
GO:0050878~regulation of body fluid levels	16	8.92E-14
GO:0007243~protein kinase cascade	22	1.96E-13
GO:0006928~cell motion	24	3.34E-13
GO:0030334~regulation of cell migration	16	1.31E-12
GO:0006954~inflammatory response	20	1.9E-12
GO:0007242~intracellular signaling cascade	35	5.15E-12
GO:0031589~cell-substrate adhesion	13	5.76E-12
GO:0040012~regulation of locomotion	16	8.37E-12
GO:0002684~positive regulation of immune system process	17	1.57E-11

GO, Gene Ontology;PPI,protein-protein interaction; Counts: number of genes

Table 3. Top10 Enriched KEGG Pathways for Crosstalk Network Nodes

Term	Count	PValue
hsa04512:ECM-receptor interaction	23	5.6E-21
hsa04510:Focal adhesion	25	1.86E-14
hsa04062:Chemokine signaling pathway	21	3.22E-11
hsa04810:Regulation of actin cytoskeleton	18	0.000000119
hsa04640:Hematopoietic cell lineage	12	0.000000211
hsa04670:Leukocyte transendothelial migration	13	0.000000732
hsa04630:Jak-STAT signaling pathway	12	0
hsa05130:Pathogenic Escherichia coli infection	7	0
hsa05200:Pathways in cancer	16	0
hsa04060:Cytokine-cytokine receptor interaction	14	0

KEGG, Kyoto Encyclopedia of Genes and Genomes;PPI, protein-protein interaction; Counts: number of genes

receptors expressed in PBMCs of breast cancer patients.

GO and pathway enrichment analyses

We used DAVID to analyze all proteins in the interaction network. A total of 359 GO functional terms were enriched, the top 20 functional terms ranked by statistical significance were listed in Table 2. We found that they were mainly involved in wound healing, inflammation, cell adhesion and migration. Top 10 KEGG pathway are shown in Table 3. ECM-receptor interaction, focal adhesion, chemokine signaling pathway were included in the list.

Discussion

Over the past two decades, the TME have collectively risen in prominence, embodied in the view that, as a systemic disease, cancer progression may be directed by the body's systemic responses to malignancy and by the involvement of organ systems located at sites distant from the site of primary tumor growth (McAllister and Weinberg, 2014b). Up to now, most analyses of such systemic instigation by tumor have concentrated on either the tumor tissues or the tumor cells, few have examined the dialogue between the tumor stroma (the local tumor environment) and the systemic tumor environment. The primary aim of this study was to decipher the secretome of breast tumor stroma cells, and the regulation of circulating cell function by stroma cell-secreted proteins.

In the present study, using mRNA expression profile, comparison of the transcriptomes of breast tumor stroma and normal control tissue was made, candidate genes potentially involved in the stroma- circulating cell dialogue were identified. There were 19 ligands that had up-regulated in breast tumor stroma, for which 26 respective receptors expressed on the peripheral blood cells of breast cancer patients. In order to explore the potential downstream target proteins of the receptors in PBMCs, we predicted PPIs of the receptors using the STRING and UniProt databases. A total of 73 interactive proteins were identified. Finally, a crosstalk network was constructed by integrating ligands-receptors-interacting proteins relationships, and was visualized using Cytoscape.

Based on the GO term and KEGG pathway analyses, the present study obtained a comprehensive understanding of the crosstalk proteins (genes). The most significant biological processes and pathway indicated an abundance of molecules contributing to a proinflammatory environment, invovled response to wounding, leukocyte activation, adhesion and migration. In cancer, TME consists of different entities, such as infiltrating immune cells (IICs), cancer-associated fibroblastic cells (CAFs), and mesenchymal stem cells (MSC) (Hanahan and Coussens, 2012; Maenhout et al., 2014). The mobilization and subsequent trafficking of these cells to tumors is thought to be due to inflammatory signaling in a tumor resembling that of an unresolved wound (Spaeth et al., 2008). For instance, among the leukocytes that infiltrate the tumor microenvironment, myeloid-derived suppressor cells (MDSCs) can be recruited to the tumor site from the peripheral blood, by a number of chemokines produced by tumors (Ley et al., 2007; Ding et al., 2015; Zhang et al., 2015). Our result indicated that, in addition to tumor cell itself , stromal cells may also contribute significantly to recruitment, activation, programming of those cells.

The hub proteins of the crosstalk network with degree more than 10 were ITGB1, ITGB3, ITGA2B, CXCR4. The first three were integrins. The integrins are expressed in many leukocytes such as T-lymphocytes, monocytes, and granulocytes and play an integral role in leukocyte migration by binding to endothelial cells and stimulating extravasation (Alon and Feigelson, 2012; Kuehn et al., 2014). Activation of integrin $\alpha 4\beta 1$, leading to the extravasation of myeloid lineage cells from the circulation, and recruit to the tumor microenvironment (Schmid et al., 2013). CXCR4 is a receptor specific for CXCL12 (also known as stromal-derived factor 1α). Increasing evidence supports a critical role of the CXCL12/CXCR4 axis for cells trafficking and recruiting to tumor site (Domanska et al., 2013). These different cell types including Endothelial Progenitor Cells (de la Puente et al., 2013), myeloid bone marrow-derived cells (BMDCs) (Hiratsuka et al., 2011), neutrophil (Seubert et al., 2015), T-regulatory cells and M2-type macrophages (Gil et al., 2014), mast cell (Ellem et al., 2014), mesenchymal stem cell (Lourenco et al., 2015), B cell (Shetty et al., 2012).

Overall, we carried out an integrated bioinformatics analysis of crosstalk between tumor stroma and the peripheral circulation. Our results provide new insight into the communication between local tumor environment and the host systemic environment. We hypothesize that secretion of these proteins by the tumor stroma and binding to their receptors in the PBMCs may play a role in exerting systemic changes which lead to the mobilization

of circulating cells. Specific therapies targeting these long-range systemic lines of molecular communication may have clinical significance. However, because the results are based on microarray data with a small sample size, more experimental validations are warranted.

Acknowledgements

Funding statement. This work was supported by the National Natural Science Foundation of China (grant no.81201702). Health and Family Planning Commission of Sichuan Province of China (grant no.120491; no.130295). Science and technology support program of Sichuan Science and Technology Department (grant no.2016GZ0364).

References

Alon R, Feigelson SW (2012). Chemokine-triggered leukocyte arrest: force-regulated bi-directional integrin activation in quantal adhesive contacts. *Curr Opin Cell Biol*, **24**, 670-6.

Barrett T, Wilhite SE, Ledoux P, et al (2013). NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res*, **41**, 991-5.

Biometric Research Branch home page (2015c).

Breast Cancer Treatment (2015b). (PDQ®) - National Library of Medicine - PubMed Health.

Christopher MJ, Rao M, Liu F, et al (2011). Expression of the G-CSF receptor in monocytic cells is sufficient to mediate hematopoietic progenitor mobilization by G-CSF in mice. *J Exp Med*, **208**, 251-60.

de la Puente P, Muz B, Azab F, et al (2013). Cell trafficking of endothelial progenitor cells in tumor progression. Clin Cancer Res, **19**, 3360-8.

Ding Y, Shen J, Zhang G, et al (2015). CD40 controls CXCR5-induced recruitment of myeloid-derived suppressor cells to gastric cancer. *Oncotarget*, **6**, 38901-11.

Domanska UM, Kruizinga RC, Nagengast WB, et al (2013). A review on CXCR4/CXCL12 axis in oncology: no place to hide. *Eur J Cancer*, **49**, 219-30.

Ellem SJ, Taylor RA, Furic L, et al (2014). A pro-tumourigenic loop at the human prostate tumour interface orchestrated by oestrogen, CXCL12 and mast cell recruitment. *J Pathol*, **234**, 86-98.

Finak G, Bertos N, Pepin F, et al (2008a). Stromal gene expression predicts clinical outcome in breast cancer. *Nat Med*, **14**, 518-27.

Franceschini A, Szklarczyk D, Frankild S, et al (2013). STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res*, **41**, 808-15.

Gil M, Komorowski MP, Seshadri M, et al (2014). CXCL12/CXCR4 blockade by oncolytic virotherapy inhibits ovarian cancer growth by decreasing immunosuppression and targeting cancer-initiating cells. *J Immunol*, **193**, 5327-37.

Graeber TG, Eisenberg D (2001). Bioinformatic identification of potential autocrine signaling loops in cancers from gene expression profiles. *Nat Genet*, **29**, 295-300.

Hanahan D, Coussens LM (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell*, **21**, 309-22.

Hattori K, Heissig B, Tashiro K, et al (2001). Plasma elevation of stromal cell-derived factor-1 induces mobilization of mature and immature hematopoietic progenitor and stem cells. *Blood*, **97**, 3354-60.

Hiratsuka S, Duda DG, Huang Y, et al (2011). C-X-C receptor type 4 promotes metastasis by activating p38 mitogen-activated protein kinase in myeloid differentiation antigen (Gr-1)-positive cells. *Proc Natl Acad Sci U S A*, **108**, 302-307.

Huang DW, Sherman BT, Lempicki RA (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, **4**, 44-57.

Irizarry RA, Hobbs B, Collin F, et al (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, **4**, 249-64.

Joyce JA, Fearon DT (2015). T cell exclusion, immune privilege, and the tumor microenvironment. *Science*, **348**, 74-80.

Kuehn C, Vermette P, Fulop T (2014). Cross talk between the extracellular matrix and the immune system in the context of endocrine pancreatic islet transplantation. A review article. *Pathol Biol (Paris)*, **62**, 67-78.

Kwast AB, Voogd AC, Menke-Pluijmers MB, et al (2014). Prognostic factors for survival in metastatic breast cancer by hormone receptor status. *Breast Cancer Res Treat*, **145**, 503-11.

LaBreche HG, Nevins JR, Huang E (2011). Integrating factor analysis and a transgenic mouse model to reveal a peripheral blood predictor of breast tumors. *Bmc Med Genomics*, **4**, 61.

Ley K, Laudanna C, Cybulsky MI, et al (2007). Getting to the site of inflammation: the leukocyte adhesion cascade updated. *Nat Rev Immunol*, **7**, 678-89.

Lourenco S, Teixeira VH, Kalber T, et al (2015). Macrophage migration inhibitory factor-CXCR4 is the dominant chemotactic axis in human mesenchymal stem cell recruitment to tumors. *J Immunol*, **194**, 3463-74.

Maenhout SK, Thielemans K, Aerts JL (2014). Location, location, location: functional and phenotypic heterogeneity between tumor-infiltrating and non-infiltrating myeloid-derived suppressor cells. Oncoimmunol, **3**, 956579.

McAllister SS, Weinberg RA (2014a). The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. *Nat Cell Biol*, **16**, 717-27.

McAllister SS, Weinberg RA (2014b). The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. *Nat Cell Biol*, **16**, 717-27.

Orimo A, Gupta PB, Sgroi DC, et al (2005). Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell*, **121**, 335-48.

Quail DF, Joyce JA (2013). Microenvironmental regulation of tumor progression and metastasis. *Nat Med*, **19**, 1423-37.

Schmid MC, Franco I, Kang SW, et al (2013). PI3-kinase gamma promotes Rap1a-mediated activation of myeloid cell integrin alpha4beta1, leading to tumor inflammation and growth. *Plos One*, **8**, 60226.

Seubert B, Grunwald B, Kobuch J, et al (2015). Tissue inhibitor of metalloproteinases (TIMP)-1 creates a premetastatic niche in the liver through SDF-1/CXCR4-dependent neutrophil recruitment in mice. *Hepatol*, **61**, 238-48.

Shaked Y, McAllister S, Fainaru O, et al (2014). Tumor dormancy and the angiogenic switch: possible implications of bone marrow- derived cells. *Curr Pharm Des*, **20**, 4920-33.

Shannon P, Markiel A, Ozier O, et al (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, **13**, 2498-504.

Shetty S, Bruns T, Weston CJ, et al (2012). Recruitment mechanisms of primary and malignant B cells to the human liver. *Hepatol*, **56**, 1521-31.

Song N, Choi JY, Sung H, et al (2015). Prediction of breast cancer survival using clinical and genetic markers by tumor subtypes. *Plos One*, **10**, 122413.

Spaeth E, Klopp A, Dembinski J, et al (2008). Inflammation and

tumor microenvironments: defining the migratory itinerary of mesenchymal stem cells. *Gene Ther*, **15**, 730-8.
Tusher VG, Tibshirani R, Chu G (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A*, **98**, 5116-21.
UniProt (2015d). a hub for protein information. *Nucleic Acids Res*, **43**, 204-12.
WHO (2015a). Breast cancer: prevention and control.
Zhang QQ, Hu XW, Liu YL, et al (2015). CD11b deficiency suppresses intestinal tumor growth by reducing myeloid cell recruitment. *Sci Rep*, **5**, 15948.