Transcriptomics profiling study of breast cancer from Kingdom of Saudi Arabia revealed altered expression of Adiponectin and Fatty Acid Binding Protein4: Is lipid metabolism associated with breast cancer?

Adnan Merdad†, Sajjad Karim2,3†, Hans-Juergen Schulten2, Manikandan Jayapal2, Ashraf Dallol2, Abdelbaset Buhmeida2, Fatima Al-Thubaity1, Mamdoooh A Gari2,3, Adeel GA Chaudhary2,3, Adel M Abuzeinadah2,3,4, Mohammed H Al-Qahtani2,3*

From 2nd International Genomic Medicine Conference (IGMC 2013) Jeddah, Kingdom of Saudi Arabia. 24-27 November 2013

Abstract

Background: Breast cancer incidence rates are increasing at an alarming rate among Saudi Arabian females. Most molecular genetic discoveries on breast cancer and other cancers have arisen from studies examining European and American patients. However, possibility of specific changes in molecular signature among cancer patients of diverse ethnic groups remains largely unexplored. We performed transcriptomic profiling of surgically-resected breast tumors from 45 patients based in the Western region of Saudi Arabia using Affymetrix Gene 1.0 ST chip. Pathway and biological function-based clustering was apparent across the tissue samples.

Results: Pathway analysis revealed canonical pathways that had not been previously implicated in breast cancer. Biological network analysis of differentially regulated genes revealed that Fatty acid binding protein 4, adipocyte (FABP4), adiponectin (ADIPOQ), and retinol binding protein 4 (RBP4) were most down regulated genes, sharing strong connection with the other molecules of lipid metabolism pathway. The marked biological difference in the signatures uncovered between the USA and Saudi samples underpins the importance of this study. Connectivity Map identified compounds that could reverse an observed gene expression signature.

Conclusions: This study describes, to our knowledge, the first genome-wide profiling of breast cancer from Saudi ethnic females. We demonstrate the involvement of the lipid metabolism pathway in the pathogenesis of breast cancer from this region. This finding also highlights the need for strategies to curb the increasing rates of incidence of this disease by educating the public about life-style risk factors such as unhealthy diet and obesity.
Background

Breast cancer (BC) is the most common cancer that affects females worldwide and is the second most frequent cause of cancer-related deaths among women in the United States [1]. According to the National Cancer Registry in the Kingdom of Saudi Arabia (KSA), BC was ranked as the most prevalent form of cancer among females, accounting for 25.1% of all newly diagnosed female cancers (5,205) in the year 2009 [2]. The ASR was 22.7/100,000 for the female population. While the median age at presentation is around 63 years in the United States and Western Europe, the median age at presentation in the KSA is 48 years [2]. The genetic variability between patients and tumors drive this clinically heterogeneous disease [3]. Much of the knowledge on the molecular genetics on BC and other cancers has been resulting from examining European and US patients. However, growing scientific knowledge suggests the likelihood of variability in molecular signature between cancers from patients of different ethnic groups. Clear differences in the patterns of p53 mutations in BC were found between Midwest US Caucasian, African- American, Austrian, and Japanese women [4,5]. Besides, varying patterns and the occurrence of germ line mutations in BRCA1 and BRCA2 between ethnic groups were also reported [6]. Recently, National Cancer Institute’s Surveillance, Epidemiology, and End Results (SEER) program show that the age-adjusted BC frequency among minority groups are significantly lower than those among white women [7]. Increase in BC mortality was also found in African American women when compared to white women [8]. All these differences continue to be mostly unexplained [9]. Moreover, changes in occurrence and clinical characteristics linked with ethnicity or race has got only limited consideration [10]. The current lack of measures to aid in stratifying BC treatment indicates the necessity for a new methodology to formulate better prognostication and therapy prediction. Although the parameters such as histological grade, stage, and tumor size are accepted as prognostic markers for BC [11], about 50% of the patients with ER-positive cancer fail on tamoxifen treatment [12,13]. Transcriptomics approaches that allow screening of thousands of genes in an experiment, has had a major impact on BC research over the past 10 years [14]. Several groups have carried out gene expression profiling of BC and classified clinically distinct subclasses of tumors [14], hereditary BC [15], and treatment prediction [16]. Many microarray studies so far reported have led to the discovery of several genes associated with BC. However, most of the gene expression profiling studies has been carried out mainly in the Caucasian population, and inclusion of the non-Caucasian population has been minimal [17,18]. Significant risk factors of BC in the western countries such as nulliparity, low parity, first time pregnancy at late age, having no history of breast feeding, etc., are usually not practiced in the Saudi society, yet BC incidence is still high among women from the KSA [19]. In addition, a study of Taiwanese population reported that women who had more than three full-term pregnancies; first full-term pregnancy at the age of below 30; and three or more years of breast feeding showed significantly reduced risk of BC [20].

These studies suggest that there is an important need to better understand the molecular mechanisms underlying BC from different populations. In view of the existence of genetic differences across regional or geographical locations, more data specific to indigenous populations are needed for comparative analysis of molecular changes in breast tumors among Saudi population. More comprehensive molecular and genomic analyses of breast cancers are important to understand the complexity and severity of the disease and to find ‘druggable’ molecular targets for the development effective treatments. In order to obtain relevant target genes associated with BC patients in the KSA, we aim to start comprehensive microarray-based gene expression profiling study of BC in the Kingdom. Transcriptomic analysis coupled with functional and pathway analysis could lead to new insight into biomarkers and signatures associated with the disease. Deregulated signaling pathways are thought to drive functional processes such as cell growth, cellular proliferation, and invasion of cancer cells. Thus, identifying such underlying driving changes will be vital for studies of tumor progression, for the identification of novel therapeutic targets.

To this end, this study aimed to survey the genes that are differentially regulated in forty five freshly frozen BC tissue samples compared to eight normal controls. The ethnic group of the cohort in this study provides unique dimension to BC research. We examined genome-wide gene expression in order to elucidate the molecular mechanisms underlying malignancy in breast tissues. Based on the differential expression signature, we found biological functions and pathways that are significantly altered in BC. Pathway-based clustering was apparent across the tissue samples. Pathway analysis revealed canonical pathways that have not been implicated in BC, previously. We also demonstrated additional utility of our BC signature by comparative analysis between the USA and the KSA samples reveals marked biological differences in signature. Furthermore, Connectivity Map (cmap) analysis offered hypotheses regarding potential treatments for differential regulation observed in the BC [21]. Thus, our finding might open up new avenues of BC research and assist discover possible new targets for BC treatment where ethnicity plays major role.

Results

The main focus of this study was to determine the transcriptomic profiles of BC representing the Saudi society.
To identify genes that are involved in BC development, we analyzed the transcriptomic profiles of nearly 28,000 annotated genes. We profiled forty-five fresh breast tissue specimens and compared them with eight normal control samples. We found 48 years as average age of incidence for BC in Saudi Arabia. We also found that 73.3% of BC patients were either obese (n=25; 55.56%) or over weight (n=8; 17.78%) whereas only 26.7% were normal weight (n=10) or under weight (n=2). We performed PCA scatter plot for visualizing the high dimensional array data. In the scatter plot, each point represents a chip. We applied PCA for identifying outliers and major effects in the data. The results of PCA of the transcriptomic data showed that the samples from the same tissue type clustered tightly together. Clear differences also observed between these cancer and normal tissues revealing distinct expression profiles for the different tissue types (Figure 1). PCA mapping showed that 30% of the overall variance in the microarray dataset is depicted by the first three principal components.

Identification of differentially expressed genes
Comparison of the genome-wide expression of breast cancer revealed 1159 differentially expressed genes, 544 genes of which were upregulated and 615 genes were downregulated (2 fold change, false discovery rate p < 0.05).

As shown in Figure 2, unsupervised hierarchical clustering revealed that the distinct gene-patterns between the two tissue types. The other sample information including histopathology and age were also overlaid at the bottom of the dendrogram. Cluster analysis also showed that the cancer tissues agglomerated in various subsets according to age, triple negative status and grade. Those of patients with triple-negative breast cancer (TNBC), young age (mean 41) and grade 3 tumor were clearly separated from the old age (mean 50), grade 2 and non-TNBC. We also investigated whether pathway based clustering is evident in the gene signature. For this, genes from each signature were extracted and analyzed for significantly altered biological processes and pathways. Our results suggest that BC patients, despite their histopathologic and age heterogeneity, were clustered into different groups defined by functions and pathways.

Functional analysis of the breast cancer-associated genes found an over expression of genes involved in cell cycle progression, DNA repair, cell death, tumor morphology and tissue developments. Specifically, genes that are known to be associated with BC, including Topoisomerase 2 alpha (TOP2A), Carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1), denticless/RA-regulated nuclear matrix associated protein (DTL), H3 histone family, member A, histone 1, H3a (HIST1H3A), the targeting protein for Xklp2 (TPX2), Hepatitis C virus NS5A-transactivated protein (KIAA0101), centromere protein-F (CENPF) and ubiquitin-conjugating enzyme E2T (UBE2T) were upregulated compared to the controls. Furthermore, genes associated with immune cell trafficking and interferon signaling including chemokine (C-X-C motif) ligand 10 (CXCL10), matrix metallopeptidase 11 (MMP11), interferon, gamma-inducible protein 6 (IFI6), chemokine (C-X-C motif) ligand 9 (CXCL9) and matrix metallopeptidase 9 and 13 (MMP9 and MMP13) were also significantly upregulated. Interestingly, the biological process, cellular movement was significantly overrepresented in both down-regulated and up-regulated gene lists pointing that the metastasis was linked to a different equilibrium of switching on and off.

Genes linked to lipid metabolism, drug metabolism, and endocrine system development were significantly down regulated in BC. The most downregulated genes including Fatty acid binding protein 4, adipocyte (FABP4), adiponectin (ADIPOQ), and retinol binding protein 4 (RBP4) were known to be involved in lipid metabolism. The downregulated genes also signify biological functions including cardiovascular system development and function, small molecules biochemistry and energy production. Both the functional and canonical pathway analysis indicate that lipid metabolism is one of the most disrupted and downregulated pathways (Table 1; Figure 3).
Pathways and networks underlying breast cancer

To understand the mechanisms by which the genes alter a wide range of physiological processes, we examined molecular networks underlying BC. Transcriptomic signatures showed significant disruption in signaling pathways associating genes of the glycerolipid metabolism, ATM signaling, ILK signaling, DNA damage and cell cycle (Table 1). Analysis by IPA shows a set of key genes that disrupt a pathway such that it then results in tumor initiation or progression. The pathway analysis revealed a strong correlation between the transcriptomic signature and the canonical glycerolipid metabolism which has not been implicated in BC before. Majority of the genes involved in the glycerolipid metabolism were downregulated. These results further support the putative role of these pathways in rendering BC susceptibility.

We further compared the genes involved in lipid metabolism with recent data published from samples from the USA by downloading publically available data of 16 infiltrating ductal carcinoma (IDC) of the breast samples from ArrayExpress accession E-GEO-D-22544(45) (Figure 4). We chose IDC as it accounts for 80% of all breast cancers.
Table 1 Canonical pathways predicted by Ingenuity Pathway Analysis

| Pathways                                      | p-value (Fisher's Exact) | Ratio | Molecules                                                                 |
|-----------------------------------------------|--------------------------|-------|---------------------------------------------------------------------------|
| Glycerolipid Metabolism                       | 6.85E+00                 | 1.36E-01 | SDCl (includes EG:20969),DGAT2,ADH1C (includes EG:11522),LYPE,GBK,PNPL2,ADH1A,GPM,PPAPDC1A,PPAPDC1B,ADACALD1,ADI2,PPAP2A,ALDH1A3,AGPAT2,ADH1A2,PL,ALD1H1B,MGLL,PPAP2C |
| Mitotic Roles of Polo-Like Kinase              | 5.26E+00                 | 2.15E-01 | KIF23,CDC25,CDC20,PTTG1,PRC1 (includes EG:233406),CDC7 (includes EG:12545),CCNB2,PLK1,CDK1,CCNB1,PK4,ANAPC5,PPP2R1B,KIF11 |
| Cardiac β-adrenergic Signaling                | 4.95E+00                 | 1.36E-01 | AKAP12,PDE2A,PLN,PPP1R1A,ADRBK2,ATP2A3,PPP1R1A,PEA1,PAK2,PALM2-APKA2,PRKAR2B,GNG1,FDE7B,PDE3,B,FDE1B,MRAS,PDE8B,GNG2,PPP1CA,PPP2R1B,APEx1 |
| ATM Signaling                                 | 4.70E+00                 | 2.22E-01 | RADS1,CDC25,C,M2,FANCD2,MAPK10,CCNB2,CREB3L4,CREB5,CDK5,CHK1,CCNB1 |
| ILK Signaling                                 | 4.50E+00                 | 1.30E-01 | FN1,MYH11,CREB5,ROH0KRT1B8,ROHOU,FGF,AKT3 (includes EG:10000),ACTG2,MYC1,CLF1,FERMT2,PKCI2,FHOJ,CREB3L4,MYL9,ROHOU,RDN3,FLNC,FLC2,ARHGEF6,MAPK10,LEP1,PPP2R1B,MYMP9 |
| Protein Kinase A Signaling                    | 3.93E+00                 | 1.03E-01 | AKAP12,PKG2,LIPE,PDE1A,CREB5,HST2,HC (includes others),PTCH2,TFGBR2,GNG1,FDE7B,PDE3B,PPP1R1A,CAJ,HST1B18,APEx1,PDE2A,PLN,GNA11,PPP1R1A,PAK4,YLPL4,CREB3L4,TTIN (includes EG:22138),PDE1C,MYL9,AKAP2,PALM2-APKA2,PRKAR2B,PGY4,ADD3,FLNC,FLC2,ARHGEF6,GNG2,TCF7L2 |
| Leukocyte Extravasation Signaling             | 3.43E+00                 | 1.15E-01 | VAV2,RAC2,CLDN11,JAM2,PIK3C2G,GN1A1,MAPK13,CLDN7,CLD1,ITGAL,RHOH,TIMP4,CLDN4,CLDN8,ERZ,CXCL12 (includes EG:20315),MAPK10,PECAM1,MAP11,ACTG2,MYMP9,CLDN3 |
| Atherosclerosis Signaling                     | 3.11E+00                 | 1.22E-01 | CMA1,CD36,MAPK13,PLAG2A2,AKT3,CLL3,PLAG2A4,AKL,BCL11,BCL1A1,BCL1A3,BCL1A2,PGMS,ADH1B,ACSL1,ALDOC |
| Glycolysis/Gluconeogenesis                    | 2.98E+00                 | 9.85E-02 | PGK1,ADH1C (includes EG:11522),ALDH1I1,ALDH2,HI1,ADH1A,ALDH1A1,ALDH1A2,FLNC,FLC2,ARHGEF6,GNG2,TCF7L2 |
| Agrin Interactions at Neuromuscular Junction  | 2.93E+00                 | 1.59E-01 | RAC2,PK3A,ARHGEF6,LAMMA2,MRAS,MAPK10,ERBB3,ACTG2,ERBB2,TGAL,EGFR |
| IL-8 Signaling                                | 2.84E+00                 | 1.04E-01 | RAC2,ANPPT1,PK3C2G,GNA1I1,RHOU/JKA3,RHOJ,CHD11,RHOQ,GNG1,FND3,MAPK10,ROHOU,MRAS,AKT3 (includes EG:20313),MAPK10,PECAM1,MAP11,ACTG2,MYMP9,CLDN3 |
| Inhibition of Angiogenesis by TSP1             | 2.65E+00                 | 1.79E-01 | TGFBR2,SDC1 (includes EG:20969),MAPK10,CD36,AKT3 (includes EG:10000),MAPK13,MYMP9 |
| Interferon Signaling                          | 2.65E+00                 | 1.94E-01 | OAS11,MYJRF9,STAT1,TAPl,BAK1,JRF1 (includes EG:16362) |
| Cell Cycle: G2/M DNA Damage Checkpoint Control | 2.60E+00                 | 1.63E-01 | CDC25,COKS2,TOPO2A,CCNB2,PLK1,CDK1,CHK1,CCNB1 |
| Cell Cycle Control of Chromosomal Replication | 2.52E+00                 | 1.94E-01 | MCM3,CDC45,CDC6 (includes EG:23834),CDC7 (includes EG:12545),MCM4,MCM7 |
| Role of BRCA1 in DNA Damage Response          | 2.47E+00                 | 1.48E-01 | RADS1,FANCD2,E2F5,BRC2A,PLK1,RP1P1,STAT1,BLM,CHTK1 |
| Integrin Signaling                            | 2.47E+00                 | 1.00E-01 | RAC2,CAPI6,PRTN6,PIK3C2G,RHOU,JTGAL,TTN (includes EG:22138),ROHOU,MNYL9,ROHO,TLN2,PAK3,ROD3,RN3,TOPO2A,CAV1,MRAS,ROHOU,AKT3 (includes EG:10000),ACTG2,TGAL7,ITGAX |
| RhoGDI Signaling                              | 2.39E+00                 | 9.45E-02 | CFL1,GNA11,RHOU,CL1,ROHOU,MNYL9,CDH2,CHD1,RHOQ,GNG1,FND3,PAK3,ROD3,ERZ,FLC2,ARHGEF6,MRAS,ROHOU,ACTG2,MYMP9 |
| Production of Nitric Oxide and Reactive Oxygen Species in Macrophages | 2.34E+00                 | 9.05E-02 | PIK3C2G,PPP1R1A4,ROHOU,MAPK13,ROHOU,JRF1 (includes EG:16362),APOC1,ALB,ROHOU,ROD3,MAPK10,ROHOU,MAPK10,STAT1,KIP1CA,PPP2R1B,APOL,RPB4 |
| Hereditary Breast Cancer Signaling            | 2.29E+00                 | 1.09E-01 | CDC25C,HDAC1,PIK3C2G,CDK1,CHK1,CCNB1,RADS1,FANCD2,MRAS,AKT3 (includes EG:10000),POL2R2B,CREC2A,MYMP9,MYMP1B |
| Bile Acid Biosynthesis                        | 2.04E+00                 | 7.55E-02 | ALDH2,ADH1A1,AKR1C1A,AKR1C2A,ALDH1A1,ADH1A3,ALDH1A2,ADH1C (includes EG:11522),ADH1B1 |
| Estrogen-Dependent Breast Cancer Signaling    | 1.77E+00                 | 1.14E-01 | HSD17B13,GF1,MRAS,PK3K2G,ARXDB6,AKT3 (includes EG:10000),CREB3L4,CREB5,EGFR |
| HER-2 Signaling                               | 1.21E+00                 | 9.88E-02 | CCNE2,MRAS,PK3K2G,PARD6B,AKT3 (includes EG:10000),ERBB3,ERBB2,EGFR |
| Breast Cancer Regulation by Stathmin1         | 7.29E-01                 | 6.67E-02 | CCNE2,GNA11,PIK3C2G,PPP1R1A4,CDK1,GNG1,GNG2,PRKAR2A,ARHGEF6,E2F5,MRAS,MYL9,UBA1C,GNG2,PPP1CA,PPP2R1B |

The table shows the significantly overrepresented canonical pathways across the whole dataset of differentially expressed genes.
Therapeutic suggestions using the Connectivity Map

Transcriptomic approaches for studying BC may be useful as it can help to identify pathways that are significantly altered in the specific population or an individual. Moreover, it will allow us to determine specific, effective drugs from the existing database. Having established the feasibility of the computational approaches, we further tried to distinguish compounds that can potentially antagonize the effects of the molecular signature of BC identified in this cohort. Cmap assesses the role of small molecules and genetic effects on transcriptome and establishes links between gene expression signatures and drugs. Based on the connectivity score, we identified 13 selected compounds with average connectivity scores less than -0.7, indicating a high negative correlation (Table 2). Interestingly, five drugs including Fulvestrant, Methotrexate, Letrozole, Doxorubicin and Tamoxifen were FDA approved BC drugs. The compounds that possibly would reverse observed BC signature and hence potential candidates for further in vitro validation comprise Hydrocortisone, Gliclazide and Gliclazide. Hydrocortisone and Gliclazide are used for aiding in fat metabolism and as anti-diabetic, respectively.

Discussion

Breast cancer therapeutic strategies today are generally based on histopathological characterization, tumor size, grading, and axillary lymph node and receptor status [1]. Even so, patients when diagnosed with similar conditions and when treated with similar drug can go through extensive differences in the development and relapse of BC. Studies show that women, who had a full-term pregnancy at an age below 30; and who went through three full-term pregnancies, and three or more years of breast feeding, were significantly protected against BC [20]. However, BC incidence is high among women from the KSA. BC risk factors including nulliparity or low parity, first full term pregnancy at very late age, no breast feeding etc., are not common in the Saudi society [19]. These factors combined with the early onset of BC among women, in this ethnicity, prompted us to study the molecular mechanism underlying the malignancies.

In the present study, we identified transcriptomic signature in BC from the KSA that is associated with clinical and histological parameters. The ethnicity of this cohort renders the power to prospectively examine, the roles of lifestyle and genetic susceptibility in the onset of BC. The genome-wide expression analysis contains several overrepresented functional gene classes and has substantial overlaps with transcriptomic signatures of metastatic human BC. The commonalities found between different populations in terms of increased cell cycle regulation and DNA integrity checkpoints were observed. A similar transcriptomic signature for many of these genes has been reported for human BC before [22-26].

Genes associated with lipid metabolism and small molecule biochemistry was significantly downregulated in the BC tissues. The role of these genes on the tumorigenesis
of BC has been reported previously [27-29]. Disruption of lipid metabolism has been implicated in tumorigenesis in several studies, especially in cancer of breast [30,31]. The differential regulation of lipid metabolism between normal and cancer subjects might reveal transformation in the metabolism of the cancer patient in this ethnicity [31]. Decreased ADIPOQ levels interrupt cellular signaling networks that are linked to cell survival, angiogenesis, proliferation, and cell-cycle regulation [32]. Adipose tissues secrete ADIPOQ and LEP-ADIPOQ axis has been well implicated in breast cancer tumorigenesis [32]. Levels of ADIPOQ have been inversely correlated with obesity.

Table 2 Therapeutic suggestions using the Connectivity Map

| Pharmaceutical perturbagen | Connectivity score | Description |
|-----------------------------|-------------------|-------------|
| Hydrocortisone              | -0.97             | Alimentary tract and metabolism (Aid in fat, protein and carbohydrate metabolism) |
| Fulvestrant *               | -0.952            | Estrogen receptor antagonist |
| Gliclazide                  | -0.934            | Potent antiatherogenic effect in type 2 diabetes |
| Deoxycortone                | -0.917            | Corticosteroids for systemic use |
| Sirolimus                   | -0.909            | Antineoplastic and immunomodulating agents |
| Glimepiride                 | -0.905            | Sulfonylurea anti-diabetic drug |
| Repaglinide                 | -0.895            | Treatment of type II diabetes |
| Methotrexate *              | -0.89             | Treatment of a number of cancers including breast |
| Letrozole *                 | -0.882            | Non-steroidal aromatase inhibitor for the treatment of hormonally-responsive breast cancer after surgery |
| Glipizide                   | -0.878            | Anti-diabetic drug from the sulfonylurea class |
| Doxorubicin *               | -0.868            | Treatment of a number of cancers including breast |
| Gilbenclamide               | -0.867            | Antidiabetic drug in a class of medications known as sulfonylureas |
| Tamoxifen *                 | -0.759            | Antagonist of the estrogen receptor in breast tissue |

Results of cmap for compounds with expression signatures opposite those of loaded gene expression levels for breast cancer.

* Approved by the FDA for breast cancer
Interestingly, studies show that women with increased \textit{ADIPQ} concentrations possess 65% reduced risk for breast cancer [27]. In another study, it has been shown that patients with low \textit{ADIPQ} levels had a significantly increased likelihood of cancer recurrence.

Unsupervised clustering analysis also revealed the gene expression signatures to be associated with histologic grade, age and triple negative status. This signature could also be used for improving stratification of patients with BC in this population. Further pathway analysis of differentially regulated genes provides novel hypotheses underlying metastatic progression of BC. In addition, cmap analysis can formulate novel therapeutics for BC, starting from a cellular approach to investigating the effects of the compound determined by cmap. It may be possible to develop stratified therapy by combining both standard chemotherapy and pathway-specific therapies.

\textit{ADIPQ}, an adipocyte secreted protein, was found to one of the most downregulated gene in our analysis and it’s lower expression levels have been shown to be associated with obesity, insulin resistance, and type 2 diabetes mellitus [33]. It has also been shown to be reduced in premenopausal women with endometrial cancer, which is tightly linked with obesity and insulin resistance. A case control study found significantly inverse relationship of serum \textit{ADIPQ} with BC [34].

Comparison of BC data from USA with KSA further revealed distinct expression levels for \textit{LEP} and other genes involved in lipid metabolism [35]. In our analysis, glycerolipid metabolism was found to be a highly significantly altered canonical pathway. However, comparative analysis of these findings is different from the recent transcriptomic profiles of BC from the USA data, as glycerolipid metabolism was ranked 41\textsuperscript{st} in the USA data [35]. In the KSA samples, glycerolipid metabolism was one the most overrepresented pathways, whereas, PPARa / RXRa activation pathway is the most disrupted one in the USA data.

To comprehend population differences in the severity of BC, studies linking both genetic and environmental factors and examining cases from different populations are important. Researchers have been investigating issues surrounding lifestyle, circadian rhythm, obesity, and adverse exposures in an attempt to identify susceptibility factors in initiating BC. Since diet and obesity contribute to changes in circulating hormone levels, they both play a key role in the development of BC. Understanding the underlying genetics and complex interactions of lifestyle, diet, and other known risk factors will remain a key area of research.

The prevalence of obesity among Saudi patients adults was 55.56\% in comparison to 33.8\% in USA adults, however combined obese and overweight percentage were almost same with 68.0\% in western and 73 \% in Saudi BC patients [36]. Obesity is primarily characterized by excess fat storage, adipocyte mass, and increase in certain types of lipids. Normally healthy women have 14–28% of fat mass in the body, but it may increase up to 60–70% in morbidly obese individuals [37]. Earlier white adipose tissue function was assigned for purely an energy storage tissue, however, recent finding has uncovered the endocrine and metabolic properties that has led to several mechanisms implicated in how obesity drives cancer prevalence and cancer deaths. Many possible mechanisms have been proposed to explain the increased risk of breast cancer associated with obesity such as increased lipids and lipid signaling, inflammatory responses, insulin resistance, adipokines, altered immune responses, and oxidative stress. Study had shown that cancer cells can access lipids from neighboring adipocyte stores and can directly use these transferred lipids as an energy source, which in turn promotes tumor growth [38]. Thus, it is critical to understand the physiological impact of obesity on breast cancer development and progression. In recent years, obesity has been identified as a significant modifiable increased risk factor for breast cancers among postmenopausal women but is unrelated or inversely related to risk among premenopausal women not in pre-menopausal women [39-43]. A study examined obesity and mortality from cancer in 495,477 US women, and found that obese women had double the death rate (relative risk, 2.12) than women with normal BMI (25 or less) [44].

Fat tissue of obese women are most important source of estrogen after ovary, secreting higher level of estrogen potentially leading to more rapid growth of estrogen-responsive breast tumors [45]. Fat cells may also promote inflammation linked cancer through tumor growth regulators and tumor necrotic factors. Obesity often cause insulin resistance, which may promote the development of breast cancer [46,47]. Adipose tissue secretes TNF-α leading to inflammation promoting cancer. Fat cells produce hormones, called adipokines, that may stimulate or inhibit cell growth [48]. For example, adiponectin, a down regulated gene in present study, has been reported to be inhibitor for proliferation and negative regulator of angiogenesis [37,49,50]. Further a significant inverse correlation between serum adiponectin levels and poor-prognosis breast cancer were confirmed. Thus, low adiponectin in obese individuals could increase risk of developing tumors with aggressive angiogenesis [34,51].

**Conclusions**

This study described, to our knowledge, the first genome-wide profiling of BC from Saudi ethnic females. Our analysis reveals appropriate biological relevance and a number of molecular pathways that may serve as targets for novel therapeutics. Our finding of 73.3\% obese vs 26.7\% normal weight clearly suggest that obesity increases the risk of...
Further studies are needed to determine the relationship between lipid metabolism dysregulation and the mechanisms underlying BC. This study may open new avenues of experimental strategies for further examination in larger cohorts of all subtypes, correlating ethnicity to molecular signature and life-style.

### Methods

#### Patients and samples

The study was performed on female BC patients from the KSA, diagnosed with invasive ductal carcinoma. The samples were collected from one of the three participating Medical Centers in Jeddah, the KSA including King Abdulaziz University Hospital, King Faisal Specialist Hospital and Research Center, and Bakshs Hospital, during years 2008-2011. For gene expression analysis, fresh tumor tissue specimens were cut from fresh surgical resections adjacent to the sites on which final histological diagnosis was performed. Fresh normal breast specimens were derived from surgically resected normal breast tissues. Patients excluded from this study had at least one of the following exclusion criteria: histopathological diagnosis was not invasive ductal carcinoma; patient history, and medical files, or specimens were not found. This left samples from 45 tumors available for transcriptomic analysis. All collected tissue specimens were immediately placed in RNAlater (Invitrogen - Life Technologies, Grand Island, NY) or RPMI 1640 medium (GIBCO, USA). Clinicopathological features such as age, tumor grade, tumor size, hormone receptor status, lymph node involvement and pathology reports were retrieved from the patients’ records after obtaining all the relevant ethical approvals. The complete clinicopathological characteristics of the 45 BC patients are summarized in Table 3. The average age in the present cohort was 48 years, with a median of 47 years (range between 27-80 years). Body mass index (BMI), a measure of weight relative to height was used to estimate Normal weight (≥18 -24.9 kg/m²) overweight (≥25-29.9 kg/m²) and obesity (≥30 kg/m²).

#### Ethical approval

All patients included in the study provided written informed consent. The study was reviewed and approved by the Center of Excellence in Genomic Medicine Research (CEGMR) ethical committee (approval number 08-CEGMR-02-ETH).

#### RNA extraction and array processing

Total RNA was extracted from fresh breast tissue specimens with the Qiagen RNeasy Mini Kit (Qiagen, Hilden, Germany) including an on-column DNase treatment according to manufacturer’s recommendations. Quality of the purified RNA was verified on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Mean value of

| Table 3 Clinicopathological characteristics of the 45 breast cancer patients |
|-----------------------------|---------|------|
| Characteristics            | N       | %    |
| Sex                         |         |      |
| Female                      | 45      | 100  |
| Age                         |         |      |
| ≤ 50                        | 32      | 71   |
| >50                         | 13      | 29   |
| Age, mean (range) years     | 48 (27-80) |    |
| BMI (Kg/m²)                 |         |      |
| Obese (≥30 Kg/m²)           | 25      | 56   |
| Over weight (≥25-29.9 Kg/m²)| 8       | 18   |
| Normal weight (≥18-24.9 Kg/m²) | 10    | 22   |
| Under weight (≤18 Kg/m²)    | 2       | 4    |
| Site                        |         |      |
| Right                       | 25      | 56   |
| Left                        | 18      | 40   |
| NA                          | 2       | 4    |
| Histological type           |         |      |
| IDC                         | 34      | 76   |
| ILC                         | 5       | 11   |
| Mixed                       | 2       | 4    |
| Others                      | 2       | 4    |
| NA                          | 2       | 4    |
| Tumor size: mean (sd) (cm)  | 3.1 (±1.4) |    |
| Grade                       |         |      |
| 1                           | 4       | 9    |
| 2                           | 17      | 38   |
| 3                           | 17      | 38   |
| NA                          | 7       | 16   |
| pN                          |         |      |
| Negative                    | 19      | 42   |
| Positive                    | 23      | 51   |
| NA                          | 3       | 7    |
| LI                          |         |      |
| Negative                    | 17      | 38   |
| Positive                    | 19      | 42   |
| NA                          | 9       | 20   |
| ER (IHC)                    |         |      |
| Negative                    | 19      | 42   |
| Positive                    | 19      | 42   |
| NA                          | 7       | 16   |
| PR (IHC)                    |         |      |
| Negative                    | 27      | 60   |
| Positive                    | 11      | 24   |
| NA                          | 7       | 16   |
| HER2 (IHC)                  |         |      |
| Negative                    | 24      | 53   |
| Positive                    | 14      | 31   |
Table 3 Clinicopathological characteristics of the 45 breast cancer patients (Continued)

| TNRS | NA | Yes | No |
|------|----|-----|----|
| 7    | 16 | 11  | 24 |

N: number of cases; BMI: Body Mass Index; NA: information not available; IDC: invasive ductal carcinoma; ILC: invasive lobular carcinoma; pNI: pathological lymph node involvement; LI: lymphatic invasion; ER: estrogen receptor; IHC: immunohistochemistry; PR: progesterone receptor; TNRS: Triple-negative receptor status

RNA integrity number (RIN) for all 50 processed samples was 8.0. RNA concentrations were determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE). 300 ng of each RNA sample was processed according to the manufacturer’s recommendation (Affymetrix, Santa Clara, CA, USA). After fragmentation and labeling, the samples were hybridized at 45°C for 17 hours to Human Gene 1.0 ST GeneChip arrays (Affymetrix, Santa Clara, CA, USA). These arrays are conceptually based on the Human Genome sequence assembly UCSC hg18, NCBI Build 36 and interrogate with a set of 764,885 probes 28,869 annotated genes.

Gene Expression Analysis
Affymetrix .CEL files and were imported to Partek Genomics Suite version 6.5 (Partek Inc., MO, USA). The data was normalized using RMA normalization. Principal component analysis (PCA) was performed on all probes to visualize high dimensional data. PCA was used to assess quality control as well as overall variance in gene expression between the disease states. Analysis of Variance (ANOVA) was applied on the complete data set and differentially expressed gene list was then generated using an FDR (Benjamini Hochberg) of 0.05 with 2 fold change cut off. Unsupervised two dimensional average linkage hierarchical clustering was performed using Spearman’s correlation as a similarity matrix. The microarray data generated in this study are in compliance with MIAME (http://www.mged.org/Workgroups/MIAME/miame.html) guidelines.

Functional and Pathway analysis
To define biological networks, interaction and functional analysis among the differentially regulated genes in BC, pathway analyses were performed using Ingenuity Pathways Analysis software (IPA) (Ingenuity Systems, Redwood City, CA). Statistically differentially expressed dataset containing 1159 genes and their corresponding probesets ID, Gene symbol, Entrez gene ID as clone identifier, p-value and fold change values were uploaded into the IPA. The functional/pathway analysis of IPA identifies the biological functions and/or diseases and pathways that are most significantly altered for the differentially expressed gene set. The significance of the connection between the expression data and the canonical pathway were calculated by ratio and/or Fisher’s exact.

Connectivity Map analysis
The differentially expressed genes identified were analyzed using the cmap in an effort to link genes associated with a phenotype with potential drug molecules. Cmap analyzes the correlation between our gene expression signature and predefined signatures of therapeutic compounds. A “connectivity score” (+1 to -1), signifying relative similarity to our differentially expressed genes, was generated using a metric based on the Kolmogorov-Smirnov statistic, as described [21]. A compound with negative connectivity scores represents compounds that potentially would reverse the molecular phenotype.

Availability of supporting data
The clinicopathological information and datasets (.CEL file) supporting the results of this article were submitted to NCBI’s Gene Expression Omnibus (GEO) under accession number GSE36295. (http://www.ncbi.nlm.nih.gov/geo/).

Conflict of interest
The authors declare that they have no potential conflicts of interest.

Authors’ contributions
AM, SK, AD, FA and MHA participated in the study design, interpretation of the results and finalization of the manuscript. HJS conducted microarray and AB conducted pathological studies. SK and JM analyzed data and drafted the manuscript. MG, AC, and AA participated in data collection, editorial support and critical review of manuscript. All authors read and approved the final manuscript.

Acknowledgements
This work was supported by the research grants from the Center of Excellence in Genomic Medicine Research (CEGMR-N08-14), Deanship of Research (HCl-1434-117-2), King Abdulaziz University, Jeddah, and King Abdulaziz City for Science and Technology, (KACST- Strategic Project ID: 10-BIO1073-03 and 10-BIO1258-03) Riyadh, Saudi Arabia. The authors would like to thank Ms. Manal Shaibad, Mrs. Manar Ata, Mrs. Nuha Alanisari, Ms. Alan AlOqami and Ms. Maha Al-Qaizi for performing RNA extraction, bioanalyzer assays and microarray experiments. We thank the patients, and the contributions of physicians, nurses, and pathologists of the King Abdulaziz University Hospital, King Faisal Specialist Hospital and Research Center, and Dr. Baksh Hospital, Jeddah, the Kingdom of Saudi Arabia.

Declarations
Publication charges for this article have been funded by Center of Excellence in Genomic Medicine Research, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia
This article has been published as part of BMC Genomics Volume 16 Supplement 1, 2015: Selected articles from the 2nd International Genomic Medical Conference (IGMC 2013): Genomics. The full contents of the supplement are available online at http://www.biomedcentral.com/ bmcgenomics/supplements/16/s1

Authors’ details
1Department of Surgery, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia. 2Center of Excellence in Genomic Medicine Research, King Abdulaziz University, Jeddah, Saudi Arabia.
References

1. American Cancer Society: Cancer Facts and Figures. 2012.
2. Cancer Incidence Report Saudi Arabia-2013. 2013 (http://www.scr.org.sa).
3. Morris SR, Carey LA. Molecular profiling in breast cancer. Reviews in endocrine & metabolic disorders 2007; 8(3):185-198.
4. Hartmann A, Blaszyk H, Saitoh S, Tsuchioka K, Tamura Y, Cunningham JM, McGovern RM, Schroeder JJ, Sommer JS, Kovach JS: High frequency of p53 gene mutations in primary breast cancers in Japanese women, a low-incidence population. British journal of cancer 1996, 73(8):896-901.
5. Hartmann A, Rosanelli G, Blaszyk H, Cunningham JM, McGovern RM, Schroeder JJ, Schaid DJ, Kovach JS, Sommer SS: Novel pattern of p53 mutation in breast cancers from Austrian women. The Journal of clinical investigation 1995, 95(2):686-689.
6. Liede A, Narod SA. Hereditary breast and ovarian cancer in Asia: genetic epidemiology of BRCA1 and BRCA2. Human mutation 2002, 20(6):413-424.
7. Siegel R, Naishadham D, Jemal A: Cancer statistics, 2013. CA a cancer journal for clinicians 2013, 63(1):11-30.
8. Newman LA, Mason J, Cote D, Vin Y, Carolin K, Bouwman D, Colditz GA: Cancer Incidence Report Saudi Arabia-2009.
9. Weir HK, Thun MJ, Hankey BF, Ries LA, Howe HL, Wingo PA, Jemal A: American Cancer Society: Cancer Facts and Figures. 2003, 290(14):1838-1842.
10. Vastag B: Staple diet and breast cancer. The New England Journal of Medicine 2003, 349(14):1838-1842.
11. Winer E, Monow M, Osborne C, Harris J: Malignant tumors of the breast. Lippincott Williams & Wilkins, 2001.
12. Clarke R, Liu MC, Bouker KB, Gu Z, Lee FY, Zhu Y, Asa SC, Gomez B, Ward E, Anderson RN, Edwards BK: Annual report to the nation on the status of cancer, 1975-2000, featuring the uses of surveillance data for cancer prevention and control. Journal of the National Cancer Institute 2003, 95(17):1276-1299.
13. Vartag B: Breast cancer racial gap examined: no easy answers to explain disparities in survival. JAMA - the journal of the American Medical Association 2003, 290(14):1838-1842.
14. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Osborne CK, Schiff R: Growth factor receptor cross-talk with estrogen receptor as a mechanism for tamoxifen resistance in breast cancer. Proceedings of the National Academy of Sciences of the United States of America 2003, 100(14):8418-8423.
15. Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, Rabinowitz D, Zhou X, Hawkins C, Meltzer PS, MacCrimmon K, Moore T, Szczylik C, Petsalakis I, cabinetoff20001, D'Amore D, Edsberg LE, Golda A, Greenberg AS, Hastie T, He Pierre, II, Hanley J, Home J, Kallioniemi A, Kallioniemi OP, Chen Y, Kohn KE, Knapp unfinished projects, 2005-2012 L, Kounnas MZ, Larson JD, Lashkari D, Lipson SE, Lister RA, Long T, Lu X, Miller CA, Mittl MT, Nasdala L, Orr C, Owada-Kato I, Perou CM, Quackenbush J, Richards WG, Rubin MA, Shah SP, Zhang L, Scherf U, Stransky N, Gaasbeek R, Kao W, Kim SK, Kononen J, Kubista J, Lemen RA, Makarov V, Nkondjoh J, Qian S, Rimm DL, Schutte AJ, Schumacher R, Serrano M, Yakhnenko O, 2011, 52(13):4492-4500. The Journal of the National Cancer Institute 2011, 103(13):1023-1037.
16. Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, Meltzer P, Gusterson B, Esteller M, Kallioniemi OP, 2002, 95(2):1276-1299.
17. Boneberg EM, Legler DF, Hoefer MM, Ohlschlegel C, Steininger H, Fuzesi L, Boneberg E, Legler DF, Hoigmaier A, Guenther R, 2005, 20(6):413-424.
18. Mantzoros C, Petridou E, Dessypris N, Chavelas C, Dalamaga M, Alexe DM, 2002, 2844-2854.
19. Pietenpol JA: Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. Cancer letters 2006, 237(1):109-114.
20. Knoop AS, Knudsen H, Ballev E, Rasmussen BB, Overgaard J, Nielsen KV, Schouaunou A, Gunnansdottir K, Olsen KE, Mouridsen H, et al: retrospective analysis of topoisomerase ii a amplifications and deletions as predictive markers in primary breast cancer patients randomly assigned to cyclophosphamide, methotrexate, and fluorouracil or cyclophosphamide, epirubicin, and fluorouracil: Danish Breast Cancer Cooperative Group. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2005, 23(30):7483-7490.
21. Ueki T, Nishidate T, Park JH, Lin ML, Shimao A, Hirata K, Nakamura Y, Katagiri T: Involvement of elevated expression of multiple cell-cycle regulator, DTL/RAMAP (dentinless/RA-regulated nuclear matrix associated protein), in the growth of breast cancer cells. Oncogene 2008, 27(43):5672-5683.
22. Beiche I, Vacher S, Lallemant F, Tozu-Kara S, Bennani H, Beuzelin M, Drouchkou K, Roulleau E, Lefebre M, Faucon F, et al: Expression analysis of mitotic spindle checkpoint genes in breast carcinoma: role of NCDC2/HEC1 in early breast tumorigenicity, and a two-gene signature for aneuploidy. Molecular cancer 2011, 10:23.
23. Kais Z, Barsky SH, Mathyssen S, Zha H, Avers D, Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, Meltzer P, 2002, 2844-2854.
24. Boneberg EM, Legler DF, Hoefer MM, Ohlschlegel C, Steininger H, Fuzesi L, Boneberg E, Legler DF, Hoigmaier A, Guenther R, 2005, 20(6):413-424.
25. Mantzoros C, Petridou E, Dessypris N, Chavelas C, Dalamaga M, Alexe DM, 2002, 2844-2854.
26. Boneberg EM, Legler DF, Hoefer MM, Ohlschlegel C, Steininger H, Fuzesi L, Boneberg E, Legler DF, Hoigmaier A, Guenther R, 2005, 20(6):413-424.
27. Boneberg EM, Legler DF, Hoefer MM, Ohlschlegel C, Steininger H, Fuzesi L, Boneberg E, Legler DF, Hoigmaier A, Guenther R, 2005, 20(6):413-424.
28. Boneberg EM, Legler DF, Hoefer MM, Ohlschlegel C, Steininger H, Fuzesi L, Boneberg E, Legler DF, Hoigmaier A, Guenther R, 2005, 20(6):413-424.
29. Boneberg EM, Legler DF, Hoefer MM, Ohlschlegel C, Steininger H, Fuzesi L, Boneberg E, Legler DF, Hoigmaier A, Guenther R, 2005, 20(6):413-424.
Molecular links between obesity and breast cancer. Endocrine-related cancer 2006, 13(2):279-292.

Nieman KM, Kenny HA, Pencicka CV, Ladanyi A, Buell-Gutbrod R, Zillhardt MR, Romero IL, Carey MS, Mills GB, Hotamisligil GS, et al. Adipocytes promote ovarian cancer metastasis and provide energy for rapid tumor growth. Nature medicine 2011, 17(11):1498-1503.

Nichols HB, Trentham-Dietz A, Egan KM, Titus-Ernstoff L, Holmes MD, Bresch AJ, Holick CN, Hampton JM, Stampfer MJ, Willett WC, et al. Body mass index before and after breast cancer diagnosis: associations with all-cause, breast cancer, and cardiovascular disease mortality. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 2009, 18(5):1403-1409.

Bianchini F, Kaaks R, Vainio H. Overweight, obesity, and cancer risk. The lancet oncology 2002, 3(9):565-574.

doi:10.1186/1471-2164-16-S1-S11

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit