Transcriptomics and metabolomics of breast cancer to improve the understanding of disease mechanisms and treatment response

Eldrid Borgan

Department of Genetics
Institute for Cancer Research
Divison of Cancer Medicine, Surgery and Transplantation
Oslo University Hospital Radiumhospitalet

UiO : University of Oslo
Oslo University Hospital Radiumhospitalet
NTNU – Trondheim Norwegian University of Science and Technology
The Research Council of Norway

© Eldrid Borgan, 2012

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo
No. 1336

ISBN 978-82-8264-271-2

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Unipub.
The thesis is produced by Unipub merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

Table of contents

Acknowledgements **5**

Aims **6**

List of papers **7**

Introduction **9**

Breast cancer characteristics *10*
Anatomy of the normal breast 10
Pathology 11
Epidemiology 11
Diagnosis and treatment 13

Breast cancer - A disease of somatic aberrations *14*
Genomic aberrations in breast cancer 16
Transcriptomics of breast cancer 16
Intrinsic molecular classification of breast cancer 17
microRNAs - role in gene regulation and cancer 19

Importance of the microenvironment *20*
Angiogenesis and VEGF signaling 21

Rewiring of metabolism in tumors *22*
The Warburg effect 24
Lipids and cholines in breast cancer 25

Targeted treatment of breast cancer *26*
Targeting angiogenesis with bevacizumab 28

Material and Methods **30**

Material *30*
Clinical breast carcinomas 30
MAS98.06 and MAS98.12 breast cancer xenograft models 30

Technology *31*
Microarray technology 31
High Resolution Magic Angle Spinning Magnetic Resonance Spectroscopy (HR MAS MRS) 33
Statistical and bioinformatic analysis 36
Analyzing differences between groups of samples 36
Dimension reduction and clustering techniques 37
Correlation analysis 37
Prior knowledge based approaches 38

Results in brief **39**

Paper I – Consequences of ischemia time on microRNA and gene expression in cancer tissue 39
Paper II - Merging transcriptomics and metabolomics – advances in breast cancer research 40
Paper III - Distinct choline metabolic profiles are associated with differences in gene expression for basal-like and luminal-like breast cancer xenograft models 41
Paper IV – Metabolic and transcriptional response to bevacizumab and doxorubicin in xenograft models of basal-like and luminal-like xenograft models 42

Discussion **43**

Methodological considerations *43*
Measuring transcript levels using microarray technology 43
Using HR MAS MRS as a metabolomic tool 44
Combining different types of transcriptomic and metabolomic data 45

Biological considerations *47*
Measuring snapshots of heterogeneous breast cancer tissue 47
Using xenografts as a model for human breast cancer 49
The future of bevacizumab as an anti breast cancer drug 50
Clinical potential of transcriptomic and metabolomic biomarkers 51

References **59**

Acknowledgements

Firstly, I want to acknowledge the Norwegian Reseach Council for funding this project, Oslo University for admitting me to the PhD program and all the patients who contributed tumor material for research, which were used in this thesis. Also, thank you to all co-authors and collaborators for your contribution to this work.

Special thanks goes to my main supervisor, Therese. From the first day we met, you have assumed that I have something to contribute, and have always shown interest in what I have to say. You have taught me many things and given valuable advise but have always left it up to me to decide whether I want to follow it or not. I admire the integrity and thoroughness you exhibit in everything you do. And I am grateful for all the extra hours you put in the last weeks before delivering this thesis!

I also want to thank my three (!) co-supervisors Anne-Lise, Ingrid and Ole Christian. My boss and co-supervisor Anne-Lise, thank you for persuading me to continue with a PhD after finishing my master thesis with you and Ingrid, and for sharing your expertise, your network and your enthusiasm. My Trondheim-supervisor Ingrid, thank you for introducing me to the world of metabolomics and always remembering me (even when I stayed in Oslo for too long). You have been very supportive, and have pushed me to have confidence in my own work. Ole Christian, my biostatistician-supervisor, thank you for treating me as a collaborator instead of a student and for being interested in my data even when there were no results (that could even be MORE exciting). I have learned tons from working with you because you always want to share your knowledge.

To all my colleagues at the department of genetics, thank you for providing a social and fun working environment. Randi and Xi, I could not wish for better office mates. Thank you for all scientific and non-scientific discussions and moral support. Hilde, you were the perfect lab tutor for me. Thank you for being always calm and indulgent, and helping me with all my microarray trouble shooting. Simen, thank you for all discussions about science, the meaning of life, and random very important things. You are a great friend and I greatly appreciate all the moral support you have given me, and for clearing your calendar to help me the last day before delivering this thesis! To Jørgen, Jo Anders and Miriam, thank you for all support, coffee breaks, beers and good times.

To all colleagues at my second work place at the MR Center at NTNU/St. Olav, thank you for always providing me with an office space and a bed (!) when I come to Trondheim, and for being great collaborators. I particularly want to thank Beathe for teaching me what I know about HR MAS MRS, and Siver for all the interesting discussions about big or small details.

To my parents, thank you for all your support and for bearing with me for being so self-absorbed. I also want to thank my friends for having the patience to listen to my (occasional?) whining, encouraging me to keep going and for understanding that I have not been able to spend as much time with them as I would have liked the last years. Most important, thank you Kjetil; for giving me the perspective that a PhD is not the most important thing in life, for your unconditional support, and for sharing your insane sense of humor with me, keeping me sane through some hard times.

Oslo, September 30th 2011

Eldrid Borgan

Aims

The overall aim of this study was to improve our understanding of breast cancer disease mechanisms by combining transcriptomic and metabolomic measurements.

Secondary aims were:

- To study transcriptional changes in mRNA and microRNA (miRNA) levels in response to ischemia time after surgery.
- To assess whether high resolution magic angle spinning magnetic resonance spectroscopy (HR MAS MRS) and gene expression microarrays could be performed on the same breast cancer tissue samples.
- To explore different strategies for combining HR MAS MRS and gene expression data.
- To explore differences within and between intrinsic subtypes of breast cancer based on transcriptomic and metabolomic data.
- To study transcriptomic and metabolomic response to treatment.

List of papers

Paper I:

Borgan E, Navon R, Moen Vollan HK, Schlichting E, Sauer T, Yakhini Z, Lingjærde OC, Sørlie T, Børresen-Dale AL: Ischemia caused by time to freezing induces systematic microRNA and mRNA responses in cancer tissue.

Molecular Oncology 2011, doi:10.1016/j.molonc.2011.08.004

Paper II:

Borgan E, Sitter B, Lingjærde OC, Johnsen H, Lundgren S, Bathen TF, Sørlie T, Børresen-Dale AL, Gribbestad IS: Merging transcriptomics and metabolomics- advances in breast cancer profiling.

BMC Cancer 2010, 10: 628

Paper III:

Moestue SA, Borgan E, Huuse E, Lindholm E, Sitter B, Børresen-Dale AL, Engebråten O, Mælandsmo G, Gribbestad IS: Distinct choline metabolic profiles are associated with differences in gene expression for basal-like and luminal-like breast cancer xenograft models.

BMC Cancer 2010, 10: 433

Paper IV:

Borgan E, Lindholm EM, Moestue SA, Mælandsmo GM, Lingjærde OC, Gribbestad IS, Børresen-Dale AL, Engebraaten O, Sørlie T: Metabolomic and transcriptomic response to bevacizumab and doxorubicin in basal-like and luminal-like breast cancer xenograft models.

Manuscript

Introduction

Breast cancer is a heterogeneous disease often regarded as a collection of diseases originating in the same tissue. Different types of genetic and epigenetic aberrations in somatic cells may lead to different types of cancer and are reflected at all molecular levels. Studying molecular profiles of all these levels are referred to as –omics. Figure 1 shows how four different molecular levels correspond to different –omics fields, genomics, transcriptomics, proteomics and metabolomics, based on the central dogma of molecular biology. In reality, the central dogma has proven too simple, as only a small fraction of the transcribed RNA in human cells is protein coding. The various roles of non-coding RNAs (ncRNAs) are just beginning to be unraveled, and this is bound to become an increasingly important aspect of transcriptomics.

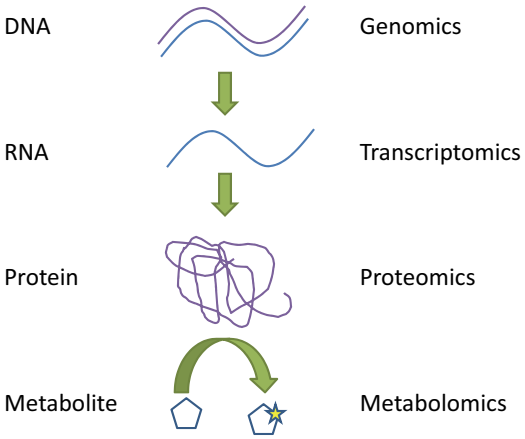


Figure 1. A simplified overview of the relationship between different molecular levels and the various –omics fields.

The importance of the microenvironment surrounding the tumor has also become increasingly evident to the research community. Tumor-stroma interactions are reflected at many levels, and especially at the metabolomic level, which is highly sensitive to properties of the microenvironment such pH and oxygen levels, and many other stimuli such as diet and drugs. Transcriptomic and metabolomic measurements provide different

information about tumors, and are not expected to be redundant. However, both similarities and differences between these two levels are interesting in order to bridge the gap between these two fields and improve classification of breast cancer into biologically relevant subgroups. Merging transcriptomic and metabolomic studies of breast cancer may increase our understanding of breast cancer disease mechanisms and treatment response.

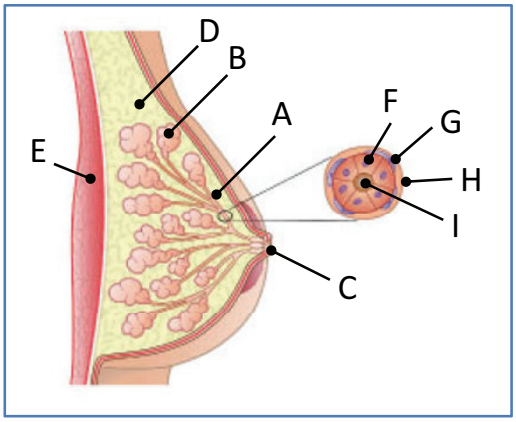
In the following sections, relevant background of the disease of breast cancer and key concepts of breast cancer biology will be covered. Central aspects of transcriptomics and metabolomics will also be covered from a biological and clinical perspective.

Breast cancer characteristics

Approximately one out of twelve Norwegian women will develop breast cancer before the age of 75 [1]. To place this thesis in context to the disease of breast cancer, some general characteristics will be briefly covered, including anatomy of normal breast and breast cancer pathology, epidemiology, diagnosis and treatment.

Anatomy of the normal breast

The anatomy of the post puberty female breast is depicted in Figure 2. The breast consists of ducts, lobules and surrounding adipose tissue. Ducts lead from the lobules, which produce milk during lactation, to the nipples. The ducts and lobules are composed of two layers of epithelial cells (luminal cells and basal/myoepithelial cells) surrounded by a basement membrane.



Breast profile:
A ducts
B lobules
C nipple
D adipose tissue
E pectoralis major muscle

Enlargement:
F epithelial cells
G myoepithelial cells
H basement membrane
I lumen (center of duct)

Figure 2. Anatomy of the breast. (Modified from Wozniak and Keely [2].)

Pathology

Premalignant (hyperplasia and carcinomas in situ) and malignant lesions (invasive carcinomas) in the breast originate from epithelial cells lining the ducts and lobules. For invasive carcinomas, neoplastic cells have invaded through the basement membrane. The majority of invasive breast cancer cases in Norway are invasive ductal carcinoma (IDC) (70-80 %) and the second largest group is invasive lobular carcinoma (ILC) (10-20 %) (http://oncolex.no). In this thesis, all studies have focused on IDC. Invasive tumors may metastasize to local and regional lymph nodes such as of the armpits (axilla), and distant metastasis most commonly occur in bone, lungs or liver (http://oncolex.no).

Epidemiology

In 2008, there were 1.38 million new female breast cancer cases and 458,400 breast cancer deaths worldwide [3]. Breast cancer is the most common cancer type in females both on a world basis and in Norway, comprising one-quarter of all female cancer cases [1,3]. There were 2745 new cases of and 671 deaths from female breast cancer in Norway 2009 [1]. Second to lung cancer, breast cancer has the highest mortality rate in females in Norway. Males can also be affected; in 2009, 15 cases of male breast cancer and 7 male breast cancer deaths were registered in Norway. A steady increase in both incidence and 5-year survival has been observed for female breast cancer since 1965 (Figure 3). The

steeper increase in incidence from 1995-2005 is likely due to the introduction of the national screening program during this period. Although incidence of breast cancer has increased up to around 2005, mortality rates have remained fairly constant since 1965. The last 10-15 years a decrease in mortality has been observed, and a decrease in incidence has also been observed since around 2005. This decrease in mortality and incidence is likely due to both improved treatment and earlier diagnosis from screening.

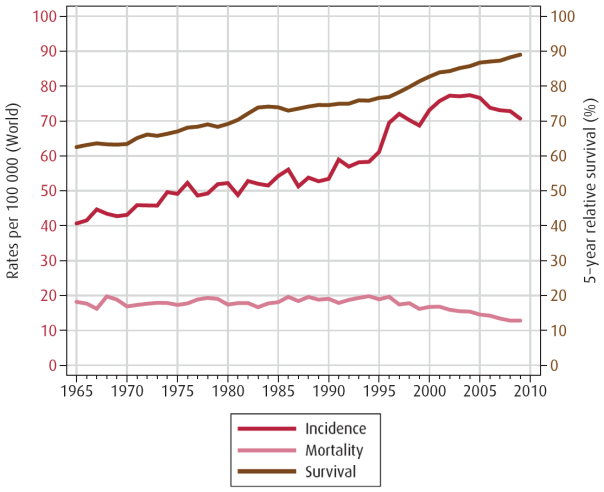


Figure 3 Trends in incidence, mortality rates and 5-year relative survival proportions for females with breast cancer. (From Cancer Registry of Norway [1].)

The incidence of breast cancer is higher in developed countries, which may result from a combination of environmental factors and the introduction of screening programs [1]. Known risk factors of breast cancer are age, hormonal exposure before age 35, hormone replacement therapy (HRT) for post-menopausal women, alcohol consumption, early menstruation, and breast cancer in nearest family. Preventive factors are giving birth at early age, high parity before age 25 and physical activity (http://oncolex.no).

Diagnosis and treatment

Breast cancer is diagnosed by combining clinical and pathological evaluations. After initial detection of breast cancer by palpation or routine mammography, diagnosis is aided using clinical examination, imaging techniques and needle biopsy. Imaging can be performed using ultrasound, positron emission tomography (PET) or magnetic resonance imaging (MRI). The needle biopsy can be performed using either fine needle aspiration or core needle biopsy. Resectable breast cancers are removed by surgery, and the surgical procedure is performed by breast-conserving surgery (lumpectomy) or removal of the entire breast (mastectomy). To assess whether the cancer has metastasized to the lymph nodes in the axilla, the sentinel node is removed and evaluated by a pathologist. The sentinel node is the first lymph node or group of nodes in the lymphatic drainage system from the breast to which breast cancer cells most often spread. If the sentinel node contains tumor cells, axillary clearance is performed.

After surgery, several methods are combined in order to guide adjuvant treatment, including clinical stage, histopathological grade, and status of histopathological markers. Staging is performed according to the TNM classification system (http://www.nbcg.no/nbcg.blaaboka). T is primarily the classification of tumor size, where invasive carcinomas of different sizes are classified as T1-T3. Additional T categories include carcinoma in situ (Tis) and invasive tumors with extension to the chest wall or skin (T4) (http://oncolex.no). N is the classification into different categories of lymph node involvement and M is the classification of no metastasis (M0) or metastasis to distant organs (M1). While TNM staging assigns cancers into stages I-IV, tumors are also divided into different grades based on histopathologic assessment of the tumor cells. IDCs are divided into grade I-III based on a combination of mitotic index, nuclear atypia and tubular formation, where high grade represents tumors that are less differentiated and more aggressive [4]. In Norway 2011, immunohistochemical staining of a small panel of markers is used as prognostic and predictive indicators, namely estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) protein or its coding gene *ERBB2* (by immunohistochemistry or fluorescence *in situ* hybridization (FISH), respectively), and the proliferation marker Ki67

(http://www.nbcg.no/nbcg.blaaboka). In addition to stage, grade and status of these predictive indicators, age also has implications for the therapeutic strategy for each patient.

Adjuvant treatment options of breast cancer in Norway consist of radiotherapy, chemotherapy, endocrine treatment or HER2 targeted therapies. Some patients receive neoadjuvant treatment, administered in order to shrink the tumor before surgery. The standard chemotherapy regimen in Norway is FEC regimen (fluorouracil, epirubicin and cyclophosphamide). For patients with tumors expressing ER, endocrine treatment in form of tamoxifen or aromatase inhibitors is given. Also, patients with tumors with HER2 overexpression/amplification receive the targeted therapy trastuzumab or lapatinib. Targeted therapies will be covered in more detail in a separate section.

Breast cancer - A disease of somatic aberrations

Over the last decades, sporadic cancer has been viewed as a disease caused by somatic mutations and genomic instability, in which a multistep process of genetic and epigenetic alterations lead to the activation of proto-oncogenes or de-activation of tumor suppressor genes [5,6]. According to this theory, cancer arises from a single normal cell that through transformation may evolve into one or several malignant clones. Accumulation of driving mutations and aberrations may ultimately lead to malignancy. Hanahan and Weinberg conceptualized hallmarks of cancer that cells need to acquire in order to become cancers [7,8]. These hallmarks are illustrated in Figure 4 and include the original “limitless replicative potential”, “evading apoptosis”, “self-sufficiency in growth signals”, “sustained angiogenesis” and “tissue invasion and metastasis” [7], the emerging hallmarks “deregulating cellular energetics” and “avoiding immune destruction” and the enabling characteristics “genome instability and mutation” and “tumor-promoting inflammation” [8].

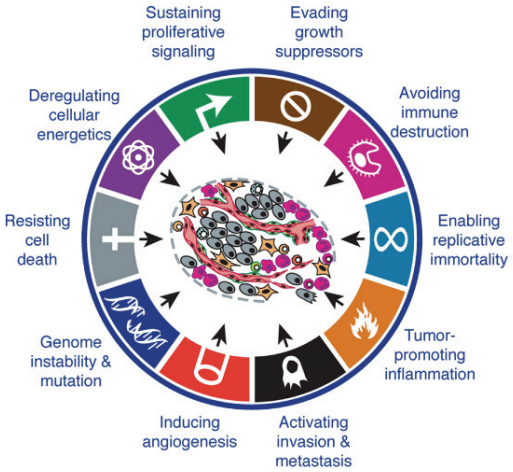


Figure 4. The updated version of hallmarks of cancer. (Modified from Hanahan and Weinberg [8].)

Different models of progression and cellular origin of breast cancer have been proposed, including the linear progression hypothesis and the cell of origin hypothesis [9]. The linear progression model suggests that transformation of a normal cell to an invasive carcinoma happens through a linear accumulation of mutations. The cell of origin hypothesis postulates that different subtypes of breast cancer may originate from different cell types, possibly different progenitor cells [9]. Additionally, according to the cancer stem cell hypothesis, only a small fraction of the cells of a tumor have proliferative capacity [10]. These cells are proposed to be less differentiated, and have the capacity of self renewal and produce the more differentiated cells of the tumor. Whatever the true cellular origins or progression paths may be, different breast cancers seem to have different “drivers” that lead to malignant transformation. A wide variety of genomic aberrations have been detected in breast tumors, which in turn affects the transcription of mRNA and ncRNA and translation into proteins [11]. All aberrations at these levels are being studied to better understand the molecular mechanisms of cancer and to ultimately offer more individualized diagnosis and treatment.

Genomic aberrations in breast cancer

The activation of proto-oncogenes or silencing or disruption of tumor suppressor genes or DNA repair genes can be caused by different DNA alterations such as mutations, copy number variation, translocation, methylation and histone modification. Frequently mutated genes in sporadic breast cancer include the tumor suppressor *TP53* and the oncogene *PIK3CA* [12]. Around 5-10 % of breast cancers are hereditary, and 80-90 % of the hereditary cases have germline mutation in the *BRCA1* or *BRCA2* genes which are involved in DNA repair [13]. Hence, these genes are important in retaining DNA integrity, and loss of function of these genes can thus lead to accumulation of genomic aberrations. New insights into the landscapes of cancer genomes from next generation sequencing techniques have established that the majority of somatic aberrations in cancer are rare variants, which has generated a debate whether cancer is caused by a handful of "drivers", or if the numerous rare variants together drive cancer development [14-17]. Copy number alterations (CNAs) are often observed in breast cancers, as deletion or amplification of whole chromosome arms or smaller fragments of DNA [18]. An example of a relatively frequent CNA in breast cancer is whole arm gain of 1q and loss of 16p, which are events associated with ER-positive tumors. Another frequent event is the amplification of a region harboring the oncogene *ERBB2* on chromosome arm 17q, and has been observed in around 18-20 % of breast cancers [19], though this number is predicted to go down in part due to more precise evaluation of equivocal cases. The mere presence and nature of genomic aberrations may be symptoms of genomic instability, and can thus have prognostic value independent of which genes are affected. Complex genomic aberrations has for example been shown to be an independent predictor of poor outcome [20].

Transcriptomics of breast cancer

Since proteins are translated from mRNA molecules, measuring levels of transcribed mRNA is more dynamic and more closely related to the activity of a gene than measuring mutations and aberrations at the DNA level. The relative ease of measuring transcript levels with microarray technology compared with measuring protein levels or activity, has made transcriptomics a dominating field with respect to molecular profiling of breast cancer. Since the first seminal breast cancer gene expression microarray paper was

published by Perou et. al in 2000 [21], a vast amount of microarray derived gene expression signatures distinguishing between biological subgroups or different prognosis groups have been published (some examples are described in [22]). A few tools for testing gene expression signatures are commercially available, including MammaPrint (Agendia) [23] and Oncotype DX (Genomic Health) [24] which can be used for risk stratification. In addition, noncoding RNAs such as microRNAs (miRNAs), have been found to play important roles in regulation of gene expression levels and protein translation, increasing the importance of transcriptomics as a field [25]. Advances in technology has recently opened up new possibilities in transcriptomics, where the whole transcriptome can be measured using RNA sequencing (RNA-Seq) in contrast to classical microarrays, where the information gained is limited by the spotted probes on the array [26].

Intrinsic molecular classification of breast cancer

The first and perhaps most influential transcriptomic approaches were performed using gene expression microarray technology to study breast cancer, resulting in a molecular taxonomy of breast cancer based on the expression of so-called intrinsic genes [21,27,28]. The expression of genes displaying high inter-variation between tumors from different patients and low intra-variation between tumors from the same patient before and after treatment was used to divide a cohort of breast cancers into distinct groups based upon hierarchical clustering. This approach identified patient groups with significantly different outcomes in terms of overall survival (OS) (Figure 5). Comparing the relative expression levels of these groups to that of cell lines of different cellular origin and healthy breast tissue, gave names to and identified subtypes, i.e. luminal A, luminal B, basal-like, normal-like and ERBB2+ (more recently referred to as HER2 enriched). A centroid-based classifier has also been published, which consists of the average expression of 534 genes for each of the five intrinsic subtypes [28]. These centroids can be correlated to expression of the same genes extracted from microarray data of independent samples, providing they are part of a representative cohort. The latter point is crucial and due to the required centering across samples for each gene prior to classification, the data must be comparable with the original cohort which the centroids are based on [29].

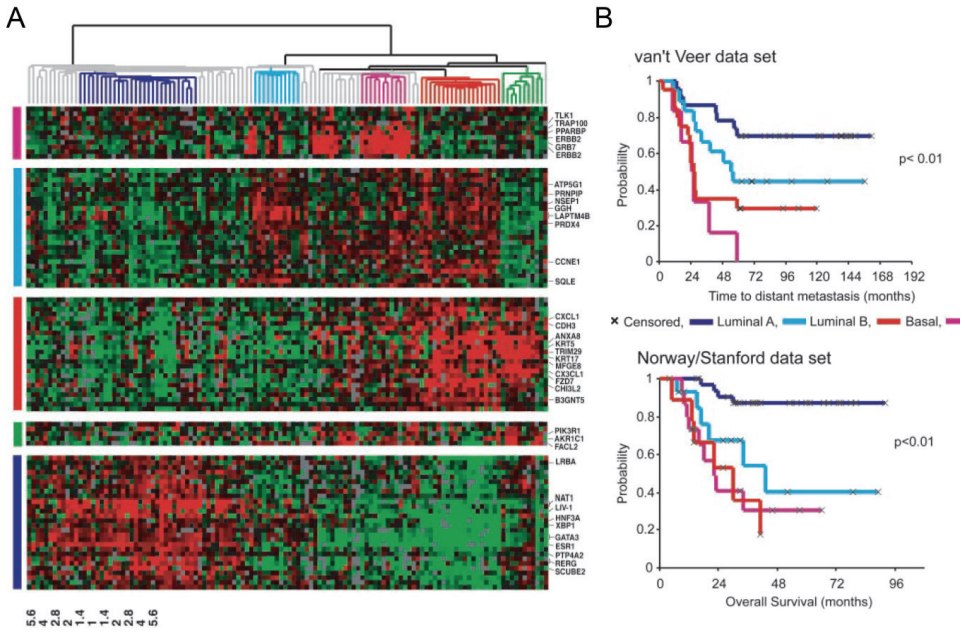


Figure 5. A) Hierarchical clustering of 115 tumor tissue samples and 7 nonmalignant tissues and an "intrinsic" gene set of 534 genes. The cluster dendrogram of the samples are colored according to intrinsic subtype (as coded in B, plus green for Normal-like) for the core samples of each suptype. These core samples were used to build the centroids that can be used to classify independent cohorts of breast carcinomas into the five subtypes. B) Kaplan-Meier curves of outcome in two different breast cancer cohorts showing differences in overall survival between the different subtypes. (Adapted from Sørlie et. al [28].)

Two alternative classifiers for the five intrinsic subtypes have also been published [30,31], and a centroid for a subgroup of the basal-like subtype, claudin-low, has also been recently reported [32]. The existence of the normal-like subtype has been questioned, based on observations that these samples are often associated with low tumor cell percentage [33]. However, this is still an unresolved issue and many studies still include this subgroup. Based on the cell of origin hypothesis, the basal-like subgroup may arise from an earlier stem cell like progenitor, while the luminal-like subtype may arise from a more differentiated luminal progenitor [9]. Contradictory to this theory, both basal-like and luminal-like cells have recently been shown to be able to reproduce both types of cells to gain a phenotypic equilibrium between cells of both subtypes [34]. The intrinsic molecular subtypes have been widely adapted in the research community, but have not yet been

implemented in the clinic. One reason might be that the combined expression of the immunohistochemical markers ER, PR and HER2 that are already utilized in the clinic do to some extent overlap with the gene expression subtypes. For instance, luminal A and B tumors are usually ER and PR positive, HER2 enriched tumors are usually HER2 positive, and basal-like tumors are usually triple negative [35]. Other surrogate markers have also been included for subtyping by immunohistochemistry, such as Ki67 and cytokeratin 5/6 to aid classification of luminal B and basal-like subtypes, respectively.

microRNAs - role in gene regulation and cancer

miRNAs are short (~22 nucleotides long) RNAs that have gained increasing attention since their first discovery in 1993 [36]. Their biogenesis and proposed functions are illustrated in Figure 6. miRNAs can regulate protein expression either by mRNA degradation or translational inhibition. The relative contribution of these mechanisms in humans have been debated, but emerging evidence points to mRNA degradation as the main mechanism of regulation [37]. miRNAs typically bind outside of the protein coding region of the target mRNA transcript, to the 3' untranslated region (UTR), with varying degree of complementarity. Numerous target prediction softwares have been developed to predict which mRNAs are targeted by each identified miRNA, and some of these miRNA-target pairs have been validated experimentally [38]. Different patterns of miRNA expression have been observed in various tissue types, cancers and subgroups of cancer [39]. Several miRNAs have been found to play a role in tumorigenesis, either as oncogenic or tumor suppressor miRNAs [40,41].

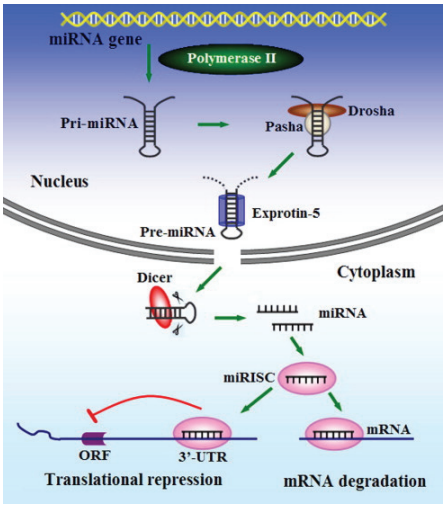


Figure 6. A simplified illustration of the biogenesis of miRNAs and their function. (From Shi and Guo [42].)

Importance of the microenvironment

Tumor-stroma-interactions have been shown to be crucial for both malignant transformation and progression in cancer [43]. In fact, a theory in cancer research describes how cancer can arise from disrupted tissue organization, which may promote or even initiate tumorigenesis in epithelial cells [44,45]. In a healthy breast, the environment surrounding the ducts and lobules consists of mostly extracellular matrix (ECM) as well as a few stromal cells such as fibroblasts and macrophages (Figure 7). This microenvironment has a protective effect on tumorigenesis, and the reversal of the carcinoma phenotype of malignant epithelial cells has been shown to be induced by injecting such cells into a healthy microenvironment [45]. The microenvironment surrounding invasive tumors has developed promotional effects on tumorigenesis, and includes tumor associated macrophages (TAMs), carcinoma associated fibroblasts (CAFs) and aberrant ECM (Figure 7). Other microenvironmental factors affecting tumor progression are pH and the availability of oxygen and nutrients, which are supplied by blood vessels developed during angiogenesis.

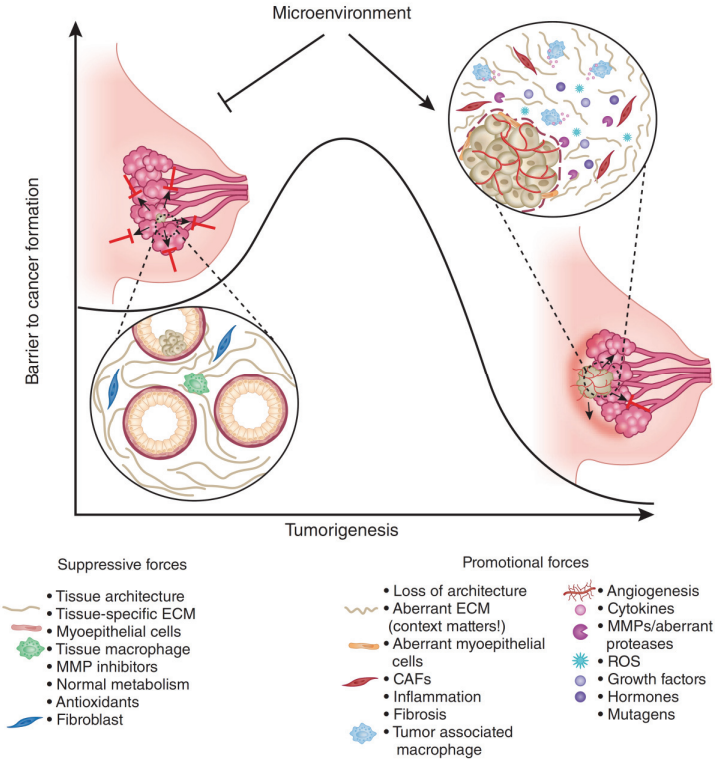


Figure 7. The microenvironment of the normal breast differs from the tumor microenvironment. In the normal situation, the stroma has suppressive force against tumorigenesis, while the stroma surrounding a tumor may be activated to have a promotional force. (From Bissel et. al [45].)

Angiogenesis and VEGF signaling

In 1971, Judah Folkman described the dependency of tumor growth on angiogenesis and how tumor cells can stimulate endothelial cells to proliferate [46]. Angiogenesis is defined as the formation of new blood vessels or neovascularization. Solid tumors start out as avascular and dormant until the so called “angiogenic switch” takes place [47]. The angiogenic switch is governed by a balance between angiogenesis activators such as vascular endothelial growth factor (VEGF) and platelet-derived growth factor beta (PDGFB), and angiogenesis inhibitors such as thrombospondin and statins [47]. The cytokine VEGF (or VEGFA) is one of the most potent pro-angiogenic factors. Hypoxia

induces VEGF expression from the tumor which is secreted and binds to VEGFR on endothelial cells, inducing tumor angiogenesis. VEGF binds to both VEGFR-1 and VEGFR-2, but the latter is responsible for the angiogenic effects of VEGF [48]. Upon VEGF binding, dimerization and autophosphorylation of VEGFR-2 activates several downstream signaling pathways, which in turn induce processes such as vascular permeability and survival, migration and proliferation of endothelial cells (Figure 8). Overexpression of VEGF is found in many tumors, and can induce abnormal and leaky blood vessels that are dependent of VEGF signaling [49].

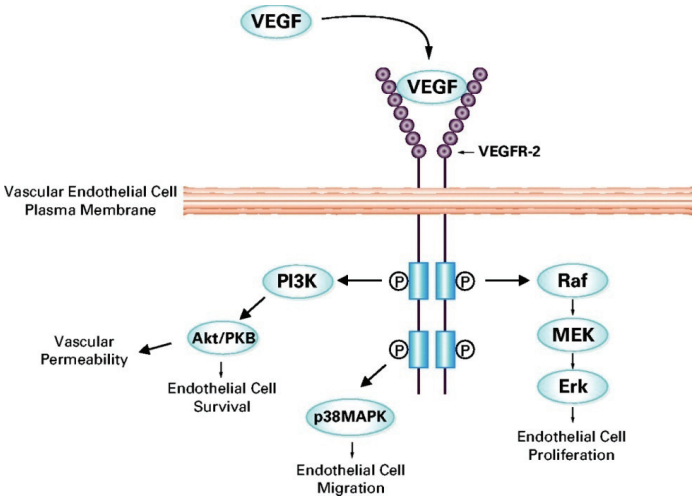


Figure 8. Vascular endothelial growth factor (VEGF) binding and regulation. VEGF binds to VEGF receptor 2 (VEGFR-2) and induces the tyrosine kinase activity of VEFGR-2 which activates downstream signaling pathways. (From Rini and Small [50].)

Rewiring of metabolism in tumors

Cell metabolism is affected by many different factors, such as nutrient availability, pH, oxygen, drugs, stroma surrounding the cells, and the genetic background of the cells. Cancer metabolism is highly aberrant compared to that of normal tissues. Where normal cells slow down production and growth when there is a lack of growth factors and

nutrients or there is high cell density, cancer cells have rewired metabolism which supports the growth and proliferation of these cells even in hostile growth conditions [51]. Some typical metabolic traits of cancers are increased uptake of glucose and glutamine and increased production of lactate, nucleotides, amino acids and lipids (Figure 9). The field of metabolomics studies the metabolome using technologies such as magnetic resonance spectroscopy (MRS), gas chromatography (GC), liquid chromatography (LC), and mass spectrometry (MS) [52]. Interest in the field of metabolomics has increased after discovering the association between aberrant metabolism and genetic mutations in genes of well studied signaling pathways [53]. The importance of metabolism was also recognized by Hanahan and Weinberg, as "deregulating cellular energetics" was added as an emerging hallmark of cancer [8].

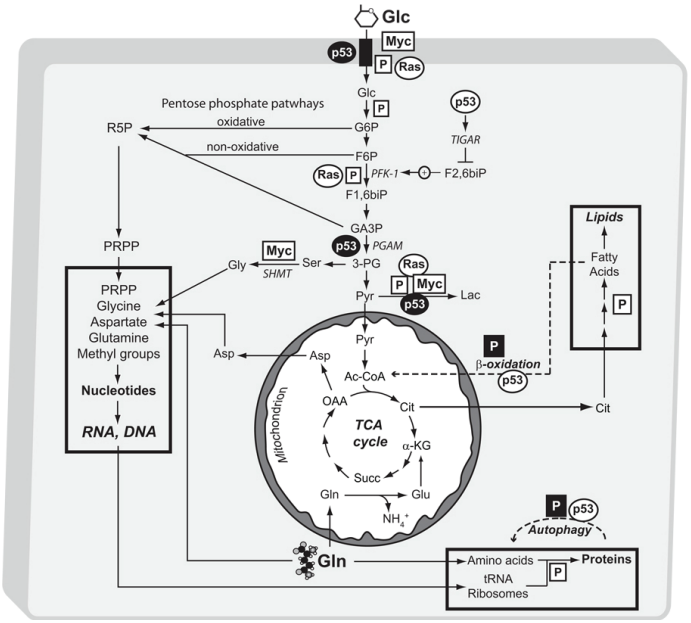


Figure 9. An overview of the central metabolic pathways involved in tumor growth, including some of the oncogenes and tumor suppressor genes that regulate these pathways (From DeBerardinis [54]).

The Warburg effect

In the 1920s Otto Warburg stated that the impairment of cellular respiration was the main cause of all cancers [55,56]. He based this statement on the observation that cancers have highly increased uptake of glucose and metabolize glucose non-oxidatively through glycolysis to lactate, even in the presence of oxygen. This theory was proven to be too simple, but decades later the interest in the Warburg effect is again increasing in the research community. The fact that tumors consume more glucose than other tissues is exploited in PET imaging which measures the uptake of the radioactive tracer 18-F fluorodeoxyglucose (FDG). The molecular mechanisms behind glucose consumption and aerobic glycolysis are still not fully understood, though genetic aberrations have been found in several genes coding for proteins regulating (e.g. hypoxia inducible factor 1 (*HIF1*) and v-myc myelocytomatosis viral oncogene homolog (*MYC*)) or directly involved in glycolysis (e.g. solute carrier family 2 (facilitated glucose transporter) (*SLC2A1*) which is also known as *GLUT1* and hexokinase 2 (*HK2*)) [57], as well as mutations in genes of the Krebs cycle and those involved in oxidative phosphorylation. By using 13-C isotope labeled glucose and 13-C MRS, cell lines and xenograft models of breast cancer have displayed higher glucose uptake and conversion to lactate than normal mammary epithelial cells [58,59]. HR MAS MRS studies on breast cancer biopsies have revealed that intracellular concentration of glucose was negatively correlated with proliferation, while lactate was positively correlated with grade [60,61]. Alanine can be synthesized from lactate, and is produced as a side product to lactate during aerobic glycolysis [59,62]. A recently published study also reported that substantial amounts of the glucose taken up in cancer cells can be rerouted to conversion to serine and glycine through phosphoglycerate dehydrogenase (PHGDH) in some cancer cells [63]. In the same study, *PHGDH* gene expression was also observed to be higher in basal-like breast cancer than in other subtypes.

In contrast to Warburg's theory, many cancer cells have functional mitochondria, and a fraction of the pyruvate produced from glycolysis is used in oxidative respiration. Maximum ATP yield from glucose in human cells is obtained by oxidative respiration, where glucose and O_2 is used to generate ATP in a series of reactions through the Krebs

cycle and oxidative phosphorylation in the mitochondria. The gain in ATP from oxidative respiration is however produced at the expense of carbon, which is lost to CO_2. Proliferating normal and cancer cells are in need of excess carbon for building macromolecules. The lower yield of ATP from aerobic glycolysis in cancer cells is compensated for by higher consumption of glucose [54]. Interestingly, studies of tumor-stroma interactions have revealed that in some aggressive cancers, fibroblasts display aerobic glycolysis while the cancer cells have oxidative respiration and take up and utilize the lactate secreted from the stromal cells [64].

Lipids and cholines in breast cancer

Lipid metabolism is often altered in cancers. Lipids serve several purposes in tumors, such as energy storage in form of fatty acids, as membrane constituents in form of phospholipids and cholesterol, and acting as lipid second messengers and mitogens [65,66]. Phospholipids found in the cell membrane, specifically the most abundant phospholipid phosphatidylcholine (PtdCho), has altered metabolism in many cancers [67]. Elevated level of free total choline (tCho) is a trademark of breast carcinomas and can be measured both *in vivo* and *ex vivo* in tumor tissue or *in vitro* in cell cultures using different magnetic resonance spectroscopy techniques. Choline (Cho) is an essential nutrient that has to be consumed through the diet [68], and choline derivatives such as phosphocholine (PCho) and glycerophosphocholine (GPC) have been found in high concentrations in milk from lactating mice and humans [69,70]. From *in vitro* studies, high PCho level or high PCho/GPC ratio have been suggested as biomarkers for malignancy [71-73].

PtdCho metabolism has been studied both on the metabolite and gene expression levels [67]. The main enzymatic reactions involved in PtdCho metabolism are shown in Figure 10. Cho enters the cell through various choline transporters, and is phosphorylated to PCho, catalyzed by choline kinase (ChoK). PCho is then converted to CDP-Cho, catalyzed by phosphate cytidylyltransferase (CT), which is further converted to PtdCho by cholinephosphotransferase (PCT). PtdCho can be catabolized to acyl-GPC (also called LPtdCho or lyso-PtdCho) and further to GPC (also called GPCho) and Cho, catalyzed by phospholipase A (PLA) and lysophospholipase (LPL) and phosphodiesterase (PD),

respectively. PtdCho can also be catabolized to PCho via phospholipase C (PLC) or to Cho via phospholipase D (PLD). All these genes have various isoforms that are more or less important in these reactions.

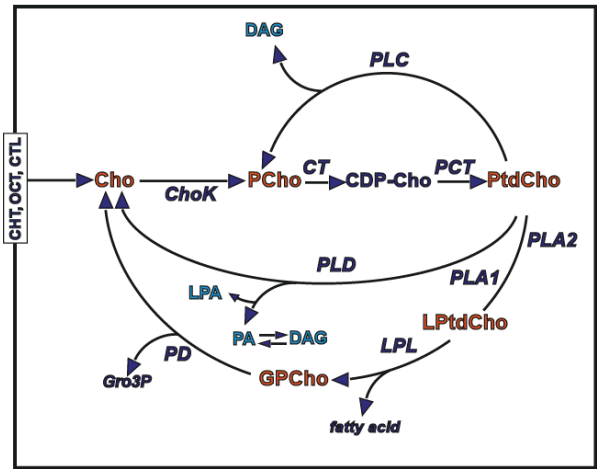


Figure 10. A simplified illustration of the phosphatidylcholine pathway. (From Podo et. al [74].)

Targeted treatment of breast cancer

In addition to surgery, classical chemotherapy and radiation therapy, several targeted treatments of breast cancer are currently being used or tested in the clinic. Drugs targeting ER and HER2 have already been discussed briefly. ER-positive tumors are hormone-dependent, and the aim of endocrine treatment is to induce cell death and reduce cell growth by limiting the levels of available estrogen. The most common anti-estrogen drug used for pre-menopausal women with ER-positive breast cancer is tamoxifen, which is an estrogen receptor antagonist that binds without activating the estrogen receptor. Aromatase inhibitors are a class of drugs mostly administered to post-menopausal women, and block the conversion of androgens to estrogen by inhibition of the enzyme aromatase. These drugs cannot inhibit *de novo* synthesis of estrogen in the ovaries, which are the primary sites of estrogen production for pre-menopausal women. After menopause, the

main source of estrogen switches from the ovaries to different fatty tissue elsewhere in the body, and post-menopausal women with ER-positive breast cancer may thus benefit from aromatase inhibitors [75].

Genes involved in signaling pathways, especially receptor tyrosine kinases, are preferred targets for many targeted therapies [76]. HER2 is overexpressed in up to 20 % of breast cancers, which makes it an attractive target in breast cancer treatment. HER2 amplification is measured routinely for all breast carcinomas in Norway by immunohistochemistry (equivocal cases are validated by FISH), and used as a predictive biomarker for assignment to trastuzumab treatment. Trastuzumab (Herceptin™) is a monoclonal antibody targeting the transmembrane receptor tyrosine kinase HER2, and was developed to treat breast cancer patients with tumors overexpressing HER2. Trastuzumab hinders dimerization of HER2 with other epidermal growth factor receptors, which prevents HER2 from utilizing its kinase potential. For breast cancer patients with HER2 positive tumors, a significant decrease in recurrence has been observed [77]. While trastuzumab targets the extracellular domain of HER2, the tyrosine kinase inhibitor lapatinib blocks the intracellular kinase domain of both EGFR/HER1 and HER2 [76]. This drug has become an important alternative for patients who develop resistance or experience adverse effects from trastuzumab.

Other targeted therapies being tested for breast cancer include PI3K inhibitors, PARP inhibitors and EGFR inhibitors [78]. Targeting cell metabolism has also been suggested as an intriguing approach for treating cancer; examples of preclinically promising approaches are targeting glycolysis by inhibiting HK2 or targeting PCho synthesis by inhibiting choline kinase [53]. Targeting the tumor stroma is also viewed as an attractive approach, as stromal therapy has been suggested to not induce resistance based on the assumption that stromal cells are not as prone to mutations as tumor cells [79]. Successful targeting of stromal components preclinically and clinically includes targeting tumor-associated fibroblasts by vaccinating against fibroblast activation protein (FAP) and targeting endothelial cells by inhibiting vascular endothelial growth factor (VEGF).

Targeting angiogenesis with bevacizumab

Targeting blood vessels as a therapeutic strategy has attracted great interest the last four decades, as sustained angiogenesis is one of the cancer hallmarks and is necessary for tumor metastasis [7]. Especially, therapies targeting VEGF signaling have been tested in clinical trials, and some have been approved for colorectal cancer, non-small cell lung cancer, glioblastoma, renal cell cancer and breast cancer [80]. VEGF induces angiogenesis by binding to VEGFR, a receptor tyrosine kinase on the surface of endothelial cells. The first targeted therapy approved by the FDA for anti-angiogenesis treatment of cancer, was bevacizumab (Avastin™) [80,81]. Bevacizumab is a VEGF monoclonal antibody, targeting the ligand instead of the tyrosine kinase receptor (VEGFR). Some of the proposed mechanisms of bevacizumab have been to inhibit growth of new vessels, induce endothelial cell apoptosis, block recruitment of bone-marrow-derived progenitor cells, elicit anti-tumor immune response by dendritic cell differentiation, and normalize tumor vasculature [80]. Results from clinical trials of bevacizumab and other anti-VEGF therapies have emphasized the latter as a more important mechanism than blocking angiogenesis (Figure 11a). The relative success of bevacizumab in combination with chemotherapy compared to bevacizumab as a monotherapy, suggests that bevacizumab facilitates the delivery of traditional chemotherapy to the tumor cells [82]. Leaky vessels resulting from over-expression of VEGF have been shown to become less permeable after targeting with anti-VEGF [83] (Figure 11b).

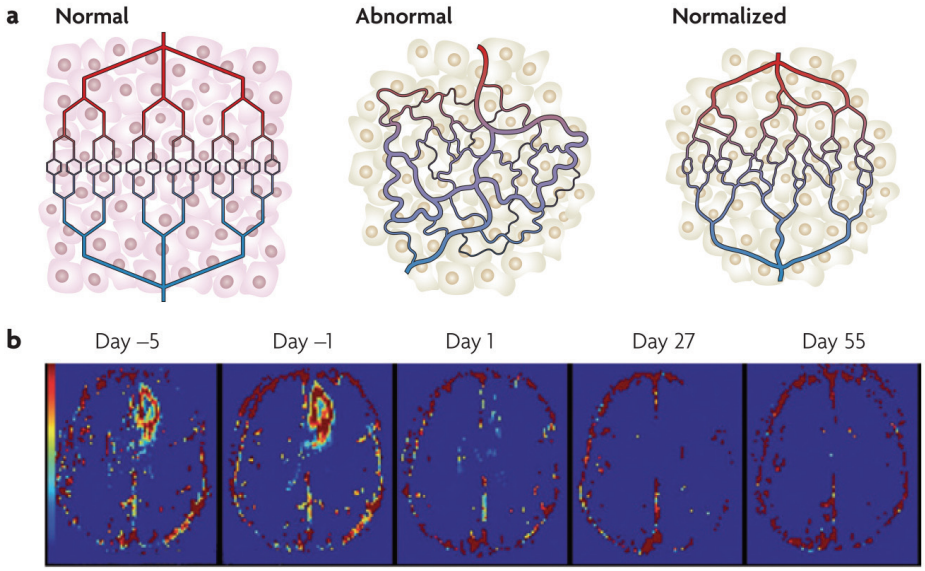


Figure 11. VEGF-targeted therapy introduces normalization of tumor vasculature. (a) An illustration of normalization of tumor vasculature. (b) Magnetic resonance imaging of a glioblastoma patient shows a rapid decrease in vascular permeability after anti-VEGF treatment. (From Ellis and Hicklin [80].)

Material and Methods

Material

Clinical breast carcinomas

The tumor samples included in Paper I are a part of an ongoing study (referred to as Oslo II) which is a collaboration between different hospitals in Norway. All tumors included in the study described in Paper I were diagnosed as IDCs and collected at Oslo University Hospital Ullevål in 2007. The tumor samples included in paper II were obtained from breast cancer patients with IDC operated at St. Olav's University Hospital in Trondheim since 2000. The tissue specimens were obtained during surgery and stored in a biobank administered by the MR Cancer Group at the Norwegian University of Science and Technology (NTNU). Additionally, previously published microarray data from a breast cancer cohort of 115 primary tumors were included in Paper IV [84]. This cohort is a subset the Oslo Micrometastasis Project, where fresh frozen breast cancer tissue was collected from 1995 to 1998.

MAS98.06 and MAS98.12 breast cancer xenograft models

The xenograft models used in paper III and IV have been established as described previously [85] and some characteristics are summarized in Table 1. In short, primary tumor tissue from 29 breast cancer patients was grafted orthotopically into severe combined immunodeficiency (SCID) mice. Two of these tumors were successfully established as xenograft models by serial transplanting tumor tissue for more than 10 passages. These xenograft models have been characterized as representing luminal-like (MAS98.06) and basal-like (MAS98.12) breast cancer [85]. Measuring gene expression using human microarrays may introduce bias to intrinsic classification of xenografts, because several genes involved in ECM were found to be expressed by mouse stromal cells rather than by the engrafted human tumors. However, an attempt of merging the microarray data from the xenografts with 40 patient tumors showed that the MAS98.06 xenograft and primary tumor clustered with tumors classified as luminal B.

Table 1. Characteristics of primary tumor and xenograft models.

	Basal-like xenograft (MAS98.12)		Luminal-like xenograft (MAS98.06)	
	Primary tumor	**Xenograft**	**Primary tumor**	**Xenograft**
Tumor grade	Grade III IDC	NA	Grade III IDC	NA
Hormone receptor status	ER-/PR+ *	ER-/PR-	ER+/PR+	ER+/PR+
ERBB2 amplification**	Negative	Negative	Negative	Negative
Intrinsic molecular subtype	Basal-like	Basal-like	Luminal-like	Luminal-like
TP53 status	Small population with codon 120 mutation	Homozygous codon 120 mutation	Heterozygous codon 179 mutation	Homozygous codon 179 mutation
Proliferation index (Ki67)	NA	28%	NA	35%
Volume doubling time	NA	1-2 days	NA	7 days

* The primary basal-like carcinoma showed very weak cytoplasmic staining for PR

** Measured at the DNA level by aCGH

Technology

Microarray technology

Microarray technology was used in all four papers of this thesis. The principle of microarray technology is to hybridize the sample RNA (or DNA) to complementary probes at fixed locations on a support material such as a glass slide or nitrocellulose surface. In this way, the levels of thousands of transcripts can be measured simultaneously. In all four studies, Agilent microarray technology was used to measure gene expression, and in Paper I, miRNA expression microarrays were also used. The basic concept of Agilent gene expression microarrays is depicted in Figure 12. mRNA in the sample is first synthesized

into double stranded cDNA using a reverse transcriptase and a DNA polymerase. The transcripts are then linearly amplified to cRNA, using an RNA polymerase. During amplification, a fluorescent dye (Cy3 was used for tumor samples) is incorporated into the cRNA transcripts by using labeled CTP nucleotides. The samples are then hybridized to microarrays overnight at the appropriate temperature, and the microarrays are scanned using a laser scanner; yielding intensity levels of each probe on the array.

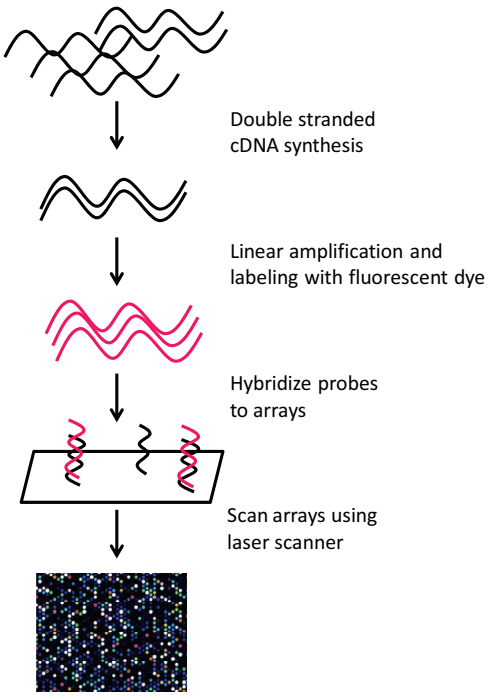


Figure 12. An illustration of the basic concept of Agilent one-color gene expression microarray technology.

These microarrays are oligonucleotide microarrays with 44,000 and 15,000 60-mer probes for mRNA and miRNA analysis, respectively. Different versions of the Feature Extraction software (Agilent) give different information about which and how many unique mRNA are represented by the 41000 unique probes on the gene expression array. Typically each unique transcript is represented by 1-3 different probes. 264 of probes are printed 10 times on the array for quality control. 723 human and 76 human viral miRNAs are

represented on the miRNA array. In Paper I, III and IV, a one color system was used. In this system, only the tumor sample RNA is Cy3 labeled and hybridized to the microarrays. In Paper II, a two color system was used with a Cy5 labeled human reference sample in addition to the Cy3 labeled tumor sample. In this system, the two samples compete for hybridization to the probes on the microarray, and the signal intensity from the two fluorescent labels is utilized to calculate a signal-ratio for each spotted probe.

High Resolution Magic Angle Spinning Magnetic Resonance Spectroscopy (HR MAS MRS)

HR MAS MRS was used in papers II-IV. HR MAS MR spectroscopy is a non-invasive nuclear magnetic resonance (NMR) technique that can be used to measure levels of free metabolites in intact tissue. Numerous spectroscopy and imaging techniques based on the NMR phenomenon have been developed for measuring anatomic, functional and biochemical properties in clinical and experimental medicine. In fact, MRI is one of the most accurate imaging tool available for cancer detection [86].

MR technology exploits the magnetic properties of certain nuclei, such as 1-H and 13-C. Magnetic nuclei have a property called spin, which when exposed to an external magnetic field B_0 aligns either with or against the magnetic field. There are always slightly more spins in the lower energy state, and this difference increases proportionally to the strength of the magnetic field. The aligned spins of the nuclei precess around the axis of B_0. The precession frequency is called the Larmor frequency and can be calculated by $\nu = -\gamma B_0/2\pi$, where γ is the gyromagnetic constant and B_0 is the external magnetic field. At an external magnetic field of B_0 = 14.1 Tesla (T), as used in this work, and the gyromagnetic constant of 1-H being γ = -26.75 x 10^7 $s^{-1}T^{-1}$, the Larmor frequency of 1-H is 600 MHz. A second magnetic field is then applied at the calculated 600 MHz to induce the magnetic resonance phenomenon by exciting the nuclei. The second magnetic field is a radio frequency (RF) pulse. When the excited nuclei relax, the MR signal can be recorded as an oscillating signal in the time domain. This signal can be converted into its frequency components by Fourier transformation.

The magic angle spinning effect in MRS was first described by Lowe [87] and Andrew [88]. The HR MAS MRS technique is based on spinning the sample rapidly (5000 MHz in this work) at an axis inclined 54.7° to the external magnetic field from the magnet (Figure 13). Conventional NMR spectroscopy of semi-solids such as tumor tissue gives poor resolution, because of anisotropic interactions between immobile nuclei. The rapid spinning in HR MAS MRS splits the broad MR bands into a narrow peak at isotropic resonance frequency and side peaks of weaker intensity. Spinning the sample at the magic angle cancels out the time independent anisotropic interactions between the spins. The resolution of HR MAS MR spectra of intact breast cancer tissue can be as good as in NMR spectra of perchloric acid extracts of tissue [89].

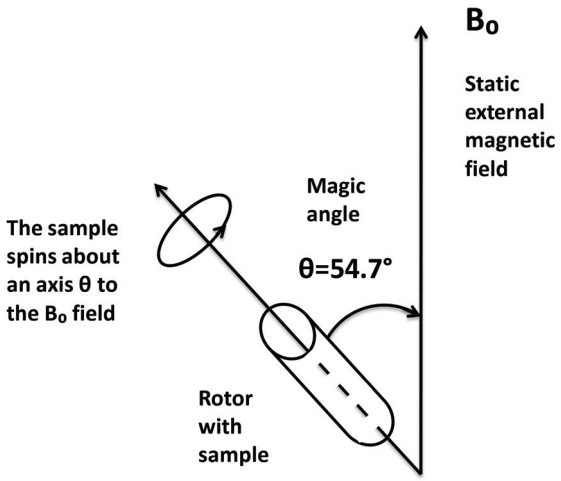


Figure 13. Illustration of how a sample spins around an axis inclined by the magic angle 54.7 ° to the applied magnetic field, B_0, in HR MAS MRS.

Different metabolites in a sample can be detected using HR MAS MRS because the nucleus of each atom experiences the external magnetic field differently according to the chemical structure in its proximity [90]. This phenomenon is called chemical shift, and is calculated in parts per million (ppm), which is the resonance frequency of a specific atom relative to a chosen reference. As an example of different chemical shifts, the protons of the amine

groups of Cho, PCho and GPC have different resonance frequencies because the amine groups are experiencing a slightly different magnetic field due to the chemical structure surrounding them. In Figure 14, a HR MAS MR spectrum of a breast cancer biopsy shows how these protons have chemical shifts at slightly different ppm. Using *in vivo* MRS, the resonance frequency of these three metabolites cannot be separated (Figure 14).

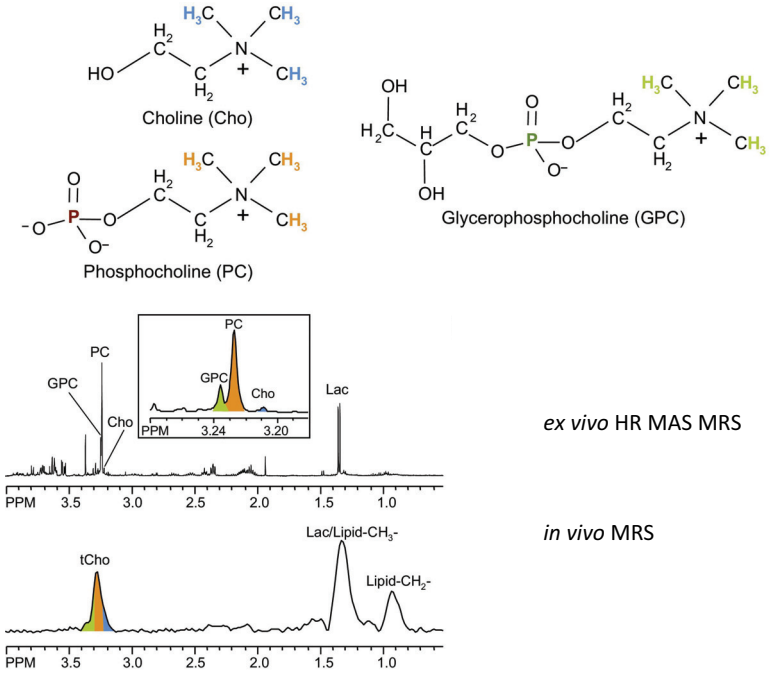


Figure 14. In 1-H MRS, the protons of small soluble molecules are detected by applying radiofrequencies of the same frequency as the nuclei of protons recess around an applied external magnetic field. Top: The choline signals observed in breast tumors originate from the protons in the amine groups of Cho (blue), PCho (orange) and GPC (green). Bottom: *ex vivo* HR MAS MRS has higher resolution than *in vivo* MRS. The total choline (tCho) signal observed *in vivo* consists of GPC, PCho and choline which can be distinguished by HR MAS MRS. (Modified from Glunde and Bhujwalla [91].

In this thesis, HR MAS MRS was performed on a Bruker Avance DRX600 spectrometer. Two different spectra were acquired for each sample, a single pulse and a spin echo spectrum. The single pulse spectra were acquired by sending a single 90 ° RF pulse, which gives

quantitative spectra. The spin echo spectra were acquired by sending a 90 ° RF pulse followed by a series of 180 ° pulses which lead to suppression of most of the lipid signals. Acquisition of both spectra included suppression of the water signal by presaturation. While the single pulse spectra were used for quantification of metabolite concentrations (Paper II), the spin echo spectra were used for multivariate analysis (Paper II). In addition, pulse acquired spectra with ERETIC were used for quantification in papers III and IV.

Statistical and bioinformatic analysis

Analyzing differences between groups of samples

A Student's t-test was applied for pairwise group comparisons of variables such as metabolite concentration (Paper III). For multiple group comparisons, an analysis of variance (ANOVA) model was applied (RIN value in relation to different time points in Paper I). For both, the null hypothesis is that the mean is identical in the groups to be compared. Both also assume that observations are independent and follow a normal distribution in each of the groups to be compared. The xenograft study described in this paper involved several animals from each xenograft model. Fitting an ANOVA model to all treatments and days for each of the xenografts in Paper IV, allowed for more degrees of freedom for the comparison of each treatment group to the controls. Similar approaches were used to analyze the gene expression and miRNA data in all papers, using linear models for microarrays (Limma) [92]. Limma has the possibility of performing t-tests, linear regression and ANOVA. This is performed for each transcript in the dataset, and the t- and F-statistics are modified using Empirical Bayes correction which makes the analysis more stable when analyzing data with a small number of samples [92]. Empirical Bayes correction "borrows" information from other transcripts to obtain a more reliable estimate of group variability which may significantly improve the accuracy of the inference. Basically, the idea is to estimate a "fudge factor" that is added to the standard deviation in the denominator of the test-statistics. In addition, p-values obtained using Limma were corrected for multiple testing by Benjamini Hochberg's False Discovery Rate (FDR) [93]. In brief, the false discovery rate is calculated by multiplying the raw p-value of each test by the number of tests performed and dividing by the number of tests that were more significant than that particular test.

Dimension reduction and clustering techniques

Hierarchical clustering is a popular method to search for structure in genomic and transcriptomic data, and was used in all papers for analyzing or visualizing results. The idea of hierarchical clustering is to group entities together in a dendrogram according to their distance in high dimensional data space. The distance measures can be calculated using metrics such as Euclidean distance or correlation distance (1 – the correlation between two entities). The distance measures can then be used to cluster different entities by using linkage methods such as single, average, complete or ward linkage. The different linkage methods use different strategies to make clusters. For example, Ward clustering tries to produce as distinct clusters as possible, while complete linkage tries to produce similar clusters. In gene expression analysis, both samples and genes are usually clustered and the gene expression values depicted in a heatmap to visualize their relative expression levels. Multidimensional scaling (MDS) is a dimension reduction technique used to explore similarities or dissimilarities in data. Metric multidimensional scaling which was used in Paper I and II uses the same distance measures that were described for hierarchical clustering. Based on these distance measures, the data are forced into a given number of dimensions (2 dimensions were used in Paper I and II).

Correlation analysis

Different approaches for calculating correlation between either miRNA or metabolites and RNA transcripts were used in Paper I and II, respectively. Pearson's correlation coefficients measure the linear dependence between two variables. The correlation coefficient rho is a value between -1 and 1, where -1 is perfect inverse correlation and 1 is perfect correlation between two variables. The non-parametric Spearman's rank correlation is performed by calculating Pearson's correlation on ranks instead of raw values. In Paper I, pairwise miRNA-mRNA Pearson correlation coefficients were calculated for each patient using the 4-5 time points. To analyze which miRNA-mRNA pairs were consistently correlated across the ten patients, a novel statistical method called rank consistency score by Navon et al. was applied using the correlation coefficients as input [94]. In Paper II, pairwise metabolite and RNA transcript correlations were calculated using Spearman's rank correlation.

Prior knowledge based approaches

Prior knowledge of transcriptomic data from Gene Ontology (GO) were used in Paper I, II and IV and miRNA-target prediction and transcription factor binding sites was used in Paper I. Gene Ontology is a controlled vocabulary that is utilized to annotate genes according to terms belonging to biological process (BP), molecular function (MF) and cellular compartment (CC) categories [95]. Within each of the main categories, the GO-terms are organized hierarchically with more specific terms further down in the hierarchy. The hierarchies are constructed of directed acyclic graphs (DAGs) which allow a descendant node to have multiple predecessors. The miRNA target prediction tool used in Paper I is a based on the miRanda database and functional data [96]. The mirSVR scores used in the analysis are based on a machine learning method to rank miRNA target sites by a score based on sequence and contextual features from miRanda as well as data from functional miRNA experiments. To analyze for enrichment of different sets of predefined transcripts among transcripts that were differentially expressed or correlated to features of interest (such as miRNAs or metabolites), the tools Gene Set Enrichment Analysis (GSEA) [97], Gene Set Analysis (GSA) (http://www-stat.stanford.edu/~tibs/GSA) and GOrilla [98,99] were used in Paper I, II and I and IV, respectively. These tools analyze whether a geneset is overrepresented among a ranked list genes of genes according to the statistic of interest, with slightly different statistical approaches. The results obtained from these enrichment analyses were also false discovery rate corrected [93].

Results in brief

Paper I – Consequences of ischemia time on microRNA and gene expression in cancer tissue

(*Molecular Oncology* 2011, doi:10.1016/j.molonc.2011.08.004)

In paper I, tumor tissue was frozen at different time points after being removed surgically to study the effect of ischemia time on the transcriptome. Tumor tissue from ten patients frozen at different time points up to 6 hours were used in mRNA and miRNA expression microarray experiments. Statistical analysis was performed to analyze if any transcripts were affected by ischemia time, and a novel statistical method was implemented to analyze miRNA-mRNA covariation. A total of 1788 mRNA and 56 miRNA transcripts were found to vary significantly with ischemia time. The most significant transcripts displayed increasing expression levels with time, such as the cancer related miRNAs hsa-miR-663 and hsa-miR-125a-3p and stress response genes such as *GADD45B*, *JUND* and *FOSB*. Such transcripts should be eliminated from any prognostic or predictive signature. By analyzing consistent miRNA-mRNA time-correlation across patients, miRNAs were associated with biological processes through their correlation with mRNAs. Several miRNAs, including hsa-miR-1228, hsa-miR-1225-5p and hsa-miR-574-5p, were associated through their correlation to mRNAs to biological processes such as “response to stimulus” and “stress response”. These miRNAs also showed enrichment of predicted targets among either their positively or negatively correlated mRNAs, which means they may play both direct and indirect roles in response to stress.

Paper II - Merging transcriptomics and metabolomics – advances in breast cancer research

(*BMC Cancer* 2010, 10: 628)

In Paper II, transcriptomics and metabolomics of breast cancer was merged to explore the potential of combining these two different types of information. Three goals of this study were pursued: First, to assess whether HR MAS MRS and gene expression microarrays could be performed on the same samples; second, whether using data from HR MAS MRS and gene expression microarrays from the same sample could contribute to a refined subclassification of breast cancer; and third, whether genes involved in specific biological processes were correlated to the concentration of various metabolites in breast carcinomas. For the first goal, paired samples were used to analyze differences in RNA integrity and gene expression between control samples and samples that had been used in HR MAS MRS. No difference in RNA integrity was found, but a subset of genes were found to show a subtle difference in expression between samples exposed to HR MAS MRS and those that were not. For the two remaining goals, a different sample set was used in HR MAS MRS and microarray experiments. The majority of samples were classified as luminal A based on the gene expression data. Within these luminal A samples, the metabolic data was used to subdivide patients into three groups. One of these groups seemed to have a higher Warburg effect, with lower glucose, higher alanine and a trend of higher lactate levels. By correlating metabolite and gene expression levels, several biological processes were found to be enriched among genes that correlated with specific metabolites. Choline was correlated to genes involved in cell cycle related processes while *myo*-inositol and taurine were correlated with genes involved in extracellular matrix processes.

Paper III - Distinct choline metabolic profiles are associated with differences in gene expression for basal-like and luminal-like breast cancer xenograft models

(*BMC Cancer* 2010, 10: 433)

In Paper III, choline metabolism was studied in orthotopic xenograft models representing basal-like and luminal-like breast cancer using HR MAS MRS and gene expression microarrays. The two xenograft models were found to have opposite choline patterns with respect to PCho and GPC. While the luminal-like tumors displayed relatively higher PCho and lower GPC levels, the basal-like tumors displayed the opposite. The basal-like also showed higher glycine levels than the luminal-like tumors. The same metabolomic differences were found in HR MAS MR spectra from patient tumors representing luminal-like (ER+/PR+) and basal-like (triple negative) subtypes, suggesting that the xenograft models are representative of human cancer. Expression levels of 105 genes involved in phospholipid metabolism were compared between the two xenograft models. Significant differences were found for 64 of these. Some of the gene expression differences were concordant with the observed metabolic differences, such as higher choline kinase (*CHKA*, *CHKB*) levels in luminal-like tumors which may explain the higher PCho levels; higher phospholipase A2, group IVA (*PLA2G4A*) levels which could explain the higher GPC concentration and higher choline dehydrogenase (*CHDH*) and sarcosine dehydrogenase (*SARDH*) levels which could explain the higher glycine levels in basal-like tumors.

Paper IV – Metabolic and transcriptional response to bevacizumab and doxorubicin in xenograft models of basal-like and luminal-like xenograft models

(*Manuscript*)

In paper IV, metabolomic and transcriptomic responses to bevacizumab and doxorubicin therapy in two xenograft models representing basal-like and luminal-like breast cancer were studied. In the xenograft models treated with bevacizumab or a combination of bevacizumab and doxorubicin, delayed tumor growth by bevacizumab was observed in the basal-like but not in the luminal-like model. In order to identify metabolomic and transcriptomic responses induced by bevacizumab, tumor samples were harvested at day 3 and 10 from animals used as controls or treated with bevacizumab, doxorubicin or the combination of these drugs, and were analyzed using gene expression microarrays and HR MAS MRS. Bevacizumab-induced molecular responses were observed in the combination treated basal-like tumors which were not observed in the non-benefiting luminal-like model. Levels of the metabolite GPC decreased after treatment in the basal-like tumors and increased after treatment in luminal-like tumors. Hence, GPC demonstrated potential as a metabolomic response biomarker. At the transcriptomic level, 14 genes showed significant differential expression between basal-like tumors receiving the combination therapy and those receiving doxorubicin monotherapy. These differences may represent the additional effect of bevacizumab treatment. Using published gene expression signatures and GO-term enrichment analysis, biological responses related to possible treatment resistance effects were explored. Additionally, differences in expression of genes involved in phospholipid metabolism between basal-like and luminal-like untreated xenografts were correlated with differences observed in human breast cancer. These finding suggest that distinct phospholipid metabolism may be intrinsic features of basal-like and luminal-like breast cancers.

Discussion

In this work, transcriptomic and metabolomic measurements of breast cancer tissue from patients and xenograft models have been studied. Methodological aspects will be discussed, including advantages and limitations of technologies and analytical approaches used in this work. Next, biological aspects will be discussed, including the relevance of the patient and xenograft tumor materials and methods used for studying breast cancer, the future of bevacizumab treatment for breast cancer, and finally the clinical potential of transcriptomic and metabolomic biomarkers.

Methodological considerations

Measuring transcript levels using microarray technology

Microarray technology has revolutionized molecular biology by allowing measurements of thousands of entities (e.g. RNA transcripts) simultaneously in a high throughput manner. However, there are several technical issues to be aware of when using microarray technology, including batch effects, cross-hybridization, and measurement accuracy [100]. Batch effects are not so obvious when comparing samples with large inter-tumor heterogeneity. However, when comparing biologically similar samples (as in Paper IV), batch effects are evident even after data normalization, and are therefore especially important to consider in such experiments. Carefully planned experimental designs of all microarray experiments performed in this work were implemented to minimize random noise and avoid systematic bias or confounding effects in the downstream analysis. Cross-hybridization of transcripts with high sequence similarity can occur when using microarray technology. Although the Agilent miRNA and mRNA expression microarrays are designed to avoid this problem, there may still be incidences of cross-hybridization for some of the probes. In paper I, probes designed to detect miRNAs from several different types of human viruses gave detectable signal levels and displayed differential expression. This result may be real, but it could also be a result of cross-hybridization. Further studies would be needed to resolve this issue. The absolute expression values measured using gene expression microarrays have been shown to have poor correlation with more

quantitative techniques such as reverse transcriptase quantitative polymerase chain reaction (RT qPCR). Comparing ratios between different groups of samples, as performed in this work, significantly improves the correlation with similar ratios obtained from RT qPCR [100]. Advances in sequencing technology over the last years have resulted in next generation sequencing techniques such as RNA-Seq, which have several advantages over microarrays in the field of transcriptomics [26]. The main advantage of RNA-Seq is that the entire transcriptome can be assessed, without the need of pre-defining transcripts of interest, as with microarray technology. This allows for the discovery of new splice variants of genes, assessing allele-specific expression and the discovery of ncRNAs that were not previously known to be transcribed. Even though RNA-Seq is superior for discovering different transcript variants, estimation of gene expression levels has been shown to agree between the two technologies [101]. Hence, microarray technology is still relevant for measuring gene expression. As of yet, RNA-Seq is many times more expensive than microarrays, has been shown to display bias in base-call error rates and genome coverage [102], and the bioinformatic and statistical frameworks for data handling and analysis are still in their infancy [101]. Additionally, measurements of miRNAs are superior when using microarrays compared with next generation sequencing [103]. Microarray technology may thus be a valuable tool in the field of transcriptomics for several years to come.

Using HR MAS MRS as a metabolomic tool

HR MAS MRS is a commonly used tool to study metabolomics of cancer specimens, mainly because it is a non-destructive technique. The advantage of keeping the tumor tissue intact after HR MAS MRS was demonstrated in this work, where gene expression analysis (Paper II) and histopathology (Paper III and IV) was performed on the same tissue that had been subjected to HR MAS MRS. Another advantage of HR MAS MRS is that it can be used as a quantitative technique, where the metabolite peaks from the single pulse spectra are compared to an internal standard (such as trimethylsilyltetradeuteropropionic acid (TSP) used in Paper I) or an electronic reference (such as the ERETIC method used in Papers III and IV). The most commonly used alternative to MRS techniques in the field of metabolomics is mass spectrometry (MS). GC/LC MS techniques have been used extensively in plant metabolomics, and have also been used to study cancer metabolomics

[52]. GC/LC MS are more sensitive techniques than MRS, and can detect metabolites at concentrations as low as 100 nM while the sensitivity of MRS is down to around 100 µM. However, sample preparation is more extensive for MS than for MRS, and may introduce bias in the analysis since not all compounds are suitable for MS. HR MAS MRS measures the free metabolites in the cytoplasm of intact cells, which gives different information than extracting all metabolites that may be membrane bound [65,104]. The main advantage of HR MAS MRS is its relevance in clinical applications, as *in vivo* MRS can be performed using standard clinical MR magnets. The 1.5 or 3 T magnets being in clinical use in Norway today cannot resolve the different choline metabolites when using 1-H *in vivo* MRS (as illustrated in Figure 14). However, with higher field strengths combined with the use of 31-P MRS [105], the GPC and PCho peaks studied in Papers III and IV may be separated also *in vivo*. Additionally, hyperpolarized 13-C MRS is a new technique that enhances the MR signal *in vivo* by injecting hyperpolarized substrates such as 13-C pyruvate [106]. With this technique, aerobic glycolysis can be studied *in vivo* by measuring the conversion of pyruvate to lactate and alanine. Hyperpolarized 13-C MRS has the potential to identify patients carrying tumors with a high Warburg effect, such as the A2 group of luminal A tumors identified in Paper II, even before a single biopsy is taken.

Combining different types of transcriptomic and metabolomic data

In this work, different types of transcriptomic profiles (miRNA and mRNA) were combined in Paper I, while mRNA transcriptomic profiles and metabolomic profiles were combined in Paper II, III and IV. Several strategies for sample handling and data analysis have been considered. For the patient tumor biopsies used in Papers I and II, the measurements of different molecular levels were from the same tissue samples. This reduces the biological variance and increases the power of statistical analysis. The xenograft tissue used for gene expression and metabolic analysis were either from different passages (Paper III and some of the samples in Paper IV) or from adjacent tumor tissue from the same passages (Paper IV), which may introduce bias when comparing transcriptomic and metabolomic levels. However, comparisons of the HR MAS MR spectra and microarray data from different passages suggest that the xenograft models are stable across passages.

Joint analysis of the transcriptomic and metabolomic data was performed using both data driven and knowledge driven approaches. In paper I, consistent time-correlation among patients was analyzed to identify miRNA-mRNA associations, and databases of prior knowledge on miRNA-mRNA relationships (target prediction) and biological functions of genes (GO) were used to generate hypotheses on what these associations could infer. Combining transcriptomic and metabolomic data (Papers II, III and IV) is not as straight forward as combining miRNA and mRNA data. Although data obtained from gene expression microarrays and HR MAS MRS are both high dimensional, they are of a vastly different nature. While microarray data consists of discrete measures which represent unique transcripts, the discrete measures from HR MAS MRS represent a continuous curve where only a subset of the data points are informative signals from metabolites. In fact, even though spectra from HR MAS MRS of breast cancer tissue contain several thousand data points, only around 30 metabolites have been allocated in these spectra [89]. In Paper II, the multivariate technique used to analyze HR MAS MR spin echo spectra (hierarchical clustering) treat each data point as a discrete point without taking into account their relative position. Multivariate analysis of gene expression data and spin echo spectra were combined in this paper to search for natural structure on the transcriptomic and metabolomic levels, in particular to identify biologically relevant subgroups among the tumor samples. It should be noted that all multivariate dimension reduction and clustering techniques are sensitive to the cohort studied as well as the data processing and choice of statistical method. Validating the same groups in independent cohorts may be difficult and classification methods such as the centroid-based used to assign intrinsic gene expression subtypes to independent samples could force incorrect group affiliations [29]. This underlines the importance of large representative cohorts for identifying relevant subgroups. In an attempt to bridge the gap between metabolomic and transcriptomic phenotypes in Paper II, prior knowledge from GO was used (comparable to the analysis of miRNA and mRNA in Paper I) to identify associations between metabolites and specific biological processes. However, the small sample size and homogenous cohort reduced the statistical power of this analysis.

In Paper III and IV, prior biological knowledge of genes involved in PtdCho metabolism was used when analyzing the gene expression and metabolomic data to select, which is a

common strategy [71,72,107,108]. This approach gives knowledge-based hypotheses of direct relationships between the levels of gene transcripts and metabolites. However, causality cannot be inferred without properly designed functional studies, which have not been the scope of this work. In paper IV, the metabolomic changes observed were not reflected at the gene expression level. This emphasizes that gene expression levels are not equivalent to enzyme levels or activity, and metabolites are affected by several different pathways as well as by the microenvironment.

Biological considerations

Measuring snapshots of heterogeneous breast cancer tissue

The molecular studies in this work were performed on bulk tissue obtained at a specific time point, representing a snapshot of the transcriptome and metabolome of these tumors. There are several important biological considerations that should be pointed out regarding this sampling procedure, with respect to the stromal component, clonal nature of the tumor and timing of the sampling. When using bulk tumor tissue, the relative contribution of cancer epithelial cells, adipocytes, fibroblasts, immune cells and endothelial cells are usually unknown and not accounted for. In Paper II, adjacent tumor tissue was used to perform histopathological evaluation of the tissue, which was correlated to the biological findings. The concentration of taurine and *myo*-inositol was found to correlate with tumor cell percentage, indicating that the tissue composition influences the metabolomic results. This bias is important to be aware of, and underlines the asset of the non-invasive nature of HR MAS MRS, allowing for evaluation of the tissue subsequent to the HR MAS MRS experiments. Tumor tissue from the xenograft models were also histopathologically evaluated after HR MAS MRS in Paper III and IV, showing that the tissue consisted predominantly of viable tumor cells. This indicated that tumor cell percentage was not as relevant in the xenograft tissue compared with the clinical samples. This may in part be explained by the relative ease of macro-dissecting xenograft tumor tissue compared to clinical samples. However, differences in stromal components between the two xenograft models have been observed, in particular more active angiogenesis and higher expression of ECM involved genes in the basal-like model [85,109]. It should be

noted that a portion of the metabolomic signals could also arise from extracellular fluids such as blood [110]. This is particularly relevant in Paper IV, where observed changes in metabolite levels after bevacizumab treatment could in theory be due to changes in tumor vascularization. However, this would also be true for response monitoring using *in vivo* MRS. It may be argued that, provided the metabolomic signals have predictive or prognostic value, dissecting their origin is not crucial.

Another aspect of tumor heterogeneity is that several different clones may be represented in a single tumor. Clonal heterogeneity can be observed by histopathology and cytogenetics [111,112]. Using high resolution genomic techniques, such as array comparative genomic hybridization (aCGH) and next generation sequencing, Navin and colleagues modeled clonal evolution within single tumors [113,114]. Notably, some tumors were observed to be monogenomic while others were polygenomic, suggesting differences in chromosomal stability between tumors [113]. The clonal nature of the patient tumor biopsies (Paper I and II) and the xenograft biopsies (Paper III and IV) are unknown, and this heterogeneity is not accounted for in the analysis of the bulk tumor samples used in this work. In addition to stromal components and the clonal nature of carcinoma cells, the timing of the sampling of tumor tissue may also affect the transcriptomic and metabolomic profiles measured. In Paper III and IV, the standardized handling of the animals and the possibility of obtaining biological replicates, ensure that the xenograft biopsies obtained are comparable. However, the treatment response on the transcriptomic and metabolomic levels observed in Paper IV might have been different if the biopsies were obtained at different time points than 3 and 10 days after treatment. For the patient tumor biopsies in Paper I and II, numerous other variables such as differences in time to diagnosis, diet and age may have affected the transcriptomic and metabolomic profiles. This increases the inter-tumor heterogeneity, but is not likely to induce any systematic bias. For paper I, sampling of several biopsies from each patient and analyzing consistent time dependent trends, eliminates some of this variation. The observed transcriptomic responses associated with ischemia time in Paper I also highlight the importance of standardized procedures in handling clinical material for subsequent molecular analysis.

Using xenografts as a model for human breast cancer

Relevant preclinical models of breast cancer are important for studies of breast cancer development, progression and for evaluation of response to treatment. There are several different model systems which can be used, including various breast cancer cell lines, genetically engineered mice and xenograft mouse models [115]. As breast cancer is a heterogeneous disease, one model of breast cancer will never capture all aspects of the disease. Mouse models have a number of advantages compared to cell lines, most notably relevant growth conditions and the possibility of studying cancer *in vivo*. Xenograft models, in comparison to genetically engineered mice, are implanted with human carcinoma cells which make these models more directly relevant to human cancer. Directly grafted orthotopic xenografts, such as the MAS98.06 and MAS98.12 xenograft models used in Paper III and IV, were made by implanting bulk tumor tissue directly into the mouse mammary fat pad. This method has the advantage that the polyclonal nature observed in many breast cancers may be represented, in contrast to injecting a monoclonal cell line. Orthotopic transplantation into the mouse mammary fat pad also provides a more relevant microenvironment than what is achieved by subcutaneous injection. It should still be noted that there are several differences between the mammary glands of the two species, including more fibroblast involvement in human mammary stroma [116]. For the xenograft models used in this work, expression of human ECM related genes observed in the primary tumors were replaced by expression of mouse ECM genes in the xenografts, indicating that mouse stromal cells are recruited by (or invade) the xenografted tumors [85]. Directly grafted orthotopic xenograft models using tumor tissue from patients are challenging to establish, illustrated by the success rate observed when establishing the two models used in this work (2/30 = 6.7 %) and the reported success rate of Marangoni and colleagues (25/200 = 12.5%) [117]. This is a technical drawback of using directly grafted xenograft models, and may introduce a selection bias towards more aggressive cancers. Luminal-like breast cancers are especially difficult to establish, exemplified by the study of Marangoni and colleagues where only one of the 25 established xenografts was ER positive [117]. Using SCID mice is essential for the stable establishment of xenografts, but the lack of immune response may also reduce their relevance as models of human breast cancer.

In spite of these drawbacks, orthotopic xenograft models are more representative model systems than most other models used for preclinical research. As an example of xenograft studies that have been shown to be representative to the study of human cancer, expression of genes identified from lung metastasis in xenograft models were found to predict lung metastasis in a patient cohort [118]. In this work, similarities were observed between basal-like and luminal-like xenografts and corresponding patient tumor subtypes. In paper III, a similar pattern of choline derivatives was observed in the xenografts and patient tumor biopsies. In Paper IV, expression of PtdCho metabolism genes were compared between the xenograft models and a patient cohort. Analyses of differential expression in basal-like compared with luminal-like tumors correlated between the xenografts and human tumors. Such similarities are reassuring for the relevance of these xenograft models as representative of basal-like and luminal-like breast cancers. These models can be used for a large array of studies including preclinical testing of new treatments and studies of metabolism using isotopically labeled compounds and *ex vivo* or *in vivo* MRS in combination with other molecular analyses such as transcriptomics. It should still be noted that the diversity within subgroups of breast cancer are high, and it is therefore essential to include more xenograft models in the future. For example, a luminal-like xenograft model without *TP53* mutation and a less proliferative phenotype would be warranted, as the MAS98.06 may represent a more luminal B type of luminal-like breast cancer.

The future of bevacizumab as an anti breast cancer drug

Based on promising results from preclinical studies, bevacizumab was predicted to be efficient cancer therapy for many different solid tumors including breast cancer. However, the benefit in progression free survival (PFS) has been limited and no improvement in OS has been observed in clinical trials of bevacizumab treatment for breast cancer, even when combined with chemotherapy [119]. In fact, recently FDA withdrew their recommendation of using bevacizumab to treat HER2-negative breast cancer [120]. Resistance to antiangiogenic therapies have been suggested to occur by either adaptive (evasive) resistance or intrinsic non-responsiveness [121]. In Paper IV, the MAS98.12 xenograft model responded by tumor growth delay to bevacizumab as monotherapy and especially

in combination with doxorubicin. However, after about 15 days, a slight increase in tumor volume was observed [109]. This may suggest that the MAS98.12 gains adaptive resistance to bevacizumab. MAS98.06 on the other hand, does not show benefit from bevacizumab as monotherapy or in combination with doxorbucin. This may suggest that the MAS98.06 represents an intrinsic non-responsive tumor. The future of bevacizumab as a breast cancer drug relies on finding robust biomarkers for selecting patients expected to benefit from this treatment, and also for monitoring treatment response and adaptive resistance effects. In Paper IV, GPC was suggested as a potential bevacizumab response biomarker, since it displayed opposite response patterns in the benefiting MAS98.12 and the non-responsive MAS98.06 model. Also, biological effects representing possible treatment resistance mechanisms were studied. Interestingly, calculating the gene expression based GGI for each sample showed that these scores decreased with bevacizumab treatment at day 10, indicating that the tumors were less aggressive after treatment, which is in contradiction to previous suggestions of bevacizumab increasing tumor aggressiveness [121].

Another issue concerning antiangiogenic therapies, is that they have been shown to have different effects with respect to duration of treatment and stage of the disease [122]. Additionally, pro-angiogenic factors other than VEGF have been shown to have higher expression levels in later stage breast cancers, which may suggest that these are more adaptive when treated with bevacizumab than earlier stage breast cancers [123]. However, bevacizumab has mostly been used in clinical trials of advanced breast cancer, which may explain some of the discrepancy between preclinical and clinical studies.

Clinical potential of transcriptomic and metabolomic biomarkers

Using RNA as biomarkers has been questioned due to the unstable nature of RNA. In Paper I, there was not significant evidence of reduction in RNA integrity (as measured by RIN values) with ischemia time up to six hours. However, for some patients, the total RNA from tumor tissue frozen at all time points showed poor RNA quality. This may suggest that RNA degradation in tumor specimens is mostly caused by other factors than ischemia time. Even though total RNA was not significantly degraded with ischemia time, several miRNA

and mRNA transcripts displayed ischemia time associated changes in their expression levels. These transcripts should preferentially be avoided as biomarkers alone or as a part of a prognostic or predictive gene signature.

The gene expression response signatures found in Paper IV in bevacizumab and doxorubicin combination treated basal-like tumors should in the future be correlated to microarray data from clinical trials of bevacizumab for validation. In Paper III and IV, mRNA transcripts coding for enzymes involved in PtCho metabolism were found to be differentially expressed between basal-like and luminal-like cancers. The most consistent finding between xenograft models and patient material was the relatively higher expression of *PLA2G4A* in basal-like than luminal-like tumors (Paper IV), which may explain the higher GPC levels observed both in basal-like xenografts and patient tumors. *PLA2G4A* codes for a phospholipase A2 which catalyzes hydrolysis of PtdCho into acyl-GPC, which can be further hydrolyzed to lyso-GPC and further to GPC and a fatty acid. acyl-GPC is in itself an interesting lipid, and can for example be converted into platelet-activating factor. *PLA2G4A* may thus be an interesting candidate for targeted treatment for patients with basal-like cancers. However, the biological mechanisms of PtdCho metabolism need to be further studied to understand the consequences of targeting *PLA2G4A* in such tumors.

The observed metabolomic phenotypes (Paper I and III) and responses to treatment (Paper IV) could theoretically, if reproduced in larger cohorts, be used in the clinic in the relatively near future. However, before the full potential of metabolomic biomarkers can be utilized, the biological mechanisms and clinical implications of increased levels of a certain metabolite must be better understood. Supplemental functional studies are thus needed to understand the biologic mechanisms behind different levels of a certain biomarker. Metabolites can have many different roles, including serving as building blocks, osmolytes, ligands, or even as second messengers, such as cAMP and PtdCho derivatives. Another unexplored field is that of global metabolite-protein interactions. Large-scale studies in yeast revealed unknown metabolite-protein interactions that were necessary for activity of the proteins [124]. Similar studies in humans may unravel yet unknown regulatory roles of metabolites, that could improve our understanding of why these molecules are abundant in cancer specimens. This could also increase the opportunities of using small molecules

mimicking key metabolites for treatment, which is a strategy that is already widely used in cancer therapy.

Conclusions and future perspectives

In this thesis, breast cancer has been studied at different molecular levels to unravel new knowledge about disease characteristics and treatment response, which may prove valuable for future patient diagnostics and developing new targeted therapies. The importance of stratification according to biologically and clinically relevant subgroups of breast cancer in experimental and clinical studies has been underlined. Transcriptomics and metabolomics have been shown to give complementary information regarding the molecular biology of breast cancer. Both knowledge driven and data driven approaches were used to combine different types of transcriptomic and metabolomic data in this thesis. New hypotheses of which biological processes are associated with higher levels of certain miRNAs and metabolites have been generated by comparing with data on expression of genes with well known functions derived from the same tumor samples or the same xenograft model. However, lack of consistency between gene expression and metabolite levels when studying specific pathways was observed in some cases and this discrepancy underlines that the informative biological signals from both types of data are not always directly associated. Using explorative data driven approaches is therefore necessary to utilize the full potential of combining these complementary data. One such approach was to subgroup patient tumors by both transcriptomic and metabolomic data. The ongoing Oslo II study, in which gene expression microarrays and HR MAS MRS are being performed on breast cancer tissue from several hundred patients will be of value both for validating the findings presented in this thesis as well as exploring new strategies of combining transcriptomic and metabolomic data. One such combinatorial strategy should be to build clinically relevant classifiers based on both types of data. This approach was beyond the scope of this thesis because of the high probability of overfitting data when using small sample sizes for building statistical models. Potential metabolomic and transcriptomic bevacizumab treatment response biomarkers presented in this thesis may in the future be correlated with microarray data obtained from the ongoing clinical trial of bevacizumab in combination with chemotherapy at Oslo University Hospital Radiumhospitalet (NCT00773695). This clinical trial will also give new information that may be used to select the most precise biomarkers for prediction of bevacizumab response and treatment monitoring. In supplement to larger cohorts of human tumors, in-house

breeding of additional xenograft models will be established in the near future, in collaboration with Marangoni and colleagues. This will allow for designed experiments to study molecular effects of new targeted therapies and to study the metabolism of isotopically labeled compounds in tumors representing different subtypes of breast cancer.

Abbreviations

aCGH	array comparative genomic hybridization
CAF	carcinoma associated fibroblast
ChoK	choline kinase
CNA	copy number alteration
CT	phosphate cytidylyltransferase
ECM	extracellular matrix
ER	estrogen receptor
FDG	fluorodeoxyglucose
FEC	fluorouracil, epirubicin and cyclophfosfphamide
FISH	fluorescence in situ hybridization
GC	gas chromatography
GO	Gene Ontology
GPC	glycerophosphocholine
HER2	human epidermal growth factor receptor 2
HK2	hexokinase 2
HR MAS	high resolution magic angle spinning
HRT	hormone replacement therapy
LC	liquid chromatography
LPL	lysophospholipase
miRNA	microRNA
MRI	magnetic resonance imaging
MRS	magnetic resonance spectroscopy
NMR	nuclear magnetic resonance
OS	overall survival
PCho	phosphocholine
PCT	cholinephosphotransferase
PET	positron emission tomography
PFS	progression free survival
PHGDH	phosphoglycerate dehydrogenase
PLA	phospholipase A
PLC	phospholipase C
PLD	phospholipase D
ppm	parts per million
PR	progesterone receptor
PtdCho	phosphatidylcholine
RF	radio frequency
RIN	RNA integrity number
RT qPCR	reverse transcriptase quantitative polymerase chain reaction
SCID	severe combined immunodeficiency
SLC2A1	solute carrier family 2 (facilitated glucose transporter)
T	Tesla
TAM	tumor associated macrophage
tCho	total choline
UTR	untranslated region
VEGF	vascular endothelial growth factor

References

1. Cancer Registry of Norway: Cancer in Norway 2009 - Cancer incidence, mortality, survival and prevalence in Norway. *Oslo: Cancer Registry of Norway* 2011;

2. Wozniak MA & Keely PJ: Use of three-dimensional collagen gels to study mechanotransduction in T47D breast epithelial cells. *Biol Proced Online* 2005; **7**: 144-161

3. Jemal A, Bray F, Center MM *et al.*: Global cancer statistics. *CA Cancer J Clin* 2011; **61**(2): 69-90

4. Elston CW & Ellis IO: Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991; **19**(5): 403-410

5. Curtis HJ: Formal discussion of: somatic mutations and carcinogenesis. *Cancer Res* 1965; **25**(8): 1305-1308

6. Bishop JM: Molecular themes in oncogenesis. *Cell* 1991; **64**(2): 235-248

7. Hanahan D & Weinberg RA: The hallmarks of cancer. *Cell* 2000; **100**(1): 57-70

8. Hanahan D & Weinberg RA: Hallmarks of cancer: the next generation. *Cell* 2011; **144**(5): 646-674

9. Polyak K: Breast cancer: origins and evolution. *J Clin Invest* 2007; **117**(11): 3155-3163

10. Kai K, Arima Y, Kamiya T *et al.*: Breast cancer stem cells. *Breast Cancer* 2010; **17**(2): 80-85

11. Chin L & Gray JW: Translating insights from the cancer genome into clinical practice. *Nature* 2008; **452**(7187): 553-563

12. Teschendorff AE & Caldas C: The breast cancer somatic 'muta-ome': tackling the complexity. *Breast Cancer Res* 2009; **11**(2): 301

13. Dumitrescu RG & Cotarla I: Understanding breast cancer risk -- where do we stand in 2005? *J Cell Mol Med* 2005; **9**(1): 208-221

14. Rubin AF & Green P: Comment on" The Consensus Coding Sequences of Human Breast and Colorectal Cancers". *Science* 2007; **317**(5844): 1500c

15. Wood LD, Parsons DW, Jones S *et al.*: The genomic landscapes of human breast and colorectal cancers. *Science* 2007; **318**(5853): 1108-1113

16. Haber DA & Settleman J: Cancer: drivers and passengers. *Nature* 2007; **446**: 145-146

17. Stephens PJ, McBride DJ, Lin ML *et al.*: Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* 2009; **462**(7276): 1005-1010

18. Chin K, DeVries S, Fridlyand J *et al.*: Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell* 2006; **10**(6): 529-541

19. Wolff AC, Hammond ME, Schwartz JN *et al.*: American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 2007; **25**(1): 118-145

20. Russnes HG, Vollan HK, Lingjaerde OC *et al.*: Genomic architecture characterizes tumor progression paths and fate in breast cancer patients. *Sci Transl Med* 2010; **2**(38): 38ra47

21. Perou CM, Sørlie T, Eisen MB *et al.*: Molecular portraits of human breast tumours. *Nature* 2000; **406**(6797): 747-752

22. Sotiriou C & Pusztai L: Gene-expression signatures in breast cancer. *N Engl J Med* 2009; **360**(8): 790-800

23. van't Veer LJ, Dai H, van de Vijver MJ *et al.*: Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; **415**(6871): 530

24. Paik S, Shak S, Tang G *et al.*: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004; **351**(27): 2817

25. Taft RJ, Pang KC, Mercer TR *et al.*: Non-coding RNAs: regulators of disease. *J Pathol* 2010; **220**(2): 126-139

26. Wang Z, Gerstein M, & Snyder M: RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009; **10**(1): 57-63

27. Sørlie T, Perou CM, Tibshirani R *et al.*: Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001; **98**(19): 10869-10874

28. Sørlie T, Tibshirani R, Parker J *et al.*: Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* 2003; **100**(14): 8418-8423

29. Sørlie T, Borgan E, Myhre S *et al.*: The importance of gene-centring microarray data. *The Lancet Oncology* 2010; **11**(8): 719-720

30. Hu Z, Fan C, Oh DS *et al.*: The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* 2006; **7**(1): 96

31. Parker JS, Mullins M, Cheang MC *et al.*: Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009; **27**(8): 1160-1167

32. Prat A, Parker JS, Karginova O *et al.*: Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* 2010; **12**(5): R68

33. Prat A & Perou CM: Deconstructing the molecular portraits of breast cancer. *Mol Oncol* 2011; **5**(1): 5-23

34. Gupta PB, Fillmore CM, Jiang G *et al.*: Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* 2011; **146**(4): 633-644

35. Tang P, Skinner KA, & Hicks DG: Molecular classification of breast carcinomas by immunohistochemical analysis: are we ready? *Diagn Mol Pathol* 2009; **18**(3): 125-132

36. Lee RC, Feinbaum RL, & Ambros V: The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 1993; **75**(5): 843-854

37. Huntzinger E & Izaurralde E: Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet* 2011; **12**(2): 99-110

38. Thomas M, Lieberman J, & Lal A: Desperately seeking microRNA targets. *Nat Struct Mol Biol* 2010; **17**(10): 1169-1174

39. Rosenfeld N, Aharonov R, Meiri E *et al.*: MicroRNAs accurately identify cancer tissue origin. *Nat Biotechnol* 2008; **26**(4): 462-469

40. Negrini M, Nicoloso MS, & Calin GA: MicroRNAs and cancer--new paradigms in molecular oncology. *Curr Opin Cell Biol* 2009; **21**(3): 470-479

41. O'Day E & Lal A: MicroRNAs and their target gene networks in breast cancer. *Breast Cancer Res* 2010; **12**(2): 201

42. Shi M & Guo N: MicroRNA expression and its implications for the diagnosis and therapeutic strategies of breast cancer. *Cancer Treat Rev* 2009; **35**(4): 328-334

43. Bhowmick NA & Moses HL: Tumor-stroma interactions. *Current opinion in genetics & development* 2005; **15**(1): 97-101

44. Soto AM & Sonnenschein C: The somatic mutation theory of cancer: growing problems with the paradigm? *Bioessays* 2004; **26**(10): 1097-1107

45. Bissell MJ & Hines WC: Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nat Med* 2011; **17**(3): 320-329

46. Folkman J: Tumor angiogenesis: therapeutic implications. *N Engl J Med* 1971; **285**(21): 1182-1186

47. Bergers G & Benjamin LE: Tumorigenesis and the angiogenic switch. *Nat Rev Cancer* 2003; **3**(6): 401-410

48. Ferrara N, Gerber HP, & LeCouter J: The biology of VEGF and its receptors. *Nat Med* 2003; **9**(6): 669-676

49. Thurston G, Suri C, Smith K *et al.*: Leakage-resistant blood vessels in mice transgenically overexpressing angiopoietin-1. *Science* 1999; **286**(5449): 2511-2514

50. Rini BI & Small EJ: Biology and clinical development of vascular endothelial growth factor-targeted therapy in renal cell carcinoma. *J Clin Oncol* 2005; **23**(5): 1028-1043

51. DeBerardinis RJ, Sayed N, Ditsworth D *et al.*: Brick by brick: metabolism and tumor cell growth. *Curr Opin Genet Dev* 2008; **18**(1): 54-61

52. Griffin JL & Shockcor JP: Metabolic profiles of cancer cells. *Nat Rev Cancer* 2004; **4**(7): 551-561

53. Kroemer G & Pouyssegur J: Tumor cell metabolism: cancer's Achilles' heel. *Cancer Cell* 2008; **13**(6): 472-482

54. DeBerardinis RJ: Is cancer a disease of abnormal cellular metabolism? New angles on an old idea. *Genet Med* 2008; **10**(11): 767-777

55. Warburg O: Über den Stoffwechsel der Carcinomzelle. *Naturwissenschaften* 1924; **12**(50): 1131-1137

56. Warburg O: On the origin of cancer cells. *Science* 1956; **123**(3191): 309-314

57. Moreno-Sanchez R, Rodriguez-Enriquez S, Marin-Hernandez A *et al.*: Energy metabolism in tumor cells. *FEBS J* 2007; **274**(6): 1393-1418

58. Kallinowski F, Vaupel P, Runkel S *et al.*: Glucose uptake, lactate release, ketone body turnover, metabolic micromilieu, and pH distributions in human breast cancer xenografts in nude rats. *Cancer Res* 1988; **48**(24 Pt 1): 7264-7272

59. Rivenzon-Segal D, Margalit R, & Degani H: Glycolysis as a metabolic marker in orthotopic breast cancer, monitored by in vivo (13)C MRS. *Am J Physiol Endocrinol Metab* 2002; **283**(4): E623-E630

60. Cheng LL, Chang IW, Smith BL *et al.*: Evaluating human breast ductal carcinomas with high-resolution magic-angle spinning proton magnetic resonance spectroscopy. *J Magn Reson* 1998; **135**(1): 194-202

61. Sitter B, Bathen TF, Singstad TE *et al.*: Quantification of metabolites in breast cancer patients with different clinical prognosis using HR MAS MR spectroscopy. *NMR Biomed* 2010;

62. Grinde MT, Moestue SA, Borgan E *et al.*: (13) C High-resolution-magic angle spinning MRS reveals differences in glucose metabolism between two breast cancer

xenograft models with different gene expression patterns. *NMR Biomed* 2011; DOI: 10.1002/nbm.1683

63. Locasale JW, Grassian AR, Melman T *et al.*: Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. *Nat Genet* 2011; **43**(9): 869-874

64. Sotgia F, Martinez-Outschoorn UE, Pavlides S *et al.*: Understanding the Warburg effect and the prognostic value of stromal caveolin-1 as a marker of a lethal tumor microenvironment. *Breast Cancer Res* 2011; **13**(4): 213

65. Delikatny EJ, Chawla S, Leung DJ *et al.*: MR-visible lipids and the tumor microenvironment. *NMR Biomed* 2011; **24**(6): 592-611

66. Vance JE & Vance DE: Phospholipid biosynthesis in mammalian cells. *Biochem Cell Biol* 2004; **82**(1): 113-128

67. Podo F: Tumour phospholipid metabolism. *NMR Biomed* 1999; **12**(7): 413-439

68. Blusztajn JK: Choline, a vital amine. *Science* 1998; **281**(5378): 794-795

69. Rohlfs EM, Garner SC, Mar MH *et al.*: Glycerophosphocholine and phosphocholine are the major choline metabolites in rat milk. *J Nutr* 1993; **123**(10): 1762-1768

70. Zeisel SH & Blusztajn JK: Choline and human nutrition. *Annu Rev Nutr* 1994; **14**: 269-296

71. Eliyahu G, Kreizman T, & Degani H: Phosphocholine as a biomarker of breast cancer: molecular and biochemical studies. *Int J Cancer* 2007; **120**(8): 1721-1730

72. Glunde K, Jie C, & Bhujwalla ZM: Molecular causes of the aberrant choline phospholipid metabolism in breast cancer. *Canc Res* 2004; **64**(12): 4270-4276

73. Katz-Brull R, Seger D, Rivenson-Segal D *et al.*: Metabolic markers of breast cancer: enhanced choline metabolism and reduced choline-ether-phospholipid synthesis. *Cancer Res* 2002; **62**(7): 1966-1970

74. Podo F, Canevari S, Canese R *et al.*: Tumour Phospholipid Metabolism. *Proc Intl Soc Mag Reson Med* 2011; **19**: ISMRM2011-8444

75. Morales L, Neven P, & Paridaens R: Choosing between an aromatase inhibitor and tamoxifen in the adjuvant setting. *Curr Opin Oncol* 2005; **17**(6): 559-565

76. Baselga J: Targeting tyrosine kinases in cancer: the second wave. *Science* 2006; **312**(5777): 1175-1178

77. Romond EH, Perez EA, Bryant J *et al.*: Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med* 2005; **353**(16): 1673

78. Curigliano G: New drugs for breast cancer subtypes: Targeting driver pathways to overcome resistance. *Cancer Treat Rev* 2011;

79. Udagawa T & Wood M: Tumor-stromal cell interactions and opportunities for therapeutic intervention. *Curr Opin Pharmacol* 2010; **10**(4): 369-374

80. Ellis LM & Hicklin DJ: VEGF-targeted therapy: mechanisms of anti-tumour activity. *Nature Reviews Cancer* 2008; **8**(8): 579-591

81. Muhsin M, Graham J, & Kirkpatrick P: Bevacizumab. *Nature Reviews Drug Discovery* 2004; **3**(12): 995-996

82. Duda DG, Jain RK, & Willett CG: Antiangiogenics: the potential role of integrating this novel treatment modality with chemoradiation for solid cancers. *J Clin Oncol* 2007; **25**(26): 4033-4042

83. Batchelor TT, Sorensen AG, di Tomaso E *et al.*: AZD2171, a pan-VEGF receptor tyrosine kinase inhibitor, normalizes tumor vasculature and alleviates edema in glioblastoma patients. *Cancer Cell* 2007; **11**(1): 83-95

84. Enerly E, Steinfeld I, Kleivi K *et al.*: miRNA-mRNA integrated analysis reveals roles for miRNAs in primary breast tumors. *PLoS One* 2011; **6**(2): e16915 doi:10.1371/journal.pone.0016915

85. Bergamaschi A, Hjortland GO, Triulzi T *et al.*: Molecular profiling and characterization of luminal-like and basal-like in vivo breast cancer xenograft models. *Mol Oncol* 2009; **3**(5-6): 469-482

86. Boetes C, Mus RD, Holland R *et al.*: Breast tumors: comparative accuracy of MR imaging relative to mammography and US for demonstrating extent. *Radiology* 1995; **197**(3): 743-747

87. Lowe IJ: Free induction decays of rotating solids. *Phys Rev Lett* 1959; **2**(7): 285-287

88. Andrew ER & Newing RA: The narrowing of nuclear magnetic resonance spectra by molecular rotation in solids. *Proceedings of the Physical Society* 1958; **72**: 959

89. Sitter B, Sonnewald U, Spraul M *et al.*: High-resolution magic angle spinning MRS of breast cancer tissue. *NMR Biomed* 2002; **15**(5): 327-337

90. Petersen SB, Muller RN, & Rinck PA: An introduction to biomedical nuclear magnetic resonance. 1985;

91. Glunde K & Bhujwalla ZM: Metabolic tumor imaging using magnetic resonance spectroscopy. *Semin Oncol* 2011; **38**(1): 26-41

92. Smyth GK: Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004; **3**: Article3

93. Benjamini Y & Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)* 1995; **57**(1): 289-300

94. Navon R, Wang H, Steinfeld I *et al.*: Novel rank-based statistical methods reveal microRNAs with differential expression in multiple cancer types. *PLoS One* 2009; **4**(11): e8003

95. Ashburner M, Ball CA, Blake JA *et al.*: Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; **25**(1): 25-29

96. Betel D, Koppal A, Agius P *et al.*: Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol* 2010; **11**(8): R90

97. Subramanian A, Tamayo P, Mootha VK *et al.*: Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005; **102**(43): 15545-15550

98. Eden E, Lipson D, Yogev S *et al.*: Discovering motifs in ranked lists of DNA sequences. *PLoS Comput Biol* 2007; **3**(3): e39

99. Eden E, Navon R, Steinfeld I *et al.*: GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 2009; **10**(1): 48

100. Draghici S, Khatri P, Eklund AC *et al.*: Reliability and reproducibility issues in DNA microarray measurements. *Trends Genet* 2006; **22**(2): 101-109

101. Malone JH & Oliver B: Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol* 2011; **9**: 34

102. Taub MA, Corrada BH, & Irizarry RA: Overcoming bias and systematic errors in next generation sequencing data. *Genome Med* 2010; **2**(12): 87

103. Willenbrock H, Salomon J, Sokilde R *et al.*: Quantitative miRNA expression analysis: comparing microarrays with next-generation sequencing. *RNA* 2009; **15**(11): 2028-2034

104. Griffin JL, Lehtimaki KK, Valonen PK *et al.*: Assignment of 1H nuclear magnetic resonance visible polyunsaturated fatty acids in BT4C gliomas undergoing ganciclovir-thymidine kinase gene therapy-induced programmed cell death. *Cancer Res* 2003; **63**(12): 3195-3201

105. Morse DL, Raghunand N, Sadarangani P *et al.*: Response of choline metabolites to docetaxel therapy is quantified in vivo by localized (31)P MRS of human breast cancer xenografts and in vitro by high-resolution (31)P NMR spectroscopy of cell extracts. *Magn Reson Med* 2007; **58**(2): 270-280

106. Golman K, Zandt RI, Lerche M *et al.*: Metabolic imaging by hyperpolarized 13C magnetic resonance imaging for in vivo tumor diagnosis. *Cancer Res* 2006; **66**(22): 10855-10860

107. Morse DL, Carroll D, Day S *et al.*: Characterization of breast cancers and therapy response by MRS and quantitative gene expression profiling in the choline pathway. *NMR Biomed* 2009; **22**(1): 114-127

108. Tzika AA, Astrakas L, Cao H *et al.*: Combination of high-resolution magic angle spinning proton magnetic resonance spectroscopy and microscale genomics to type brain tumor biopsies. *Int J Mol Med* 2007; **20**(2): 199-208

109. Lindholm E, Kristian A, Nalwoga H *et al.*: Effect of antiangiogenic therapy on tumor growth, vasculature and kinase activity in basal- and luminal-like breast cancer xenografts. *Submitted to Br J Cancer* 2011;

110. Bell JD, Brown JCC, & Sadler PJ: NMR studies of body fluids. *NMR Biomed* 1989; **2**(5 6): 246-256

111. Komaki K, Sano N, & Tangoku A: Problems in histological grading of malignancy and its clinical significance in patients with operable breast cancer. *Breast Cancer* 2006; **13**(3): 249-253

112. Teixeira MR, Pandis N, Bardi G *et al.*: Clonal heterogeneity in breast cancer: karyotypic comparisons of multiple intra- and extra-tumorous samples from 3 patients. *Int J Cancer* 1995; **63**(1): 63-68

113. Navin N, Krasnitz A, Rodgers L *et al.*: Inferring tumor progression from genomic heterogeneity. *Genome Res* 2010; **20**(1): 68-80

114. Navin N, Kendall J, Troge J *et al.*: Tumour evolution inferred by single-cell sequencing. *Nature* 2011; **472**(7341): 90-94

115. Vargo-Gogola T & Rosen JM: Modelling breast cancer: one size does not fit all. *Nat Rev Cancer* 2007; **7**(9): 659-672

116. Hovey RC, McFadden TB, & Akers RM: Regulation of mammary gland growth and morphogenesis by the mammary fat pad: a species comparison. *J Mammary Gland Biol Neoplasia* 1999; **4**(1): 53-68

117. Marangoni E, Vincent-Salomon A, Auger N *et al.*: A new model of patient tumor-derived breast cancer xenografts for preclinical assays. *Clin Cancer Res* 2007; **13**(13): 3989-3998

118. Minn AJ, Gupta GP, Siegel PM *et al.*: Genes that mediate breast cancer metastasis to lung. *Nature* 2005; **436**(7050): 518-524

119. Burstein HJ: Bevacizumab for advanced breast cancer: all tied up with a RIBBON? *J Clin Oncol* 2011; **29**(10): 1232-1235

120. Lenzer J: FDA committee votes to withdraw bevacizumab for breast cancer. *BMJ* 2011; **343**: d4244

121. Bergers G & Hanahan D: Modes of resistance to anti-angiogenic therapy. *Nat Rev Cancer* 2008; **8**(8): 592-603

122. Ebos JM, Lee CR, Cruz-Munoz W *et al.*: Accelerated metastasis after short-term treatment with a potent inhibitor of tumor angiogenesis. *Cancer Cell* 2009; **15**(3): 232-239

123. Relf M, LeJeune S, Scott PA *et al.*: Expression of the angiogenic factors vascular endothelial cell growth factor, acidic and basic fibroblast growth factor, tumor growth factor beta-1, platelet-derived endothelial cell growth factor, placenta growth factor, and pleiotrophin in human primary breast cancer and its relation to angiogenesis. *Cancer Res* 1997; **57**(5): 963-969

124. Li X, Gianoulis TA, Yip KY *et al.*: Extensive in vivo metabolite-protein interactions revealed by large-scale systematic analyses. *Cell* 2010; **143**(4): 639-650

Errata

Changes in the text are marked in bold.

- All headings are changed to black font to reduce costs when printing
- Page 4:
  Line 5: "miRNA" → "mi**cro**RNA **(miRNA)**"
- Page 14:
  Line 7: "genes which are **gene** involved" → "genes which are involved"
- Page 23:
  Line 16: "Choline is" → "Choline **(Cho)** is"
  Line 24: "Cho**line** enters the cell" → "Cho enters the cell"
  Line 29: "lysophospholipase (LPL) phosphodiesterase (PD)" → "lysophospholipase (LPL) **and** phosphodiesterase (PD)"
- Page 29:
  Table 1: Format table layout to greytones
- Page 30:
  Line 13 :"probes on the array" → "probes on the **gene expression** array"
- Page 31:
  Line 2: "on th**is** array" → "on th**e miRNA** array"
  Line 3: "RNA is**s**" → "RNA is"
- Page 32 :
  Line 8-9: "spinning in HR MAS MR" → "spinning in HR MAS MR**S**"
  Figure 13 legend: "magnetic field, B_0." → "magnetic field, B_0**, in HR MAS MRS**."
- Page 33:
  Figure 14 legend: "and GPC (green) **as shown at the**. " → "and GPC (green)."
- Page 34:
  Line 10-11: "such metabolite concentration" → "such **as** metabolite concentration"
  Line 12: "Paper I) **or**." → "Paper I)."
  Line 15-17: **"ANOVA may be applied to any number of groups; in particular it may be applied to the special case of two groups and was used for the pairwise comparisons in Paper IV."** → "" (Delete whole sentence because this is not true for Paper IV)
  Line 17: "in this paper" → "in **Paper IV**"
- Page 35:
  Line 16: "**Hierarchical clustering was performed in all papers."** → ""
- Page 37:
  Line 9: "mi**cro**RNA" → "miRNA"
- Page 39:
  Line 9-10: "high PCho and low GPC levels" → "**relatively** high**er** PCho and low**er** GPC levels"
  Line 20-21: "(*CHDH*)**;** and sarcosine dehydrogenase (*SARDH*)" → "(*CHDH*) and sarcosine dehydrogenase (*SARDH*) **levels**"
- Page 49:
  Line 8: "dox**ro**bucin" → " dox**or**bucin"

I

II

Borgan et al. BMC Cancer 2010, 10:628
http://www.biomedcentral.com/1471-2407/10/628

BMC Cancer

RESEARCH ARTICLE Open Access

Merging transcriptomics and metabolomics - advances in breast cancer profiling

Eldrid Borgan[1,2,5†], Beathe Sitter[2†], Ole Christian Lingjærde[3], Hilde Johnsen[1], Steinar Lundgren[2,4], Tone F Bathen[2], Therese Sørlie[1,3], Anne-Lise Børresen-Dale[1,5], Ingrid S Gribbestad[2*]

Abstract

Background: Combining gene expression microarrays and high resolution magic angle spinning magnetic resonance spectroscopy (HR MAS MRS) of the same tissue samples enables comparison of the transcriptional and metabolic profiles of breast cancer. The aim of this study was to explore the potential of combining these two different types of information.

Methods: Breast cancer tissue from 46 patients was analyzed by HR MAS MRS followed by gene expression microarrays. Two strategies were used to combine the gene expression and metabolic data; first using multivariate analyses to identify different groups based on gene expression and metabolic data; second correlating levels of specific metabolites to transcripts to suggest new hypotheses of connections between metabolite levels and the underlying biological processes. A parallel study was designed to address experimental issues of combining microarrays and HR MAS MRS.

Results: In the first strategy, using the microarray data and previously reported molecular classification methods, the majority of samples were classified as luminal A. Three subgroups of luminal A tumors were identified based on hierarchical clustering of the HR MAS MR spectra. The samples in one of the subgroups, designated A2, showed significantly lower glucose and higher alanine levels than the other luminal A samples, suggesting a higher glycolytic activity in these tumors. This group was also enriched for genes annotated with Gene Ontology (GO) terms related to cell cycle and DNA repair. In the second strategy, the correlations between concentrations of *myo*-inositol, glycine, taurine, glycerophosphocholine, phosphocholine, choline and creatine and all transcripts in the filtered microarray data were investigated. GO-terms related to the extracellular matrix were enriched among the genes that correlated the most to *myo*-inositol and taurine, while cell cycle related GO-terms were enriched for the genes that correlated the most to choline. Additionally, a subset of transcripts was identified to have slightly altered expression after HR MAS MRS and was therefore removed from all other analyses.

Conclusions: Combining transcriptional and metabolic data from the same breast carcinoma sample is feasible and may contribute to a more refined subclassification of breast cancers as well as reveal relations between metabolic and transcriptional levels.
See Commentary: http://www.biomedcentral.com/1741-7015/8/73

Background

Transcriptomics and metabolomics in cancer research have traditionally been considered as two separate fields. Different levels of the molecular processes are studied, aiming at improving cancer treatment by understanding the underlying mechanisms of the disease. Breast cancer treatment decisions today are mainly based on tumor size, histological characterization, grading and receptor status, as well as axillary lymph node status, and age of the patients [1]. However, patients with similar diagnosis and treatment can experience large differences in the progression and relapse of their disease. The various -omics fields, transcriptomics in particular, have provided an understanding of breast cancer as a group of molecularly distinct neoplastic disorders [2]. Clinical use

* Correspondence: ingrid.s.gribbestad@ntnu.no
† Contributed equally
[2]Department of Circulation and Medical Imaging, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
Full list of author information is available at the end of the article

BioMed Central

© 2010 Borgan et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

of molecular characterization of breast cancer has the potential to stratify breast cancer patients for more individual treatment, but has so far only been implemented to a limited extent.

The field of transcriptomics, using DNA microarrays that enable measurements of thousands of RNA transcripts in a single experiment, has had a huge impact on breast cancer research over the last decade [2]. One of the important findings has been the classification of breast cancer into five subtypes (luminal A, luminal B, basal-like, ERBB2 enriched and normal-like) based on gene expression profiles of so called intrinsic genes [3,4]. This molecular subtyping of breast cancer, has been reproduced in several studies and is also associated with clinical outcome across datasets [5,6].

Metabolomics studies the metabolites and how they are affected by specific cellular processes. The possibility of using *in vivo* magnetic resonance spectroscopy (MRS) as a diagnostic, prognostic or predictive tool in the clinic simultaneously with an MR imaging (MRI) examination, makes MRS techniques attractive methods for molecular classification of disease. Metabolic profiling of intact biological samples using high resolution magic angle spinning magnetic resonance spectroscopy (HR MAS MRS) enables measurement of multiple cellular metabolites simultaneously. The method has been utilized in a wide range of biological applications [7], and studies of cancers have proven HR MAS MRS to be a promising tool in cancer diagnosis and treatment monitoring [8]. Importantly, the sample is kept intact throughout the HR MAS MRS analysis and can subsequently be analyzed by gene expression analysis.

Profiling gene expression and metabolite content in the same breast carcinoma samples enables comparisons of molecular findings at different levels. Gene expression data and metabolite data from MRS techniques of different samples from the same breast cancer cell line or xenograft model have been combined previously [9-12], but these studies have mainly focused on specific genes involved in choline metabolism, known as the Kennedy pathway. Combining transcriptomic and metabolomic profiling of the same sample allow us to capture a comprehensive picture at a given moment in time. Such studies could reveal differences and similarities between groups of samples at different molecular levels and provide a fundament for enhanced knowledge of the biological dynamics of breast cancers.

The aim of this study was to combine gene expression microarrays and HR MAS MRS for more refined profiling of breast tumors, and to explore some of the potentials and limitations of the experimental procedures. This study focuses on the most prevalent type of breast cancer, invasive ductal carcinomas (IDC) with oestrogen (ER) receptor positive disease, and the largest molecular subgroup within these tumors, luminal A. ER positive breast cancer accounts for approximately 2/3 of the cases, and although they overall have a relatively good prognosis, some patients experience relapse and do not respond to antieostrogen treatment. So far there are no biomarkers available to identify those patients and no targets for therapy, and identification of molecular markers of such tumors, possibly by metabolic profiling, are therefore of high relevance. Identifying subgroups of patients within this group is an important goal to make it possible to further individualize cancer diagnosis and treatment.

Methods

Patients

Tissue samples were selected from a local breast cancer tissue bank, obtained from patients with palpable breast lesions who underwent scheduled surgery for breast cancer at St. Olavs University Hospital in Trondheim, Norway (March 2000 - March 2004). Samples were placed in cryogenic vials and immersed in liquid nitrogen immediately after dissection. In this study we selected samples from patients diagnosed with IDC, who did not receive any treatment prior to surgery. Clinical data of the patients, such as diagnosis, tumor grade, tumor size, hormone receptor status and lymph node involvement were obtained from patient records, including pathology reports, and are summarized in Table 1. The Regional Committee for Medical and Health Research Ethics approved the study protocol, and all patients provided written informed consent.

Sample preparation and experimental design

An illustration of the experimental workflow as well as the data analysis for this study is shown in Figure 1. Tumor tissue samples from 46 patients were cut from each tumor and analyzed by HR MAS MRS. Proceeding HR MAS MRS, the tissue was directly snap frozen and later used to extract RNA for microarray analysis. For 11 of the tissue samples, the quality of RNA obtained after HR MAS MRS analysis was of insufficient quality, and adjacent tissue from the same tumors, not subjected to HR MAS MRS, was used for RNA extraction and microarray analysis. Additionally, histopathology was performed on adjacent tumor tissue for each sample, as described previously [13], to assess the percentage of cancer cells in each sample. The samples contained on average 23% tumor cells with a range of 0-80%.

HR MAS MRS experiments

Sample preparation and HR MAS MRS experiments were performed as previously described [13]. In brief, all samples were cut to fit a 4 mm o.d. rotor with inserts (total sample volume 50 μL). Samples analyzed by HR

Table 1 Clinical data on the 46 patients included in the main study

Characteristics	Summary
Age, mean (range) years	64 (30 - 91)
Tumor size, mean (sd) (cm)	2.5 (± 1.4)
Grade	
I	5
II	17
III	21
NA	3
Axillary lymph node status	
Negative	22
Positive	20
NA	4
Oestrogen receptor (ER) status	
Negative	5
Positive	41
Progesterone receptor (PgR) status	
Negative	9
Positive	36
NA	1

All patients were diagnosed with invasive ductal carcinoma (IDC). All characteristics, except for age and tumor size, are summarized by number of patients in each group. ER and PgR were determined by immunohistochemistry, and samples with ≥10% staining cancer cells were considered receptor positive.

MAS MRS weighed 18.6 mg in average (7.7 to 27.1 mg). The remaining rotor volume was filled with buffer (phosphate buffered saline (PBS) in D_2O containing trimethylsilyl tetradeuteropropionic acid (TSP)). HR MAS MRS experiments were performed on a Bruker Avance DRX600 spectrometer equipped with a $^{1}H/^{13}C$ MAS probe (Bruker BioSpin GmbH, Germany). Samples were spun at 5 kHz while keeping the temperature at 4°C. Two types of one-dimensional ^{1}H spectra were recorded, a single pulse experiment with water suppression and a spin-echo experiment (CPMG) for suppression of broad signals from macromolecules. Spectral assignments were performed based on previous HR MAS MRS studies of breast cancer lesions [13,14].

RNA extraction and microarrays

Total RNA was extracted from the biopsies that were used in the HR MAS MRS procedure from each patient, following the protocol for the RNeasy Mini Kit (Qiagen, USA). High quality RNA, measured by BioAnalyzer 2100 (Agilent Technologies, USA) was used in microarray experiments, using 44k two-color Agilent Human Whole Genome Oligo Microarrays. The microarray analysis was performed according to the manufacturer's protocol, using Cy5 labeling for the tumor RNA and Cy3 for the reference RNA, Universal Human Reference (Stratagene, USA). Data were extracted from the microarray images using Feature Extraction Software v8.5 (Agilent Technologies, USA). The final dataset included gene expression microarray data from 46 patients, 35 obtained from RNA extracted from the HR MAS MRS analyzed tissue and 11 from adjacent tumor tissue, as illustrated in Figure 1.

Preprocessing and normalizing the microarray data

The two-color microarray data were preprocessed and analyzed in R (v2.8) and Bioconductor, using the Feature Extraction text files. The red (sample) and green (reference) signals were set to be g(r)MeanSignal, and the data were normalized using the Limma package [15]. Each microarray was background corrected by subtracting the g(r)SpatialDetrendSurfaceValue values, and normalized with loess normalization. Normalization between arrays was performed with Gquantile normalization, which ensures the same empirical distribution for the green channel and adjusts the red channel accordingly. The data were log2 transformed and missing data were imputed using local least squares imputation from the pcaMethods package [16]. Control probes and probes which were flagged as outliers on more than 20% of the arrays were removed. The average of duplicate probes was calculated, and the averaged probe value with the highest IQR was picked to represent each transcript, when the gene symbol (supplied by Agilent) was represented by different probes. A subset of 1199 probes that were identified to have an altered gene expression after HR MAS MRS in a parallel study described below were also removed.

Subtyping with microarray and HR MAS MRS data

Using gene median centered microarray data, the expression of "intrinsic" genes of the tumors from all 46 patients were analyzed for Spearman's correlation to published centroids of the five intrinsic subtypes [17]. The samples were classified according to the highest correlation value to the centroids. Prior to subtyping, probes of bad quality and 9 transcripts identified to show altered gene expression after HR MAS MRS were removed from the centroids, resulting in a match of 286 out of 306 intrinsic genes. The threshold for classification into a subtype was set to be correlation above 0.1 (One sample showed the same correlation to two subtypes and was regarded as unclassified). The high dimensional spin echo spectra (spectral region 4.8 - 1.45 ppm, excluding lipid containing regions at 2.8 ppm and 2.05 ppm, described by 5993 data points) of the 31 tumors classified as luminal A and with tumor cell percentage above 5%, were used to assess if adding the metabolic data revealed additional structure within this group of tumors. Normalization of the selected spectral regions was performed by mean normalization.

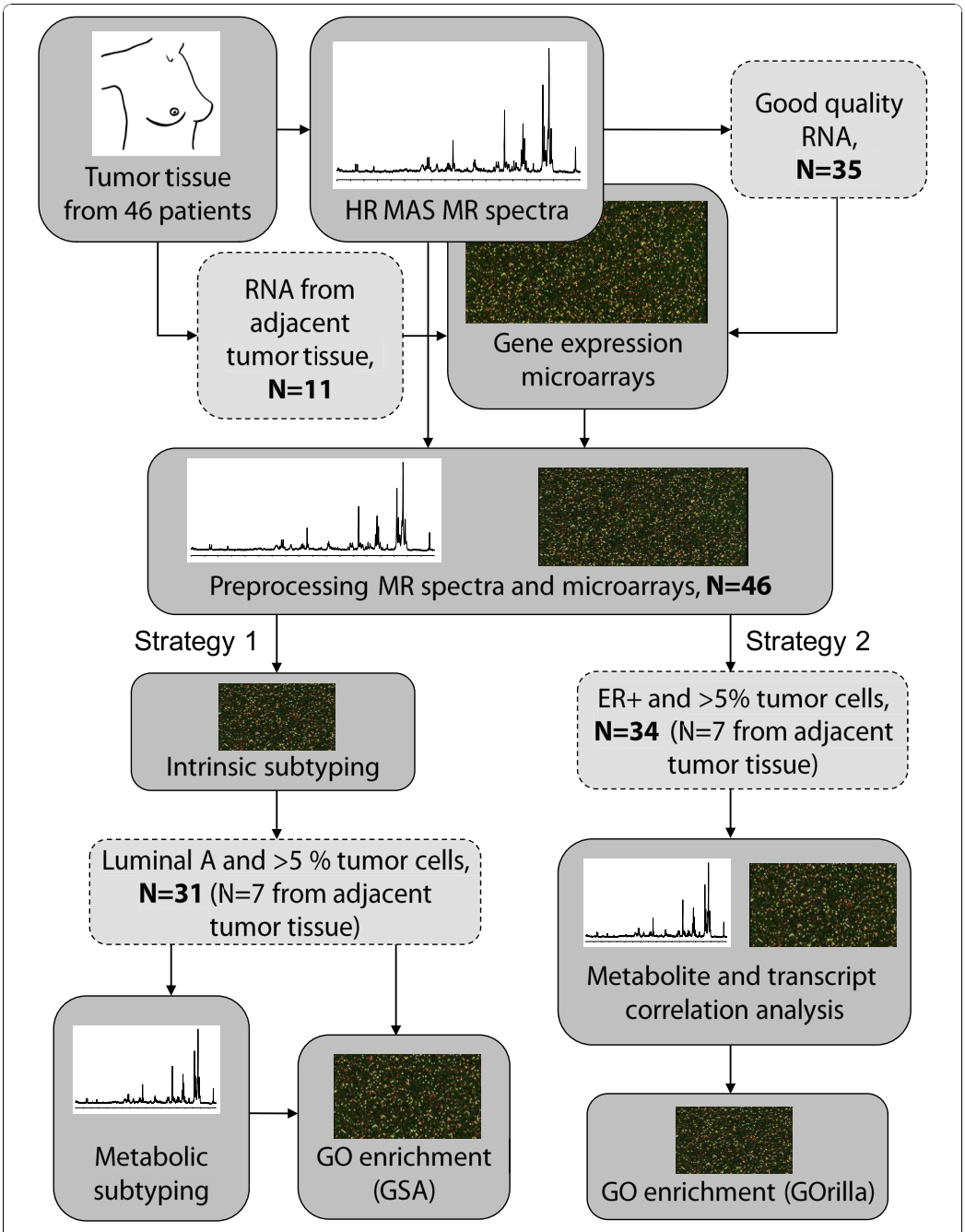


Figure 1 An illustration of the experimental and data analysis workflow. Experimental design and the two strategies of combining microarray and HR MAS MRS data are illustrated schematically.

Unsupervised hierarchical clustering with complete linkage was performed using Spearman's correlation as similarity measure, identifying three luminal A groups by cutting the dendrogram at height 0.39, which was decided as a threshold that gave convincing separation between groups. Multidimensional scaling (MDS) using Spearman's correlation as similarity measure was also used to assess if the same samples grouped together when using a different multivariate analysis. To test for metabolic differences between the three groups, t-tests were performed between pairs of points in the spectra. Differential gene expression analysis between the three identified luminal A groups was performed using the Limma package (R/Bioconductor) [15]. Modified t-tests were performed on each gene, with Empirical Bayes correction of the test statistics and Benjamini and Hochberg adjusted p-values. To assess if sets of genes known to be involved in the same process, share the same function or location were concordantly differentially expressed between the three luminal A groups, Gene Ontology (GO) [18] enrichment analyses were performed using Gene Set Analysis (GSA) [19]. GSA is a refined version of the original Gene Set Enrichment Analysis, which can be used to test if sets of genes are differentially expressed between two groups, using a maxmean statistic and permutation analysis to calculate the p-values. The GSA analysis was run with 10000 permutations, and the raw p-values were used. The Benjamini Hochberg false discovery rate was calculated separately.

Correlation analysis between metabolite concentration and gene expression

Microarray and HR MAS MRS data from the same samples of the ER+ tumors with tumor cell percentage > 5% (N = 34) were used to test for nonzero correlations between metabolite and gene expression levels (For seven of the samples, the microarray data were obtained from adjacent tumor tissue). Quantification of eight tissue metabolites was performed based on peak areas in the single-pulse spectra and TSP concentrations as in [13]. After preprocessing the HR MAS MR spectra (WINNMR, Bruker Biospin GmbH, Germany), peak areas were calculated for the signals arising from *myo*-inositol, taurine, choline, GPC, PCho, glucose, creatine and glycine using the Voigt (combined Lorentzian and Gaussian) area function in PeakFit (Seasolve, MA USA). Quantification of lactate was not performed since lipid signals concealed the lactate signals (signals from lipid -CH_2 overlap with lactate CH_3 and glycerol backbone of triglycerides overlap with lactate CH) in two thirds of the spectra. Probes on the microarray with IQR < 0.8 were removed, and Spearman correlation tests between the 8 metabolites and all of the 13875 unique probes were performed. For each metabolite, all transcripts were ranked according to Spearman's rank correlation coefficient. The resulting genelists were analyzed for enriched GO-terms using GOrilla, which searches for GO-terms that are enriched in the top of the list compared to the rest of the list using the mHG statistics [20,21].

Assessing the effect of HR MAS MRS on RNA integrity and transcription

No apparent differences were observed between the microarray data from the 35 samples that did undergo HR MAS MRS compared to the 11 samples that did not. However, to explore this issue more thoroughly, an additional study was designed to assess the feasibility of performing gene expression microarrays and HR MAS MRS on the same sample as well as identify transcripts that should be filtered out from microarray analyses if they have been systematically affected by the HR MAS MRS procedure. This additional study included 18 pairs of tumor samples, obtained from the local breast cancer tissue bank described above. One sample from each pair was analyzed by HR MAS MRS, RNA was isolated from both samples from each of the 18 pairs, and gene expression microarray analysis was performed. The HR MAS MRS experiments were performed as described above, while the RNA extraction, microarray experiments and preprocessing of the microarray data were performed using slightly different protocols, as described in Additional file 1: Supplementary documentation of exploring the effect of HR MAS MRS on RNA integrity and transcription. The effect of HR MAS MRS on RNA integrity was tested using a paired t-test on the RIN-values measured using Bioanalyzer 2100 (Agilent Technologies). Unsupervised hierarchical clustering of the normalized microarray data from the 18 pairs of samples and modified t-tests were performed [15] to test for differential gene expression caused by the HR MAS MRS procedure. The transcripts that showed significantly higher or lower (fdr < 0.01) expression in the samples analyzed by HR MAS MRS, were tested for enriched GO-terms [22]. Transcripts identified as having significantly altered expression after HR MAS MRS were removed from the microarray data in the main study. A more detailed description of the data analysis from this additional study can be found in Additional file 1: Supplementary documentation of exploring the effect of HR MAS MRS on RNA integrity and transcription.

Results

Two main strategies of combining information from the microarray data with the HR MAS MR spectra are presented, as shown in the workflow illustrated in Figure 1. The first strategy uses multidimensional gene expression

microarray data and the HR MAS MR spectra to identify molecular subgroups of tumors. The second strategy compares the quantified concentration of specific metabolites to the expression level of each transcript in the filtered microarray data. In a parallel study, we performed additional experiments to assess the impact of sample treatment during HR MAS MRS on RNA integrity and gene expression.

Transcriptional and metabolic subtyping based on high-dimensional data

Each of the 46 tumor samples (Figure 1), were subtyped according to the highest Spearman's correlation of the expression of intrinsic genes to published centroids from the five subtypes [17]. This subtyping method resulted in the tumors being classified into 36 luminal A, 7 normal-like, 1 ERBB2 enriched, 1 basal-like, and 1 unclassified. The samples classified as luminal A were all ER positive. The ER-negative samples were classified as 1 basal-like and 4 normal-like. The HR MAS MR spectra from the tumors classified as luminal A (n = 31 with tumor cell percentage >5%), were used to search for diversity within this subgroup in the metabolic data. Hierarchical clustering based on the HR MAS MR spin echo spectra revealed three clusters that were convincingly separated when selecting a threshold of the dendrogram at height 0.39 (Figure 2). The samples belonging to each of these three subgroups of luminal A samples, designated A1, A2 and A3, also clustered together when applying MDS to the HR MAS MR spectra (Figure 3). Differences in the metabolic and gene expression profiles between the A1-A3 groups of luminal A tumors are illustrated in Figure 4. The metabolites allocated at the points that significantly differed between the profiles ($p < 0.001$) include α-glucose, β-glucose, amino acids (signals from α-H which is the Hydrogen bonded to the α-C of all amino acids), *myo*-inositol, alanine and lipid residues [14]. Glucose signals were significantly lower in the A2 group compared to the A1 and A3 luminal tumors. The α-H amino acid signals were significantly lower in A1 and higher in A3, compared to A2. Alanine signals were significantly higher in A2 compared to A3. Signals from lipid residues were significantly higher and signals from *myo*-inositol were significantly lower in A1 than in A2 and A3. Differential gene expression analysis between the three luminal A groups, using Limma, resulted in no significant differences in expression of single genes with fdr < 0.01. However, when analyzing differences in gene expression between sets of genes annotated with the same GO-terms, using GSA, different significantly enriched GO-terms (fdr < 0.01) were revealed by comparing the three luminal A groups (Figure 4). The A2 tumors were enriched for biological processes related to the cell cycle and DNA repair, namely "Meiosis I", "Meiotic recombination" and "Double strand break repair" compared to the A1 tumors. The A3 group was enriched for the molecular function "Protein C terminus binding" compared to the A2 group. There was not significant evidence of differences in tumor percentage between the three groups.

Correlating metabolic and transcriptional profiles

The levels of 8 specific metabolites were compared with the transcriptional activity in each tumor in the ER+ samples (n = 34 with tumor cell percentage >5%). These samples include the 31 luminal A samples mentioned above plus three samples that were classified as normal-like, but also correlated to the luminal A centroid. Tissue concentrations of some of the 8 quantified metabolites co-varied, see Additional file 2: Scatterplot of metabolite concentrations and tumor percentage. Taurine, *myo*-inositol and choline correlated significantly with each other, as did creatine, PCho, GPC and glycine. Glucose showed a negative or no significant correlation to all the other metabolites. Lists of all transcripts on the microarray with IQR > 0.8 (13875 unique probes) were ranked according to correlation, from high to low, to each of the 8 metabolites. Plots showing the co-variation of *myo*-inositol and choline concentration with the expression of the transcripts that correlated the most to these metabolites are shown in Figure 5. The ranked lists of transcripts were used to test for GO enrichment towards the top of the list, using GOrilla [20,21]. Significantly enriched (fdr < 0.003) GO-terms in the biological process (BP), cellular component (CC) and molecular function (MF) GO-categories are given for each metabolite in Table 2. GO-terms related to the extracellular matrix are enriched towards the top of the lists of transcripts ranked according to correlation to *myo*-inositol and taurine. GO-terms related to the cell cycle, such as "cell cycle process" and "chromosome segregation" are enriched towards the top of the list of transcripts ranked according to correlation to choline. The lists of transcripts ranked according to correlation to glucose, creatine and glycine only gave one enriched GO-term each, "immune system process", "mannosyltransferase activity" and "respiratory chain", respectively. There were no significantly enriched GO-terms for the lists of transcripts ranked according to correlation to GPC and PCho.

Effect of HR MAS MRS on RNA quality and gene expression

The effect of the HR MAS MRS procedure on RNA integrity and transcription in the breast cancer tissue was assessed separately on 18 pairs of samples. There was not significant evidence that total RNA integrity, quantified by the RIN-value (measured with Bioanalyzer

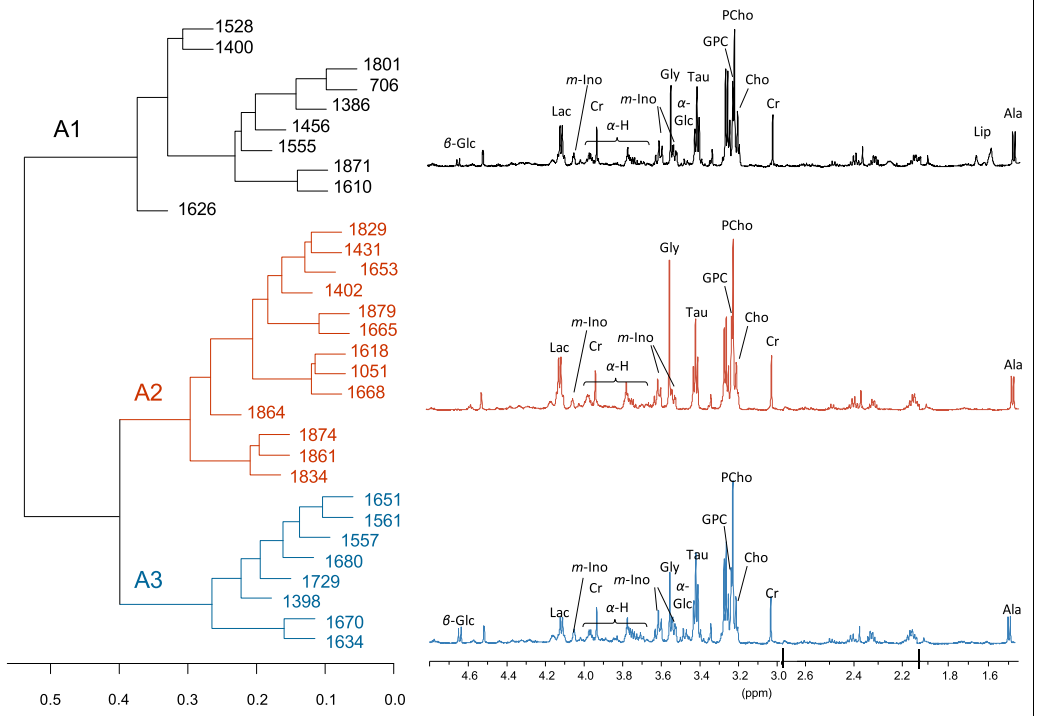


Figure 2 Metabolic subtyping of luminal A samples using hierarchical clustering. The HR MAS MR spectra of the samples classified as luminal A (N = 31), cluster in three groups using hierarchical clustering with Spearman's correlation as the distance measure and complete linkage. The mean spectra of each luminal A group are shown. Selected metabolite allocations are indicated with abbreviations (β-Glc: β-glucose; Lac: lactate; Cr: creatine; *m*-Ino: *myo*-inositol; α-H: α-H of amino acids; Gly: glycine; α-Glc: α-glucose; Tau: taurine; Cho: choline; GPC: glycerophosphocholine; PCho: phosphocholine; Ala: alanine). The spectral chemical shift-scale is non-continuous due to deletion of signals from lipid residues at 2.8 ppm and 2.05 ppm.

2100), was affected by HR MAS MRS (p-value = 0.86). The results from the microarray analysis are illustrated in Additional file 3: Plots illustrating the effect of HR MAS MRS on the transcriptome. Hierarchical clustering of the microarray data show that the samples cluster in pairs. Using a false discovery rate of 0.01, 1199 transcripts were significantly differentially expressed in samples that had been subjected to the HR MAS MRS procedure (865 lower and 334 higher, detailed results in Additional file 4: Significantly lower expressed transcripts after HR MAS MRS and Additional file 5: Significantly higher expressed transcripts after HR MAS MRS). Only 15 of the 1199 transcripts have an estimated log2 fold change larger than 1. Among the 1199 transcripts, several GO-terms were enriched, like "RNA metabolic process" and "Regulation of gene expression" for the transcripts higher expressed after HR MAS MRS and "Protein localization", "Vesicle mediated transport" and "Tricarboxylic acid cycle" for the transcripts lower expressed after HR MAS MRS. The 1199 significantly differentially expressed transcripts were removed from the microarray data of the 46 samples in the main study.

Discussion

In this study, we have shown the feasibility of merging transcriptomics and metabolomics data from the very same tumor tissue sample. Two strategies of combining microarray data and HR MAS MR spectra are presented, providing a framework for how information from these different molecular levels can be combined

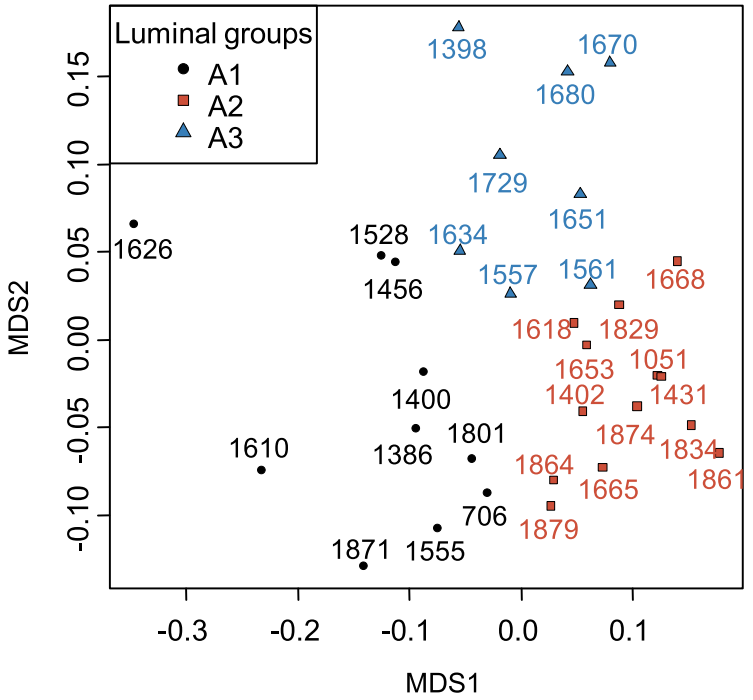
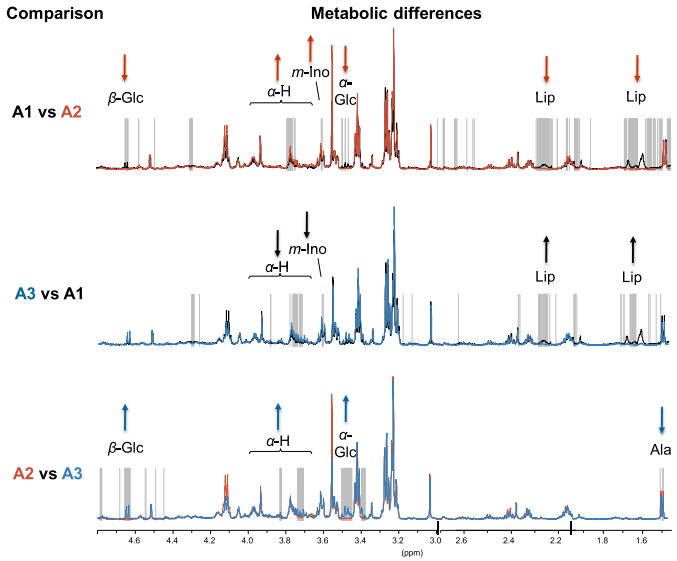


Figure 3 Multidimensional scaling of the HR MAS MR spectra of the luminal A samples. Multidimensional scaling of the HR MAS MR spectra of 31 luminal A samples, colored according to the A1-A3 subgroups found by hierarchical clustering.

and analyzed. We also identified a set of transcripts which showed slightly altered expression after the HR MAS MRS procedure, but overall this variation was smaller than the biological variation in tumors from one patient to another.

In the first strategy to combine gene expression microarray data and HR MAS MR spectra, the expression of "intrinsic" genes was used to classify the samples into established molecular subtypes. The HR MAS MR spectra of the majority of tumor samples, which were classified as luminal A, were further explored to investigate whether metabolic characteristics could define subgroups within a transcriptionally homogenous set of samples. The use of spin echo acquired spectral profiles ensured a more extensive use of metabolic information than using calculated tissue metabolite concentrations, which is limited by several peak areas being non-quantifiable. Three subgroups of luminal A tumors were identified (Figure 2 and 3). The fact that samples cluster together differently with respect to the transcriptional and metabolic profiles (results not shown), indicates that microarrays and HR MAS MRS reflect different traits of the tumors. Lower levels of glucose, which may reflect high energy consumption, and higher levels of alanine in A2 compared with the other luminal A samples indicate that the A2 subgroup has a higher Warburg effect [23]. The lactate signal in A2 also appears to be higher than in the other groups, although not at the significance threshold level. From the GO enrichment analysis, the A2 group was found to be enriched for processes related to cell cycle and DNA repair, compared to A1. The presented subclassification of luminal A might have identified a subgroup of patients (A2) with a more aggressive breast cancer, based on the metabolic and transcriptional profile. However, since this group is small and no long term clinical follow-up is yet available for these patients, a larger cohort with clinical data needs to be analyzed in order to validate whether this finding has clinical impact. It should be noted that intrinsic molecular subtyping is sensitive for selection bias in the cohort analyzed, because of the required gene centring of the microarray data prior to classification. All samples were therefore included in the intrinsic molecular classification, resulting in 10% ER negative samples which is slightly lower than the typical ER negative frequency in IDC. All samples classified as



Enriched GO-terms

GO-category	GO-term	GSA score	p-value
BP	MEIOSIS_I	1.80	0
	MEIOTIC_RECOMBINATION	1.80	0
	DOUBLE_STRAND_BREAK_REPAIR	2.51	0.0001
No significant GO-terms			
MF	PROTEIN_C_TERMINUS_BINDING	-0.91	0

Figure 4 Differences in metabolic and gene expression profiles of identified groups of luminal A samples. Mean spectra of the A1 (black), A2 (red) and A3 (blue) groups of luminal A tumors are plotted with the significantly different points ($p < 0.001$) indicated with grey vertical lines. At ppm values with allocated metabolites and significant differences in one group versus the other, the direction of the arrows indicate if the latter group has the highest (up) or lowest (down) mean value. GO-terms that were significantly enriched (fdr < 0.01) in the different groups are indicated for each comparison, calculated using Gene Set Analysis (GSA) on the microarray data.

luminal A were ER positive and the majority were PgR positive (3 samples were PgR negative and 1 sample had no IHC data for PgR), which supports the classification since luminal A samples are mostly ER/PgR positive and typically 40-50% of IDCs are classified as luminal A [5]. Even though these preliminary results revealing metabolic subgroups within luminal A tumors need to be reproduced in a larger cohort, they suggest that microarray and HR MAS MRS data complement each other, which can be exploited both in subclassification and for constructing predictors of outcome or treatment response.

The second strategy to combine gene expression microarrays and HR MAS MRS was performed by correlating metabolite concentration and gene expression. In this approach, we have not focused on any specific pathways, but correlated the metabolite concentrations to all transcripts on the microarray that showed some variation across samples. We excluded samples with ER negative status from this analysis to avoid detecting associations related to ER-status, which is known to have a profound effect on the transcriptional profile [24]. The three ER positive samples that were not classified as luminal A were classified as normal-like. Since the gene expression of these three samples also correlated to the published luminal A centroid [17], they were included in the correlation analyses to increase power. The gene transcripts that correlated the most to the concentration of taurine and *myo*-inositol were enriched for GO-terms associated to extracellular processes, which could reflect a tumor-stroma interaction. "Cell adhesion" was also one of the enriched GO-terms for the gene transcripts that correlated to *myo*-inositol, which supports this hypothesis. Taurine and *myo*-inositol are known to be involved in osmoregulation and volume regulation [25]. It should be noted that the concentrations of taurine and *myo*-inositol correlated negatively to tumor percentage (Additional file 2: Scatterplot of metabolite concentrations and tumor percentage), which could contribute to the apparent association of these metabolites to extracellular processes. Gene transcripts that correlated the most to choline concentration were enriched for cell cycle related GO-terms which indicate that the choline level in these samples reflects proliferation. Choline is involved in glycerophospholipid metabolism and is a nutrient taken up by the cells as well as a breakdown product from phosphatidylcholine. The total choline signal can be detected by *in vivo* MRS,

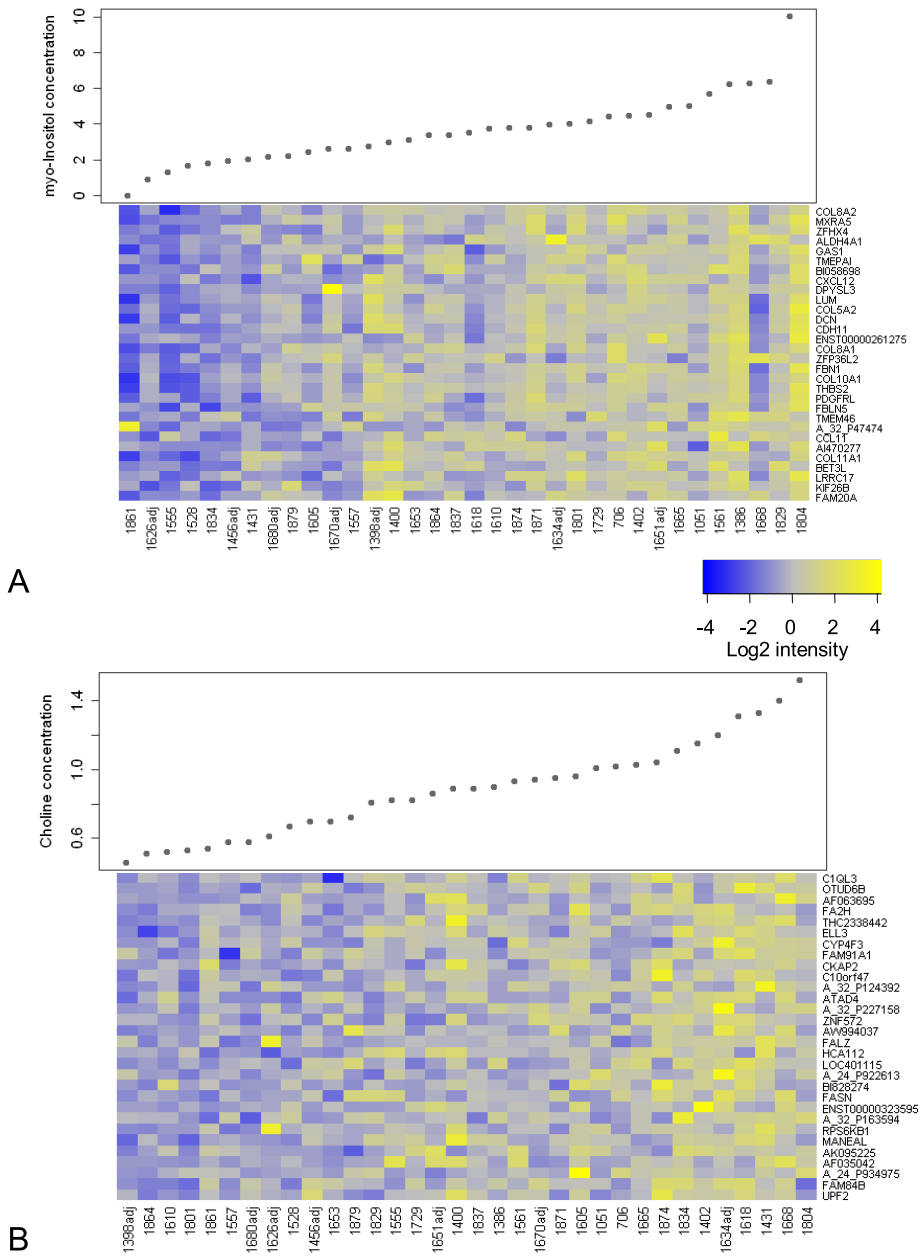


Figure 5 Correlation analysis for *myo*-inositol and choline. The concentration of (A) *myo*-inositol and (B) choline is plotted for the ER positive samples (n = 34), which are ordered according to concentration. The expression of the 30 transcripts that correlated most to these metabolites are shown as heatmaps (see color bar for scale), ordered according to Spearman's rank correlation coefficient. The samples for which the microarray data were obtained from RNA extracted from adjacent tumor tissue are labeled with the prefix "adj".

Table 2 GO enrichment analysis of genes that correlated to the quantified metabolites, using GOrilla

Metabolite/GO-category	GO-Term	Description	P-value	Enrichment	(N,B,n,b)
myo-Inositol					
BP	GO:0032502	developmental process	6.40E-10	1.45	(7307,1397,864,240)
BP	GO:0030198	extracellular matrix organization	1.00E-09	7.16	(7307,56,310,17)
BP	GO:0007155	cell adhesion	1.20E-08	2.45	(7307,377,412,52)
BP	GO:0022610	biological adhesion	1.40E-08	2.44	(7307,378,412,52)
BP	GO:0043062	extracellular structure organization	2.20E-07	4.92	(7307,86,311,18)
MF	GO:0005201	extracellular matrix structural constituent	2.10E-09	7.39	(7307,51,310,16)
CC	GO:0005576	extracellular region	3.80E-21	2.76	(7307,885,311,104)
CC	GO:0044421	extracellular region part	7.10E-17	2.28	(7307,513,757,121)
CC	GO:0031012	extracellular matrix	5.30E-15	5.34	(7307,159,310,36)
CC	GO:0005578	proteinaceous extracellular matrix	2.90E-14	5.57	(7307,142,305,33)
CC	GO:0044420	extracellular matrix part	2.30E-10	3.44	(7307,71,986,33)
CC	GO:0005581	collagen	1.60E-08	21.83	(7307,26,103,8)
Taurine					
BP	GO:0032502	developmental process	8.20E-08	1.48	(7307,1397,630,178)
BP	GO:0048731	system development	1.90E-07	2.23	(7307,314,554,53)
MF	GO:0005201	extracellular matrix structural constituent	6.20E-10	16.25	(7307,51,97,11)
CC	GO:0044421	extracellular region part	8.30E-14	2.44	(7307,513,485,83)
CC	GO:0005576	extracellular region	3.90E-12	2.35	(7307,885,267,76)
CC	GO:0031012	extracellular matrix	2.60E-10	3.43	(7307,159,482,36)
CC	GO:0044420	extracellular matrix part	1.80E-09	12.73	(7307,71,97,12)
CC	GO:0005578	proteinaceous extracellular matrix	2.10E-09	3.52	(7307,142,482,33)
CC	GO:0005581	collagen	5.10E-08	9.48	(7307,26,326,11)
Choline					
BP	GO:0090304	nucleic acid metabolic process	2.00E-09	1.75	(7307,497,990,118)
BP	GO:0022402	cell cycle process	2.30E-08	2.09	(7307,239,952,65)
BP	GO:0007059	chromosome segregation	3.90E-07	4.65	(7307,25,943,15)
BP	GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	8.80E-07	1.54	(7307,640,994,134)
CC	GO:0044428	nuclear part	9.80E-11	1.7	(7307,632,999,147)
CC	GO:0005634	nucleus	5.80E-09	1.3	(7307,2062,990,364)

Table 2 GO enrichment analysis of genes that correlated to the quantified metabolites, using GOrilla *(Continued)*

CC	GO:0000775	chromosome, centromeric region	2.50E-08	10.2	(7307,26,303,11)
Glucose					
BP	GO:0002376	immune system process	1.60E-09	1.83	(7307,420,999,105)
Creatine					
MF	GO:0000030	mannosyltransferase activity	5.80E-08	73.81	(7307,4,99,4)
Glycine					
CC	GO:0070469	respiratory chain	5.80E-07	4.74	(7307,22,981,14)

GO-terms in the Biological Process (BP), Molecular Function (MF) and Cellular Compartment (CC) categories that were significantly enriched (p-value < 10E-7) towards the top of the lists of transcripts ranked according to correlation to each of the quantified metabolites. Enrichment is defined as (b/n)/(B/N), where N is the total number of genes, B is the total number of genes associated with a specific GO term, n is the number of genes in the "target set", b is the number of genes in the "target set" associated with a specific GO term [21]. The p-value threshold corresponds to fdr = 8 metabolites * (877 CC GO-terms*10^-7 + 2416 MF GO-terms*10^-7 + 6307 BP GO-terms*10^-7)/31 significant GO terms = 0.003. There were no significantly enriched GO-terms associated with GPC or PCho.

and is elevated in breast cancer compared to normal mammary tissue and benign lesions [26]. No significantly enriched GO-terms were found in the genes that correlated the most to the two other choline metabolites involved in the total choline peak, PCho and GPC. Glucose correlated to genes that were significantly enriched for the GO-term "immune system process". Glucose concentration has been shown to have an inverse relationship to the number of proliferating cells [27] and tumor cell density [28]. For creatine, only the GO-term "mannosyltransferase activity", which is a glycosylation process, was significantly enriched among the genes that correlated to the metabolite. The genes that correlated to the amino acid glycine, were also only significantly enriched for one GO-term, "respiratory chain", suggesting a possible association between aerobic respiration and glycine levels in these tumor samples. However, glycine was the only metabolite that showed significant positive correlation to tumor percentage (Additional file 2: Scatterplot of metabolite concentrations and tumor percentage.), which could have influenced this relationship. It is worth noting that few transcripts showed a strong correlation to the eight metabolites, as can be seen in the examples in Figure 5. Since metabolite concentrations reflect the sum of many different pathways, correlating the expression of single genes to metabolites is probably not the optimal way to compare the transcriptional and metabolic profiles of tumors. The fact that the most correlated genes were not directly associated with the metabolic pathways of the metabolites they correlated to also emphasizes the complexity of the relationships between gene expression levels and metabolite concentrations. Improved quantification of tissue metabolite concentrations using ERETIC [27] or dieretic [29] and more refined approaches for data analysis, possibly involving curated metabolic pathways, should be explored in the future when larger datasets of microarray and HR MAS MRS data from the same tumor samples can be obtained, with corresponding clinical information.

MRS and microarray experiments have not previously been performed on the same breast cancer sample from the same patient. A study by Tzika *et al.* combined gene expression microarrays and HR MAS MRS on the same sample of brain tumor and control biopsies [30]. However, no results were reported of combining these two types of information, except for stating that a number of transcripts correlated well to the measured metabolites.

Breast cancer biology is highly complex, which is reflected at many different molecular levels. Using gene expression microarray and HR MAS MRS data from the very same tumor sample can reduce the biological variance which gives a higher power to study the transcriptional and metabolic levels in a combined approach. Even though HR MAS MRS leaves the tumor tissue intact, the procedure exposes the tissue to several potential stresses, including hypoxic conditions and lack of nutrients by being embedded in a surrounding buffer at 4°C for approximately an hour, as well as high centrifugal force and magnetic field during the HR MAS MRS acquisition. In our parallel study to address this issue, total RNA integrity was not significantly affected by HR MAS MRS (p-value = 0.86), and findings in a similar evaluation in prostate tissue support this result [31]. The pairs of tumor samples from each patient that had or had not been analyzed by HR MAS MRS cluster together (Additional file 3: Plots illustrating the effect of HR MAS MRS on the transcriptome), indicating that patient to patient variation is larger than the effect of HR MAS MRS. Even though 1199 transcripts were defined as differentially expressed (fdr < 0.01) by HR MAS MRS, these transcripts showed a small fold change. In their study of prostate tissue, Santos *et al.* reported no differential expression caused by HR MAS

MRS [31]. However, Santos *et al*. used cDNA microarray data from unpaired samples to test for differential expression. The patient heterogeneity might have been too high to achieve the power needed to detect possible changes in gene expression in their study.

Conclusions

Performing HR MAS MRS and microarray analysis on the same sample is feasible, and the effect of HR MAS MRS on the transcriptome was shown to be subtle. Three subgroups of samples within the most prevalent intrinsic subgroup of breast cancer, luminal A, were found using multivariate analyses of HR MAS MRS spectra. One of the subgroups of luminal A samples, designated A2, had metabolic and transcriptional features indicating a higher Warburg effect and more proliferation than the other luminal A groups. Using a different strategy, enrichment analysis of genes with expression levels that correlated to metabolite concentrations revealed different enriched GO-terms associated with specific metabolites. GO-terms related to the extracellular matrix were enriched among the genes that correlated the most to *myo*-inositol and taurine, while cell cycle related GO-terms were enriched among the genes that correlated the most to choline. We have shown that combining transcriptional and metabolic data from the same breast carcinoma sample can contribute to a more refined subclassification of breast cancers as well as reveal relationships between these molecular levels. This study has paved the way for further studies in larger patient cohorts of all subtypes, correlating metabolic subgroups to histopathological characteristics, treatment response and clinical outcome.

List of Abbreviations

Abbreviations used: CHO: choline; GPC: glycerophosphocholine; PCHO: phosphocholine; IDC: invasive ductal carcinoma; HR MAS MRS: high resolution magic angle spinning magnetic resonance spectroscopy; PBS: phosphate buffered saline; IQR: inter quartile range; PCA: principal component analysis; TSP: trimethylsilyl tetradeuteropropionic acid; FDR: false discovery rate; GO: Gene Ontolgy; MDS: multidimensional scaling; GSA: Gene Set Analysis

Additional material

Additional file 1: Supplementary documentation of exploring the effect of HR MAS MRS on RNA integrity and transcription. Additional documentation of the experimental procedure and data analysis involving the 18 pairs of samples used in the parallel study to assess the feasibility of performing HR MAS MRS and gene expression microarrays on the same samples.

Additional file 2: Scatterplot of metabolite concentrations and tumor percentage. The log2 of the tissue concentrations (μmol/gram) of glycerophosphocholine (GPC), glycine, phosphocholine (PCho), creatine, choline, taurine, *myo*-inositol and glucose as well as the tumor percentage (Tumor) are plotted against each other in the upper diagonal panel. Spearman's rank correlation coefficients are given in the lower diagonal panel. Significant ($p < 0.05$) correlations are indicated by bold font. The values on the axes represent log2 tissue concentrations (μmol/gram), except for the Tumor axis which represents percentage.

Additional file 3: Plots illustrating the effect of HR MAS MRS on the transcriptome. (A) Hierarchical clustering of the 18 pairs of tumors before (control) and after HR MAS (HRMAS). (B) A plot of the number of significantly differentially expressed (DE) genes caused by HR MAS MRS as a function of false discovery rate. Bonferroni corrected p-value = 0.05 is indicated in the plot. (C) A volcanoplot of significance versus the estimated fold change caused by HR MAS MRS. Transcripts with fdr < 0.01 are colored red or blue to indicate higher or lower expression after HR MAS MRS, respectively. The top Biological Process GO-terms of the differentially expressed genes are listed on each side with the same color-code.

Additional file 4: Significantly lower expressed transcripts after HR MAS MRS. A summary of the 865 transcripts with significantly lower expression after HR MAS MRS, identified using Limma (R/Bioconductor). The columns ProbeUID, AgilentProbeID, GeneSymbol, SystematicName are annotations as given by Agilent. The columns logFoldChange, AverageExpression, t.statistic, p.value and adjusted.p.value are the results from the modified paired t-tests performed by Limma.

Additional file 5: Significantly higher expressed transcripts after HR MAS MRS. A summary of the 334 transcripts with significantly higher expression after HR MAS MRS, identified using Limma (R/Bioconductor). The columns ProbeUID, AgilentProbeID, GeneSymbol, SystematicName are annotations as given by Agilent. The columns logFoldChange, AverageExpression, t.statistic, p.value and adjusted.p.value are the results from the modified paired t-tests performed by Limma.

Acknowledgements
Grant sponsors: The Norwegian Research Council, grant numbers 175459 and 183379, and 163027. St. Olavs University Hospital Cancer Foundation.

Author details
[1]Department of Genetics, Institute for Cancer Research, Division of Surgery and Cancer, Oslo University Hospital Radiumhospitalet, Oslo, Norway. [2]Department of Circulation and Medical Imaging, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [3]Biomedical Research Group, Department of Informatics, University of Oslo, Oslo, Norway. [4]Department of Oncology, St. Olavs University Hospital Trondheim, Trondheim, Norway. [5]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Norway.

Authors' contributions
EB carried out HR MAS MRS and microarray experiments and performed microarray preprocessing, data analysis and interpretation of HR MAS MRS and microarray data, and drafted the manuscript. BS performed HR MAS MRS and microarray experiments, did the preprocessing of the HR MAS MRS data, contributed to analysis and interpretation of HR MAS MRS and microarray data and drafting of the manuscript. OCL contributed to analysis and interpretation of HR MAS MRS and microarray data. HJ carried out microarray experiments. TB contributed to interpretation of HR MAS MRS and microarray data. TS contributed to analysis and interpretation of microarray data. SL provided tumor material and clinical data. ALBD conceived of the study, and participated in its design and helped drafting the manuscript. ISG conceived of the study and participated in its design and coordination. All authors revised the manuscript and approved the final version.

Competing interests
The authors declare that they have no competing interests.

Received: 11 May 2010 Accepted: 16 November 2010
Published: 16 November 2010

References

1. Cianfrocca M, Goldstein LJ: **Prognostic and predictive factors in early-stage breast cancer.** *Oncologist* 2004, **9**:606.
2. Sotiriou C, Piccart MJ: **Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care?** *Nat Rev Cancer* 2007, **7**:545-553.
3. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D: **Molecular portraits of human breast tumours.** *Nature* 2000, **406**:747-752.
4. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de RM, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein LP, Borresen-Dale AL: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc Natl Acad Sci USA* 2001, **98**:10869-10874.
5. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lonning PE, Brown PO, Borresen-Dale AL, Botstein D: **Repeated observation of breast tumor subtypes in independent gene expression data sets.** *Proc Natl Acad Sci USA* 2003, **100**:8418-8423.
6. Sorlie T, Wang YL, Xiao CL, Johnsen H, Naume B, Samaha RR, Borresen-Dale AL: **Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms.** *BMC Genomics* 2006, **7**:127.
7. Lindon JC, Beckonert OP, Holmes E, Nicholson JK: **High-resolution magic angle spinning NMR spectroscopy: Application to biomedical studies.** *Prog Nucl Magn Reson Spectrosc* 2009, **55**:79-100.
8. Sitter B, Bathen T, Tessem MB, Gribbestad IS: **High-resolution magic angle spinning (HR MAS) MR spectroscopy in metabolic characterization of human cancer.** *Prog Nucl Magn Reson Spectrosc* 2009, **54**:239-254.
9. Eliyahu G, Kreizman T, Degani H: **Phosphocholine as a biomarker of breast cancer: molecular and biochemical studies.** *Int J Cancer* 2007, **120**:1721-1730.
10. Glunde K, Jie C, Bhujwalla ZM: **Molecular causes of the aberrant choline phospholipid metabolism in breast cancer.** *Canc Res* 2004, **64**:4270-4276.
11. Moestue S, Borgan E, Huuse E, Lindholm E, Sitter B, Borresen-Dale AL, Engebraaten O, Maelandsmo G, Gribbestad I: **Distinct choline metabolic profiles are associated with differences in gene expression for basal-like and luminal-like breast cancer xenograft models.** *BMC Cancer* 2010, **10**:433.
12. Morse DL, Carroll D, Day S, Gray H, Sadarangani P, Murthi S, Job C, Baggett B, Raghunand N, Gillies RJ: **Characterization of breast cancers and therapy response by MRS and quantitative gene expression profiling in the choline pathway.** *NMR Biomed* 2009, **22**:114-127.
13. Sitter B, Lundgren S, Bathen TF, Halgunset J, Fjosne HE, Gribbestad IS: **Comparison of HR MAS MR spectroscopic profiles of breast cancer tissue with clinical parameters.** *NMR Biomed* 2006, **19**:30-40.
14. Sitter B, Sonnewald U, Spraul M, Fjosne HE, Gribbestad IS: **High-resolution magic angle spinning MRS of breast cancer tissue.** *NMR Biomed* 2002, **15**:327-337.
15. Smyth GK: **Linear models and empirical bayes methods for assessing differential expression in microarray experiments.** *Stat Appl Genet Mol Biol* 2004, **3**:Article 3.
16. Stacklies W, Redestig H, Scholz M, Walther D, Selbig J: **pcaMethods–a bioconductor package providing PCA methods for incomplete data.** *Bioinformatics* 2007, **23**:1164-1167.
17. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L: **The molecular portraits of breast tumors are conserved across microarray platforms.** *BMC Genomics* 2006, **7**:96.
18. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, **25**:25-29.
19. **Gene Set Analysis.** [http://www-stat.stanford.edu/~tibs/GSA/].
20. Eden E, Lipson D, Yogev S, Yakhini Z: **Discovering motifs in ranked lists of DNA sequences.** *PLoS Comput Biol* 2007, **3**:e39.
21. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z: **GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists.** *BMC Bioinformatics* 2009, **10**:48.
22. Falcon S, Gentleman R: **Using GOstats to test gene lists for GO term association.** *Bioinformatics* 2007, **23**:257-258.
23. Vander Heiden MG, Cantley LC, Thompson CB: **Understanding the Warburg effect: the metabolic requirements of cell proliferation.** *Sci Signal* 2009, **324**:1029.
24. Gruvberger S, Ringnér M, Chen Y, Panavally S, Saal LH, Borg A, Ferno M, Peterson C, Meltzer PS: **Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns.** *Canc Res* 2001, **61**:5979.
25. Griffin JL, Shockcor JP: **Metabolic profiles of cancer cells.** *Nat Rev Cancer* 2004, **4**:551-561.
26. Kvistad KA, Bakken IJ, Gribbestad IS, Ehrnholm B, Lundgren S, Fjosne HE, Haraldseth O: **Characterization of neoplastic and normal human breast tissues with in vivo 1H MR spectroscopy.** *J Magn Reson Imaging* 1999, **10**:159-164.
27. Sitter B, Bathen TF, Singstad TE, Fjosne HE, Lundgren S, Halgunset J, Gribbestad IS: **Quantification of metabolites in breast cancer patients with different clinical prognosis using HR MAS MR spectroscopy.** *NMR Biomed* 2010, **23**:424-31.
28. Lyng H, Sitter B, Bathen TF, Jensen LR, Sundfør K, Kristensen GB, Gribbestad IS: **Metabolic mapping by use of high-resolution magic angle spinning 1 H MR spectroscopy for assessment of apoptosis in cervical carcinomas.** *BMC Cancer* 2007, **7**:11.
29. Wider G, Dreier L: **Measuring protein concentrations by NMR spectroscopy.** *J Am Chem Soc* 2006, **128**:2571-2576.
30. Tzika AA, Astrakas L, Cao H, Mintzopoulos D, Andronesi OC, Mindrinos M, Zhang J, Rahme LG, Blekas KD, Likas AC, Galatsanos NP, Carroll RS, Black PM: **Combination of high-resolution magic angle spinning proton magnetic resonance spectroscopy and microscale genomics to type brain tumor biopsies.** *Int J Mol Med* 2007, **20**:199-208.
31. Santos CF, Kurhanewicz J, Tabatabai ZL, Simko JP, Keshari KR, Gbegnon A, Santos RD, Federman S, Shinohara K, Carroll PR, Haqq CM, Swanson MG: **Metabolic, pathologic, and genetic analysis of prostate tissues: quantitative evaluation of histopathologic and mRNA integrity after HR-MAS spectroscopy.** *NMR Biomed* 2009, **23**:391-8.
32. Strand C, Enell J, Hedenfalk I, Ferno M: **RNA quality in frozen breast cancer samples and the influence on gene expression analysis–a comparison of three evaluation methods using microcapillary electrophoresis traces.** *BMC Mol Biol* 2007, **8**:38.
33. Falcon S, Gentleman R: **Using GOstats to test gene lists for GO term association.** *Bioinformatics* 2007, **23**:257-258.
34. Efron B, Tibshirani R: **Empirical Bayes methods and false discovery rates for microarrays.** *Genet Epidemiol* 2002, **23**:70-86.

Pre-publication history

The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2407/10/628/prepub

doi:10.1186/1471-2407-10-628
Cite this article as: Borgan *et al.*: Merging transcriptomics and metabolomics - advances in breast cancer profiling. *BMC Cancer* 2010 **10**:628.

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

BioMed Central

Additional files

Additional file 1:
Supplementary documentation of exploring the effect of HR MAS MRS on RNA integrity and transcription. Additional documentation of the experimental procedure and data analysis involving the 18 pairs of samples used in the parallel study to assess the feasibility of performing HR MAS MRS and gene expression microarrays on the same samples.

Additional file 2:
Scatterplot of metabolite concentrations and tumor percentage. The log2 of the tissue concentrations (µmol/gram) of glycerophosphocholine (GPC), glycine, phosphocholine (PCho), creatine, choline, taurine, myo-inositol and glucose as well as the tumor percentage (Tumor) are plotted against each other in the upper diagonal panel. Spearman's rank correlation coefficients are given in the lower diagonal panel. Significant ($p < 0.05$) correlations are indicated by bold font. The values on the axes represent log2 tissue concentrations (µmol/gram), except for the Tumor axis which represents percentage.

Additional file 3:
Plots illustrating the effect of HR MAS MRS on the transcriptome. (A) Hierarchical clustering of the 18 pairs of tumors before (control) and after HR MAS (HRMAS). (B) A plot of the number of significantly differentially expressed (DE) genes caused by HR MAS MRS as a function of false discovery rate. Bonferroni corrected p-value = 0.05 is indicated in the plot. (C) A volcanoplot of significance versus the estimated fold change caused by HR MAS MRS. Transcripts with fdr < 0.01 are colored red or blue to indicate higher or lower expression after HR MAS MRS, respectively. The top Biological Process GO-terms of the differentially expressed genes are listed on each side with the same color-code.

Additional file 4:
Significantly lower expressed transcripts after HR MAS MRS. A summary of the 865 transcripts with significantly lower expression after HR MAS MRS, identified using Limma (R/Bioconductor). The columns ProbeUID, AgilentProbeID, GeneSymbol, SystematicName are annotations as given by Agilent. The columns logFoldChange, AverageExpression, t.statistic, p.value and adjusted.p.value are the results from the modified paired t-tests performed by Limma.

Available at: http://www.biomedcentral.com/1471-2407/10/628/additional/

Additional file 5:
Significantly higher expressed transcripts after HR MAS MRS. A summary of the 334 transcripts with significantly higher expression after HR MAS MRS, identified using Limma (R/Bioconductor). The columns ProbeUID, AgilentProbeID, GeneSymbol, SystematicName are annotations as given by Agilent. The columns logFoldChange, AverageExpression, t.statistic, p.value and adjusted.p.value are the results from the modified paired t-tests performed by Limma.

Available at: http://www.biomedcentral.com/1471-2407/10/628/additional/

Additional File 1

Supplementary documentation of exploring the effect of HR MAS MRS on RNA integrity and transcription

Sample preparation and experimental design

An additional study was designed to study the effect of HR MAS MRS on RNA quality and gene expression. Fresh frozen tumor tissue from 18 breast cancer patients was split in two similarly sized samples. The first part of the sample was snap frozen in liquid nitrogen and used directly for RNA extraction, while the second part was analyzed by HR MAS MRS and then snap frozen until RNA extraction.

RNA extraction and microarrays

Total RNA was extracted from the fresh frozen tissue following the protocol of the TRIzol reagent (Invitrogen, USA). Total RNA integrity was measured using BioAnalyzer 2100 (Agilent Technologies, USA). The pair of samples from each patient was further processed when the RNA quality of one sample exceeded a RIN-value of 6, suggested by Strand et al. [32]. Total RNA (400 ng) from 18 pairs of samples were used in one-color microarray experiments using 4x44k Agilent Human Whole Genome Oligo Microarrays. The microarrays were scanned using a G2565A scanner and the images were extracted using Agilent Feature Extraction software (v 9.3).

Preprocessing and normalizing the microarray data

The Feature Extraction text files were used directly and analyzed using R (v 2.8) and Bioconductor. The gMeanSignal was used and corrected for multiplicative detrending effects by dividing by the gMultDetrendSignal. The arrays were quantile normalized and log2

transformed using the Limma package (R/Bioconductor) [15]. Control probes, probes which were flagged as outliers on more than 20 % of the arrays, and probes which were flagged as present on less than 50 % of the arrays were removed. The average of duplicate probes was taken, and the probe with the highest IQR was picked when a gene was represented by different probes, leaving 27114 probes.

Testing the effect of HR MAS MRS on RNA integrity and transcription

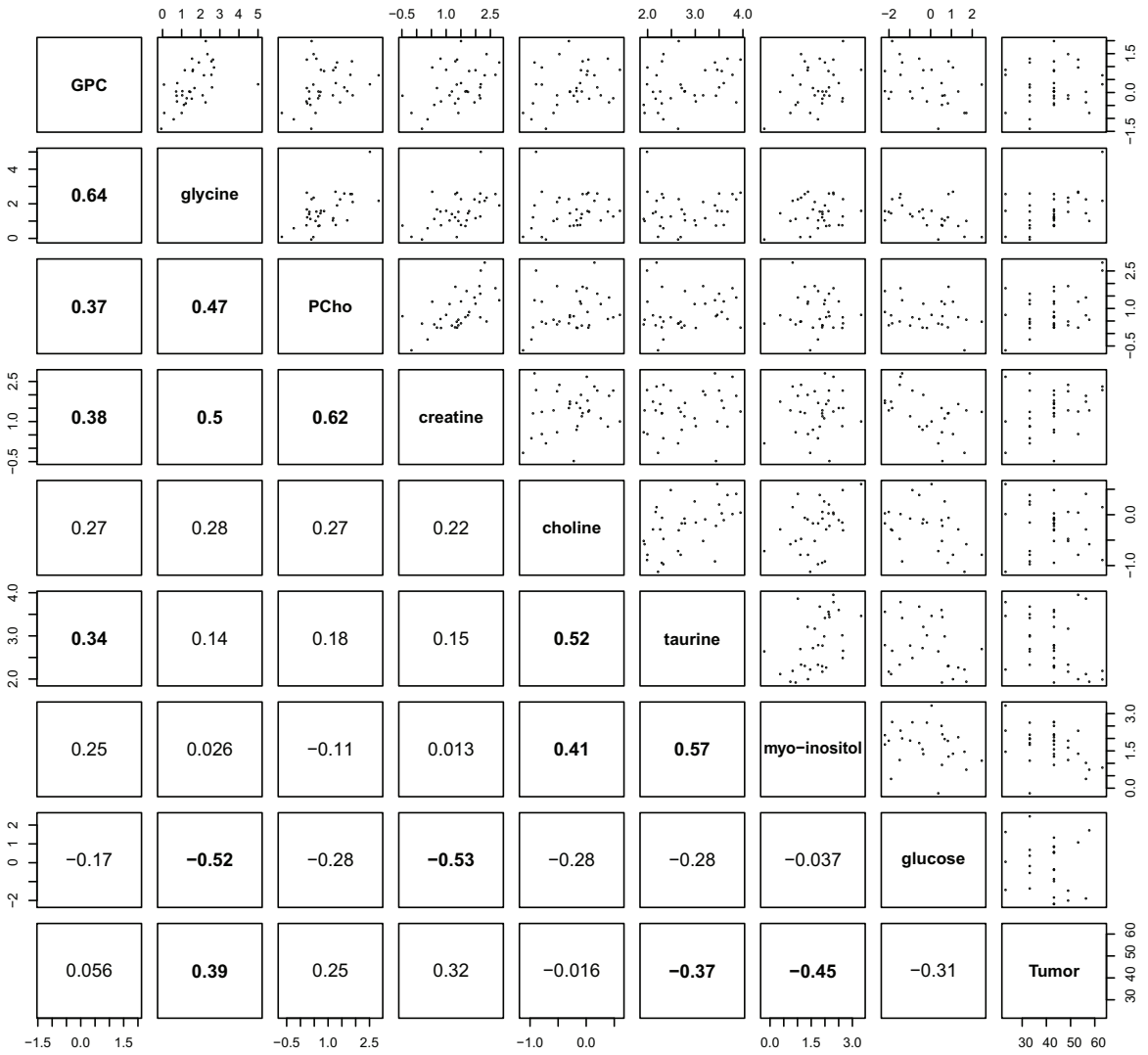
The effect of HR MAS MRS on RNA integrity was tested using a paired t-test on the RIN-values (measured using Bioanalyzer 2100). Unsupervised hierarchical clustering of the microarray data from the 18 pairs of samples was performed using Eucledian distance and complete linkage. Tests for the effect of HR MAS on gene expression, were performed using Limma (R/ Bioconductor) [15]. The genes that had significantly higher or lower (fdr<0.01) expression in the samples analyzed by HR MAS MRS, were tested for enriched gene ontology (GO) terms [18], using the GOstats package (R/ Bioconductor) [33]. Each of the two genelists were compared to all genes in the filtered microarray data with assigned GO-terms, by using conditional hypergemometric tests to find GO terms that were represented more frequently than expected by chance.

The genes that showed significantly lower expression after HR MAS MRS were enriched for GO-terms related to metabolism, such as “acetyl-CoA metabolic process”, “tricarboxylic acid cycle” and “generation of precursor of precursor metabolites and energy” (see Additional file 3: Plots illustrating the effect of HR MAS MRS on the transcriptome), which suggest a shutdown of metabolism. It is not known if this happens before or during the actual HR MAS MRS acquisition. However, the genes with known function that were higher expressed in samples exposed to HR MAS MRS, were enriched for biological processes such as RNA

splicing and regulation of gene expression. This may indicate that lower expression of several transcripts in the samples exposed to HR MAS MRS is caused by regulation of gene expression in the cells of the tumor tissue, as opposed to random degradation of RNA transcripts.

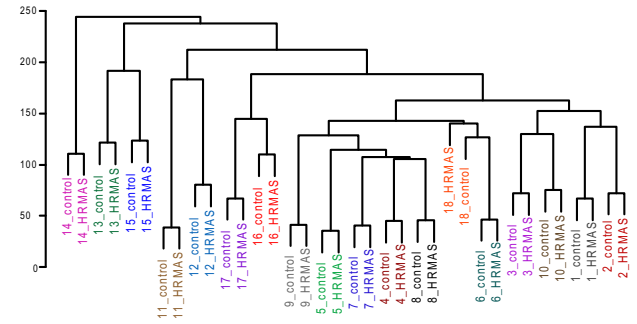
One limitation to this study is that several samples with or without the HR MAS MRS procedure showed RNA quality below a recommended threshold of 6 [32]. This was done to avoid bias in the analysis of the effect of HR MAS MRS on gene expression, which could arise when excluding samples with poor RNA quality from before or after HR MAS MRS. The low RNA quality in some samples may have introduced additional noise to the data. The false discovery rate calculations are based on the assumption that there is no difference in expression between the samples, which can give an underestimation of the false discovery rate [34]. However, it should be noted that the same test for differential expression was applied to the paired samples with scrambled labels for which sample had been exposed to HR MAS MRS. This test resulted in zero genes passing the fdr<0.05 threshold (results not shown), which supports that the false discovery rate calculations do not simply reflect the heterogeneity of neighboring tissue from the same tumor.

Additional File 2

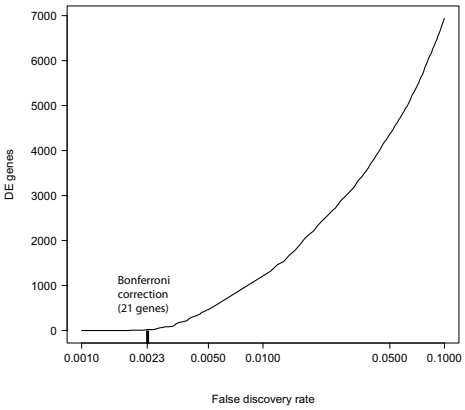


Additional File 3

A

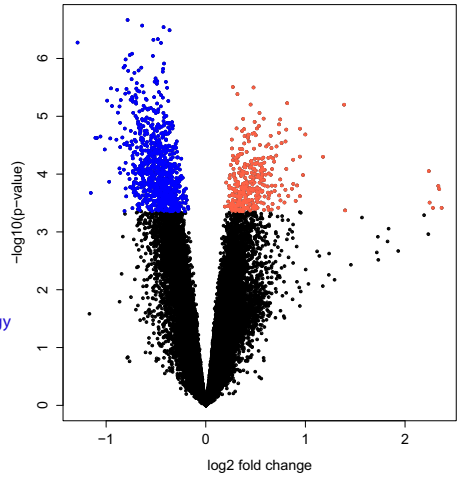


B



C

Protein localization
Antigen processing and presentation
of peptide antigen via MHC class I
Interspecies interaction between organisms
ER to Golgi vesicle-mediated transport
Vesicle-mediated transport
Intracellular protein transport
acetyl-CoA metabolic process
Establishment of localization
Cellular localization
Protein transport
Nucleoside triphosphate biosynthetic process
Tricarboxylic acid cycle
Generation of precursor metabolites and energy
Protein modification process



RNA metabolic process
RNA splicing
Macromolecule metabolic process
Embryonic hindlimb morphogenesis
Gene expression
Regulation of gene expression
Primary metabolic process
Regulation of transcription, DNA-dependent
mRNA processing
Negative regulation of transcription
RNA biosynthetic process
Cellular metabolic process
Regulation of myeloid leukocyte differentiation
Regulation of metabolic process

III

Moestue et al. BMC Cancer 2010, 10:433
http://www.biomedcentral.com/1471-2407/10/433

BMC Cancer

RESEARCH ARTICLE Open Access

Distinct choline metabolic profiles are associated with differences in gene expression for basal-like and luminal-like breast cancer xenograft models

Siver A Moestue[1*], Eldrid Borgan[1,2], Else M Huuse[1], Evita M Lindholm[3], Beathe Sitter[1], Anne-Lise Børresen-Dale[2,4], Olav Engebraaten[3,4], Gunhild M Mælandsmo[3], Ingrid S Gribbestad[1]

Abstract

Background: Increased concentrations of choline-containing compounds are frequently observed in breast carcinomas, and may serve as biomarkers for both diagnostic and treatment monitoring purposes. However, underlying mechanisms for the abnormal choline metabolism are poorly understood.

Methods: The concentrations of choline-derived metabolites were determined in xenografted primary human breast carcinomas, representing basal-like and luminal-like subtypes. Quantification of metabolites in fresh frozen tissue was performed using high-resolution magic angle spinning magnetic resonance spectroscopy (HR MAS MRS).

The expression of genes involved in phosphatidylcholine (PtdCho) metabolism was retrieved from whole genome expression microarray analyses.

The metabolite profiles from xenografts were compared with profiles from human breast cancer, sampled from patients with estrogen/progesterone receptor positive (ER+/PgR+) or triple negative (ER-/PgR-/HER2-) breast cancer.

Results: In basal-like xenografts, glycerophosphocholine (GPC) concentrations were higher than phosphocholine (PCho) concentrations, whereas this pattern was reversed in luminal-like xenografts. These differences may be explained by lower choline kinase (*CHKA*, *CHKB*) expression as well as higher PtdCho degradation mediated by higher expression of phospholipase A2 group 4A (*PLA2G4A*) and phospholipase B1 (*PLB1*) in the basal-like model. The glycine concentration was higher in the basal-like model. Although glycine could be derived from energy metabolism pathways, the gene expression data suggested a metabolic shift from PtdCho synthesis to glycine formation in basal-like xenografts. In agreement with results from the xenograft models, tissue samples from triple negative breast carcinomas had higher GPC/PCho ratio than samples from ER+/PgR+ carcinomas, suggesting that the choline metabolism in the experimental models is representative for luminal-like and basal-like human breast cancer.

Conclusions: The differences in choline metabolite concentrations corresponded well with differences in gene expression, demonstrating distinct metabolic profiles in the xenograft models representing basal-like and luminal-like breast cancer. The same characteristics of choline metabolite profiles were also observed in patient material from ER+/PgR+ and triple-negative breast cancer, suggesting that the xenografts are relevant model systems for studies of choline metabolism in luminal-like and basal-like breast cancer.

* Correspondence: siver.a.moestue@ntnu.no
[1]Department of Circulation and Medical Imaging, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
Full list of author information is available at the end of the article

BioMed Central

© 2010 Moestue et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background

Optimal treatment of individual breast cancer patients is still a major challenge in oncology. An approach to improve and individualize the treatment beyond the markers and stratification tools used at present, is through molecular subtyping of breast cancer [1]. Based on variation in gene expression profiles, five molecular subtypes have been identified [1-3]. The gene expression patterns of these subtypes are similar across multiple samples from the same tumor, shows no treatment-related changes and have been reproduced in a number of patient populations [2-7]. However, the current use of these molecular subgroups in clinical practice remains limited. Further understanding of the differences in biology between the various subtypes is needed in order to predict therapeutic response and provide individual treatment strategies based on gene expression profiles [8].

Elevated levels of choline metabolites is a known feature of breast cancer, and it has been shown that drugs targeting choline metabolism have selective *in vivo* and *in vitro* cytotoxic efficacy against a variety of cancer types [9-13]. Magnetic resonance spectroscopy (MRS) is a valuable tool for studies of choline metabolism both in patients and in experimental systems [14]. High resolution magic angle spinning (HR MAS) MRS of *ex vivo* tissue samples has been particularly useful, as it allows assessment of individual choline metabolites in intact tissue specimens. Increased concentrations of choline, phosphocholine (PCho) and glycerophosphocholine (GPC) has been demonstrated both in cultured breast cancer cells [15-17] and in human breast cancer biopsies [18-21]. It has also been shown that choline metabolism is altered following chemotherapy, suggesting the possibility of using MRS for therapy monitoring [22-24]. However, to utilize these findings in diagnosis and individualized therapy monitoring of breast cancer patients, a better understanding of the choline metabolism abnormalities on a molecular level is needed.

Several studies investigating expression of genes involved in metabolism of phosphatidylcholine (PtdCho) have been performed using breast cancer cell lines [16,17,25,26]. PtdCho is an important cellular membrane lipid, and this metabolic pathway directly involves choline, PCho and GPC. In addition, genes involved in transmembrane choline transport and conversion of choline to glycine have been suggested to be important for the observed choline concentrations in breast cancer cells [16,27]. The choline metabolism profiles observed in cultured breast cancer cells are more homogenous than those seen in human biopsies. In order to bridge the gap between *in vitro* research and clinical breast cancer, there is a great need for animal models representing different types of breast cancer for use in functional and mechanistic studies. Serial orthotopic transplantation of clinical tumor isolates in immunodeficient mice is considered a promising tool for investigation of human breast cancer biology [28]. Establishment of relevant experimental models of basal-like breast cancer is especially important both in order to understand the special characteristics of this subtype, to find potential new molecular targets for therapy and to establish potential biomarkers for monitoring response to therapy.

The aim of this study was to compare the choline metabolite patterns in animal models of basal-like and luminal-like subtypes of breast cancer, and to study the expression of genes related to choline metabolism in order to explain the differences between the two breast cancer subtypes. The two orthotopic xenograft models used, MAS98.12 and MAS98.06, represent basal-like and luminal-like subtypes of breast cancer, respectively [29]. Both models have been established by direct inoculation of primary human tumor material into immunodeficient animals. The content of creatine, choline, PCho, GPC, taurine and glycine in the xenografts as well as human breast cancer tissue samples was determined using HR MAS MRS. The molecular basis of the observed differences in choline metabolism was studied using gene expression microarray data. In order to evaluate if the xenograft models are representative for human disease, the metabolic profiles were compared to corresponding profiles from patients with ER+/PgR+ or triple negative breast cancer.

Methods

Animal model

The MAS98.12 and MAS98.06 tumor models were established by orthotopic implantation of biopsy tissues from primary mammary carcinomas in SCID mice as previously described [29]. Both the primary carcinomas and the xenograft models have been characterized using gene expression profiling. These analyses demonstrated that the primary carcinomas could be classified as luminal-like and basal-like subtypes of breast cancer, and that these molecular subtypes were retained in the MAS98.06 (luminal-like) and MAS98.12 (basal-like) xenografts. Relevant characteristics of the models are presented in Table 1. The tumors are serially transplanted. Tissue used for HR MAS MRS was from passage 47 (MAS98.12) and 28 (MAS98.06), and tissue used for RNA microarray analysis was from passage 45 (MAS98.12) and 25 (MAS98.06).

The animals were kept under pathogen-free conditions. Housing conditions included temperature between 19°C and 22°C, humidity between 50% and 60%, 20 air

Table 1 Summary of xenograft characteristics

	Basal-like xenograft (MAS98.12)		Luminal-like xenograft (MAS98.06)	
	Primary tumor	**Xenograft**	**Primary tumor**	**Xenograft**
Tumor grade	Grade III IDC	NA	Grade III IDC	NA
Lymph node status	No metastasis	NA	Metastasis to 12 of 25 nodes No distant metastases	NA
Differentiation	Poorly differentiated	Poorly differentiated	Well differentiated	Poorly differentiated
Hormone receptor status	ER-/PgR+**	ER-/PgR-	ER+/PgR+	ER+/PgR+
ERBB2 amplification*	Negative	Negative	Negative	Negative
Intrinsic molecular subtype	Basal-like	Basal-like	Luminal-like	Luminal-like
TP53 status	Wildtype	Mutated	Mutated	Mutated
Volume doubling time	NA	1-2 days	NA	7 days
Proliferation index (Ki67)	Missing	28%	Missing	35%

Summary of characteristics related to genotype and phenotype of the xenograft models

* Measured at the DNA level by array Comparative Genomic Hybridization (aCGH)

** The primary basal-like carcinoma had very weak cytoplasmic staining for PgR[29].

changes/hr and a 12 hr light/dark cycle. The animals were fed RM1 diet (Scanbur BK, Norway) and distilled tap water *ad libitum*. The drinking water was supplemented with 17-β-estradiol at a concentration of 4 μg/ml in order to ensure stimulation of the estrogen receptors and promote tumor growth in the MAS98.06 xenografts. With respect to tumor growth rate, this estrogen supplement correspond to the use of s.c. continuous release 17-β-estradiol pellets (1.7 mg/pellet), which were used during establishment of the animal models [29]. To provide equal experimental conditions, the MAS98.12 xenografts also received estradiol supplement. This could in theory cause non-ER-mediated effects on choline metabolite profile. However, the similarities between human tissue samples and xenograft tissue suggest that such effects are insignificant in ER-breast cancer.

Following sacrifice by cervical dislocation, tumor tissue was harvested from 10 animals from each model for the HR MAS MRS analyses and for 6 animals from each model for gene expression microarray analyses, at tumor diameters of approximately 13-15 mm. Samples were put in cryogenic vials and immersed in liquid nitrogen immediately after dissection and stored under cryogenic conditions until analysis. All procedures and experiments involving animals were approved by The National Animal Research Authority, and carried out according to the European Convention for the Protection of Vertebrates used for Scientific Purposes.

Human tissue samples

For comparison of xenograft models with human breast cancer tissue, biopsies from 22 breast cancer patients were identified in our internal database based on histopathology/immunohistochemistry data. Patients with either ER+/PgR+ (n = 14) or triple negative (n = 8) phenotype were included. Biopsy material was obtained during surgery and immediately frozen in liquid nitrogen. Histopathology and immunohistochemistry data for the selected patients was obtained from hospital records. Patient and tumor characteristics are presented in Table 2. The biopsy material was subject to HR MAS MRS analysis and subsequent histopathological evaluation using hematoxylin/eosin (HE) staining. The use of patient material was approved by the Regional Committee for Medical and Health Research Ethics, and informed written consent was obtained from all included patients.

HR MAS MRS of xenograft tissue

Storage time before HR MAS MRS analysis was less than one month for all 20 samples. Macroscopically viable tumor tissue was cut to fit a 30 μl disposable insert (Bruker Biospin Corp.), prefilled with 3 μl PBS made on D_2O containing 98.8 mM trimethylsilyltetradeuteropropionic acid (TSP) for chemical shift referencing. The average sample weight was 15 ± 3 mg (mean ± SD). HR MAS MR spectra were recorded using a Bruker AVANCE DRX600 spectrometer equipped with a $^{1}H/^{13}C$ HR MAS probe (Bruker BioSpin Corp.). Samples were spun at 5 kHz with an instrumental temperature setting of 4°C. A pulse-acquired experiment including the ERETIC sequence (ereticpr.drx; Bruker) was performed for all samples. The ERETIC signal was positioned at -1.0 ppm. The water resonance was saturated for 15 seconds (60 dB continuous wave), followed by a 60-degree pulse for excitation. Signals were collected over a sweep width of 16.7 ppm. 128 FIDs were acquired into 64K points during 3.28 seconds. Spectra were Fourier transformed into 128K after 0.3 Hz exponential line broadening and chemical shifts were calibrated to the TSP singlet at 0 ppm. Spectral

Table 2 Summary of patient characteristics

Subtype	n	Patient age (years)	Phenotype	Tumor grade1/2/3	Tumor size (cm)	Mean tumor fraction (%)	Mean connective tissue fraction (%)	Mean fatty tissue fraction (%)
ER+/PgR+	14	57 ± 16	ER+/PgR+	1/10/3	2.3 ± 1.3	23 ± 11	72 ± 11	5 ± 7
Triple negative	8	57 ± 17	ER-/PgR-/HER2-	0/3/5	2.2 ± 1.0	38 ± 32	55 ± 31	6 ± 7

Summary of patient and sample characteristics of the different subgroups of human tissue samples (mean ± SD)

assignments were performed based on a previous HR MAS MRS study of breast cancer lesions [30]. One HR MAS MRS spectrum from the MAS98.06 animals was lost due to technical error.

The regions from 0.20 to -0.20 ppm (TSP), -0.85 to -1.15 ppm (ERETIC) and 3.60 to 2.90 ppm (glycine, taurine, GPC, PCho, choline, and creatine) were selected for quantification in all spectra. Peak areas were calculated by curve fitting (PeakFit v 4, Systat Software Inc) using a combination of Gaussian and Lorentzian lineshapes (Voigt function). The correlation coefficient of the fit (r^2) was > 0.95 for all spectra. The ERETIC signal was quantified to 3.17×10^{-7} moles using a series of creatine calibration standards as previously described [31]. Concentrations of tissue metabolites ([*MET*]) were calculated relative to the ERETIC signal using equation (1):

$$[MET] = \frac{A_{MET}}{A_{ERETIC}} \times \frac{1}{k_{MET}} \times \frac{n_{ERETIC}}{m_{sample}} \quad (1)$$

A_{MET} and A_{ERETIC} are the calculated areas of the metabolite and the ERETIC signals, respectively; k_{MET} is the number of protons giving rise to the metabolite signal; n_{ERETIC} is the number of moles the ERETIC signal represents; and m_{sample} is the mass of the sample. The metabolite concentrations measured using the ERETIC signal were compared using a 2-sided Student's t-test with a significance level of $p < 0.05$ using Sigmaplot 11.0 (Systat Software Inc.).

HR MAS MRS of human tissue samples

Human tissue samples were prepared for HR MAS MRS analysis using the same procedure as the xenograft samples. Spectra were acquired using a spin-echo Carr-Purcell-Meiboom-Gill sequence (cpmgpr; Bruker) with 2 s water suppression prior to a 90° excitation pulse. The spin-echo sequence for suppression of broad peaks was performed using a delay of 1 ms repeated 136 times, resulting in an effective echo time of 285 ms. A total of 128 scans over a spectral region of 10 kHz were collected into 32k points during 1.64 s. The spectra were Fourier transformed into 128 K after 0.3 Hz exponential line broadening, and the metabolite region from 3.60 to 3.00 ppm was selected for further evaluation. The spectra were normalized by scaling the spectral data of all samples to achieve an equal total area for each spectrum. Metabolite peak areas were then obtained by curve fitting as described above.

Histopathology

Following HR MAS MRS analysis, the xenograft samples were fixed in 10% neutral buffered formalin and embedded in paraffin. One histopathological section were prepared from each sample, stained with hematoxylin/eosin/saffron (HES) according to standard protocol and evaluated microscopically. A visual evaluation with respect to the presence of viable tumor tissue and the extent of necrosis was performed. Tumor grade, hormone receptor status and HER2 expression of human tissue samples was obtained from hospital records. In addition, specimens analysed by HR MAS MRS were HES-stained and the relative areas of normal and neoplastic epithelial tissue, necrotic tissue, fat and fibrous connective tissue were scored.

Gene expression analysis

Gene expression analysis was performed on tumor tissue from 6 animals from each of the two xenograft models, using a one-color microarray-based platform (Agilent). Total RNA was isolated from snap frozen tumor tissue using TRIzol (Invitrogen) and resuspended in RNase-free water. Total RNA (700 ng) was amplified, labelled with Cy3, and 1.65 μg cRNA was hybridized to 4 × 44 k Agilent Whole Human Genome Oligo Microarrays at 60°C and 10 rpm for 17 hours, according to the manufacturer's protocol. The arrays were scanned using an Agilent G2565A DNA microarray scanner and extracted using Feature Extraction (v 10.1.1.1, Agilent). One microarray from the MAS98.06 model was removed due to poor array quality. The microarray data was normalized and analysed using R (v 2.9.0) and the LIMMA Bioconductor package [32]. The raw signals were corrected for multiplicative detrending effects and the arrays were quantile normalized and log2 transformed. Probes which were flagged as outliers by the Feature Extraction software or were present in less than 30% of the samples, were removed. The signal intensities were averaged between duplicate probes, and the probe with the highest interquartile range was selected to represent each unique transcript.

A total of 119 genes were selected for further analysis. The selection criteria were a) genes involved in KEGG *homo sapiens* glycerophospholipid pathway hsa:00564 [33], or b) genes coding for proteins reported to be directly involved in choline transport and choline and glycine metabolism [16,34-36]. Of the selected genes, 117 were represented on the microarray (full gene list supplied as additional file 1).

Testing for differential expression between the xenograft models was performed using t-tests with Empirical Bayesian correction of the test statistics [32]. To account for multiple testing, an adjusted p-value of 0.05 (using Benjamini & Hochberg's false discovery rate) was defined as the threshold for significant differential expression between the xenograft models. The microarray data from the significantly differentially expressed genes was centered across genes and clustered across genes and samples using hierarchical clustering with Euclidian distance and complete linkage. The relationship between gene expression and metabolite concentrations was explored using Ingenuity Pathways Analysis (Ingenuity Systems), and an illustration was adapted from the canonical Glycerophospholipid Metabolism and Glycine, Serine and Threonine Metabolism pathways [33]. The abovementioned gene list was also extracted from microarray data from previously described passages of the same xenograft models [29], to ensure that gene expression remain stable throughout serial transplantation of the xenografts.

Results

Histopathology

All xenograft samples were found to contain mainly viable tumor tissue and stromal connective tissue, shown previously to be recruited mouse stromal tissue [29], with negligible necrosis (< 10% area) in 18 of 19 samples. The HR MAS MRS data was therefore considered to be representative of the metabolite concentrations in the solid tumors. One sample in the MAS98.06 group contained a necrotic area, microscopically estimated to 25% area of the specimen. However, the metabolite concentrations measured in this sample differed from the group mean by less than ± 2 SD, and the sample was therefore not excluded from the data set. The mean fractions of tumor and connective tissue in the human tissue samples are presented in Table 2.

HR MAS MRS of xenograft samples

The HR MAS MRS analyses revealed several significant differences in the metabolic profiles of the two xenograft models. Mean ^{1}H HR MAS MRS spectra from the two models are shown in Figure 1 (spectral region 3.6 - 3.0 ppm). The metabolites assigned in Figure 1 were quantifiable in all spectra. The metabolite concentrations calculated using the ERETIC reference signal are presented in Table 3. There was no significant difference in choline concentration between the models. However, the concentrations of GPC and PCho were significantly higher than the choline concentration in both the basal-like and the luminal-like model. While all the samples from basal-like xenografts showed higher concentration of GPC than PCho, the samples from luminal-like xenografts invariably showed lower concentrations of GPC than PCho. The differences in GPC and PCho concentrations between the two xenograft models were statistically significant ($p < 0.001$ and $p < 0.01$, respectively). The concentration of glycine was significantly higher in the basal-like than in the luminal-like model ($p < 0.002$).

HR MAS MRS of human tissue samples

The HR MAS MRS spectra from the tissue samples were retrieved from our internal database, and mean spectra from the two groups are shown in Figure 1. The mean metabolite profiles demonstrated that triple negative breast cancer tissue had high GPC and low PCho concentrations, whereas tissue from ER+/PgR+ patients had low GPC and high PCho. There was a significant difference in the GPC/PCho peak area ratio between ER+/PgR+ and triple negative samples (0.8 ± 0.5 and 1.5 ± 0.7, respectively. $p = 0.01$), corresponding to the findings from the xenograft models. The mean spectra from human tissue samples also suggested that the glycine concentration was higher in triple negative breast cancer tissue samples. Using the glycine/total peak area ratio as marker for glycine content, this trend was not statistically significant ($p = 0.19$). The relative choline peak area was significantly higher in triple negative tissue ($p < 0.00003$) and the relative creatine peak area was significantly lower ($p = 0.024$).

Gene expression analysis of xenograft tissue

Of the 119 investigated genes, 67 were differentially expressed between the xenograft models at a 5% adjusted (false discovery rate) significance level. Microarray data from earlier passages of the same xenograft models [29] showed similar trends of differential expression (data not shown). The complete results from the gene expression analysis are available as additional file. A heatmap of the differentially expressed genes is presented in Figure 2, with hierarchical clustering of genes and samples. In the following sections, only genes directly involved in synthesis and degradation of PtdCho from choline are considered.

Among the five selected genes coding for proteins known to be involved in transmembrane choline transport, only solute carrier family 44, member 1 (*SLC44A1*) showed significantly different expression between the two models. The expression of this transporter, also known as choline transporter-like protein 1 (*CTL1*), was

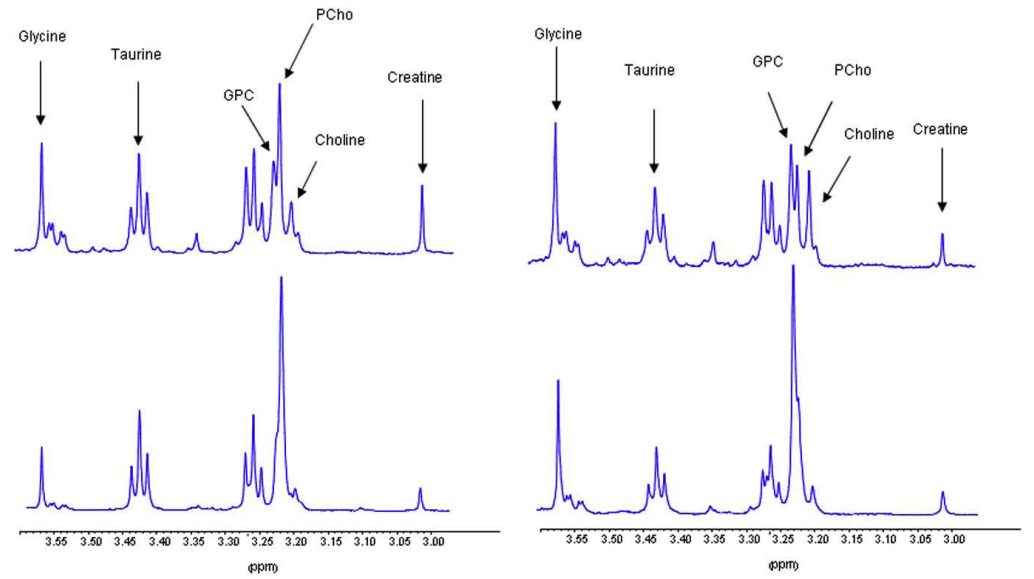


Figure 1 Mean HR MAS MRS spectra from human tissue samples and xenograft tissue. HR MAS MRS spectra (spectral region 3.6 to 3.0 ppm) of ER+/PgR+ (top left) and triple negative (top right) human tissue samples, and luminal-like (bottom left) and basal-like (bottom right) xenograft samples. Spectral assignments are provided for peaks used in quantification.

lower in basal-like than luminal-like xenografts. Solute carrier family 22, member 1 (*SLC22A1*) and solute carrier family 44, member 2 (*SLC44A2*) were similarly expressed in the two models, whereas solute carrier family 5 (choline transporter), member 7 (*SLC5A7*) and solute carrier family 22 (organic cation transporter), member 2 (*SLC22A2*), were expressed below the limit of detection. *SLC5A7* is also known as choline transporter 1 (*CHT1*), a high-affinity choline-specific transporter protein, whereas *SLC44A2* is also known as choline-transporter like protein 2 (*CTL2*).

Genes directly involved in choline metabolism which were differentially expressed between the xenograft models are listed in Table 4 and 5. A schematic overview of intracellular choline metabolite concentrations and the comparative gene expression (the anabolic Kennedy pathway, PtdCho breakdown and conversion of choline to glycine) between the xenograft models is presented in Figure 3. As shown in Figure 3, choline is converted to PCho through the action of two isoforms of the same enzyme, choline kinase alpha and beta. The expression of the genes (*CHKA*, *CHKB*) coding for both isoforms was significantly lower in the basal-like than in the luminal-like model.

Table 3 Metabolite concentrations

	MAS98.12 (n = 10)	MAS98.06 (n = 9)
Creatine	4.1 ± 1.4	3.4 ± 1.7
Choline	1.2 ± 0.7	0.9 ± 0.6
Phosphocholine *	4.5 ± 2.1	9.1 ± 4.4
Glycerophosphocholine **	9.8 ± 2.5	2.7 ± 1.7
Taurine	14.7 ± 4.1	19.1 ± 9.1
Glycine *	8.2 ± 3.0	4.0 ± 1.8

Metabolite concentrations in basal-like (MAS98.12) and luminal-like (MAS98.06) xenografts calculated from HR MAS MRS spectra using the ERETIC method (μmol/g, mean ± SD, * $p < 0.01$, ** $p < 0.001$)

Conversion of PCho to CDP-choline is mediated through the alpha and beta isoforms of phosphate cytidylyl transferase 1 (*PCYT1A* and *PCYT1B*). The expression of *PCYT1B* was significantly higher, and *PCYT1A* was significantly lower in the basal-like than the luminal-like model. The gene coding for choline phosphotransferase 1 (*CHPT1*), which converts CDP-PCho to PtdCho, had a significantly lower expression level in the basal-like than in the luminal-like model.

PtdCho is degraded by several different phospholipases. Enzymes in the phosphoplipase A2 group (*PLA2*) convert PtdCho to acyl-GPC. Of the 13 *PLA2* isoforms studied, two were significantly higher expressed in the basal-like model, three were significantly lower expressed, five showed no significant difference in expression and three were below the limit of detection.

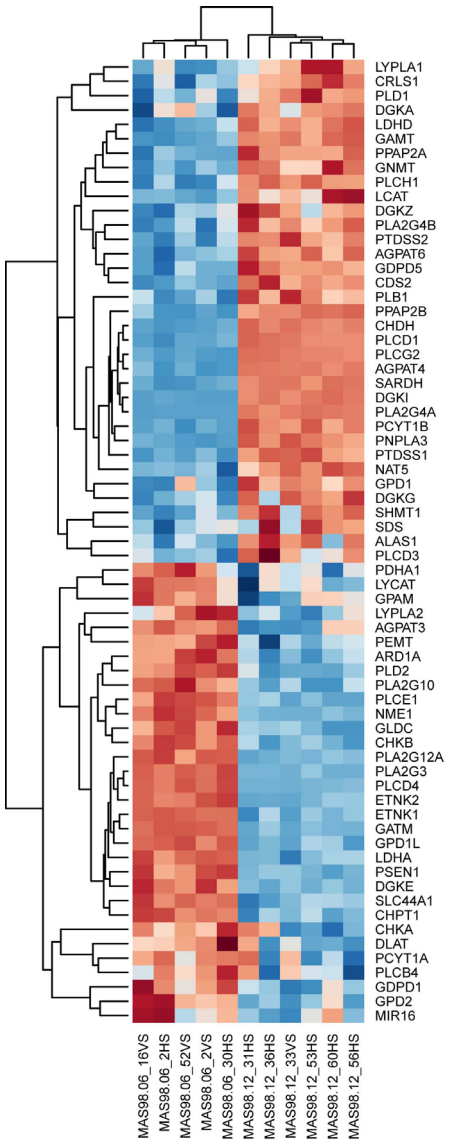


Figure 2 Heatmap of differentially expressed genes in xenograft models. Hierarchical clustering of the 67 differentially expressed genes (false discovery rate < 0.05) involved in the KEGG *homo sapiens* glycerophospholipid pathway hsa:00564, choline transport or directly involved in conversion from choline to glycine (Red: high expression compared to mean expression in xenograft samples. Blue: low expression compared to mean expression in xenograft samples). The microarray data was centred across genes and clustered across genes and samples using Euclidian distance and complete linkage.

The largest difference in gene expression between the two models was found for *PLA2G4A*, where a log2-fold difference of 6.4 in gene expression was observed. Phospholipase B1 (*PLB1*) is involved in both deacetylation steps from PtdCho to GPC, and was significantly higher expressed in the basal-like than in the luminal-like model.

Phospholipase D, with the two isoforms *PLD1* and *PLD2*, converts PtdCho to choline. The expression of *PLD1* was significantly higher and the expression of *PLD2* was significantly lower in the basal-like compared to the luminal-like model. Other genes related to the degradation of PtdCho, such as lecithin-cholesterol acetyltransferase (*LCAT*) and phosphatidylserine synthase 1 (*PTDSS1*) also had significantly higher expression levels in the basal-like than in the luminal-like model. The *GDPD5* gene, coding for glycerophosphodiester phosphodiesterase (GDPD) was significantly higher expressed in the basal-like model than in the luminal-like model, indicating that GPC degradation may occur at a higher rate in basal-like xenografts. However, an isoform of this gene, *GDPD1*, was higher expressed in the luminal-like xenografts. As shown in Figure 3, choline dehydrogenase (*CHDH*) mediates the irreversible conversion of choline to betaine, which is a key precursor in the synthesis of glycine. The expression of *CHDH* was significantly higher in the basal-like than in the luminal-like model. Sarcosine dehydrogenase (*SARDH*), involved in the conversion of betaine to glycine, also had significantly higher expression levels in the basal-like model.

Discussion

The HR MAS MRS data demonstrated significant differences in choline metabolite pattern between the basal-like and luminal-like xenograft models. In particular, the difference in GPC and PCho concentrations is an interesting finding, as the pattern seen in the basal-like model does not correspond to typical *in vitro* choline metabolite patterns [15,17]. In addition, expression data showed that several genes directly associated with choline metabolism differed significantly between the two models. Differences in expression of genes involved in choline metabolism corresponded to differences in metabolite concentrations, suggesting that transcriptional differences between the models are reflected in the HR MAS MRS spectra. The relative amounts of GPC and PCho in human tissue samples from triple negative and ER+/PgR+ subtypes of breast cancer corresponded well with the data from the xenografts.

In order to evaluate if the choline metabolism in the xenograft models is representative for basal-like and luminal-like breast cancer in humans, they were compared to data from triple negative and ER+/PgR+ breast cancer patients. It is assumed that the triple-negative

Table 4 Differentially expressed genes

Entrez ID	Probe name	Gene name	Encoded protein	Log2-fold difference	Adjusted p-value (false discovery rate)
5321	A_23_P11685	PLA2G4A	Phospholipase A2, group IV A	6.4	$4.4E^{-16}$
55349	A_23_P69293	CHDH	Choline dehydrogenase	3.3	$4.0E^{-13}$
1757	A_24_P35400	SARDH	Sarcosine dehydrogenase	2.5	$7.6E^{-12}$
9468	A_24_P941353	PCYT1B	Phosphate cytidylyltransferase 1, choline, beta	1.7	$3.7E^{-9}$
31896	A_23_P87401	GDPD5	Glycerophosphodiester phosphodiesterase domain containing 5	1.0	$4.2E^{-6}$
8681	A_23_P403424	PLA2G4B	Phospholipase A2, group IV B	0.9	$9.8E^{-5}$
9791	A_23_P168868	PTDSS1	Phosphatidylserine synthase I	0.9	$9.6E^{-7}$
10434	A_23_P19192	LYPLA1	Lysophospholipase 1	0.9	0.001
3931	A_23_P218237	LCAT	Lecithin-cholesterol acyltransferase	0.8	0.0002
151056	A_23_P56356	PLB1	Phospholipase B1	0.7	0.0001
5337	A_23_P155335	PLD1	Phospholipase D1	0.7	0.0005

Genes directly involved in choline metabolism with significantly higher expression in basal-like (MAS98.12) than luminal-like (MAS98.06) tumors

phenotype is a valid surrogate marker for basal-like breast cancer, as approximately 90% of triple-negative breast carcinomas can be classified as basal-like based on the intrinsic molecular subtyping developed by Sørlie *et al* [3,37]. On the other hand, expression of estrogen and/or progesterone receptors is a typical feature of luminal A and B subtypes, whereas the ERBB2 and basal-like subtypes of breast cancer rarely express hormone receptors [38,39]. Therefore, the ER+/PgR+ phenotype is considered to be a valid surrogate marker for luminal-like subtypes of breast cancer.

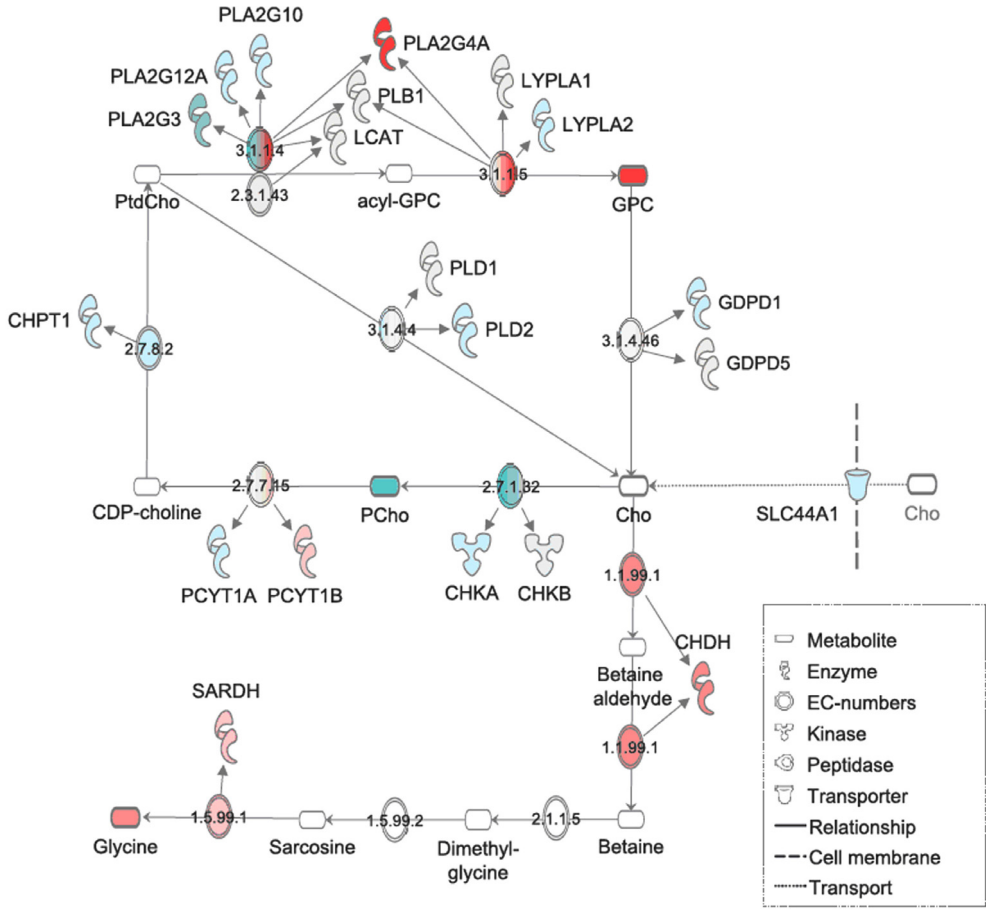
Using gene expression profiling, the molecular causes for the differences in choline metabolism was further explored in the xenograft models. The heatmap of all 64 significantly differentially expressed genes in Figure 2, clearly shows that different sets of genes related to phospholipid metabolism are higher expressed the basal-like model compared to the luminal-like models. This indicates that the regulation of choline metabolism differ between the two xenograft models. Although this study does not provide data on metabolic flux, the methods used are suitable for highlighting key steps in choline metabolism. Comparison of these two disease models does not, however, give any information with respect to the difference between choline metabolism in normal breast versus breast cancer tissue. Nevertheless, gene expression profiling of the xenograft models showed significant differences in the expression of genes directly involved in choline metabolism, suggesting that these genes may play key roles in regulation of choline metabolite concentrations in human breast cancer.

Increased choline transport has been associated with the abnormally high concentrations of PCho observed in breast cancer [16,27,40]. In our study, the influx of choline in the two models could not be fully evaluated from the gene expression data, as only one of five investigated

Table 5 Differentially expressed genes

Entrez ID	Probe name	Gene name	Encoded protein	Log2-fold difference	Adjusted p-value (false discovery rate)
50487	A_23_P17814	PLA2G3	Phospholipase A2, group III	-3.1	$1.1E^{-12}$
81579	A_23_P30020	PLA2G12A	Phospholipase A2, group XII A	-1.6	$2.5E^{-9}$
56994	A_23_P105571	CHPT1	Choline phosphotransferase 1	-1.4	$1.6E^{-7}$
23446	A_23_P216630	SLC44A1	Solute carrier family 44, member 1 (CTL1)	-1.1	$7.9E^{-7}$
5338	A_23_P4308	PLD2	Phospholipase D2	-1.1	$1.2E^{-6}$
8399	A_23_P88767	PLA2G10	Phospholipase A2, group X	-1.1	$3.1E^{-6}$
1119	A_23_P314120	CHKB	Choline kinase beta	-0.8	$2.0E^{-6}$
11313	A_24_P276490	LYPLA2	Lysophospholipase II	-0.4	0.005
5130	A_23_P252681	PCYT1A	Phosphate cytidylyltransferase 1, choline, alpha	-0.4	0.035
24657	A_23_P84666	GDPD1	Glycerophosphodiester phosphodiesterase domain containing I	-0.4	0.005
1119	A_23_P124742	CHKA	Choline kinase, alpha	-0.3	0.047

Genes directly involved in choline metabolism with significantly lower expression in basal-like (MAS98.12) than luminal-like (MAS98.06) tumors



© 2000-2010 Ingenuity Systems, Inc. All rights reserved.

Figure 3 Differences in gene expression and metabolite concentrations. Differences in intracellular choline metabolism, choline transport and glycine formation between basal-like and luminal-like xenografts. Red: higher gene expression/metabolite concentration in basal-like xenografts. Blue: higher gene expression/metabolite concentration in luminal-like xenografts. White: Non-significant difference in expression/concentration. Log2-fold difference between models is indicated by color intensity. Only significantly differentially expressed genes are illustrated at each EC number, which represents specific enzymatic reactions. The metabolites which were measured with HR MAS MRS are outlined with a thicker line.

choline transporters was differentially expressed. Choline transport has been shown to be less important than PtdCho turnover for total choline metabolite concentrations [17]. Differences in choline uptake may still have impact on the choline metabolite concentrations, and specific studies using isotopically labelled choline could possibly allow accurate evaluation of choline transport rate in the two xenograft models.

In breast cancer cells, the intracellular metabolism of choline is divided in two major pathways as shown in Figure 3: Betaine production or PtdCho synthesis [27,34]. In the betaine synthesis pathway, choline is oxidized to betaine through the action of choline dehydrogenase (*CHDH*). Betaine is then demethylated to glycine. *In vitro* studies of MCF7-cells have shown that PtdCho synthesis is the pathway predominantly accountable for choline turnover [34]. The first step in the PtdCho synthesis pathway is the phosphorylation of choline through choline kinase, yielding PCho (Figure 3). It has been shown that increased expression of *CHKA* is

critical for proliferation both of mammary epithelial cells and breast cancer [41], but *in vitro* studies of different breast cancer cell lines have not conclusively demonstrated a correlation between *CHKA* expression and PCho concentration [16,26]. In our study, the expression of *CHKA* and *CHKB* was significantly lower in the basal-like than in the luminal-like model, although some variability in expression was observed (Figure 2). This is consistent with the lower PCho concentrations measured in the basal-like model. Betaine production is thought to contribute only slightly to the overall conversion of choline, and neither choline transport nor GPC degradation is conclusive with respect to their contribution to the choline pool. As normal breast tissue or benign breast lesions rarely exhibit increased choline metabolite levels, the xenograft models are believed to represent typical choline metabolism abnormalities of breast carcinomas [42,43]. Therefore, it should be stressed that *CHKA* and *CHKB* expression is likely to be upregulated in both xenograft models compared to normal breast tissue. The lower PCho concentrations in the basal-like xenografts may also in part be a result of higher *CHDH* expression. This suggests that conversion of choline to betaine is upregulated, shifting the metabolic flux in favour of glycine formation. *SARDH*, related to conversion of betaine to glycine, was also significantly higher expressed in the basal-like model. The concentration of glycine in the basal-like model was indeed higher than in the luminal-like model, suggesting that there is a difference in choline routing and glycine production between the two breast cancer subtypes. An association between tumor aggressiveness and glycine concentration has been noted also in clinical breast cancer tissue biopsies [21]. Abnormalities in cancer energy metabolism are widely recognized, and differences in glycine concentration between the two xenograft models in this study could well be an indirect result of this phenomenon.

Degradation of PtdCho is the primary source of GPC. The expression of *PLA2G4A*, *PLA2G4B*, *LCAT*, *LYPLA1* and *PLB1*, which all are associated with this pathway, was higher in the basal-like model. Other genes (*PLA2G3*, *PLA2G12A*, *PLA2G10*, *LYPLA2*) were lower expressed in the basal-like model, and a clear association between PtdCho degradation and GPC concentration could not be concluded. However, *in vitro* studies have suggested that GPC concentrations are associated with *PLA2G4A* levels, which is consistent with our findings [17]. A lower rate of GPC degradation could account for the higher GPC concentration observed in the basal-like xenograft model. The expression of *GDPD5* was, however, higher in basal-like xenografts. The observed differences between the two models in the relative expression of different genes assigned to the abovementioned enzymatic steps could be reflecting the relative importance of different gene products coding for proteins with the same enzymatic activity in the two models.

By associating choline metabolite concentrations with tumor cell phenotype, it has been proposed that PCho concentration increase with the malignancy of the tumor cell line when grown in culture [15]. However, other *in vitro* studies have failed to show a correlation between malignancy and choline metabolite concentrations [16]. It has been suggested that differences in experimental design, particularly the stage of cell growth, are accountable for these discrepancies [26]. In all the abovementioned *in vitro* studies of breast cancer cells, PCho concentrations were significantly higher than GPC concentrations. However, both in xenograft models of breast cancer and in clinical tissue samples, GPC concentrations higher than PCho concentrations have been observed [21,44]. GPC concentration has been shown to be negatively correlated with estrogen receptor content in breast carcinomas, which agrees with the relatively high GPC content in the basal-like xenograft [45]. Our data show that GPC concentration is significantly lower and PCho concentration is significantly higher in the luminal-like animal model, which represents a less aggressive disease than the basal-like model. This suggests that the relationship between choline metabolite concentrations and malignancy of solid tumors is more complex than indicated by studies of breast cancer cell lines. Discrepancies between *in vitro* data and clinical data may be attributed to the microenvironment of solid tumors. It has recently been shown that the metabolic profiles change when the same breast cancer cell lines are studied both *in vitro* and *in vivo* [46]. In addition, *in vitro* simulation of microenvironmental factors in solid tumors has demonstrated that PCho and GPC concentrations respond to changes in acidity, oxygenation level and glucose accessibility [44].

The relevance of the basal-like and luminal-like xenografts used in this study was further supported by comparing the choline metabolite pattern with that of human tissue samples from ER+/PgR+ and triple negative breast cancer. Evaluation of metabolite levels through relative peak areas demonstrated that the mean GPC/PCho ratio was significantly higher in triple negative breast cancer than in ER+/PgR+ breast cancer. The relative PCho area was significantly higher in ER+/PgR+ samples than in samples from triple negative breast cancer. A trend towards higher glycine concentration was also found in triple negative tissue samples. Interestingly, the choline concentration in triple negative breast cancer was higher than in ER+/PgR+ breast cancer. Overall, the striking similarity between xenografts and human tissue samples with respect to GPC and PCho levels suggest that the xenografts have maintained genetic and/or microenvironmental features from the primary carcinomas which are relevant for the choline

metabolite pattern. The spectra from human tissue samples also suggest that PCho concentrations alone are not a reliable prognostic biomarker. The triple negative samples represent disease with poor prognosis, yet the PCho level in these samples appear to be significantly lower than in ER+/PgR+ samples. This finding encourages large-scale studies of the metabolite pattern and gene expression in the different molecular subtypes of breast cancer, as this may reveal new drug targets or suggest strategies for individualised therapy using drugs targeting the choline metabolism pathways.

When interpreting the gene expression data from the two xenograft models, it should be kept in mind that gene expression not always represents the actual enzymatic activity. Isoforms of the same enzyme may exhibit differences in transcriptional regulation, and mRNA concentrations do not account for post-translational modification of enzymes. In addition, the concentrations of all investigated choline-containing compounds are determined by more than one metabolic reaction. Thus, a simplistic model for correlating gene expression with metabolite concentration is not applicable. The net rate of all relevant metabolic reactions governs the metabolite concentrations, and the relative importance of each metabolic reaction is unknown. This must be kept in mind when interpreting the data. However, the gene expression data provide significant information in terms of highlighting the reactions that are most likely to be relevant for the observed differences in metabolic pattern. Hypotheses generated on the basis of microarray data should be evaluated by tracking the flux of metabolites through the different pathways.

Comparing our data with pre-existing studies of choline metabolism in cultured cells and *in vivo* models with data from human biopsies, we suggest that primary tumor xenografts are more relevant model systems than cell cultures with respect to investigation of metabolic profiles in different breast cancer subtypes, and may be a better approach to studies of therapeutic efficacy in the different breast cancer subtypes. As the choline metabolite profile of the xenograft models used in the study appear representative of basal-like and luminal-like human breast cancer, the models are considered valuable tools for testing of targeted drugs and for monitoring response to treatment in these subtypes of breast cancer.

Conclusions

HR MAS MRS analyses of a basal-like and a luminal-like xenograft model demonstrated significant differences in choline metabolite concentrations. In the more aggressive basal-like tumor, GPC concentrations were higher than PCho concentrations, whereas this pattern was reversed in the luminal-like model. Glycine concentration was also significantly higher in the basal-like model. These differences could at least in part be explained by lower choline kinase expression and increased PtdCho degradation in the basal-like model. The gene expression data also suggested a possible shift in metabolic flux from PtdCho synthesis to glycine formation in the basal-like model. The choline metabolism pattern in the xenografts corresponded well with spectra from tissue samples from triple negative and ER+/PgR+ human breast carcinomas, suggesting that the basal-like and luminal-like xenografts may be a relevant model system for studies of choline metabolism in these two subtypes of human breast cancer.

Additional material

Additional file 1: Differential gene expression. Excel spreadsheet containing results from the differential gene expression analysis of the 119 investigated genes.

Acknowledgements

The authors wish to thank Unn Granli and Borgny Ytterhus for staining of histology sections and Alexandr Kristian for assistance in animal surgery. The work presented is sponsored by the Research Council of Norway, grants no. 175459/V50, 186479/V50, 183379/S10 and 183621/S10.

Author details

[1]Department of Circulation and Medical Imaging, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [2]Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway. [3]Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway. [4]Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway.

Authors' contributions

SAM conceived and designed the study, performed the HR MAS MRS analysis, performed the histopathological analysis, interpreted the data and wrote the manuscript. EB performed the gene expression analysis and interpreted the microarray data. EMH performed the HR MAS MRS analysis. EL carried out the *in vivo* experiments. BS established the HR MAS MRS protocol and supervised the analyses. ALBD contributed with expertise in molecular biology techniques. OE designed and coordinated the *in vivo* experiments. GMM participated in design and coordination of the study. ISG designed and coordinated the study and gave final approval of the manuscript. All co-authors critically revised the manuscript and approved the final version.

Competing interests

The authors declare that they have no competing interests.

Received: 15 April 2010 Accepted: 17 August 2010
Published: 17 August 2010

References

1. Perou CM, Sorlie T, Eisen MB, van de RM, Jeffrey SS, Rees CA, *et al*: **Molecular portraits of human breast tumours.** *Nature* 2000, **406**:747-752.
2. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, *et al*: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc Natl Acad Sci USA* 2001, **98**:10869-10874.
3. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, *et al*: **Repeated observation of breast tumor subtypes in independent gene expression data sets.** *Proc Natl Acad Sci USA* 2003, **100**:8418-8423.
4. Calza S, Hall P, Auer G, Bjohle J, Klaar S, Kronenwett U, *et al*: **Intrinsic molecular signature of breast cancer in a population-based cohort of 412 patients.** *Breast Cancer Res* 2006, **8**:R34.

5. Kurebayashi J, Moriya T, Ishida T, Hirakawa H, Kurosumi M, Akiyama F, *et al*: **The prevalence of intrinsic subtypes and prognosis in breast cancer patients of different races.** *Breast* 2007, **16(Suppl 2)**:S72-S77.
6. Mullins M, Perreard L, Quackenbush JF, Gauthier N, Bayer S, Ellis M, *et al*: **Agreement in breast cancer classification between microarray and quantitative reverse transcription PCR from fresh-frozen and formalin-fixed, paraffin-embedded tissues.** *Clin Chem* 2007, **53**:1273-1279.
7. Sorlie T, Wang Y, Xiao C, Johnsen H, Naume B, Samaha RR, *et al*: **Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms.** *BMC Genomics* 2006, **7**:127.
8. Stadler ZK, Come SE: **Review of gene-expression profiling and its clinical use in breast cancer.** *Crit Rev Oncol Hematol* 2009, **69**:1-11.
9. Banez-Coronel M, de Molina AR, Rodriguez-Gonzalez A, Sarmentero J, Ramos MA, Garcia-Cabezas MA, *et al*: **Choline kinase alpha depletion selectively kills tumoral cells.** *Curr Cancer Drug Targets* 2008, **8**:709-719.
10. Hernandez-Alcoceba R, Fernandez F, Lacal JC: **In vivo antitumor activity of choline kinase inhibitors: a novel target for anticancer drug discovery.** *Cancer Res* 1999, **59**:3112-3118.
11. Lacal JC: **Choline kinase: a novel target for antitumor drugs.** *IDrugs* 2001, **4**:419-426.
12. Rodriguez-Gonzalez A, Ramirez de MA, Fernandez F, Ramos MA, del Carmen NM, Campos J, *et al*: **Inhibition of choline kinase as a specific cytotoxic strategy in oncogene-transformed cells.** *Oncogene* 2003, **22**:8803-8812.
13. Rodriguez-Gonzalez A, Ramirez de MA, Banez-Coronel M, Megias D, Lacal JC: **Inhibition of choline kinase renders a highly selective cytotoxic effect in tumour cells through a mitochondrial independent mechanism.** *Int J Oncol* 2005, **26**:999-1008.
14. Tozaki M: **Proton MR spectroscopy of the breast.** *Breast Cancer* 2008, **15**:218-223.
15. Aboagye EO, Bhujwalla ZM: **Malignant transformation alters membrane choline phospholipid metabolism of human mammary epithelial cells.** *Cancer Res* 1999, **59**:80-84.
16. Eliyahu G, Kreizman T, Degani H: **Phosphocholine as a biomarker of breast cancer: molecular and biochemical studies.** *Int J Cancer* 2007, **120**:1721-1730.
17. Glunde K, Jie C, Bhujwalla ZM: **Molecular causes of the aberrant choline phospholipid metabolism in breast cancer.** *Cancer Res* 2004, **64**:4270-4276.
18. Gillies RJ, Morse DL: **In vivo magnetic resonance spectroscopy in cancer.** *Annu Rev Biomed Eng* 2005, **7**:287-326.
19. Mackinnon WB, Barry PA, Malycha PL, Gillett DJ, Russell P, Lean CL, *et al*: **Fine-needle biopsy specimens of benign breast lesions distinguished from invasive cancer ex vivo with proton MR spectroscopy.** *Radiology* 1997, **204**:661-666.
20. Negendank W: **Studies of human tumors by MRS: a review.** *NMR Biomed* 1992, **5**:303-324.
21. Sitter B, Lundgren S, Bathen TF, Halgunset J, Fjosne HE, Gribbestad IS: **Comparison of HR MAS MR spectroscopic profiles of breast cancer tissue with clinical parameters.** *NMR Biomed* 2006, **19**:30-40.
22. Meisamy S, Bolan PJ, Baker EH, Bliss RL, Gulbahce E, Everson LI, *et al*: **Neoadjuvant chemotherapy of locally advanced breast cancer: predicting response with in vivo (1)H MR spectroscopy–a pilot study at 4 T.** *Radiology* 2004, **233**:424-431.
23. Morse DL, Raghunand N, Sadarangani P, Murthi S, Job C, Day S, *et al*: **Response of choline metabolites to docetaxel therapy is quantified in vivo by localized (31)P MRS of human breast cancer xenografts and in vitro by high-resolution (31)P NMR spectroscopy of cell extracts.** *Magn Reson Med* 2007, **58**:270-280.
24. Neeman M, Eldar H, Rushkin E, Degani H: **Chemotherapy-induced changes in the energetics of human breast cancer cells; 31P- and 13C-NMR studies.** *Biochim Biophys Acta* 1990, **1052**:255-263.
25. Glunde K, Jie C, Bhujwalla ZM: **Mechanisms of indomethacin-induced alterations in the choline phospholipid metabolism of breast cancer cells.** *Neoplasia* 2006, **8**:758-771.
26. Morse DL, Carroll D, Day S, Gray H, Sadarangani P, Murthi S, *et al*: **Characterization of breast cancers and therapy response by MRS and quantitative gene expression profiling in the choline pathway.** *NMR Biomed* 2009, **22**:114-127.
27. Katz-Brull R, Seger D, Rivenson-Segal D, Rushkin E, Degani H: **Metabolic markers of breast cancer: enhanced choline metabolism and reduced choline-ether-phospholipid synthesis.** *Cancer Res* 2002, **62**:1966-1970.
28. Vargo-Gogola T, Rosen JM: **Modelling breast cancer: one size does not fit all.** *Nat Rev Cancer* 2007, **7**:659-672.
29. Bergamaschi A, Hjortland GO, Triulzi T, Sorlie T, Johnsen H, Ree AH, *et al*: **Molecular profiling and characterization of luminal-like and basal-like in vivo breast cancer xenograft models.** *Mol Oncol* 2009, **3**:469-482.
30. Sitter B, Sonnewald U, Spraul M, Fjosne HE, Gribbestad IS: **High-resolution magic angle spinning MRS of breast cancer tissue.** *NMR Biomed* 2002, **15**:327-337.
31. Sitter B, Bathen TF, Singstad TE, Fjosne HE, Lundgren S, Halgunset J, *et al*: **Quantification of metabolites in breast cancer patients with different clinical prognosis using HR MAS MR spectroscopy.** *NMR Biomed* 2010, **23**:424-431.
32. Smyth GK: **Linear models and empirical bayes methods for assessing differential expression in microarray experiments.** *Stat Appl Genet Mol Biol* 2004, **3, Article3.**
33. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes.** *Nucleic Acids Res* 2000, **28**:27-30.
34. Katz-Brull R, Margalit R, Degani H: **Differential routing of choline in implanted breast cancer and normal organs.** *Magn Reson Med* 2001, **46**:31-38.
35. Michel V, Yuan Z, Ramsubir S, Bakovic M: **Choline transport for phospholipid synthesis.** *Exp Biol Med (Maywood)* 2006, **231**:490-504.
36. Gallazzini M, Ferraris JD, Burg MB: **GDPD5 is a glycerophosphocholine phosphodiesterase that osmotically regulates the osmoprotective organic osmolyte GPC.** *Proc Natl Acad Sci USA* 2008, **105**:11026-11031.
37. Kreike B, van KM, Horlings H, Weigelt B, Peterse H, Bartelink H, *et al*: **Gene expression profiling and histopathological characterization of triple-negative/basal-like breast carcinomas.** *Breast Cancer Res* 2007, **9**:R65.
38. Sorlie T: **Molecular portraits of breast cancer: tumour subtypes as distinct disease entities.** *Eur J Cancer* 2004, **40**:2667-2675.
39. Bhargava R, Striebel J, Beriwal S, Flickinger JC, Onisko A, Ahrendt G, *et al*: **Prevalence, morphologic features and proliferation indices of breast carcinoma molecular classes using immunohistochemical surrogate markers.** *Int J Clin Exp Pathol* 2009, **2**:444-455.
40. Katz-Brull R, Degani H: **Kinetics of choline transport and phosphorylation in human breast cancer cells; NMR application of the zero trans method.** *Anticancer Res* 1996, **16**:1375-1380.
41. Ramirez de MA, Banez-Coronel M, Gutierrez R, Rodriguez-Gonzalez A, Olmeda D, Megias D, *et al*: **Choline kinase activation is a critical requirement for the proliferation of primary human mammary epithelial cells and breast tumor progression.** *Cancer Res* 2004, **64**:6732-6739.
42. Gribbestad IS, Petersen SB, Fjosne HE, Kvinnsland S, Krane J: **1H NMR spectroscopic characterization of perchloric acid extracts from breast carcinomas and non-involved breast tissue.** *NMR Biomed* 1994, **7**:181-194.
43. Kvistad KA, Bakken IJ, Gribbestad IS, Ehrnholm B, Lundgren S, Fjosne HE, *et al*: **Characterization of neoplastic and normal human breast tissues with in vivo (1)H MR spectroscopy.** *J Magn Reson Imaging* 1999, **10**:159-164.
44. Eliyahu G, Maril N, Margalit R, Degani H: **Choline Metabolism in Breast Cancer; The Influence of the Microenvironmental conditions [Abstract].** *Proc Intl Soc Mag Reson Med* 2007, **15.**
45. Giskeodegard GF, Grinde MT, Sitter B, Axelson DE, Lundgren S, Fjosne HE, *et al*: **Multivariate modeling and prediction of breast cancer prognostic factors using MR metabolomics.** *J Proteome Res* 2010, **9**:972-979.
46. Mori N, Glunde K, Takagi T, Xiong L, Wides F, Bhujwalla Z: **Tumor microenvironmental alterations of lipid metabolism detected by comparing cancer cells with tumors [Abstract].** *Proc Intl Soc Mag Reson Med* 2009, **17**:2310.

Pre-publication history

The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2407/10/433/prepub

doi:10.1186/1471-2407-10-433
Cite this article as: Moestue *et al.*: Distinct choline metabolic profiles are associated with differences in gene expression for basal-like and luminal-like breast cancer xenograft models. *BMC Cancer* 2010 **10**:433.

Additional files

Additional file 1:
Differential gene expression. Excel spreadsheet containing results from the differential gene expression analysis of the 119 investigated genes.

The complete file is available at: http://www.biomedcentral.com/1471-2407/10/433/additional/

Additional File 1

Feature Num	ProbeUID	ProbeName	GeneName	Systematic Name	Description	logFC (estimate of the log2-fold-change in MAS_98.12 compared to MAS_98.06)	Average expression	t-value	P-Value	adjusted p-value (false discovery rate)
17467	15927	A_23_P385105	PLCD4	NM_032726	Homo sapiens phospholipase C, delta 4 (PLCD4), mRNA [NM_032726]	-5.2	7.4	-35	1.0E-15	2.7E-14
30806	28226	A_23_P129064	GATM	NM_001482	Homo sapiens glycine amidinotransferase (L-arginine:glycine amidinotransferase) (GATM), nuclear gene encoding mitochondrial protein, mRNA [NM_001482]	-4.7	8.8	-31	6.6E-15	1.4E-13
16765	15265	A_23_P17814	PLA2G3	NM_015715	Homo sapiens phospholipase A2, group III (PLA2G3), mRNA [NM_015715]	-3.1	8.7	-26	8.2E-14	1.1E-12
15920	14494	A_23_P46025	ETNK2	NM_018208	Homo sapiens ethanolamine kinase 2 (ETNK2), mRNA [NM_018208]	-4.4	7.7	-21	1.7E-12	1.6E-11
24754	22586	A_23_P318284	GPD1L	NM_015141	Homo sapiens glycerol-3-phosphate dehydrogenase 1-like (GPD1L), mRNA [NM_015141]	-1.9	11.6	-15	3.0E-10	2.1E-09
43391	39696	A_23_P30020	PLA2G12A	NM_030821	Homo sapiens phospholipase A2, group XIIA (PLA2G12A), mRNA [NM_030821]	-1.6	11.0	-14	3.9E-10	2.5E-09
43888	40094	A_23_P35617	PLCE1	NM_016341	Homo sapiens phospholipase C, epsilon 1 (PLCE1), mRNA [NM_016341]	-2.2	11.1	-14	5.5E-10	3.2E-09
3994	3647	A_23_P419239	ETNK1	NM_018638	Homo sapiens ethanolamine kinase 1 (ETNK1), transcript variant 1, mRNA [NM_018638]	-1.7	9.3	-13	1.2E-09	6.6E-09
45188	41053	A_23_P123596	GLDC	NM_000170	Homo sapiens glycine dehydrogenase (decarboxylating) (GLDC), mRNA [NM_000170]	-3.2	9.8	-11	1.0E-08	5.4E-08
42464	38814	A_23_P152804	NME1	NM_198175	Homo sapiens non-metastatic cells 1, protein (NM23A) expressed in (NME1), transcript variant 1, mRNA [NM_198175]	-1.1	15.3	-11	1.2E-08	6.1E-08
30230	27628	A_32_P44568	LDHA	NM_005566	Homo sapiens lactate dehydrogenase A (LDHA), mRNA [NM_005566]	-1.5	12.1	-11	1.4E-08	6.8E-08
33472	30619	A_23_P105571	CHPT1	NM_020244	Homo sapiens choline phosphotransferase 1 (CHPT1), mRNA [NM_020244]	-1.4	10.8	-10	3.5E-08	1.6E-07
10075	9259	A_24_P916927	DGKE	BC022297	Homo sapiens diacylglycerol kinase, epsilon 64kDa, mRNA (cDNA clone IMAGE:4839690), complete cds. [BC022297]	-1.2	5.6	-10	4.1E-08	1.8E-07
8139	7416	A_23_P216630	SLC44A1	NM_080546	Homo sapiens solute carrier family 44, member 1 (SLC44A1), mRNA [NM_080546]	-1.1	13.2	-9.0	2.0E-07	7.9E-07
45099	41044	A_23_P106174	PSEN1	AJ008005	Homo sapiens PSN1 gene, alternative transcript. [AJ008005]	-0.9	7.4	-8.9	2.3E-07	8.8E-07
32905	30051	A_23_P4308	PLD2	NM_002663	Homo sapiens phospholipase D2 (PLD2), mRNA [NM_002663]	-1.1	6.6	-8.7	3.4E-07	1.2E-06
31691	29075	A_23_P314120	CHKB	NM_005198	Homo sapiens choline kinase beta (CHKB), transcript variant 1, mRNA [NM_005198]	-0.8	9.2	-8.3	6.1E-07	2.0E-06
43994	40135	A_23_P88767	PLA2G10	NM_003561	Homo sapiens phospholipase A2, group X (PLA2G10), mRNA [NM_003561]	-1.1	6.0	-8.0	9.8E-07	3.1E-06
25824	23604	A_23_P148546	ARD1A	NM_003491	Homo sapiens ARD1 homolog A, N-acetyltransferase (S. cerevisiae) (ARD1A), mRNA [NM_003491]	-0.8	11.1	-6.7	7.5E-06	2.1E-05
1461	1142	A_24_P344711	AGPAT3	NM_020132	Homo sapiens 1-acylglycerol-3-phosphate O-acyltransferase 3 (AGPAT3), transcript variant 1, mRNA [NM_020132]	-2.3	8.8	-5.6	5.7E-05	1.4E-04
39665	36335	A_23_P163955	PEMT	NM_007169	Homo sapiens phosphatidylethanolamine N-methyltransferase (PEMT), nuclear gene encoding mitochondrial protein, transcript variant 2, mRNA [NM_007169]	-1.0	10.0	-5.6	5.5E-05	1.4E-04
32069	29301	A_23_P119812	GPD2	NM_000408	Homo sapiens glycerol-3-phosphate dehydrogenase 2 (mitochondrial) (GPD2), nuclear gene encoding mitochondrial protein, transcript variant 2, mRNA [NM_000408]	-0.7	7.3	-5.0	1.5E-04	3.4E-04
43697	39916	A_32_P92505	LYCAT	NM_182551	Homo sapiens lysocardiolipin acyltransferase (LYCAT), transcript variant 1, mRNA [NM_182551]	-1.0	10.2	-4.3	6.4E-04	1.3E-03
7586	6866	A_24_P281923	DLAT	NM_001931	Homo sapiens dihydrolipoamide S-acetyltransferase (E2 component of pyruvate dehydrogenase complex) (DLAT), mRNA [NM_001931]	-1.2	8.6	-4.0	1.2E-03	2.4E-03
11406	10447	A_24_P276490	LYPLA2	NM_007260	Homo sapiens lysophospholipase II (LYPLA2), mRNA [NM_007260]	-0.4	6.7	-3.6	2.5E-03	4.6E-03

Feature Num	ProbeUID	ProbeName	GeneName	Systematic Name	Description	logFC (estimate of the log2-fold-change in MAS_98.12 compared to MAS_98.06)	Average expression	t-value	P-Value	adjusted p-value (false discovery rate)
26908	24657	A_23_P84666	GDPD1	NM_182569	Homo sapiens glycerophosphodiester phosphodiesterase domain containing 1 (GDPD1), mRNA [NM_182569]	-0.4	5.3	-3.6	2.7E-03	4.9E-03
25477	23276	A_24_P227069	GPAM	NM_020918	Homo sapiens glycerol-3-phosphate acyltransferase, mitochondrial (GPAM), mRNA [NM_020918]	-0.5	8.8	-3.2	5.7E-03	1.0E-02
44156	40256	A_23_P251095	PDHA1	NM_000284	Homo sapiens pyruvate dehydrogenase (lipoamide) alpha 1 (PDHA1), mRNA [NM_000284]	-0.4	13.2	-3.1	7.5E-03	1.3E-02
6277	5702	A_23_P54758	MIR16	NM_016641	Homo sapiens membrane interacting protein of RGS16 (MIR16), mRNA [NM_016641]	-0.5	9.5	-2.7	1.6E-02	2.7E-02
40022	36674	A_23_P252681	PCYT1A	NM_005017	Homo sapiens phosphate cytidylyltransferase 1, choline, alpha (PCYT1A), mRNA [NM_005017]	-0.4	8.4	-2.6	2.2E-02	3.5E-02
26531	24273	A_23_P28898	PLCB4	NM_000933	Homo sapiens phospholipase C, beta 4 (PLCB4), transcript variant 1, mRNA [NM_000933]	-0.7	8.4	-2.5	2.3E-02	3.7E-02
12261	11224	A_23_P124742	CHKA	NM_001277	Homo sapiens choline kinase alpha (CHKA), transcript variant 1, mRNA [NM_001277]	-0.3	12.8	-2.4	3.0E-02	4.7E-02
2201	1869	A_23_P11685	PLA2G4A	NM_024420	Homo sapiens phospholipase A2, group IVA (cytosolic, calcium-dependent) (PLA2G4A), mRNA [NM_024420]	6.4	8.4	53	2.4E-18	2.5E-16
14300	13107	A_23_P169909	DGKI	NM_004717	Homo sapiens diacylglycerol kinase, iota (DGKI), mRNA [NM_004717]	6.3	8.2	47	1.5E-17	7.8E-16
38259	35024	A_23_P106675	PLCG2	NM_002661	Homo sapiens phospholipase C, gamma 2 (phosphatidylinositol-specific) (PLCG2), mRNA [NM_002661]	4.9	8.7	38	3.0E-16	1.1E-14
35111	32057	A_23_P69293	CHDH	AJ272267	Homo sapiens partial mRNA for choline dehydrogenase (chdh gene). [AJ272267]	3.3	11.1	29	1.8E-14	3.1E-13
34778	31764	A_32_P141768	AGPAT4	NM_020133	Homo sapiens 1-acylglycerol-3-phosphate O-acyltransferase 4 (lysophosphatidic acid acyltransferase, delta) (AGPAT4), mRNA [NM_020133]	3.9	7.6	27	3.6E-14	5.4E-13
14428	13185	A_23_P80739	PLCD1	NM_006225	Homo sapiens phospholipase C, delta 1 (PLCD1), mRNA [NM_006225]	3.1	10.3	24	2.1E-13	2.4E-12
26815	24482	A_24_P35400	SARDH	NM_007101	Homo sapiens sarcosine dehydrogenase (SARDH), mRNA [NM_007101]	2.5	7.2	23	5.8E-13	6.1E-12
21456	19651	A_23_P17914	PNPLA3	NM_025225	Homo sapiens patatin-like phospholipase domain containing 3 (PNPLA3), mRNA [NM_025225]	4.3	9.4	21	1.8E-12	1.6E-11
23970	21927	A_23_P201808	PPAP2B	NM_003713	Homo sapiens phosphatidic acid phosphatase type 2B (PPAP2B), transcript variant 1, mRNA [NM_003713]	2.5	8.3	17	5.7E-11	4.6E-10
26424	24234	A_23_P108143	GAMT	NM_000156	Homo sapiens guanidinoacetate N-methyltransferase (GAMT), transcript variant 1, mRNA [NM_000156]	2.6	7.8	16	6.5E-11	4.9E-10
23495	21530	A_24_P941353	PCYT1B	NM_004845	Homo sapiens phosphate cytidylyltransferase 1, choline, beta (PCYT1B), mRNA [NM_004845]	1.7	6.8	14	4.1E-10	2.5E-09
18128	16521	A_23_P54918	LDHD	NM_153486	Homo sapiens lactate dehydrogenase D (LDHD), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA [NM_153486]	1.7	9.2	9.8	6.7E-08	2.8E-07
29759	27237	A_23_P154698	CDS2	NM_003818	Homo sapiens CDP-diacylglycerol synthase (phosphatidate cytidylyltransferase) 2 (CDS2), mRNA [NM_003818]	1.1	8.5	9.5	1.1E-07	4.3E-07
18309	16692	A_23_P168868	PTDSS1	NM_014754	Homo sapiens phosphatidylserine synthase 1 (PTDSS1), mRNA [NM_014754]	0.9	14.8	8.8	2.6E-07	9.6E-07
10292	9347	A_23_P70060	PPAP2A	NM_176895	Homo sapiens phosphatidic acid phosphatase type 2A (PPAP2A), transcript variant 2, mRNA [NM_176895]	1.0	10.6	8.5	4.6E-07	1.6E-06
22216	20265	A_24_P103686	AGPAT6	NM_178819	Homo sapiens 1-acylglycerol-3-phosphate O-acyltransferase 6 (lysophosphatidic acid acyltransferase, zeta) (AGPAT6), mRNA [NM_178819]	1.0	11.0	7.8	1.2E-06	3.8E-06
31896	29138	A_23_P87401	GDPD5	NM_030792	Homo sapiens glycerophosphodiester phosphodiesterase domain containing 5 (GDPD5), mRNA [NM_030792]	1.0	7.4	7.7	1.4E-06	4.3E-06

Feature Num	ProbeUID	ProbeName	GeneName	Systematic Name	Description	logFC (estimate of the log2-fold-change in MAS_98.12 compared to MAS_98.06)	Average expression	t-value	P-Value	adjusted p-value (false discovery rate)
36235	33179	A_23_P102770	NAT5	NM_016100	Homo sapiens N-acetyltransferase 5 (NAT5), transcript variant 1, mRNA [NM_016100]	0.9	12.1	7.3	2.9E-06	8.4E-06
20435	18708	A_23_P339240	PLCH1	NM_014996	Homo sapiens phospholipase C, eta 1 (PLCH1), mRNA [NM_014996]	0.9	8.4	7.0	4.1E-06	1.2E-05
36976	33750	A_32_P353798	PTDSS2	NM_030783	Homo sapiens phosphatidylserine synthase 2 (PTDSS2), mRNA [NM_030783]	0.7	12.1	6.3	1.5E-05	4.1E-05
27768	25471	A_23_P403424	PLA2G4B	NM_005090	Homo sapiens phospholipase A2, group IVB (cytosolic) (PLA2G4B), mRNA [NM_005090]	0.9	10.5	5.9	2.9E-05	7.5E-05
41522	38017	A_23_P56356	PLB1	NM_153021	Homo sapiens phospholipase B1 (PLB1), mRNA [NM_153021]	0.7	6.0	5.6	4.9E-05	1.2E-04
18563	17004	A_23_P204736	GPD1	NM_005276	Homo sapiens glycerol-3-phosphate dehydrogenase 1 (soluble) (GPD1), mRNA [NM_005276]	1.5	6.6	5.4	8.0E-05	1.9E-04
29686	27110	A_23_P218237	LCAT	NM_000229	Homo sapiens lecithin-cholesterol acyltransferase (LCAT), mRNA [NM_000229]	0.8	6.7	5.3	9.4E-05	2.2E-04
35060	32117	A_24_P205364	SHMT1	NM_004169	Homo sapiens serine hydroxymethyltransferase 1 (soluble) (SHMT1), transcript variant 1, mRNA [NM_004169]	0.7	8.7	4.9	2.1E-04	4.8E-04
34063	31232	A_23_P155335	PLD1	NM_002662	Homo sapiens phospholipase D1, phosphatidylcholine-specific (PLD1), mRNA [NM_002662]	0.7	9.3	4.8	2.5E-04	5.4E-04
27086	24832	A_23_P7957	GNMT	NM_018960	Homo sapiens glycine N-methyltransferase (GNMT), mRNA [NM_018960]	0.5	5.7	4.8	2.6E-04	5.5E-04
18507	16901	A_23_P19192	LYPLA1	NM_006330	Homo sapiens lysophospholipase I (LYPLA1), mRNA [NM_006330]	0.9	12.1	4.4	4.9E-04	1.0E-03
15223	13844	A_23_P400435	CRLS1	NM_019095	Homo sapiens cardiolipin synthase 1 (CRLS1), mRNA [NM_019095]	0.5	12.9	4.3	6.1E-04	1.2E-03
20364	18581	A_24_P191588	ALAS1	NM_000688	Homo sapiens aminolevulinate, delta-, synthase 1 (ALAS1), transcript variant 1, mRNA [NM_000688]	0.7	10.8	4.2	8.6E-04	1.7E-03
6560	5915	A_23_P127627	DGKZ	NM_003646	Homo sapiens diacylglycerol kinase, zeta 104kDa (DGKZ), transcript variant 2, mRNA [NM_003646]	0.6	9.6	4.0	1.1E-03	2.2E-03
6878	6346	A_23_P40926	DGKG	NM_001346	Homo sapiens diacylglycerol kinase, gamma 90kDa (DGKG), transcript variant 1, mRNA [NM_001346]	0.4	5.2	3.6	2.9E-03	5.2E-03
26959	24592	A_23_P105307	DGKA	NM_201444	Homo sapiens diacylglycerol kinase, alpha 80kDa (DGKA), transcript variant 1, mRNA [NM_201444]	0.4	7.2	3.1	7.4E-03	1.3E-02
10394	9526	A_24_P304439	SDS	NM_006843	Homo sapiens serine dehydratase (SDS), mRNA [NM_006843]	0.6	8.3	2.9	1.0E-02	1.7E-02
42936	39201	A_23_P351757	PLCD3	NM_133373	Homo sapiens phospholipase C, delta 3 (PLCD3), mRNA [NM_133373]	0.3	5.8	2.9	1.1E-02	1.9E-02

IV

Metabolomic and transcriptomic response to bevacizumab and doxorubicin in basal-like and luminal-like breast cancer xenograft models

Eldrid Borgan[1,4,5], Evita M Lindholm[2], Siver Moestue[4], Gunhild M Mælandsmo[2,6], Ole Christian Lingjærde[7], Ingrid S Gribbestad[4], Anne-Lise Børresen-Dale[1,5], Olav Engebraaten[3,5], Therese Sørlie[1]

1 Departments of Genetics, 2 Tumor Biology, and 3 Oncology, Division for Surgery and Cancer Medicine, Oslo University Hospital Radiumhospitalet

4 Department of Circulation and Medical Imaging, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

5 Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Norway

6 Department of Pharmacy, University of Tromsø, Tromsø, Norway

7 Biomedical Research Group, Department of Informatics, University of Oslo, Oslo, Norway

Abstract

BACKGROUND: The VEGF targeted antiangiogenic drug bevacizumab has shown varying results in clinical trials of breast cancer. Identifying robust biomarkers for selection of patients that may benefit from bevacizumab treatment is important for the future use of this drug.

METHODS: Two established xenograft models representing basal-like and luminal-like breast cancer were used to study metabolomic and transcriptomic responses. Controls and xenografts treated with bevacizumab, doxorubicin or the combination of the two drugs were sacrificed at day 3 and 10 after treatment. Gene expression microarrays and high resolution magic angle spinning magnetic resonance spectroscopy (HR MAS MRS) was performed on all tumor samples. In addition, microarray data from human breast tumors were used to compare expression of genes involved in phospholipid metabolism with the observed differences in the xenograft models.

RESULTS: Metabolomic and transcriptomic bevacizumab induced changes were observed in the benefiting basal-like tumors which were not observed in the non-benefiting luminal-like model. The metabolite with highest potential as response marker was glycerophosphocholine (GPC) which showed opposite response patterns in the two differently responding xenografts; lower in the treated basal-like tumors and higher in luminal-like tumors compared with controls. Phosphocholine (PCho) on the other hand, showed either no difference or increase in concentration after treatment in both models, diluting the total choline (tCho) signal. On the global transcriptomic level, the largest changes in gene expression compared with controls were observed for the most efficient treatment in each of the two xenograft models. Of particular interest were 14 genes showing significant differential expression between basal-like tumors receiving combination therapy and doxorubicin monotherapy, including higher expression of very low density lipoprotein receptor (*VLDR*), and lower expression of hemoglobin, theta 1 (*HBQ1*). Using published gene expression signatures and GO-term enrichment analysis, biological responses including possible treatment resistance effects were explored. When evaluating expression of genes involved in choline and phospholipid metabolism, a concordant response to the metabolomic changes was not observed. However, differences in expression of these genes between basal-like and luminal-like untreated xenografts were correlated with differences observed in human breast cancer, suggesting that distinct phospholipid metabolism patterns may be intrinsic features of these breast cancer subtypes.

CONCLUSIONS: This study underlines the importance of characterizing biological differences between subtypes of breast cancer to identify personalized biomarkers for selecting patients for bevacizumab treatment and evaluating response to therapy.

Introduction

Targeting tumor stroma has evolved as an attractive strategy to treat solid cancers. One implemented strategy for several cancers is antiangiogenic treatment utilizing a VEGF targeting monoclonal antibody (bevacizumab). Limited benefit in progression free survival (PFS) and lack of benefit in overall survival (OS) reported from clinical trials have induced questions about the relevance of bevacizumab for breast cancer [1]. However, recent results indicate that prolonged bevacizumab administration in metastatic breast cancer patients may give a considerable benefit in OS [2]. The usefulness of bevacizumab as an anti breast cancer drug will depend on identification of precise biomarkers for subgroups of patients that with high probability will respond to this drug [3]. Several potential biomarkers have been proposed, including VEGFR polymorphisms, and VEGF, PDGFR-β and VCAM1 expression levels [4-7], but none have been established as reproducible. In addition to predicting which patients may benefit from bevacizumab, identifying biomarkers for monitoring treatment response could accelerate the evaluation of individual therapeutic regimes. The aim of this study was to identify response biomarkers and study the biology of treatment response and possible resistance effects by combining metabolomics and transcriptomics of breast cancer xenograft models.

The tumor metabolome is known to be highly affected by extracellular factors such as the microenvironment, pH, oxygen, nutrients and drugs. Thus, metabolomic techniques have the potential to be more sensitive in monitoring treatment than other approaches such as measuring levels of RNA or proteins. Several metabolic markers of treatment response have been suggested, such as total choline (tCho) from *in vivo* MRS [8,9], and from the higher resolution MRS spectra also metabolites such as phosphocholine (PCho), glycerophosphocholine (GPC), and lactate [10,11]. Preclinical tumor model systems are valuable tools for the study of treatment response mechanisms in a controlled manner. In particular, xenografts are attractive models since human carcinomas can be studied *in vivo* surrounded by a relevant microenvironment [12]. In addition, direct grafting of human tumor tissue is more representative of human cancer than using cultured cells. Two directly grafted orthotopic xenograft models representing basal-like and luminal-like breast cancer have previously been established and characterized at the transcriptomic and metabolomic levels [13-15]. Distinct patterns of choline derivatives observed in the xenografts were also found in clinical tumor samples, suggesting that these two models are relevant for studies of metabolism and treatment response in the two types of breast cancer [15]. Recently, the xenografts were treated with bevacizumab, doxorubicin and combination therapy, and the basal-like model showed significantly

improved response to combination therapy compared with doxorubicin alone, while doxorubicin monotherapy or a combination of the two were equally effective in the luminal-like model [16]. Tumor tissue was harvested at day 3 and 10 from the treatment groups, and in the present study, metabolomics and transcriptomics of tumor tissue from these experiments was performed using high resolution magic angle spinning magnetic resonance spectroscopy (HR MAS MRS) and gene expression microarrays. Comparison of these differently responding models with respect to potential response markers and the underlying biology may generate hypotheses for future studies in selecting breast cancer patients who may benefit from antiangiogenic therapy with bevacizumab.

We demonstrated that glycerophosphocholine (GPC) is a promising biomarker on the metabolomic level and that several gene transcripts were associated with responses only in the bevacizumab benefiting basal-like tumors. Additionally, microarray data from a set of human breast cancers were been used to show that differences in the expression of genes involved in phospholipid metabolism may be an intrinsic characteristic of basal-like and luminal-like breast cancer.

Material and methods

Xenograft models and treatment

Two orthotopic xenograft models, MAS98.12 and MAS98.06 have been established by directly grafted orthotopic implantation of primary breast cancer tissue in SCID mice and are serially transplanted, as previously described [13]. Using gene expression microarray, the MAS98.12 and MAS98.06 xenografts were characterized as representing basal-like and luminal-like breast cancer, respectively [13]. These models were used to test response to bevacizumab and doxorubicin treatment, showing that the MAS98.12 xenografts had the greatest benefit (in terms of tumor growth delay) from a combination of bevacizumab and doxorubicin, while the MAS98.06 xenografts showed no additional benefit from adding bevacizumab [16]. MAS98.12 also responded significantly to bevacizumab monotherapy, though not to the same extent as to the combination therapy. An overview of the experimental procedure in the current work is illustrated in Figure 1. Locally bred athymic nude mice (NCr-Foxn1^{nu}) were kept under pathogen-free conditions, at constant temperature (21.5±0.5 ºC) and humidity (55±5 %), 20 air changes/hr and a 12 hr light/dark cycle. The mice were fed RM3 diet (Scanbur BK, Norway) and distilled tap water *ad libitum*. The water was supplemented with 4 µg/ml 17-β-estradiol, which the luminal-like xenograft is dependent on. Animals from the two xenograft models were randomly assigned to different treatment groups after the tumor diameter reached approximately 5 mm. For each of the two xenograft models, tumors were collected from animals used as controls or treated with bevacizumab (5 mg/kg), doxorubicin (8 mg/kg) or a combination of the two therapies (n = 6 tumors for each group). Bevacizumab treatment was repeated at day 4 and day 7. Animals were sacrificed and tissue harvested at either day 3 or 10 after treatment, in triplicates within each treatment group. This resulted in 24 tumor samples from each of the models. In addition, controls and bevacizumab treated luminal-like xenografts not fed with estradiol were included for comparison (n = 12). Tumor tissue from all animals was stored in liquid nitrogen and independent tumor pieces were used for HR MAS MRS and gene expression microarray analysis. Out of the 48 samples from the estrogen fed animals analyzed with both experimental techniques, 33 were from the same tumor, 11 were from the same mouse but from different tumors (from bilateral implantations), and 4 were from different mice. The 12 tumors from animals not receiving estrogen supplement were included in the gene expression analysis for normalization purposes. All procedures and experiments involving animals are approved by the National Animal Research Authority and were conducted according to the European Convention for the Protection of Vertebrates used for Scientific Purposes.

HR MAS MRS experiments and data processing

HR MAS MRS experiments and data preprocessing were performed on a Bruker Avance DRX600 spectrometer equipped with a 1H/13C HR MAS probe (Bruker BioSpin Corp.). Macroscopically viable tumor tissue was cut to fit a 30 µl disposable insert (Bruker Biospin Corp.), prefilled with 3 µl PBS made on D_2O containing 98.8 mM trimethylsilyltetradeuteropropionic acid (TSP) for chemical shift referencing. Sample weight was 13.6 ± 4.3 mg (mean ± SD). Samples were spun at 5 kHz and 4°C. A pulse-acquired experiment including the ERETIC sequence (ereticpr.drx; Bruker) was performed for all samples. The ERETIC signal was positioned at -1.0 ppm. The water resonance was saturated for 15 seconds (60 dB continuous wave), followed by a 60-degree pulse for excitation. Signals were collected over a sweep width of 16.7 ppm. 128 FIDs were acquired into 64K points during 3.28 seconds. Spectra were Fourier transformed into 128K after 0.3 Hz exponential line broadening and chemical shifts were calibrated to the TSP singlet at 0 ppm. Spectral assignments were performed based on a previous HR MAS MRS study of breast cancer lesions [17]. The regions representing TSP, ERETIC, creatine, choline, phosphocholine (PCho), glycerophosphocholine (GPC), taurine (Tau), glycine (Gly), and lactate (Lac), were selected for quantification in all spectra. Peak areas were calculated by polynomial curve fitting (PeakFit v 4, Systat Software Inc) with a combination of Gaussian and Lorentzian line-shapes (Voigt function). Concentrations of tissue metabolites were calculated relative to the ERETIC signal as previously described [15].

Gene expression microarray experiments and normalization

Total RNA from all 60 tumor pieces was extracted using the TRIzol reagent (Invitrogen) according to the manufacturer's protocol. Total RNA concentration was measured using NanoDrop (NanoDrop Technologies) and quality assessed using 2100 Bioanalyzer (Agilent Technologies). 700 ng total RNA was amplified and labeled using low-RNA input Linear Amplification RNA kit (Agilent Technologies) and cy3-CTP, and hybridized to one-color Agilent Technologies 4×44K Whole Human Genome Microarrays, according to the manufacturer's protocol One-Color Microarray-Based Gene Expression Analysis v5.5. The amount of cy3-labelled cRNA used for hybridization was 1.65 µg. The microarrays were scanned using Agilent Technologies Microarray Scanner G2505C. Data were extracted from the scanned images using Feature Extraction software (Agilent Technologies) version 10.1.1.1 and protocol GE1-v5_10_Apr08 for mRNA and further processed and normalized using R (v2.10.1) and Bioconductor [18]. Data from all 60 samples were included in normalization and statistical analysis to

increase the power of the statistical analysis. The microarray data were detrended for multiplicative effects and log2 transformed. Data from control probes, probes with inferior quality for more than 20 percent of the arrays (such spots were defined as feature outliers from the FE file), and probes that were flagged as present on less than 20 percent of the arrays were omitted from the analysis. After setting spots flagged as outliers to NA, the average value of duplicate probes was used to represent each unique probe. Missing data were imputed using k-nearest neighbors [19] and quantile normalized. For transcripts (based on GeneName, as provided by Agilent Technologies) represented by multiple different probes, the probe with the highest interquartile range (IQR) was chosen to represent each transcript. The normalized dataset included data for 28150 unique mRNA transcripts on 60 microarrays from biopsies; 24 from MAS98.12 animals, 24 from MAS98.06 animals fed with estradiol and 12 from MAS98.06 animals not treated with estradiol.

Statistical analysis

Analysis of metabolic profiles

The quantified concentrations of specific metabolites were used to assess whether any of these metabolites were affected by the treatments. Data for each metabolite were modeled separately for each of the xenograft models using a two way ANOVA with interaction, the covariates being treatment and days treated as factors. The ANOVA analysis was used to compare each treatment with controls at day 3 and day 10, respectively.

Analysis of gene expression profiles

The gene expression data were analyzed for differences between treatment groups in the two xenograft models. Exploiting the factorial design of the experiments, the Bioconductor package Limma was used to model the expression levels of each transcript in the filtered microarray data as a function of the different days and treatments, equivalent of a two way ANOVA [20]. Data from the two xenograft models were modeled separately, and treatment and day were used as factors. An observed batch effect in the microarray data, corresponding to amplification date, was also corrected for in the model by using the blocking factor implemented in Limma [21]. All differences between treatments compared with controls from the same day were analyzed by applying different contrasts to the model as described by Smyth [22]. Additionally, the combination treatment was

compared with doxorubicin monotherapy. The test statistics were Empirical Bayes modified and the p-values were adjusted by Benjamini Hochberg correction [20,23]. This analysis was performed separately on all gene transcripts as input and on a gene set of 105 genes representing phospholipid metabolism [15]. To analyze whether any biological processes were enriched among the genes most differentially expressed between treatments and controls, GOrilla was used [24,25]. Lists of all 28150 transcripts in the normalized microarray data were ranked according to log fold change to analyze GO-term enrichment towards the top of the lists.

Using gene expression signatures as surrogate measures

Published microarray based gene expression signatures including genomic grade index (GGI) [26], a hypoxia classifier [27] and a vascular invasion signature [28] were used to study differences in biological features of interest in relation to bevacizumab treatment effects and the possibility of induced acquired resistance. Using scores from each of these signatures, calculated as described below, comparisons between the two xenograft models and between the different treatments were performed using ANOVA, in the same manner as described above for assessing differences in metabolite concentrations. To analyze differences in aggressiveness of the tumors between the controls and the different treatments, the gene expression based genomic grade index (GGI) was used as a surrogate measure [26]. GGI was originally developed to differentiate histological grade II tumors into grade I or III. A signature of 97 genes representing grade III or I tumors is used to calculate the GGI score. The raw GGI score, which is a weighted average of the expression levels of the 97 genes were used as a score for aggressiveness in the tumors. As a surrogate marker for hypoxia, the gene expression based hypoxia classifier from Chi et. al [27] was applied. The original probes were matched using the R-package BiomaRt [29]. For each probe for each sample, the median expression value across all samples was subtracted (median centering). As described by Chi et. al, a hypoxia score was calculated by taking the median expression of the hypoxia related genes for each sample [27]. Another aggressive feature of cancers is vascular invasion of tumor epithelial cells, which may be enhanced in leaky vessels. A recently published vascular invasion signature was used as surrogate marker for this phenomenon [28]. This 18-gene signature represents differences in gene expression between vascular invasion positive and negative endometrial tumors, based on staining for presence of tumor cells in blood vessels [30]. Data for the selected genes were mean centered across all samples, and the log fold changes of the vascular invasion signature were analyzed for nonzero correlations to the expression of the matched genes for each sample. The

correlation coefficients were used as scores to compare vascular invasion between the xenograft models and treatments.

Analysis of phospholipid metabolism gene expression in a patient cohort

To assess whether the distinct phospholipid metabolism gene expression patterns observed in the two xenograft models are representative of basal-like and luminal-like cancer, a microarray dataset from a patient cohort of 115 breast carcinomas was used [31]. Since the patient and xenograft microarray data were from the same platform (Agilent Technologies 4×44K Whole Human Genome Microarrays), the probes representing the 105 phospholipid metabolism genes were exactly matched. The assigned intrinsic subtypes of the tumors were used to select the 16 basal-like tumors and 61 luminal tumors in the dataset (including both luminal A and luminal B). Data were centered across genes and a heatmap was generated by clustering of the 105 genes and 77 samples using Pearson's correlation and average linkage. Limma was used to test for differences between the two groups in the patient data, and generated (modified) t-statistics were compared with those from the xenograft data (as previously published [15]) using Pearson's correlation test.

Results

Metabolic response to treatment

Quantified metabolite concentrations were used to analyze differences in Cho, GPC, PCho, tCho, Tau, Cre, Gly and Lac between treated and control tumors from both models (See Supplementary Figure 1). The basal-like xenograft represented a bevacizumab responder while the luminal-like xenograft represented a non-responder to this drug. Therefore, changes observed in bevacizumab treated basal-like and not in luminal-like tumors were of particular interest as these could have the potential of differentiating responders from non-responders. Indeed, GPC is potentially such a marker, as it showed opposite patterns of response to treatment in the two differently responding xenograft models (Figure 2). In basal-like tumors, GPC concentrations were significantly lower after any of the three treatment modalities at day 10, and there was a similar trend at day 3. This was in contrast to luminal-like tumors, where GPC concentrations were significantly *higher* in the doxorubicin and combination treated animals compared with controls at day 10. PCho on the other hand, showed either no change in concentrations or tended to increase in treated animals of both models (Figure 2). However, the latter was significant only for bevacizumab treated basal-like tumors and combination treated luminal-like tumors. Importantly, for the basal-like model, the observed decrease in GPC and increase in PCho after bevacizumab or combination treatment resulted in no significant differences in the concentration of tCho. This was in contrast to the luminal-like model where tCho showed a significant increase after the combination treatment compared with controls. This shows that the value of tCho as a response marker by using *in vivo* MRS may be reduced due these differences in patterns of choline derivatives between tumor subgroups. Table 1 summarizes all metabolites that displayed either higher or lower levels after the three different treatment regimes at day 10. Notably, Gly, Cre and Tau all displayed significantly lower concentrations in the doxorubicin and bevacizumab treated basal-like tumors, but not in the combination treated tumors even though this was the most effective treatment.

Gene expression response to the most efficient treatment

The most efficient treatment, measured as growth delay in each of the two models, revealed the largest differences in gene expression, at day 10 (Figure 3). Hence, for the basal-like model, the largest transcriptional response was seen in tumors from animals given the combination therapy, and 40 transcripts showed significantly different levels compared with controls; 38 at higher expression levels including *HOXB13*, *PPP1R3C*, *TMEM45A* and 2 genes were lower expressed (*SLC2A12* and

KRT5) (see Supplementary Table 1 for details). Furthermore, when comparing combination treated with doxorubicin treated basal-like tumors, 14 transcripts showed significantly different expression; 8 higher expressed, including *VLDLR*, *PPFIA4* and *MCHR1,* and 6 lower expressed, including *HBQ1*, *DAB1* and *ACCN1* (Table 2). The GO-term "muscle contraction" was enriched among the transcripts higher expressed in the combination treated versus both the doxorubicin treated and untreated basal-like tumors (Supplementary Table 2). In addition, several immune process related GO-terms such as "immune response", "inflammatory response" and "response to cytokine stimulus" were enriched among the transcripts that were lower expressed in the combination treated when compared with doxorubicin treated basal-like tumors. However, expression changes of such genes in xenograft tumors are difficult to interpret, since the host animals were immunocompromised.

For the luminal-like model, the largest response was observed in the tumors given doxorubicin monotherapy, and 380 transcripts showed significantly different expression compared with the control animals. Of these, 201 were higher expressed, including *EGF*, *RARRES3* and *CAPSL*. Lower expression levels were seen in 179 genes, including *HBA2*, *HBA1* and *KCNK5* (see Supplementary Table 3 for details). In doxorubicin treated versus control luminal-like tumors, GO-terms such as "type I interferon-mediated signaling pathway", "muscle contraction" and "muscle filament sliding" were significantly enriched among the transcripts that were higher expressed. Among the transcripts that were lower expressed, the terms "DNA replication", "DNA strand elongation involved in DNA replication" and "ncRNA processing" were significantly enriched (Supplementary Table 2), which is in agreement with the proposed mechanism of doxorubicin of targeting topoisomerase II.

Using gene expression scores as surrogate markers

Evasive or adaptive resistance mechanisms to antiangiogenic treatment have been suggested, such as increased aggressiveness/invasiveness or recruitment of bone marrow-derived cells due to increased hypoxia and lack of nutrients [32]. The published gene expression signatures GGI [26], hypoxia [27] and vascular invasion [28] were used as surrogate markers for the extent of aggressiveness, hypoxia and vascular invasion by tumor epithelial cells, respectively. A score for each of the three signatures was calculated for each sample, and these scores were compared between the two models and the different treatments (Table 3).

There was not significant evidence of differences in GGI between the two models. Lower GGI scores were observed after treatment with bevacizumab in both xenografts, and for the basal-like tumors a more significant decrease was observed after combination treatment. Interestingly, a significantly higher GGI was observed in the doxorubicin treated basal-like tumors at day 3 compared with controls. The luminal-like xenografts showed significantly higher baseline hypoxia scores than the basal-like xenografts. Furthermore, hypoxia scores increased in both models after treatment with either bevacizumab or the combination. This increase was most profound for the bevacizumab treated basal-like tumors. Comparing vascular invasion scores between the two models, basal-like tumors showed significantly higher baseline scores. While no significant differences between treatments were observed for the basal-like model, vascular invasion was significantly higher for all treated luminal-like tumors compared with controls.

Phospholipid metabolism gene expression in xenograft models and patient cohort

In a previous study, we showed that genes involved in choline transport, phospholipid metabolism, and glycine metabolism were differentially expressed between these xenograft models, and levels of expression corresponded to metabolomic changes [15]. Using these same genes to analyze differences between treated animals and controls, lecithin-cholesterol acyltransferase (*LCAT*) showed significantly higher levels of expression in combination treated basal-like tumors compared with controls at day 10. No other detectable gene expression changes could be associated with the observed treatment responses in choline derivatives. The different responses on the metabolomic level may reflect intrinsic diverse phospholipid metabolism between the models. To assess whether the metabolic phenotypes of the xenograft models are intrinsic properties of basal-like and luminal-like tumors, a human breast tumor microarray dataset was used to compare expression differences of genes involved in phospholipid metabolism between these two biologically distinct subgroups of breast cancer (Supplementary Figure 2A). A clustering analysis revealed great similarity between basal-like tumors, while the luminal-like tumors were more heterogeneous in their expression of genes involved in phospholipid metabolism. Analyses of differential expression of these genes between basal-like and luminal-like tumors were compared between human and xenograft tumors [15]. The correlation between these two analyses was found to be significant (rho = 0.34, p-value < 4.1×10^{-4}) (Supplementary Figure 2B and C). The most significant differences in gene expression observed both in the xenograft and human basal-like versus luminal-like tumors included higher expression of phospholipase A2, group IVA (*PLA2G4A*), phospholipase C, gamma 2 (*PLCG2*) and 1-

acylglycerol-3-phosphate O-acyltransferase 4 (lysophosphatidic acid acyltransferase, delta) (*AGPAT4*); and lower expression of glycerol-3-phosphate dehydrogenase 1-like (*GPD1L*), phospholipase C, delta 4 (*PLCD4*) and ethanolamine kinase 2 (*ETNK2*).

Discussion

Antiangiogenic therapy such as bevacizumab may be more efficient for certain subgroups of breast cancer. To exploit the potential of bevacizumab as a breast cancer drug, individualized biomarkers are needed to select patients and assess treatment response. Metabolomic biomarkers are especially attractive because they have the potential of being measured *in vivo* allowing treatment monitoring and possibilities for adapting treatment regimes according to response. Several *in vitro* studies of breast cancer cell lines have reported high PCho levels or PCho/GPC ratio as malignant markers [33]. Previously, we have shown that the luminal-like model has a high PCho/GPC ratio while the basal-like model shows the opposite, and also that PCho/GPC ratios were lower in triple negative human cancers [15]. Together this suggests that PtdCho metabolism in breast cancer may be more complicated than previously assumed. The difference in GPC content between the two xenograft models may be associated with the way the tumors respond to treatment. While the GPC levels in the bevacizumab responding basal-like tumors decrease after treatment, they increase in the non-responding luminal-like tumors. PCho on the other hand, showed no differences or increased in both luminal-like and basal-like tumors after treatment. This underlines the importance of characterizing subtypes of breast cancer separately when evaluating response markers. Clinical 1-H *in vivo* MRS of breast tumors detects tCho which is the sum of GPC, PCho and Cho signals. Opposite behavior of GPC and PCho in response to treatment, as observed using *ex vivo* HR MAS MRS for the basal-like tumors, may reduce the utility of tCho as a response marker. However, with increasing field strengths and the possibility of implementing 31-P probes in clinical magnets, the GPC and PCho peaks could be separated in the spectra obtained *in vivo* [34]. Thus, metabolomic biomarkers such as GPC could in the future be measured without taking a single biopsy. Since GPC displays opposite patterns of baseline concentrations and response to treatment in a benefiting and a non-benefiting model, this metabolite may have the potential as a biomarker both to select patients for bevacizumab treatment and to monitor treatment response. However, the role of GPC and PCho in different subtypes of cancer should be studied more extensively to understand the behavior of these metabolites in breast cancer subtypes and in response to treatment. Functional studies in different models of breast cancer will be of great importance to unravel the mechanisms behind changes in GPC content.

On the gene expression level, this study suggests that the most significant gene expression changes occur in tumors which benefit the most from a given treatment, i.e. the combination of bevacizumab and doxorubicin for basal-like and doxorubicin for luminal-like tumors. A similar trend was also found when studying the same tumors at the proteomic level, i.e. the most effective treatment was linked to higher kinase activity at day 10 [16]. Those transcripts expressed at significantly different levels between combination and doxorubicin treated tumors in the basal-like xenograft model (Table 2) are particularly interesting since these may suggest mechanistic explanations for the additional effect of bevacizumab in these tumors. Some of the genes with higher expression in the combination treated tumors compared with controls have previously been associated with angiogenesis, such as *VLDLR* [35] and *CHI3L1* [36,37]. Of the transcripts whose expression was reduced after combination treatment, two hemoglobins (*HBQ1* and *HBD*) were represented. It could be speculated that the genes showing higher expression may represent adaptive resistance processes while the lower expressed genes could represent differences in composition of different cell types in the tumors due to the antiangiogenic effects of bevacizumab.

Focusing solely on genes involved in phospholipid metabolism [15], *LCAT* was the only gene found to be significantly differentially expressed in response to treatment. LCAT has both cholesterol and PtdCho (also called lechitin) acetyltransferase acitvity and may thus catalyze the same reaction as phospholipase A2. The increase in expression of *LCAT* in combination treated basal-like tumors is contradictory to the decrease in GPC as measured by HR MAS MRS. However, the increase in *LCAT* levels may be due to feedback mechanisms rather than exerting a direct effect on the levels of GPC, or that the LCAT enzyme in fact does not catalyze this reaction in the basal-like tumors. The apparent lack of transcriptional response of genes involved in phospholipid metabolism shows that the observed responses in choline derivatives are not necessarily coupled to changes at the gene expression level. This may be because changes in enzymatic activity are regulated at the post-translational level. Also, the distinct gene expression patterns of phospholipid metabolism genes found in the untreated basal-like and luminal-like xenografts correlated with those of human breast tumors (Supplementary Figure 2), suggesting that the differences in phospholipid metabolism may be an intrinsic characteristic to basal-like and luminal-like breast cancer. Interestingly, *PLA2G4A* was one of the most significantly differentially expressed between basal-like (higher) and luminal-like (lower) in for both xenografts and human tumors. PLA2G4A catalyzes the hydrolysis of membrane lipids and may catalyze the breakdown of PtdCho to acyl-GPC. *PLA2G4A* may catalyze the breakdown of PtdCho to acyl-GPC and has been reported to be underexpressed in breast cancer cells concomitant with low

GPC levels compared with normal epithelial breast cells [38]. Thus, targeting *PLA2G4A* could perhaps be an interesting approach in treatment of basal-like cancers.

The higher GPC/PCho ratio and corresponding gene expression levels observed in the basal-like xenograft and patient tumors have not previously been reported in similar studies of PtdCho metabolism in triple negative or ER-positive breast cancer cell lines [33,38,39]. Such *in vitro* studies will however not account for the effects of tumor heterogeneity and the tumor microenvironment. Furthermore, metabolomic profiles generated from *in vitro* studies are sensitive to cell culturing conditions, including glucose availability and growth phase [40]. The differences in GPC/PCho ratios observed between the two types of breast cancer *in vivo* may be due to diversity in genetic aberrations, hormonal effects or impact of the microenvironment affecting PtdCho metabolism. Levels of GPC have been shown to be inversely correlated with ER-status in MRS studies of breast cancer [41,42]. Differences in the microenvironment for these xenograft models have been observed, most notably more active angiogenesis and higher expression of genes involved in ECM (from both tumor and host murine cells) in the basal-like model [13,16]. It should be noted that all measurements in this study are based on bulk tumor, meaning that the observed signals arise from a mix of different types of cells and extracellular constituents.

Bevacizumab induced changes in expression of genes involved in biological processes associated with adaptive treatment resistance are of particular interest in light of the limited success of bevacizumab in breast cancer clinical trials. The observed bevacizumab specific decrease in GGI is contradictory to claims that antiangiogenic therapy may increase aggressiveness [43,44]. However, bevacizumab associated increases in hypoxia scores were observed in both xenograft models, and this increase was higher in the combination treated basal-like tumors. Increased hypoxia has been linked to invasiveness and metastasis [45], and is proposed to induce adaptive resistance to antiangiogenic therapies [46]. It is worth noting that aldehyde dehydrogenase 2 family (*LDHA2*), which was among the most significantly higher expressed genes after combination treatment in basal-like xenografts, has, in addition to being involved in many metabolic pathways, been linked to hypoxia [47]. In the untreated xenografts, a vascular invasion score was significantly higher in basal-like tumors than in luminal-like tumors, but the vascular invasion phenotype was increased after treatment in luminal-like tumors. A potential for increased vascular invasion after treatment in luminal-like tumors may thus be inferred, however, this increase did not reach the level of vascular invasion found in the in the untreated basal-like tumors (results not shown). It should be noted that this gene signature was

based on vascular invasion in endometrial cancer [28]. Another possible indication of bevacizumab induced invasiveness was enrichment of the GO-term “muscle contraction” (including genes such as myosin and actin) when comparing the combination treated basal-like tumors with controls or with doxorubicin monotherapy. This could perhaps also be linked to adaptive resistance by tumor cell migration to more vascularized areas [46,48].

This study has highlighted metabolomic and transcriptomic responses to bevacizumab and doxorubicin therapy, suggesting differential effects in luminal-like and basal-like breast cancer. The biological connection between response at the metabolomic and transcriptomic level was not evident, suggesting complementary information retrieved from these two types of analyses. Importantly, the results from this study show the importance of patient stratification with respect to biologically relevant subgroups when searching for biomarkers in preclinical as well as clinical breast cancer trials. However, results from different clinical trials of bevacizumab in breast cancer, have reported that both ER positive cancers [49] and triple negative cancers [50] may benefit from the addition of bevacizumab to chemotherapy. Thus, biological features apart from the classical subtype definitions may determine whether a tumor is responsive to bevacizumab. Biomarkers with the potential to predict or monitor treatment response may in fact be found across subtypes. Thus, new xenograft models representing the tumor heterogeneity of basal-like and luminal-like cancers and carefully planned clinical trials will be important for investigating the biology of bevacizumab response in breast cancer.

Acknowledgements

We want to thank Hilde Johnsen for performing some of the microarray experiments, Xi Zhao for sharing her expertise in R and gene signatures, and Guro Giskeødegård and Beathe Sitter for help with processing the HR MAS spectra and allocating metabolites, respectively. This work was supported by the Norwegian Research Council (163027 to Therese Sørlie, and FUGE project 183379), and Oslo University Hospital and a generous donation from Monica Nordal memorial fund to Olav Engebraaten.

Tables

Table 1. Overview of metabolites that displayed significantly ($p < 0.05$) different concentrations in treated animals compared with controls at day 10. Arrows indicate whether the concentration of each metabolite was higher (↑) or lower (↓) after treatment. Abbreviations: GPC: glycerophosphocholine, tCho: total choline, PCho: phosphocholine, Tau: taurine, Cre: creatine, Gly: glycine.

Xenograft	Metabolites	Doxorubicin	Bevacizumab	Bevacizumab + Doxorubicin
Basal-like	Cholines	GPC↓***, tCho↓*	PCho↑*, GPC↓*	GPC↓*
	Other metabolites	Gly↓***, Cre↓***, Tau↓	Gly↓*, Cre↓*, Tau↓*	
Luminal-like	Cholines	GPC↑**, tCho↑*		PCho↑*, GPC↑*, tCho↑**
	Other metabolites			Tau↓*

*p < 0.05, ** p < 0.01, *** p < 0.001

Table 2. Transcripts with the most significant difference in expression levels between combination and doxorubicin treated basal-like tumors. Abbreviations: logFC: log2 fold change.

Higher expression levels						
Probe Name	**Gene Name**	**Description**	**logFC**	**t**	**p-value**	**FDR**
A_23_P43476	VLDLR	Very low density lipoprotein receptor, transcript variant 1	1.4	7.8	1.7E-07	0.0047
A_23_P420692	PPFIA4	Protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 4	1.0	6.6	1.9E-06	0.013
A_23_P139166	CB959193	CB959193 AGENCOURT_13778270 NIH_MGC_184 cDNA clone IMAGE:30351353 5'	1.9	6.2	4.7E-06	0.019
A_23_P211543	MCHR1	Melanin-concentrating hormone receptor 1	0.8	6.0	6.2E-06	0.022
A_23_P35414	PPP1R3C	Protein phosphatase 1, regulatory (inhibitor) subunit 3C	1.6	5.9	8.2E-06	0.026
A_23_P137665	CHI3L1	Chitinase 3-like 1 (cartilage glycoprotein-39)	0.93	5.8	9.7E-06	0.027
A_32_P128586	DA727827	DA727827 NT2RM1 cDNA clone NT2RM1001124 5'	0.86	5.8	1.2E-05	0.029
A_23_P117851	CPLX3	Complexin 3	0.93	5.6	1.6E-05	0.032
Lower expression levels						
ProbeName	**GeneName**	**Description**	**logFC**	**t**	**p-value**	**FDR**
A_23_P23850	DAB1	Disabled homolog 1 (Drosophila)	-1.7	-6.6	1.9E-06	0.013
A_23_P49254	HBQ1	Hemoglobin, theta 1	-1.6	-6.7	1.5E-06	0.013
A_32_P159612	ACCN1	Amiloride-sensitive cation channel 1, neuronal (degenerin)	-1.7	-6.4	3.1E-06	0.017
A_24_P416097	ENST00000336999	UNC45 homolog A (UNC-45A) (Smooth muscle cell-associated protein 1)	-1.2	-6.3	3.6E-06	0.017
A_24_P75190	HBD	Hemoglobin, delta	-2.0	-5.7	1.3E-05	0.030
A_23_P29939	SNCA	Synuclein, alpha (non A4 component of amyloid precursor), transcript variant NACP112	-2.0	-5.7	1.4E-05	0.030

Table 3. Scores from three gene expression signatures (genomic grade index (GGI) [26], hypoxia [27], and vascular invasion [28]) were compared between the two xenograft models and between treatments within each model. Abbreviations: Dox: doxorubicin, Bev: bevacizumab.

Gene signature	Differences in controls	Treatment differences basal-like	Treatment differences luminal-like
GGI	Not significant difference	Day 3: Dox↑** Day 10: Bev↓**, Bev+Dox↓**	Day 3: Bev ↓*
Hypoxia	Luminal-Like↑***	Day 3: Bev↑**, Bev+Dox↑* Day 10: Bev↑*** , Bev+Dox↑**	Day 3: Bev↑* Day 10: Bev↑**
Vascular invasion	Basal-like↑***	Not significant difference	Day 10: Dox↑*, Bev↑*, Bev+Dox↑*

*p < 0.05, ** p < 0.01, *** p < 0.001

Figure legends

Figure 1. Illustration of the experimental design of treatment experiments and metabolomic and transcriptomic analysis. Con: control; Dox: doxorubicin; Bev: bevacizumab; wo/est: without estrogen; rep: replicates; HR MAS MRS: high resolution magic angle spinning magnetic resonance spectroscopy.

Figure 2. Concentrations of glycerophosphocholine (GPC) and phosphocholine (PCho) at day 10 is plotted for each xenograft model and colored according to treatment (n = 3 animals per group). *p-value < 0.05.

Figure 3. Heatmaps of significantly differentially expressed genes (FDR < 0.05) at day 10 between controls and the most efficient treatment for each model: (A) combination treatment for basal-like and (B) doxorubicin for luminal-like xenografts (only the top 100 transcripts are shown). Data were centered across genes and clustering of samples and genes was performed using Euclidian distance and complete linkage. Red is relatively higher and blue is relatively lower expression levels. Transcripts that were also significantly differentially expressed when comparing combination treated with doxorubicin treated basal-like tumors are indicated with bold gene symbols in A.

Supplementary material

Supplementary Figure 1. Concentrations of creatine, choline, phosphocholine (PCho), glycerophosphocholine (GPC), taurine, glycine, and lactate are plotted separately for each xenograft model at day 3 and 10 and are colored according to treatment. *p-value < 0.05.

Supplementary Figure 2. Expression of genes involved in phospholipid metabolism in a patient cohort. A) Heatmap of basal-like and luminal-like human tumor samples and 105 phospholipid metabolism genes. Samples and genes are clustered using Pearson's correlation distance and average linkage. B) The t-statistics from comparisons of basal-like and luminal-like groups are plotted for the patient and xenograft data. C) log10 of adjusted p-values (FDR) for the t-statistics in B are plotted. For the genes higher expressed in basal-like, the log10(FDR) value is multiplied by -1 (returning positive values).

Supplementary Table 1. Significantly differentially expressed genes between combination treated and untreated basal-like tumors at day 10. The columns ProbeName, GeneName, SystematicName and Description are annotation from Agilent, while the columns logFC (log2 fold change), AveExpr (average expression), t (t-statistic), P.Value (p-value), adj.P.Val (FDR) are output from differential expression analysis using Limma.

Supplementary Table 2. Enriched GO-terms obtained using GOrilla [25] ($p < 10^{-9}$). Enrichment is defined as (b/n) / (B/N), where N is the total number of genes, B is the total number of genes associated with a specific GO term, n is the number of genes in the "target set", b is the number of genes in the "target set" associated with a specific GO-term.

Supplementary Table 3. Significantly differentially expressed genes between doxorubicin treated and untreated luminal-like tumors at day 10. The columns ProbeName, GeneName, SystematicName and Description are annotation from Agilent, while the columns logFC (log2 fold change), AveExpr (average expression), t (t-statistic), P.Value (p-value), adj.P.Val (FDR) are output from differential expression analysis using Limma.

References

1. Burstein HJ: Bevacizumab for advanced breast cancer: all tied up with a RIBBON? *J Clin Oncol* 2011; **29**(10): 1232-1235

2. Smith I, Pierga JY, Biganzoli L *et al.*: Final overall survival results and effect of prolonged (>/=1 year) first-line bevacizumab-containing therapy for metastatic breast cancer in the ATHENA trial. *Breast Cancer Res Treat* 2011; DOI: 10.1007/s10549-011-1695-8

3. Schneider BP & Sledge GW, Jr.: Anti-vascular endothelial growth factor therapy for breast cancer: can we pick the winners? *J Clin Oncol* 2011; **29**(18): 2444-2447

4. Schneider BP, Wang M, Radovich M *et al.*: Association of vascular endothelial growth factor and vascular endothelial growth factor receptor-2 genetic polymorphisms with outcome in a trial of paclitaxel compared with paclitaxel plus bevacizumab in advanced breast cancer: ECOG 2100. *J Clin Oncol* 2008; **26**(28): 4672-4678

5. Yang SX, Steinberg SM, Nguyen D *et al.*: Gene expression profile and angiogenic marker correlates with response to neoadjuvant bevacizumab followed by bevacizumab plus chemotherapy in breast cancer. *Clin Cancer Res* 2008; **14**(18): 5893-5899

6. Jubb AM, Miller KD, Rugo HS *et al.*: Impact of exploratory biomarkers on the treatment effect of bevacizumab in metastatic breast cancer. *Clin Cancer Res* 2011; **17**(2): 372-381

7. Baar J, Silverman P, Lyons J *et al.*: A vasculature-targeting regimen of preoperative docetaxel with or without bevacizumab for locally advanced breast cancer: impact on angiogenic biomarkers. *Clin Cancer Res* 2009; **15**(10): 3583-3590

8. Jagannathan NR, Kumar M, Seenu V *et al.*: Evaluation of total choline from in-vivo volume localized proton MR spectroscopy and its response to neoadjuvant chemotherapy in locally advanced breast cancer. *Br J Cancer* 2001; **84**(8): 1016-1022

9. Meisamy S, Bolan PJ, Baker EH *et al.*: Neoadjuvant chemotherapy of locally advanced breast cancer: predicting response with in vivo (1)H MR spectroscopy--a pilot study at 4 T. *Radiology* 2004; **233**(2): 424-431

10. Beloueche-Babari M, Chung YL, Al-Saffar NM *et al.*: Metabolic assessment of the action of targeted cancer therapeutics using magnetic resonance spectroscopy. *Br J Cancer* 2010; **102**(1): 1-7

11. Podo F, Canevari S, Canese R *et al.*: MR evaluation of response to targeted treatment in cancer cells. *NMR Biomed* 2011; **24**(6): 648-672

12. Vargo-Gogola T & Rosen JM: Modelling breast cancer: one size does not fit all. *Nat Rev Cancer* 2007; **7**(9): 659-672

13. Bergamaschi A, Hjortland GO, Triulzi T *et al.*: Molecular profiling and characterization of luminal-like and basal-like in vivo breast cancer xenograft models. *Mol Oncol* 2009; **3**(5-6): 469-482

14. Grinde MT, Moestue SA, Borgan E *et al.*: (13) C High-resolution-magic angle spinning MRS reveals differences in glucose metabolism between two breast cancer xenograft models with different gene expression patterns. *NMR Biomed* 2011; DOI: 10.1002/nbm.1683

15. Moestue SA, Borgan E, Huuse E *et al.*: Distinct choline metabolic profiles are associated with differences in gene expression for basal-like and luminal-like breast cancer xenograft models. *BMC Cancer* 2010; **10**(1): 433

16. Lindholm E, Kristian A, Nalwoga H *et al.*: Effect of antiangiogenic therapy on tumor growth, vasculature and kinase activity in basal- and luminal-like breast cancer xenografts. *submitted to Br J Cancer* 2011;

17. Sitter B, Sonnewald U, Spraul M *et al.*: High-resolution magic angle spinning MRS of breast cancer tissue. *NMR Biomed* 2002; **15**(5): 327-337

18. Gentleman RC, Carey VJ, Bates DM *et al.*: Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004; **5**(10): R80

19. Troyanskaya O, Cantor M, Sherlock G *et al.*: Missing value estimation methods for DNA microarrays. *Bioinformatics* 2001; **17**(6): 520-525

20. Smyth GK: Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004; **3**: Article3

21. Smyth GK, Michaud J, & Scott HS: Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics* 2005; **21**(9): 2067-2075

22. Smyth,G.K. Limma: linear models for microarray data in *Gentleman R., Carey V., Dudoit S., Irizarry R., Huber W. (eds.), Bioinformatics and Computational Biology Solutions using R and Bioconductor* 397-420 (Springer, New York, 2011).

23. Benjamini Y & Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)* 1995; **57**(1): 289-300

24. Eden E, Lipson D, Yogev S *et al.*: Discovering motifs in ranked lists of DNA sequences. *PLoS Comput Biol* 2007; **3**(3): e39

25. Eden E, Navon R, Steinfeld I *et al.*: GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 2009; **10**(1): 48

26. Sotiriou C, Wirapati P, Loi S *et al.*: Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 2006; **98**(4): 262-272

27. Chi JT, Wang Z, Nuyten DS *et al.*: Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. *PLoS Med* 2006; **3**(3): e47 doi:10.1371/journal.pmed.0030047

28. Mannelqvist M, Stefansson IM, Bredholt G *et al.*: Gene expression patterns related to vascular invasion and aggressive features in endometrial cancer. *Am J Pathol* 2011; **178**(2): 861-871

29. Durinck S, Moreau Y, Kasprzyk A *et al.*: BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* 2005; **21**(16): 3439-3440

30. Stefansson IM, Salvesen HB, Immervoll H *et al.*: Prognostic impact of histological grade and vascular invasion compared with tumour cell proliferation in endometrial carcinoma of endometrioid type. *Histopathology* 2004; **44**(5): 472-479

31. Enerly E, Steinfeld I, Kleivi K *et al.*: miRNA-mRNA integrated analysis reveals roles for miRNAs in primary breast tumors. *PLoS One* 2011; **6**(2): e16915 doi:10.1371/journal.pone.0016915

32. Bergers G & Hanahan D: Modes of resistance to anti-angiogenic therapy. *Nat Rev Cancer* 2008; **8**(8): 592-603

33. Eliyahu G, Kreizman T, & Degani H: Phosphocholine as a biomarker of breast cancer: molecular and biochemical studies. *Int J Cancer* 2007; **120**(8): 1721-1730

34. Morse DL, Raghunand N, Sadarangani P *et al.*: Response of choline metabolites to docetaxel therapy is quantified in vivo by localized (31)P MRS of human breast cancer xenografts and in vitro by high-resolution (31)P NMR spectroscopy of cell extracts. *Magn Reson Med* 2007; **58**(2): 270-280

35. Oganesian A, Armstrong LC, Migliorini MM *et al.*: Thrombospondins use the VLDL receptor and a nonapoptotic pathway to inhibit cell division in microvascular endothelial cells. *Mol Biol Cell* 2008; **19**(2): 563-571

36. Saidi A, Javerzat S, Bellahcene A *et al.*: Experimental anti-angiogenesis causes upregulation of genes associated with poor survival in glioblastoma. *Int J Cancer* 2008; **122**(10): 2187-2198

37. Nishikawa KC & Millis AJ: gp38k (CHI3L1) is a novel adhesion and migration factor for vascular cells. *Exp Cell Res* 2003; **287**(1): 79-87

38. Glunde K, Jie C, & Bhujwalla ZM: Molecular causes of the aberrant choline phospholipid metabolism in breast cancer. *Canc Res* 2004; **64**(12): 4270-4276

39. Katz-Brull R, Seger D, Rivenson-Segal D *et al.*: Metabolic markers of breast cancer: enhanced choline metabolism and reduced choline-ether-phospholipid synthesis. *Cancer Res* 2002; **62**(7): 1966-1970

40. Morse DL, Carroll D, Day S *et al.*: Characterization of breast cancers and therapy response by MRS and quantitative gene expression profiling in the choline pathway. *NMR Biomed* 2009; **22**(1): 114-127

41. Barzilai A, Horowitz A, Geier A *et al.*: Phosphate metabolites and steroid hormone receptors of benign and malignant breast tumors. A Nuclear Magnetic Resonance study. *Cancer* 1991; **67**(11): 2919-2925

42. Giskeodegard GF, Grinde MT, Sitter B *et al.*: Multivariate modeling and prediction of breast cancer prognostic factors using MR metabolomics. *J Proteome Res* 2010; **9**(2): 972-979

43. Kerbel RS: Issues regarding improving the impact of antiangiogenic drugs for the treatment of breast cancer. *Breast* 2009; **18 Suppl 3**: S41-S47

44. Paez-Ribes M, Allen E, Hudock J *et al.*: Antiangiogenic therapy elicits malignant progression of tumors to increased local invasion and distant metastasis. *Cancer Cell* 2009; **15**(3): 220-231

45. Young SD & Hill RP: Effects of reoxygenation on cells from hypoxic regions of solid tumors: anticancer drug sensitivity and metastatic potential. *J Natl Cancer Inst* 1990; **82**(5): 371-380

46. Du R, Lu KV, Petritsch C *et al.*: HIF1alpha induces the recruitment of bone marrow-derived vascular modulatory cells to regulate tumor angiogenesis and invasion. *Cancer Cell* 2008; **13**(3): 206-220

47. Milosevic J, Maisel M, Wegner F *et al.*: Lack of hypoxia-inducible factor-1 alpha impairs midbrain neural precursor cells involving vascular endothelial growth factor signaling. *J Neurosci* 2007; **27**(2): 412-421

48. Rubenstein JL, Kim J, Ozawa T *et al.*: Anti-VEGF antibody treatment of glioblastoma prolongs survival but results in increased vascular cooption. *Neoplasia* 2000; **2**(4): 306-314

49. Bear HD, Tang G, Rastogi P *et al.*: The effect on pCR of bevacizumab and/or antimetabolites added to standard neoadjuvant chemotherapy: NSABP protocol B-40. *J Clin Oncol* 2011; **29**: (suppl; abstr LBA1005)

50. Gerber B, Eidtmann H, Rezai M *et al.*: Neoadjuvant bevacizumab and anthracycline–taxane-based chemotherapry in 686 triple-negative primary breast cancers: Seconday endpoint analysis of the GeparQuinto study (GBG 44). *J Clin Oncol* 2011; **29**: (suppl; abstr 1006)

Figure 1

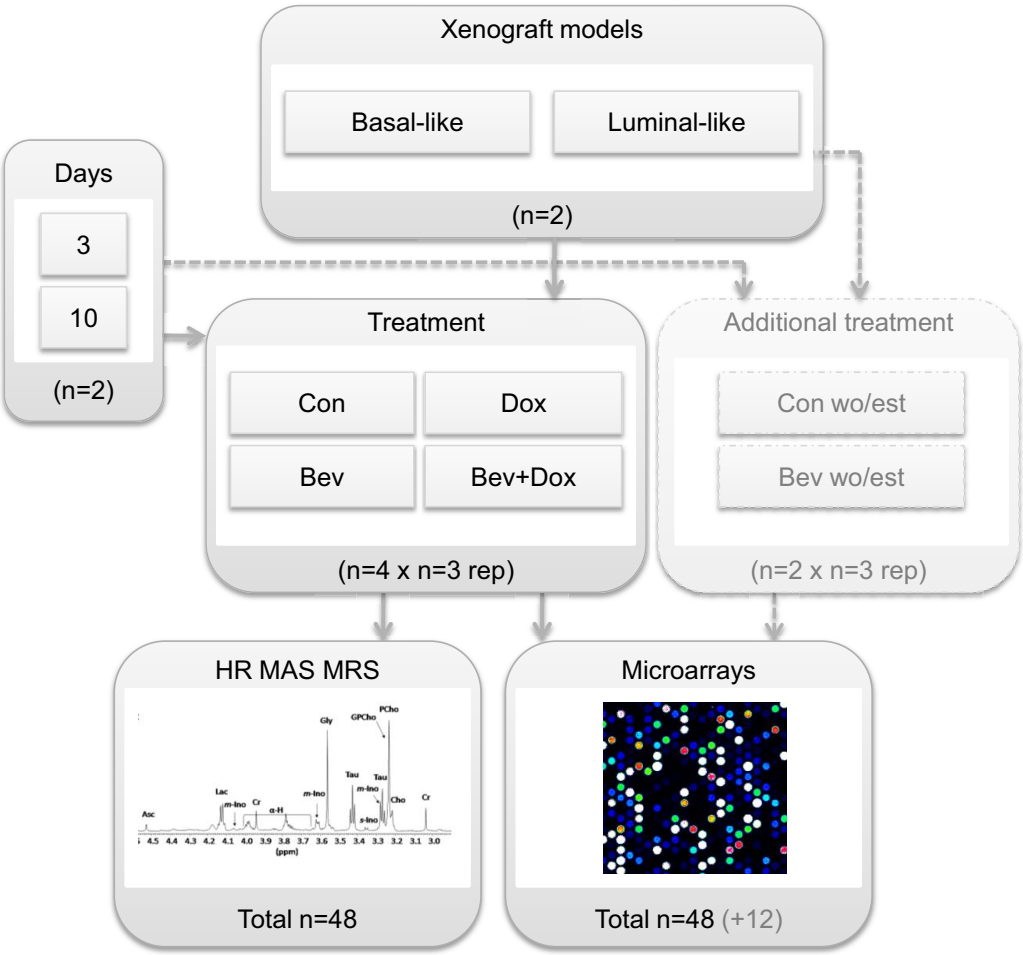

Xenograft models
Basal-like
Luminal-like
(n=2)
Days
3
10
(n=2)
Treatment
Con
Dox
Bev
Bev+Dox
(n=4 x n=3 rep)
Additional treatment
Con wo/est
Bev wo/est
(n=2 x n=3 rep)
HR MAS MRS
Gly
GPCho
PCho
Tau
Tau
Lac
m-Ino
Cr
m-Ino
α-H
m-Ino
Cho
Cr
Asc
s-Ino
4.5 4.4 4.3 4.2 4.1 4.0 3.9 3.8 3.7 3.6 3.5 3.4 3.3 3.2 3.1 3.0
(ppm)
Total n=48
Microarrays
Total n=48 (+12)

Figure 2

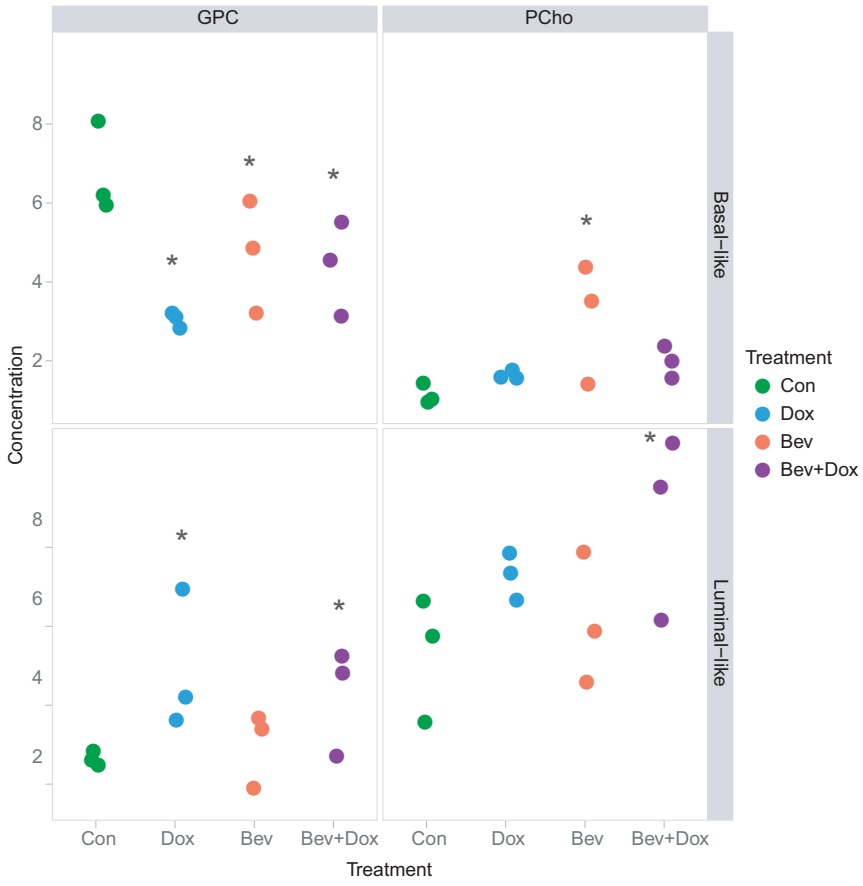
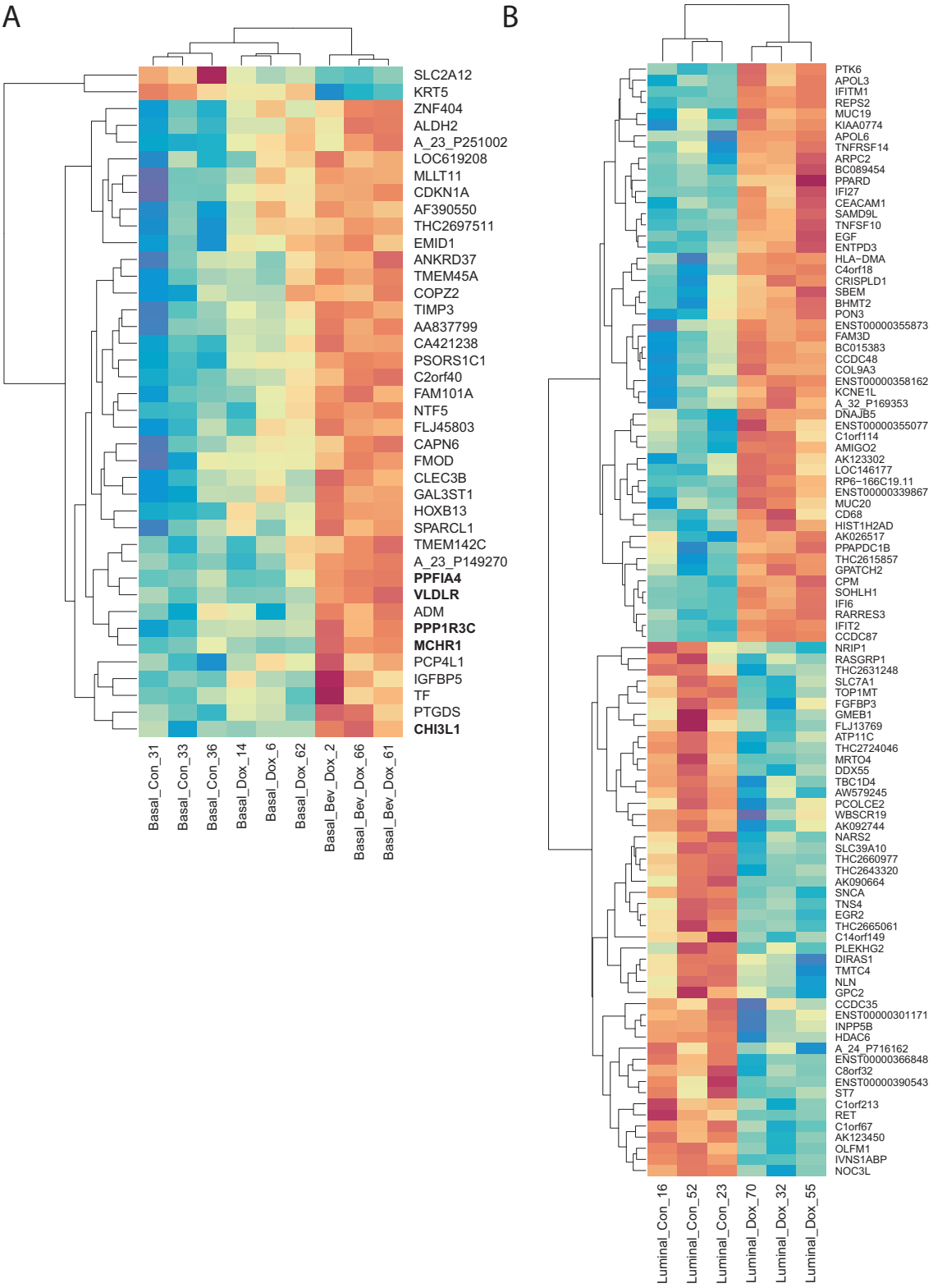
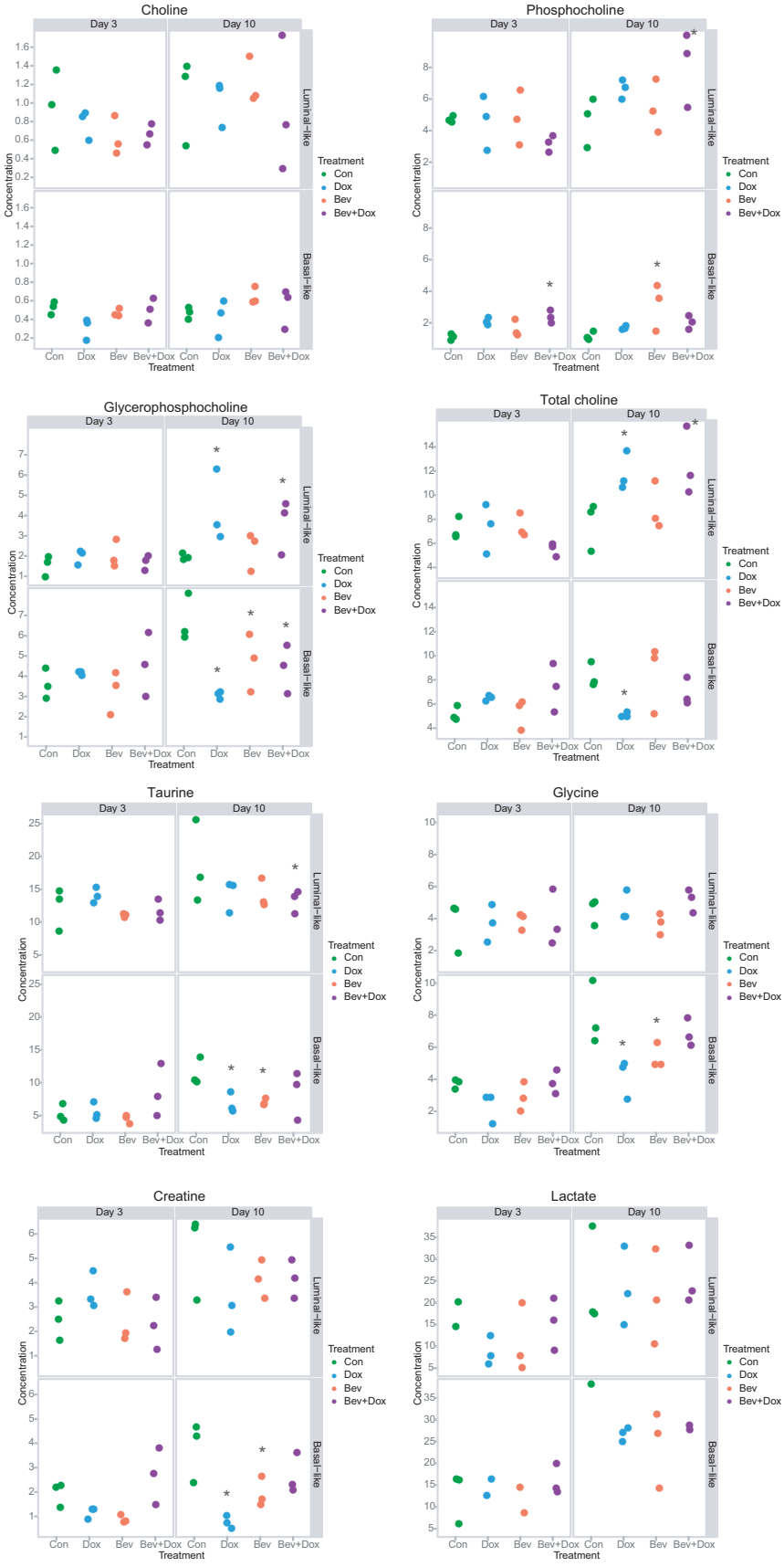

GPC
PCho
Basal–like
Luminal–like
Concentration
8
6
4
2
Con
Dox
Bev
Bev+Dox
Treatment
Treatment
Con
Dox
Bev
Bev+Dox

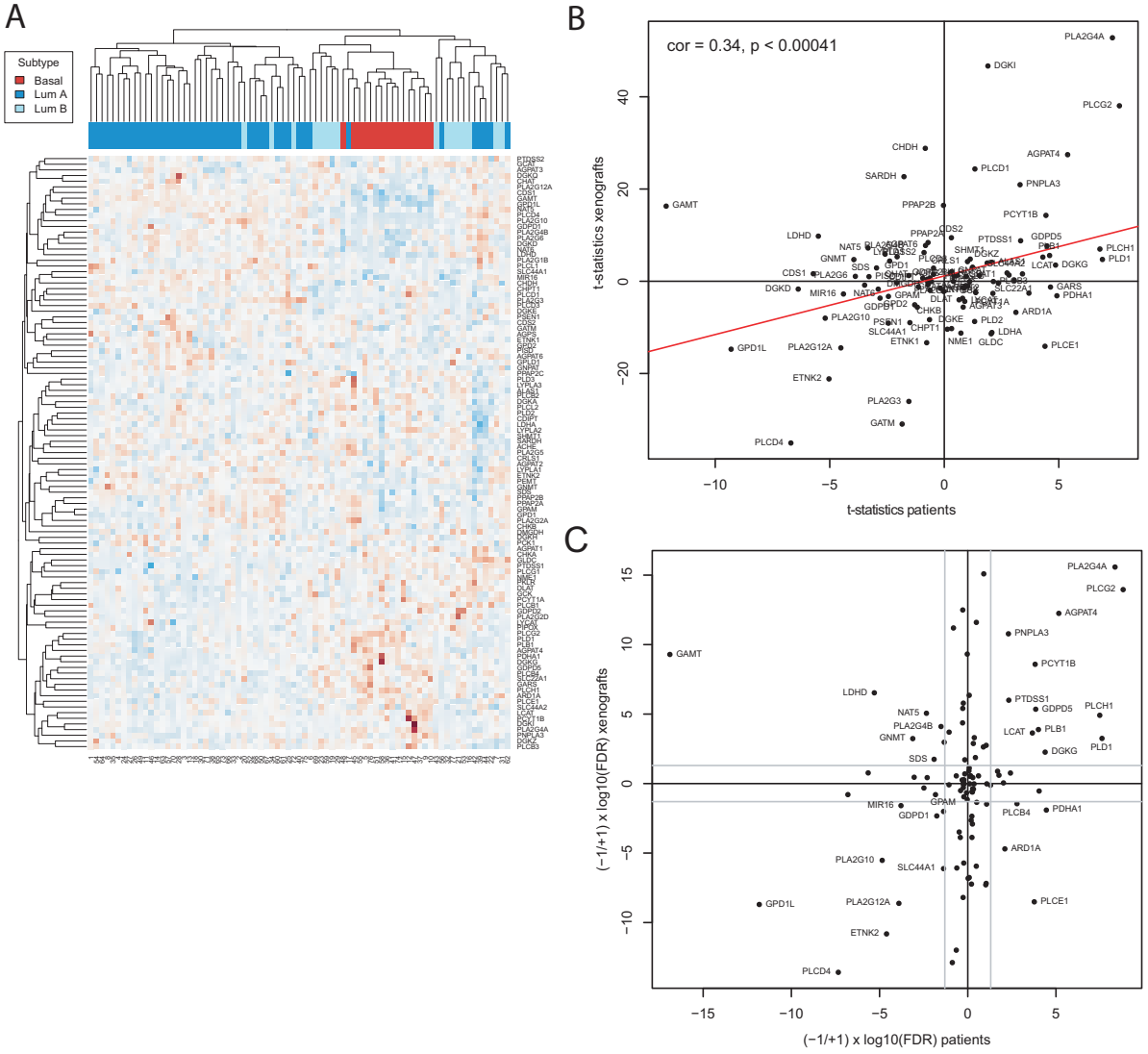
Figure 3


A
SLC2A12
KRT5
ZNF404
ALDH2
A_23_P251002
LOC619208
MLLT11
CDKN1A
AF390550
THC2697511
EMID1
ANKRD37
TMEM45A
COPZ2
TIMP3
AA837799
CA421238
PSORS1C1
C2orf40
FAM101A
NTF5
FLJ45803
CAPN6
FMOD
CLEC3B
GAL3ST1
HOXB13
SPARCL1
TMEM142C
A_23_P149270
PPFIA4
VLDLR
ADM
PPP1R3C
MCHR1
PCP4L1
IGFBP5
TF
PTGDS
CHI3L1
Basal_Con_31
Basal_Con_33
Basal_Con_36
Basal_Dox_14
Basal_Dox_6
Basal_Dox_62
Basal_Bev_Dox_2
Basal_Bev_Dox_66
Basal_Bev_Dox_61
B
PTK6
APOL3
IFITM1
REPS2
MUC19
KIAA0774
APOL6
TNFRSF14
ARPC2
BC089454
PPARD
IFI27
CEACAM1
SAMD9L
TNFSF10
EGF
ENTPD3
HLA−DMA
C4orf18
CRISPLD1
SBEM
BHMT2
PON3
ENST00000355873
FAM3D
BC015383
CCDC48
COL9A3
ENST00000358162
KCNE1L
A_32_P169353
DNAJB5
ENST00000355077
C1orf114
AMIGO2
AK123302
LOC146177
RP6−166C19.11
ENST00000339867
MUC20
CD68
HIST1H2AD
AK026517
PPAPDC1B
THC2615857
GPATCH2
CPM
SOHLH1
IFI6
RARRES3
IFIT2
CCDC87
NRIP1
RASGRP1
THC2631248
SLC7A1
TOP1MT
FGFBP3
GMEB1
FLJ13769
ATP11C
THC2724046
MRTO4
DDX55
TBC1D4
AW579245
PCOLCE2
WBSCR19
AK092744
NARS2
SLC39A10
THC2660977
THC2643320
AK090664
SNCA
TNS4
EGR2
THC2665061
C14orf149
PLEKHG2
DIRAS1
TMTC4
NLN
GPC2
CCDC35
ENST00000301171
INPP5B
HDAC6
A_24_P716162
ENST00000366848
C8orf32
ENST00000390543
ST7
C1orf213
RET
C1orf67
AK123450
OLFM1
IVNS1ABP
NOC3L
Luminal_Con_16
Luminal_Con_52
Luminal_Con_23
Luminal_Dox_70
Luminal_Dox_32
Luminal_Dox_55

Supplementary Figure 1



Supplementary Figure 2



Supplementary Table 1

ProbeName	GeneName	SystematicName	Description	logFC	AveExpr	t	P.Value	adj.P.Val
A_23_P4212	HOXB13	NM_006361	Homo sapiens homeobox B13 (HOXB13), mRNA [NM_006361]	1.7	7.1	9.3	0.0000	0.0003
A_23_P35414	PPP1R3C	NM_005398	Homo sapiens protein phosphatase 1, regulatory (inhibitor) subunit 3C (PPP1R3C), mRNA [NM_005398]	2.2	7.4	8.1	0.0000	0.0013
A_23_P113212	TMEM45A	NM_018004	Homo sapiens transmembrane protein 45A (TMEM45A), mRNA [NM_018004]	1.9	9.2	7.9	0.0000	0.0013
A_23_P43476	VLDLR	NM_003383	Homo sapiens very low density lipoprotein receptor (VLDLR), transcript variant 1, mRNA [NM_003383]	1.4	7.5	7.7	0.0000	0.0013
A_23_P69497	CLEC3B	NM_003278	Homo sapiens C-type lectin domain family 3, member B (CLEC3B), mRNA [NM_003278]	2.4	12.1	7.3	0.0000	0.0026
A_23_P420692	PPFIA4	NM_015053	Homo sapiens protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 4 (PPFIA4), mRNA [NM_015053]	1.1	7.8	6.8	0.0000	0.0059
A_23_P137665	CHI3L1	NM_001276	Homo sapiens chitinase 3-like 1 (cartilage glycoprotein-39) (CHI3L1), mRNA [NM_001276]	1.1	12.7	6.6	0.0000	0.0073
A_23_P56433	C2orf40	NM_032411	Homo sapiens chromosome 2 open reading frame 40 (C2orf40), mRNA [NM_032411]	1.4	6.5	6.3	0.0000	0.0114
A_23_P4899	NTF5	NM_006179	Homo sapiens neurotrophin 5 (neurotrophin 4/5) (NTF5), mRNA [NM_006179]	1.5	6.8	6.3	0.0000	0.0119
A_23_P399078	TIMP3	NM_000362	Homo sapiens TIMP metallopeptidase inhibitor 3 (Sorsby fundus dystrophy, pseudoinflammatory) (TIMP3), mRNA [NM_000362]	2.9	8.4	6.2	0.0000	0.0123
A_23_P211543	MCHR1	NM_005297	Homo sapiens melanin-concentrating hormone receptor 1 (MCHR1), mRNA [NM_005297]	0.7	5.5	5.9	0.0000	0.0215
A_24_P14634	EMID1	NM_133455	Homo sapiens EMI domain containing 1 (EMID1), mRNA [NM_133455]	1.0	9.0	5.8	0.0000	0.0254
A_23_P113351	SPARCL1	NM_004684	Homo sapiens SPARC-like 1 (mast9, hevin) (SPARCL1), mRNA [NM_004684]	2.1	8.1	5.7	0.0000	0.0260
A_24_P237586	ANKRD37	NM_181726	Homo sapiens ankyrin repeat domain 37 (ANKRD37), mRNA [NM_181726]	1.3	9.0	5.7	0.0000	0.0260
A_23_P154115	IGFBP5	NM_000599	Homo sapiens insulin-like growth factor binding protein 5 (IGFBP5), mRNA [NM_000599]	3.1	7.1	5.7	0.0000	0.0260
A_23_P218047	KRT5	NM_000424	Homo sapiens keratin 5 (epidermolysis bullosa simplex, Dowling-Meara/Kobner/Weber-Cockayne types) (KRT5), mRNA [NM_000424]	-1.7	9.2	-5.7	0.0000	0.0260
A_32_P330691	AF390550	AF390550	Homo sapiens cervical cancer suppressor gene 5 mRNA, complete cds. [AF390550]	0.9	6.8	5.6	0.0000	0.0283
A_23_P211468	AA837799	AA837799	oe06h09.s1 NCI_CGAP_Ov2 Homo sapiens cDNA clone IMAGE:1385153, mRNA sequence [AA837799]	2.6	9.7	5.5	0.0000	0.0283
A_23_P36753	ALDH2	NM_000690	Homo sapiens aldehyde dehydrogenase 2 family (mitochondrial) (ALDH2), nuclear gene encoding mitochondrial protein, mRNA [NM_000690]	1.6	9.9	5.5	0.0000	0.0283
A_23_P106898	TMEM142C	NM_152288	Homo sapiens transmembrane protein 142C (TMEM142C), mRNA [NM_152288]	1.0	8.3	5.5	0.0000	0.0283
A_23_P90333	ZNF404	NM_001033719	Homo sapiens zinc finger protein 404 (ZNF404), mRNA [NM_001033719]	0.6	8.0	5.5	0.0000	0.0283
A_32_P214665	PCP4L1	ENST00000309878	Homo sapiens cDNA clone IMAGE:4792407. [BC067884]	1.7	8.1	5.5	0.0000	0.0283
A_32_P200697	FAM101A	NM_181709	Homo sapiens family with sequence similarity 101, member A (FAM101A), mRNA [NM_181709]	2.8	7.3	5.4	0.0000	0.0283
A_23_P127948	ADM	NM_001124	Homo sapiens adrenomedullin (ADM), mRNA [NM_001124]	2.7	11.9	5.4	0.0000	0.0283
A_23_P133902	PSORS1C1	NM_014068	Homo sapiens psoriasis susceptibility 1 candidate 1 (PSORS1C1), mRNA [NM_014068]	1.0	6.6	5.3	0.0000	0.0337
A_23_P251002	A_23_P251002	A_23_P251002	Unknown	0.7	8.1	5.3	0.0000	0.0356
A_23_P149270	A_23_P149270	A_23_P149270	Unknown	0.9	6.1	5.3	0.0000	0.0364
A_23_P101093	COPZ2	NM_016429	Homo sapiens coatomer protein complex, subunit zeta 2 (COPZ2), mRNA [NM_016429]	1.1	8.2	5.3	0.0000	0.0364
A_24_P756657	LOC619208	NM_001033564	Homo sapiens hypothetical protein LOC619208 (LOC619208), mRNA [NM_001033564]	0.8	9.3	5.2	0.0000	0.0376
A_23_P59210	CDKN1A	NM_000389	Homo sapiens cyclin-dependent kinase inhibitor 1A (p21, Cip1) (CDKN1A), transcript variant 1, mRNA [NM_000389]	1.3	8.7	5.2	0.0000	0.0376
A_32_P22750	CA421238	CA421238	CA421238 UI-H-FG0-bct-h-02-0-UI.s1 NCI_CGAP_EN1_2 Homo sapiens cDNA clone UI-H-FG0-bct-h-02-0-UI 3', mRNA sequence [CA421238]	0.9	6.1	5.2	0.0000	0.0376
A_23_P146554	PTGDS	NM_000954	Homo sapiens prostaglandin D2 synthase 21kDa (brain) (PTGDS), mRNA [NM_000954]	1.8	8.9	5.2	0.0000	0.0376
A_23_P395001	SLC2A12	NM_145176	Homo sapiens solute carrier family 2 (facilitated glucose transporter), member 12 (SLC2A12), mRNA [NM_145176]	-0.7	6.6	-5.2	0.0000	0.0376
A_23_P114883	FMOD	NM_002023	Homo sapiens fibromodulin (FMOD), mRNA [NM_002023]	1.7	10.9	5.2	0.0000	0.0376
A_23_P120863	GAL3ST1	NM_004861	Homo sapiens galactose-3-O-sulfotransferase 1 (GAL3ST1), mRNA [NM_004861]	0.9	6.3	5.1	0.0000	0.0376
A_23_P217570	CAPN6	NM_014289	Homo sapiens calpain 6 (CAPN6), mRNA [NM_014289]	1.2	6.5	5.1	0.0001	0.0413
A_23_P23346	MLLT11	NM_006818	Homo sapiens myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 11 (MLLT11), mRNA [NM_006818]	1.1	9.6	5.1	0.0001	0.0413
A_32_P180336	FLJ45803	NM_207429	Homo sapiens FLJ45803 protein (FLJ45803), mRNA [NM_207429]	1.6	9.7	5.1	0.0001	0.0413
A_32_P49854	THC2697511	THC2697511	Unknown	0.8	7.5	5.0	0.0001	0.0465
A_23_P212500	TF	NM_001063	Homo sapiens transferrin (TF), mRNA [NM_001063]	2.1	13.2	5.0	0.0001	0.0477

Supplementary Table 2

MAS98.12 bev+dox versus Con

Higher in bev+dox

GO Term	Description	P-value	Enrichment	N	B	n	b
GO:0003012	muscle system process	1.07E-11	6.75	15472	171	295	22
GO:0006936	muscle contraction	6.28E-11	6.95	15472	151	295	20

Lower in bev+dox

GO Term	Description	P-value	Enrichment	N	B	n	b
None							

MAS98.12 bev+dox versus dox

Higher in bev+dox

GO Term	Description	P-value	Enrichment	N	B	n	b
GO:0003012	muscle system process	4.89E-11	7.61	15472	171	226	19
GO:0006936	muscle contraction	5.20E-11	8.16	15472	151	226	18
GO:0003008	system process	7.19E-11	3.34	15472	764	255	42

Lower in bev+dox

GO Term	Description	P-value	Enrichment	N	B	n	b
GO:0002376	immune system process	1.05E-13	5.58	15472	992	81	29
GO:0006952	defense response	6.41E-12	6.82	15472	616	81	22
GO:0051240	positive regulation of multicellular organismal process	1.21E-11	5.35	15472	347	225	27
GO:0009607	response to biotic stimulus	2.84E-11	4.97	15472	402	217	28
GO:0034097	response to cytokine stimulus	7.60E-11	6.22	15472	320	171	22
GO:0051239	regulation of multicellular organismal process	2.15E-10	2.76	15472	1258	232	52
GO:0006954	inflammatory response	2.68E-10	10.91	15472	245	81	14
GO:0034612	response to tumor necrosis factor	3.13E-10	40.1	15472	63	49	8
GO:0006955	immune response	3.75E-10	4.17	15472	580	192	30
GO:0009611	response to wounding	7.12E-10	8.24	15472	371	81	16

MAS98.06 Dox versus Con

Higher in dox

GO Term	Description	P-value	Enrichment	N	B	n	b
GO:0071357	cellular response to type I interferon	8.29E-19	9.57	15472	58	753	27
GO:0060337	type I interferon-mediated signaling pathway	8.29E-19	9.57	15472	58	753	27
GO:0034340	response to type I interferon	1.49E-18	9.4	15472	59	753	27
GO:0019221	cytokine-mediated signaling pathway	3.41E-14	4.53	15472	186	753	41
GO:0071345	cellular response to cytokine stimulus	4.80E-14	4.15	15472	223	753	45
GO:0034097	response to cytokine stimulus	2.96E-13	4.3	15472	320	427	38
GO:0003012	muscle system process	5.82E-13	8.33	15472	171	228	21
GO:0006936	muscle contraction	6.42E-12	8.54	15472	151	228	19
GO:0030049	muscle filament sliding	1.04E-11	162.86	15472	38	15	6
GO:0033275	actin-myosin filament sliding	1.04E-11	162.86	15472	38	15	6
GO:0070252	actin-mediated cell contraction	1.45E-11	154.72	15472	40	15	6
GO:0030048	actin filament-based movement	8.37E-11	119.02	15472	52	15	6
GO:0006941	striated muscle contraction	9.27E-11	19.14	15472	39	228	11

Lower in dox

GO Term	Description	P-value	Enrichment	N	B	n	b
GO:0090304	nucleic acid metabolic process	9.20E-16	1.53	15472	2109	1773	369
GO:0006259	DNA metabolic process	1.77E-14	1.98	15472	604	1773	137
GO:0006260	DNA replication	1.25E-13	2.93	15472	170	1771	57
GO:0034641	cellular nitrogen compound metabolic process	2.70E-13	1.36	15472	2991	1776	468
GO:0006139	nucleobase, nucleoside, nucleotide & nucleic acid metabolic process	2.97E-13	1.43	15472	2602	1776	426
GO:0022403	cell cycle phase	4.25E-13	2.14	15472	461	1788	114
GO:0022402	cell cycle process	5.38E-13	1.87	15472	739	1788	160
GO:0034470	ncRNA processing	6.10E-13	2.76	15472	192	1751	60
GO:0006807	nitrogen compound metabolic process	8.96E-13	1.35	15472	3089	1776	478
GO:0022616	DNA strand elongation	1.07E-11	6.26	15472	33	1498	20
GO:0006310	DNA recombination	2.72E-11	2.83	15472	161	1700	50
GO:0044260	cellular macromolecule metabolic process	4.13E-11	1.28	15472	3940	1757	571
GO:0034660	ncRNA metabolic process	5.02E-11	2.33	15472	265	1751	70
GO:0007049	cell cycle	6.28E-11	1.78	15472	677	1788	139
GO:0006271	DNA strand elongation involved in DNA replication	3.86E-10	6	15472	31	1498	18

Supplementary Table 3

ProbeName	GeneName	SystematicName	Description	logFC	AveExpr	t	P.Value	adj.P.Val
A_23_P155979	EGF	NM_001963	Homo sapiens epidermal growth factor (beta-urogastrone) (EGF), mRNA [NM_001963]	1.6	6.6	7.6	0.0000	0.0004
A_23_P1962	RARRES3	NM_004585	Homo sapiens retinoic acid receptor responder (tazarotene induced) 3 (RARRES3), mRNA [NM_004585]	1.6	8.4	7.5	0.0000	0.0004
A_23_P133536	CAPSL	NM_144647	Homo sapiens calcyphosine-like (CAPSL), transcript variant 1, mRNA [NM_144647]	1.5	8.0	7.2	0.0000	0.0007
A_23_P77731	CRYM	NM_001888	Homo sapiens crystallin, mu (CRYM), transcript variant 1, mRNA [NM_001888]	2.5	8.1	6.9	0.0000	0.0009
A_23_P26457	HBA2	NM_000517	Homo sapiens hemoglobin, alpha 2 (HBA2), mRNA [NM_000517]	-3.6	10.9	-6.9	0.0000	0.0009
A_23_P37856	HBA1	NM_000558	Homo sapiens hemoglobin, alpha 1 (HBA1), mRNA [NM_000558]	-3.5	10.8	-6.7	0.0000	0.0011
A_32_P33083	VCX2	NM_016378	Homo sapiens variable charge, X-linked 2 (VCX2), mRNA [NM_016378]	1.5	13.8	6.6	0.0000	0.0011
A_23_P121253	TNFSF10	NM_003810	Homo sapiens tumor necrosis factor (ligand) superfamily, member 10 (TNFSF10), mRNA [NM_003810]	1.5	7.4	6.6	0.0000	0.0011
A_24_P23979	VCX3A	NM_016379	Homo sapiens variable charge, X-linked 3A (VCX3A), mRNA [NM_016379]	1.5	14.0	6.5	0.0000	0.0016
A_24_P245976	VCX	NM_013452	Homo sapiens variable charge, X-linked (VCX), mRNA [NM_013452]	1.5	13.7	6.3	0.0000	0.0021
A_32_P196520	AK023159	AK023159	Homo sapiens cDNA FLJ13097 fis, clone NT2RP3002173. [AK023159]	-0.8	6.9	-6.3	0.0000	0.0021
A_24_P529168	MIRH1	BC040320	Homo sapiens chromosome 13 open reading frame 25, mRNA (cDNA clone IMAGE:4838522). [BC040320]	-1.2	7.8	-6.2	0.0000	0.0023
A_32_P352358	OR7E125P	NM_001040066	Homo sapiens similar to olfactory receptor 873 (PJCG6), mRNA [NM_001040066]	0.8	7.1	6.2	0.0000	0.0023
A_23_P151915	GCNT3	ENST00000267857	glucosaminyl (N-acetyl) transferase 3, mucin type [Source:RefSeq_peptide;Acc:NP_004742] [ENST00000267857]	1.0	6.2	6.1	0.0000	0.0028
A_23_P319423	KCNK5	NM_003740	Homo sapiens potassium channel, subfamily K, member 5 (KCNK5), mRNA [NM_003740]	-1.8	8.5	-6.1	0.0000	0.0028
A_23_P127484	CCDC87	NM_018219	Homo sapiens coiled-coil domain containing 87 (CCDC87), mRNA [NM_018219]	0.7	6.7	6.0	0.0000	0.0028
A_24_P158193	A_24_P158193	A_24_P158193	Unknown	0.8	5.4	6.0	0.0000	0.0028
A_23_P86931	SLC37A2	NM_198277	Homo sapiens solute carrier family 37 (glycerol-3-phosphate transporter), member 2 (SLC37A2), mRNA [NM_198277]	1.3	6.1	6.0	0.0000	0.0028
A_23_P201459	IFI6	NM_022873	Homo sapiens interferon, alpha-inducible protein 6 (IFI6), transcript variant 3, mRNA [NM_022873]	3.0	9.1	6.0	0.0000	0.0028
A_23_P48513	IFI27	NM_005532	Homo sapiens interferon, alpha-inducible protein 27 (IFI27), mRNA [NM_005532]	1.8	6.3	5.9	0.0000	0.0029
A_23_P11390	VCY	NM_004679	Homo sapiens variable charge, Y-linked (VCY), mRNA [NM_004679]	1.4	11.0	5.9	0.0000	0.0029
A_23_P30315	TRIM7	NM_033342	Homo sapiens tripartite motif-containing 7 (TRIM7), transcript variant 6, mRNA [NM_033342]	-1.3	9.2	-5.9	0.0000	0.0032
A_23_P23869	LHX4	NM_033343	Homo sapiens LIM homeobox 4 (LHX4), mRNA [NM_033343]	-1.3	6.4	-5.9	0.0000	0.0032
A_23_P115417	RGL1	NM_015149	Homo sapiens ral guanine nucleotide dissociation stimulator-like 1 (RGL1), mRNA [NM_015149]	1.7	8.0	5.8	0.0000	0.0036
A_23_P21207	UBE1L	NM_003335	Homo sapiens ubiquitin-activating enzyme E1-like (UBE1L), mRNA [NM_003335]	1.4	7.0	5.8	0.0000	0.0036
A_32_P424761	C12orf28	NM_182530	Homo sapiens chromosome 12 open reading frame 28 (C12orf28), mRNA [NM_182530]	1.1	8.5	5.8	0.0000	0.0036
A_24_P147849	A_24_P147849	A_24_P147849	Unknown	-4.0	11.7	-5.7	0.0000	0.0039
A_32_P8952	LMBRD2	NM_001007527	Homo sapiens LMBR1 domain containing 2 (LMBRD2), mRNA [NM_001007527]	-0.7	6.3	-5.7	0.0000	0.0039
A_23_P22013	RSPO1	NM_001038633	Homo sapiens R-spondin homolog (Xenopus laevis) (RSPO1), mRNA [NM_001038633]	-1.5	8.2	-5.7	0.0000	0.0039
A_24_P576445	TBC1D8	AK098124	Homo sapiens cDNA FLJ40805 fis, clone TRACH2009060. [AK098124]	-0.9	6.9	-5.6	0.0000	0.0041
A_23_P500501	FGFR3	NM_000142	Homo sapiens fibroblast growth factor receptor 3 (achondroplasia, thanatophoric dwarfism) (FGFR3), transcript variant 1, mRNA [NM_000142]	-1.1	11.9	-5.6	0.0000	0.0041
A_23_P210158	THC2676186	THC2676186	YG2T_YEAST (P53262) Protein YGR106C precursor, partial (7%) [THC2676186]	-0.6	9.5	-5.6	0.0000	0.0041
A_32_P224327	LOC730091	AK056401	Homo sapiens cDNA FLJ31839 fis, clone NT2RP7000086. [AK056401]	1.3	11.0	5.6	0.0000	0.0041
A_24_P923483	SLC6A13	BC020867	Homo sapiens solute carrier family 6 (neurotransmitter transporter, GABA), member 13, mRNA (cDNA clone IMAGE:4594185), complete cds. [BC020867]	1.3	7.6	5.6	0.0000	0.0041
A_23_P217339	PRKX	NM_005044	Homo sapiens protein kinase, X-linked (PRKX), mRNA [NM_005044]	-0.7	8.0	-5.6	0.0000	0.0041
A_23_P46936	EGR2	NM_000399	Homo sapiens early growth response 2 (Krox-20 homolog, Drosophila) (EGR2), mRNA [NM_000399]	-1.2	6.0	-5.6	0.0000	0.0042
A_24_P769672	KRTHB5	AL359596	Homo sapiens mRNA; cDNA DKFZp547P055 (from clone DKFZp547P055). [AL359596]	-0.6	9.6	-5.5	0.0000	0.0046
A_23_P74449	HPDL	NM_032756	Homo sapiens 4-hydroxyphenylpyruvate dioxygenase-like (HPDL), mRNA [NM_032756]	-1.2	11.3	-5.5	0.0000	0.0050
A_32_P121674	A_32_P121674	A_32_P121674	Unknown	-0.8	8.5	-5.5	0.0000	0.0053
A_23_P307860	SVOPL	NM_174959	Homo sapiens SV2 related protein homolog (rat)-like (SVOPL), mRNA [NM_174959]	0.9	7.0	5.5	0.0000	0.0053
A_23_P131646	RPIA	NM_144563	Homo sapiens ribose 5-phosphate isomerase A (ribose 5-phosphate epimerase) (RPIA), mRNA [NM_144563]	-0.6	12.5	-5.4	0.0000	0.0061
A_23_P34757	GPATCH2	NM_018040	Homo sapiens G patch domain containing 2 (GPATCH2), mRNA [NM_018040]	0.8	8.9	5.4	0.0000	0.0064
A_24_P126851	A_24_P126851	A_24_P126851	Unknown	-0.5	8.3	-5.4	0.0000	0.0064
A_32_P214860	A_32_P214860	A_32_P214860	Unknown	1.5	8.4	5.4	0.0000	0.0064
A_23_P210164	HOXD8	NM_019558	Homo sapiens homeobox D8 (HOXD8), mRNA [NM_019558]	-0.8	8.9	-5.3	0.0000	0.0064
A_24_P302070	RP6-166C19.11	NM_173571	Homo sapiens cancer/testis CT47 family, member 11 (CT47.11), mRNA [NM_173571]	1.3	6.8	5.3	0.0000	0.0064
A_23_P205519	ABHD4	NM_022060	Homo sapiens abhydrolase domain containing 4 (ABHD4), mRNA [NM_022060]	0.8	9.6	5.3	0.0000	0.0066
A_23_P308150	FAM123B	NM_152424	Homo sapiens family with sequence similarity 123B (FAM123B), mRNA [NM_152424]	-0.7	7.2	-5.3	0.0000	0.0066
A_24_P931443	GPR68	NM_003485	Homo sapiens G protein-coupled receptor 68 (GPR68), mRNA [NM_003485]	-1.4	9.1	-5.3	0.0000	0.0066
A_24_P941167	APOL6	NM_030641	Homo sapiens apolipoprotein L, 6 (APOL6), mRNA [NM_030641]	1.8	7.8	5.3	0.0000	0.0066
A_24_P205137	HDAC6	BC011498	Homo sapiens histone deacetylase 6, mRNA (cDNA clone IMAGE:4179066), complete cds. [BC011498]	-0.6	5.7	-5.3	0.0000	0.0071
A_23_P24433	CTSF	NM_003793	Homo sapiens cathepsin F (CTSF), mRNA [NM_003793]	0.8	12.6	5.2	0.0000	0.0075
A_32_P206549	BC035156	BC035156	Homo sapiens cDNA clone IMAGE:5265210. [BC035156]	-0.7	7.6	-5.2	0.0000	0.0077
A_23_P101374	CYP2S1	NM_030622	Homo sapiens cytochrome P450, family 2, subfamily S, polypeptide 1 (CYP2S1), mRNA [NM_030622]	-0.8	5.4	-5.2	0.0000	0.0077
A_24_P506977	C7orf40	AK096179	Homo sapiens cDNA FLJ38860 fis, clone MESAN2011977. [AK096179]	-0.7	11.7	-5.2	0.0000	0.0077
A_23_P17420	BCAS1	NM_003657	Homo sapiens breast carcinoma amplified sequence 1 (BCAS1), mRNA [NM_003657]	1.5	10.9	5.2	0.0000	0.0077
A_23_P166566	CCDC48	NM_024768	Homo sapiens coiled-coil domain containing 48 (CCDC48), mRNA [NM_024768]	1.2	8.8	5.2	0.0000	0.0078
A_23_P128447	LRRK2	NM_198578	Homo sapiens leucine-rich repeat kinase 2 (LRRK2), mRNA [NM_198578]	1.6	6.9	5.2	0.0000	0.0078
A_24_P914479	SNX5	BC002724	Homo sapiens sorting nexin 5, mRNA (cDNA clone IMAGE:3629947), complete cds. [BC002724]	-0.8	8.3	-5.2	0.0000	0.0078
A_23_P57570	A4GALT	NM_017436	Homo sapiens alpha 1,4-galactosyltransferase (globotriaosylceramide synthase) (A4GALT), mRNA [NM_017436]	-1.0	7.2	-5.2	0.0000	0.0078
A_23_P468	C1orf114	NM_021179	Homo sapiens chromosome 1 open reading frame 114 (C1orf114), mRNA [NM_021179]	0.8	5.9	5.2	0.0000	0.0078
A_23_P71415	C8orf32	NM_018024	Homo sapiens chromosome 8 open reading frame 32 (C8orf32), mRNA [NM_018024]	-0.6	10.2	-5.2	0.0000	0.0078
A_23_P300070	IQCG	NM_032263	Homo sapiens IQ motif containing G (IQCG), mRNA [NM_032263]	0.9	7.6	5.2	0.0000	0.0078

ProbeName	GeneName	SystematicName	Description	logFC	AveExpr	t	P.Value	adj.P.Val
A_23_P47839	DDX55	NM_020936	Homo sapiens DEAD (Asp-Glu-Ala-Asp) box polypeptide 55 (DDX55), mRNA [NM_020936]	-0.7	8.4	-5.2	0.0000	0.0079
A_24_P406601	OLFM1	NM_014279	Homo sapiens olfactomedin 1 (OLFM1), transcript variant 1, mRNA [NM_014279]	-0.7	8.2	-5.1	0.0000	0.0079
A_24_P50245	HLA-DMA	NM_006120	Homo sapiens major histocompatibility complex, class II, DM alpha (HLA-DMA), mRNA [NM_006120]	1.1	8.3	5.1	0.0000	0.0080
A_23_P120472	TFAP2C	NM_003222	Homo sapiens transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma) (TFAP2C), mRNA [NM_003222]	-0.5	10.8	-5.1	0.0000	0.0080
A_23_P218928	C4orf18	NM_016613	Homo sapiens chromosome 4 open reading frame 18 (C4orf18), transcript variant 2, mRNA [NM_016613]	1.6	7.2	5.1	0.0000	0.0080
A_23_P65262	RP11-298P3.3	U50529	Human BRCA2 region, mRNA sequence CG016. [U50529]	1.0	9.8	5.1	0.0000	0.0082
A_23_P69383	PARP9	NM_031458	Homo sapiens poly (ADP-ribose) polymerase family, member 9 (PARP9), mRNA [NM_031458]	1.2	9.6	5.1	0.0000	0.0082
A_32_P195924	BU661610	BU661610	cl74d04.z1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo sapiens cDNA clone cl74d04 5', mRNA sequence [BU661610]	-0.7	10.0	-5.1	0.0000	0.0082
A_24_P686965	SH2D5	AK124869	Homo sapiens cDNA FLJ42879 fis, clone BRHIP3001283. [AK124869]	-0.8	6.6	-5.1	0.0000	0.0082
A_24_P134850	PLEKHH1	NM_020715	Homo sapiens pleckstrin homology domain containing, family H (with MyTH4 domain) member 1 (PLEKHH1), mRNA [NM_020715]	-0.9	6.4	-5.1	0.0000	0.0088
A_23_P52266	IFIT1	NM_001548	Homo sapiens interferon-induced protein with tetratricopeptide repeats 1 (IFIT1), transcript variant 2, mRNA [NM_001548]	3.8	9.0	5.0	0.0000	0.0095
A_24_P937231	AK092744	AK092744	Homo sapiens cDNA FLJ35425 fis, clone SMINT2001615, highly similar to FERRITIN LIGHT CHAIN. [AK092744]	-0.7	6.1	-5.0	0.0000	0.0097
A_24_P224998	THC2525241	THC2525241	Q3SZ59_BOVIN (Q3SZ59) Ribosomal protein L36a, partial (94%) [THC2525241]	-0.7	7.5	-5.0	0.0000	0.0097
A_24_P4816	GABARAPL1	NM_031412	Homo sapiens GABA(A) receptor-associated protein like 1 (GABARAPL1), mRNA [NM_031412]	0.8	9.9	4.9	0.0000	0.0114
A_24_P612200	BF246504	BF246504	601854749F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4074517 5', mRNA sequence [BF246504]	-0.6	8.9	-4.9	0.0000	0.0115
A_23_P370666	GLYATL1	NM_080661	Homo sapiens glycine-N-acyltransferase-like 1 (GLYATL1), mRNA [NM_080661]	1.2	8.0	4.9	0.0000	0.0127
A_23_P9152	RCL1	NM_005772	Homo sapiens RNA terminal phosphate cyclase-like 1 (RCL1), mRNA [NM_005772]	-0.6	11.9	-4.9	0.0000	0.0127
A_23_P20683	KIAA0020	NM_014878	Homo sapiens KIAA0020 (KIAA0020), mRNA [NM_014878]	-0.6	11.1	-4.9	0.0000	0.0127
A_23_P160200	ZNF643	NM_023070	Homo sapiens zinc finger protein 643 (ZNF643), mRNA [NM_023070]	-0.7	8.1	-4.9	0.0000	0.0127
A_23_P322845	PPAPDC1B	BC033025	Homo sapiens phosphatidic acid phosphatase type 2 domain containing 1B, mRNA (cDNA clone MGC:32924 IMAGE:5267610), complete cds. [BC033025]	0.7	12.1	4.9	0.0000	0.0129
A_24_P859032	THC2660977	THC2660977	Unknown	-0.7	6.3	-4.9	0.0000	0.0129
A_24_P278853	TMTC4	NM_032813	Homo sapiens transmembrane and tetratricopeptide repeat containing 4 (TMTC4), transcript variant 1, mRNA [NM_032813]	-0.7	7.2	-4.9	0.0000	0.0129
A_23_P27688	LOC199800	BC022233	Homo sapiens hypothetical protein LOC199800, mRNA (cDNA clone IMAGE:4401286), with apparent retained intron. [BC022233]	-0.8	6.5	-4.9	0.0000	0.0129
A_24_P918147	AK023391	AK023391	Homo sapiens cDNA FLJ13329 fis, clone OVARC1001795. [AK023391]	1.4	9.0	4.9	0.0000	0.0131
A_24_P466102	AL390143	AL390143	Homo sapiens mRNA; cDNA DKFZp547N074 (from clone DKFZp547N074) [AL390143]	-0.8	6.5	-4.8	0.0000	0.0137
A_23_P73429	HCLS1	NM_005335	Homo sapiens hematopoietic cell-specific Lyn substrate 1 (HCLS1), mRNA [NM_005335]	0.8	7.1	4.8	0.0000	0.0141
A_23_P319783	SPAG17	NM_206996	Homo sapiens sperm associated antigen 17 (SPAG17), mRNA [NM_206996]	0.7	6.9	4.8	0.0000	0.0145
A_23_P24004	IFIT2	NM_001547	Homo sapiens interferon-induced protein with tetratricopeptide repeats 2 (IFIT2), mRNA [NM_001547]	2.2	8.3	4.8	0.0000	0.0148
A_23_P425925	KRT222P	NM_152349	Homo sapiens keratin 222 pseudogene (KRT222P), mRNA [NM_152349]	1.0	6.6	4.8	0.0000	0.0148
A_23_P88095	TBC1D4	NM_014832	Homo sapiens TBC1 domain family, member 4 (TBC1D4), mRNA [NM_014832]	-0.6	7.7	-4.8	0.0001	0.0154
A_23_P32175	LHX6	NM_014368	Homo sapiens LIM homeobox 6 (LHX6), transcript variant 1, mRNA [NM_014368]	-1.0	10.2	-4.8	0.0001	0.0154
A_32_P84580	AK130486	AK130486	Homo sapiens cDNA FLJ26976 fis, clone SLV02257. [AK130486]	-0.7	7.9	-4.8	0.0001	0.0154
A_32_P169353	A_32_P169353	A_32_P169353	Unknown	0.8	6.8	4.7	0.0001	0.0160
A_23_P138655	CYP26A1	NM_057157	Homo sapiens cytochrome P450, family 26, subfamily A, polypeptide 1 (CYP26A1), transcript variant 2, mRNA [NM_057157]	0.9	6.2	4.7	0.0001	0.0160
A_23_P156025	IRX2	ENST00000302057	Iroquois-class homeodomain protein IRX-2 (Iroquois homeobox protein 2) (Homeodomain protein IRXA2). [Source:Uniprot/SWISSPROT;Acc:Q9BZI1] [ENST00000302057]	-1.0	9.4	-4.7	0.0001	0.0160
A_32_P69849	NGFRAP1L1	NM_001012978	Homo sapiens NGFRAP1-like 1 (NGFRAP1L1), mRNA [NM_001012978]	0.8	7.4	4.7	0.0001	0.0163
A_23_P69908	GLRX	NM_002064	Homo sapiens glutaredoxin (thioltransferase) (GLRX), mRNA [NM_002064]	1.1	10.7	4.7	0.0001	0.0163
A_32_P2050	THC2631248	THC2631248	Q7WVU3_ACTPL (Q7WVU3) Collagen adhesin (Fragment), partial (5%) [THC2631248]	-0.7	7.9	-4.7	0.0001	0.0163
A_24_P382319	CEACAM1	NM_001712	Homo sapiens carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein) (CEACAM1), transcript variant 1, mRNA [NM_001712]	0.9	6.5	4.7	0.0001	0.0164
A_23_P132718	SEMA3B	NM_004636	Homo sapiens sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3B (SEMA3B), transcript variant 1, mRNA [NM_004636]	-1.3	11.2	-4.7	0.0001	0.0164
A_23_P8013	HIST1H2BL	NM_003519	Homo sapiens histone cluster 1, H2bl (HIST1H2BL), mRNA [NM_003519]	0.6	13.6	4.7	0.0001	0.0164
A_32_P16204	LOC375295	ENST00000295549	Homo sapiens hypothetical gene supported by BC013438, mRNA (cDNA clone IMAGE:3899073), partial cds. [BC013438]	-1.0	11.4	-4.7	0.0001	0.0164
A_23_P46447	BX538051	BX538051	Homo sapiens mRNA; cDNA DKFZp686F09156 (from clone DKFZp686F09156). [BX538051]	1.0	6.6	4.7	0.0001	0.0165
A_23_P391506	IVNS1ABP	NM_016389	Homo sapiens influenza virus NS1A binding protein (IVNS1ABP), transcript variant 2, mRNA [NM_016389]	-0.5	10.7	-4.7	0.0001	0.0169
A_23_P160992	FMO4	NM_002022	Homo sapiens flavin containing monooxygenase 4 (FMO4), mRNA [NM_002022]	0.6	6.5	4.7	0.0001	0.0169
A_23_P128967	ALDH6A1	NM_005589	Homo sapiens aldehyde dehydrogenase 6 family, member A1 (ALDH6A1), nuclear gene encoding mitochondrial protein, mRNA [NM_005589]	0.6	9.0	4.7	0.0001	0.0171
A_32_P130641	STARD9	ENST00000290607	StAR-related lipid transfer protein 9 (StARD9) (START domain- containing protein 9) (Fragment). [Source:Uniprot/SWISSPROT;Acc:Q9P2P6] [ENST00000290607]	-0.7	6.7	-4.7	0.0001	0.0176
A_23_P355244	SAMD9	NM_017654	Homo sapiens sterile alpha motif domain containing 9 (SAMD9), mRNA [NM_017654]	1.1	6.5	4.7	0.0001	0.0176
A_24_P504050	TMEM170	BX648484	Homo sapiens mRNA; cDNA DKFZp686O1555 (from clone DKFZp686O1555). [BX648484]	-0.6	9.9	-4.6	0.0001	0.0179
A_32_P167592	ENST00000339867	ENST00000339867	similar to Interferon-induced transmembrane protein 3 (Interferon-inducible protein 1-8U) (LOC650205), mRNA [Source:RefSeq_dna;Acc:XR_018421] [ENST00000339867]	1.2	7.5	4.6	0.0001	0.0179
A_24_P925292	AI590892	AI590892	AI590892 tw26f04.x1 NCI_CGAP_Ov35 Homo sapiens cDNA clone IMAGE:2260831 3', mRNA sequence [AI590892]	-0.6	5.9	-4.6	0.0001	0.0179
A_23_P142497	ANKRD41	NM_152363	Homo sapiens ankyrin repeat domain 41 (ANKRD41), mRNA [NM_152363]	-0.8	6.3	-4.6	0.0001	0.0179
A_23_P121545	GPM6A	NM_201591	Homo sapiens glycoprotein M6A (GPM6A), transcript variant 2, mRNA [NM_201591]	0.8	6.4	4.6	0.0001	0.0179
A_24_P271323	SLC2A10	NM_030777	Homo sapiens solute carrier family 2 (facilitated glucose transporter), member 10 (SLC2A10), mRNA [NM_030777]	0.9	9.2	4.6	0.0001	0.0180
A_32_P310219	AK026422	AK026422	Homo sapiens cDNA: FLJ22769 fis, clone KAIA1316. [AK026422]	-0.8	6.0	-4.6	0.0001	0.0180
A_23_P434374	MUC19	AK093065	Homo sapiens cDNA FLJ35746 fis, clone TESTI2004399, moderately similar to APOMUCIN. [AK093065]	1.7	6.2	4.6	0.0001	0.0181
A_23_P64828	OAS1	NM_002534	Homo sapiens 2',5'-oligoadenylate synthetase 1, 40/46kDa (OAS1), transcript variant 2, mRNA [NM_002534]	2.2	5.9	4.6	0.0001	0.0183
A_23_P88411	SEL1L	NM_005065	Homo sapiens sel-1 suppressor of lin-12-like (C. elegans) (SEL1L), mRNA [NM_005065]	0.8	8.7	4.6	0.0001	0.0183
A_23_P112982	SVOP	NM_018711	Homo sapiens SV2 related protein homolog (rat) (SVOP), mRNA [NM_018711]	0.8	6.2	4.6	0.0001	0.0184
A_32_P70315	TIMP4	NM_003256	Homo sapiens TIMP metallopeptidase inhibitor 4 (TIMP4), mRNA [NM_003256]	0.6	5.5	4.6	0.0001	0.0184
A_23_P29237	APOL3	NM_145641	Homo sapiens apolipoprotein L, 3 (APOL3), transcript variant beta/a, mRNA [NM_145641]	0.8	7.4	4.6	0.0001	0.0184
A_23_P206324	HSDL1	NM_031463	Homo sapiens hydroxysteroid dehydrogenase like 1 (HSDL1), mRNA [NM_031463]	-0.5	7.1	-4.6	0.0001	0.0187

ProbeName	GeneName	SystematicName	Description	logFC	AveExpr	t	P.Value	adj.P.Val
A_23_P61371	TMEM173	NM_198282	Homo sapiens transmembrane protein 173 (TMEM173), mRNA [NM_198282]	1.0	9.5	4.6	0.0001	0.0187
A_23_P104471	DUSP13	NM_001007271	Homo sapiens dual specificity phosphatase 13 (DUSP13), transcript variant 1, mRNA [NM_001007271]	0.8	6.6	4.6	0.0001	0.0187
A_23_P126908	TNFRSF14	NM_003820	Homo sapiens tumor necrosis factor receptor superfamily, member 14 (herpesvirus entry mediator) (TNFRSF14), mRNA [NM_003820]	0.9	8.5	4.6	0.0001	0.0196
A_23_P127964	PRCP	NM_199418	Homo sapiens prolylcarboxypeptidase (angiotensinase C) (PRCP), transcript variant 2, mRNA [NM_199418]	0.7	13.0	4.5	0.0001	0.0206
A_32_P354945	BC015383	BC015383	Homo sapiens cDNA clone IMAGE:4426835, partial cds. [BC015383]	0.7	6.0	4.5	0.0001	0.0206
A_23_P211007	NRIP1	NM_003489	Homo sapiens nuclear receptor interacting protein 1 (NRIP1), mRNA [NM_003489]	-0.9	10.3	-4.5	0.0001	0.0210
A_23_P433218	OR7E91P	NR_002185	Homo sapiens olfactory receptor, family 7, subfamily E, member 91 pseudogene (OR7E91P) on chromosome 2 [NR_002185]	0.8	8.6	4.5	0.0001	0.0210
A_23_P160318	COL16A1	NM_001856	Homo sapiens collagen, type XVI, alpha 1 (COL16A1), mRNA [NM_001856]	0.9	10.0	4.5	0.0001	0.0210
A_24_P311926	HLA-G	NM_002127	Homo sapiens HLA-G histocompatibility antigen, class I, G (HLA-G), mRNA [NM_002127]	0.8	13.4	4.5	0.0001	0.0214
A_23_P250164	HGD	NM_000187	Homo sapiens homogentisate 1,2-dioxygenase (homogentisate oxidase) (HGD), mRNA [NM_000187]	1.4	7.0	4.5	0.0001	0.0214
A_24_P212549	LOC727941	XR_015169	PREDICTED: Homo sapiens similar to solute carrier family 25 member 24 isoform 1 (LOC727941), mRNA [XR_015169]	-0.5	7.1	-4.5	0.0001	0.0214
A_23_P215735	ST7	NM_018412	Homo sapiens suppression of tumorigenicity 7 (ST7), transcript variant a, mRNA [NM_018412]	-0.5	9.8	-4.5	0.0001	0.0215
A_23_P326319	C16orf45	NM_033201	Homo sapiens chromosome 16 open reading frame 45 (C16orf45), mRNA [NM_033201]	1.1	9.2	4.5	0.0001	0.0216
A_32_P182388	ZNF77	NM_021217	Homo sapiens zinc finger protein 77 (ZNF77), mRNA [NM_021217]	-0.4	9.6	-4.5	0.0001	0.0216
A_23_P333951	ENST00000366848	ENST00000366848	C1orf67 protein. [Source:Uniprot/SPTREMBL;Acc:Q0VDD9] [ENST00000366848]	-0.8	9.1	-4.5	0.0001	0.0216
A_23_P209987	POLR1B	NM_019014	Homo sapiens polymerase (RNA) I polypeptide B, 128kDa (POLR1B), mRNA [NM_019014]	-0.6	8.8	-4.5	0.0001	0.0216
A_32_P14721	DNHD2	AK128592	Homo sapiens cDNA FLJ46751 fis, clone TRACH3022960, weakly similar to Dynein beta chain, ciliary. [AK128592]	1.1	6.3	4.5	0.0001	0.0216
A_23_P17663	MX1	NM_002462	Homo sapiens myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse) (MX1), mRNA [NM_002462]	2.9	8.5	4.5	0.0001	0.0220
A_23_P74914	KIAA0133	NM_014777	Homo sapiens KIAA0133 (KIAA0133), mRNA [NM_014777]	-0.6	9.1	-4.5	0.0001	0.0228
A_23_P259071	AREG	NM_001657	Homo sapiens amphiregulin (schwannoma-derived growth factor) (AREG), mRNA [NM_001657]	-1.9	11.2	-4.5	0.0001	0.0228
A_24_P307964	SOHLH1	NM_001012415	Homo sapiens spermatogenesis and oogenesis specific basic helix-loop-helix 1 (SOHLH1), mRNA [NM_001012415]	1.3	6.9	4.5	0.0001	0.0228
A_24_P368023	BC039457	BC039457	Homo sapiens cDNA clone IMAGE:5312122. [BC039457]	0.8	8.5	4.5	0.0001	0.0228
A_23_P145874	SAMD9L	NM_152703	Homo sapiens sterile alpha motif domain containing 9-like (SAMD9L), mRNA [NM_152703]	1.3	6.1	4.5	0.0001	0.0228
A_24_P803885	ENST00000355873	ENST00000355873	CDNA FLJ43187 fis, clone FCBBF3023443. [Source:Uniprot/SPTREMBL;Acc:Q6ZUZ4] [ENST00000355873]	0.8	7.8	4.5	0.0001	0.0228
A_32_P48279	BX117479	BX117479	BX117479 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGp998I075985, mRNA sequence [BX117479]	0.8	6.2	4.5	0.0001	0.0229
A_32_P67303	A_32_P67303	A_32_P67303	Unknown	-1.0	7.5	-4.4	0.0001	0.0230
A_23_P14083	AMIGO2	NM_181847	Homo sapiens adhesion molecule with Ig-like domain 2 (AMIGO2), mRNA [NM_181847]	1.2	12.7	4.4	0.0001	0.0234
A_32_P224234	AK123450	AK123450	Homo sapiens cDNA FLJ41456 fis, clone BRSTN2012320. [AK123450]	-0.7	7.9	-4.4	0.0001	0.0234
A_24_P332711	ENST00000358162	ENST00000358162	Unknown	0.6	6.4	4.4	0.0001	0.0238
A_32_P131633	CR598849	CR598849	full-length cDNA clone CS0DI035YF03 of Placenta Cot 25-normalized of Homo sapiens (human) [CR598849]	-0.9	7.4	-4.4	0.0001	0.0238
A_23_P251387	REPS2	NM_004726	Homo sapiens RALBP1 associated Eps domain containing 2 (REPS2), transcript variant 1, mRNA [NM_004726]	0.5	10.1	4.4	0.0001	0.0241
A_23_P15394	CD68	NM_001251	Homo sapiens CD68 molecule (CD68), transcript variant 1, mRNA [NM_001251]	0.9	6.7	4.4	0.0001	0.0241
A_23_P102122	ARPC2	NM_152862	Homo sapiens actin related protein 2/3 complex, subunit 2, 34kDa (ARPC2), transcript variant 1, mRNA [NM_152862]	0.4	15.0	4.4	0.0001	0.0241
A_23_P422851	CABLES1	NM_138375	Homo sapiens Cdk5 and Abl enzyme substrate 1 (CABLES1), mRNA [NM_138375]	-0.6	12.2	-4.4	0.0001	0.0242
A_23_P206901	NDE1	NM_017668	Homo sapiens nudE nuclear distribution gene E homolog 1 (A. nidulans) (NDE1), mRNA [NM_017668]	0.9	13.2	4.4	0.0001	0.0247
A_23_P379945	KIAA1509	NM_001080414	Homo sapiens KIAA1509 (KIAA1509), mRNA [NM_001080414]	-0.7	8.3	-4.4	0.0001	0.0249
A_32_P86682	THC2698927	THC2698927	Unknown	1.2	8.1	4.4	0.0001	0.0251
A_32_P65157	BQ071182	BQ071182	BQ071182 AGENCOURT_6853099 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5927550 5', mRNA sequence [BQ071182]	-0.6	9.3	-4.4	0.0001	0.0252
A_23_P204133	GALNT6	NM_007210	Homo sapiens UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 6 (GalNAc-T6) (GALNT6), mRNA [NM_007210]	1.0	10.1	4.4	0.0001	0.0256
A_23_P115636	TRDMT1	NM_004412	Homo sapiens tRNA aspartic acid methyltransferase 1 (TRDMT1), transcript variant a, mRNA [NM_004412]	-0.6	7.4	-4.4	0.0002	0.0269
A_23_P124642	RASGRP1	NM_005739	Homo sapiens RAS guanyl releasing protein 1 (calcium and DAG-regulated) (RASGRP1), mRNA [NM_005739]	-0.7	5.5	-4.4	0.0002	0.0269
A_24_P927639	BC032755	BC032755	Homo sapiens, Similar to peroxisome biogenesis factor 13, clone IMAGE:5581685, mRNA. [BC032755]	-0.8	6.4	-4.4	0.0002	0.0270
A_23_P103282	TMEM59	NM_004872	Homo sapiens transmembrane protein 59 (TMEM59), mRNA [NM_004872]	0.4	13.9	4.3	0.0002	0.0270
A_23_P103812	FAM5C	NM_199051	Homo sapiens family with sequence similarity 5, member C (FAM5C), mRNA [NM_199051]	0.7	11.8	4.3	0.0002	0.0270
A_23_P107412	P4HB	NM_000918	Homo sapiens procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), beta polypeptide (P4HB), mRNA [NM_000918]	0.6	14.7	4.3	0.0002	0.0270
A_32_P204676	FABP5	NM_001444	Homo sapiens fatty acid binding protein 5 (psoriasis-associated) (FABP5), mRNA [NM_001444]	-0.9	9.8	-4.3	0.0002	0.0270
A_24_P77082	KMO	NM_003679	Homo sapiens kynurenine 3-monooxygenase (kynurenine 3-hydroxylase) (KMO), mRNA [NM_003679]	1.4	9.1	4.3	0.0002	0.0270
A_23_P202496	NOC3L	NM_022451	Homo sapiens nucleolar complex associated 3 homolog (S. cerevisiae) (NOC3L), mRNA [NM_022451]	-0.7	10.5	-4.3	0.0002	0.0270
A_32_P46840	AK090664	AK090664	Homo sapiens cDNA FLJ33345 fis, clone BRACE2003713. [AK090664]	-0.8	9.4	-4.3	0.0002	0.0271
A_23_P35099	MRTO4	NM_016183	Homo sapiens mRNA turnover 4 homolog (S. cerevisiae) (MRTO4), mRNA [NM_016183]	-0.5	11.6	-4.3	0.0002	0.0273
A_23_P212469	ENTPD3	NM_001248	Homo sapiens ectonucleoside triphosphate diphosphohydrolase 3 (ENTPD3), mRNA [NM_001248]	1.0	9.4	4.3	0.0002	0.0273
A_23_P424734	EXOC3L	NM_178516	Homo sapiens exocyst complex component 3-like (EXOC3L), mRNA [NM_178516]	-0.6	5.9	-4.3	0.0002	0.0277
A_24_P298409	HLA-C	NM_002117	Homo sapiens major histocompatibility complex, class I, C (HLA-C), mRNA [NM_002117]	0.7	13.5	4.3	0.0002	0.0277
A_23_P112241	DNAJB5	NM_012266	Homo sapiens DnaJ (Hsp40) homolog, subfamily B, member 5 (DNAJB5), mRNA [NM_012266]	0.6	10.0	4.3	0.0002	0.0290
A_23_P118722	ASGR1	NM_001671	Homo sapiens asialoglycoprotein receptor 1 (ASGR1), mRNA [NM_001671]	0.6	7.2	4.3	0.0002	0.0290
A_23_P360354	FAM81B	NM_152548	Homo sapiens family with sequence similarity 81, member B (FAM81B), mRNA [NM_152548]	0.7	5.6	4.3	0.0002	0.0290
A_23_P202245	RET	NM_020975	Homo sapiens ret proto-oncogene (RET), transcript variant 2, mRNA [NM_020975]	-1.1	7.3	-4.3	0.0002	0.0291
A_32_P154911	PRR15	NM_175887	Homo sapiens proline rich 15 (PRR15), mRNA [NM_175887]	1.6	9.6	4.3	0.0002	0.0291
A_23_P40108	COL9A3	NM_001853	Homo sapiens collagen, type IX, alpha 3 (COL9A3), mRNA [NM_001853]	0.9	11.5	4.3	0.0002	0.0297
A_23_P35912	CASP4	NM_033306	Homo sapiens caspase 4, apoptosis-related cysteine peptidase (CASP4), transcript variant gamma, mRNA [NM_033306]	0.6	11.2	4.3	0.0002	0.0301
A_23_P69339	ACAA1	NM_001607	Homo sapiens acetyl-Coenzyme A acyltransferase 1 (peroxisomal 3-oxoacyl-Coenzyme A thiolase) (ACAA1), nuclear gene encoding mitochondrial protein, mRNA [NM_001607]	0.5	12.6	4.3	0.0002	0.0301
A_23_P10488	AK057704	AK057704	Homo sapiens cDNA FLJ33142 fis, clone UTERU1000192. [AK057704]	0.8	9.1	4.3	0.0002	0.0306
A_23_P117782	LARP6	NM_018357	Homo sapiens La ribonucleoprotein domain family, member 6 (LARP6), transcript variant 1, mRNA [NM_018357]	0.8	10.5	4.2	0.0002	0.0315
A_23_P59960	CRISPLD1	NM_031461	Homo sapiens cysteine-rich secretory protein LCCL domain containing 1 (CRISPLD1), mRNA [NM_031461]	0.8	8.3	4.2	0.0002	0.0315

ProbeName	GeneName	SystematicName	Description	logFC	AveExpr	t	P.Value	adj.P.Val
A_24_P399942	ATP11C	NM_173694	Homo sapiens ATPase, Class VI, type 11C (ATP11C), transcript variant 1, mRNA [NM_173694]	-0.6	7.2	-4.2	0.0002	0.0315
A_23_P119794	CR608907	CR608907	full-length cDNA clone CS0DM002YA18 of Fetal liver of Homo sapiens (human) [CR608907]	-0.7	8.3	-4.2	0.0002	0.0315
A_32_P21848	THC2716178	THC2716178	ALU5_HUMAN (P39192) Alu subfamily SC sequence contamination warning entry, partial (5%) [THC2716178]	0.6	5.4	4.2	0.0002	0.0315
A_24_P24371	ENST00000390543	ENST00000390543	Immunoglobulin heavy chain C gene segment [Source:IMGT/GENE_DB;Acc:IGHG4] [ENST00000390543]	-0.5	5.3	-4.2	0.0002	0.0315
A_24_P99071	RANBP5	NM_002271	Homo sapiens RAN binding protein 5 (RANBP5), mRNA [NM_002271]	-0.6	9.5	-4.2	0.0002	0.0315
A_23_P4286	XAF1	NM_017523	Homo sapiens XIAP associated factor-1 (XAF1), transcript variant 1, mRNA [NM_017523]	2.7	6.1	4.2	0.0002	0.0315
A_24_P56557	ATPBD4	NM_080650	Homo sapiens ATP binding domain 4 (ATPBD4), mRNA [NM_080650]	-0.6	8.9	-4.2	0.0002	0.0315
A_24_P745352	LOC643447	XR_016392	PREDICTED: Homo sapiens similar to Keratin, type II cytoskeletal 8 (Cytokeratin-8) (CK-8) (Keratin-8) (K8) (LOC643447), mRNA [XR_016392]	-0.7	9.0	-4.2	0.0002	0.0315
A_24_P673968	ENST00000371276	ENST00000371276	Tetratricopeptide repeat protein 22 (TPR repeat protein 22). [Source:Uniprot/SWISSPROT;Acc:Q5TAA0] [ENST00000371276]	0.7	6.3	4.2	0.0002	0.0315
A_23_P215549	PON3	NM_000940	Homo sapiens paraoxonase 3 (PON3), mRNA [NM_000940]	0.8	10.8	4.2	0.0002	0.0315
A_23_P78892	FBL	NM_001436	Homo sapiens fibrillarin (FBL), mRNA [NM_001436]	-0.4	14.8	-4.2	0.0002	0.0316
A_23_P118203	LOC124220	NM_145252	Homo sapiens similar to common salivary protein 1 (LOC124220), mRNA [NM_145252]	0.8	12.0	4.2	0.0002	0.0317
A_23_P86838	SLC36A4	NM_152313	Homo sapiens solute carrier family 36 (proton/amino acid symporter), member 4 (SLC36A4), mRNA [NM_152313]	-0.5	8.4	-4.2	0.0002	0.0317
A_23_P111961	RBM13	NM_032509	Homo sapiens RNA binding motif protein 13 (RBM13), mRNA [NM_032509]	-0.5	9.1	-4.2	0.0002	0.0317
A_23_P53370	RND1	NM_014470	Homo sapiens Rho family GTPase 1 (RND1), mRNA [NM_014470]	1.4	7.4	4.2	0.0002	0.0317
A_23_P428738	ANG	NM_001145	Homo sapiens angiogenin, ribonuclease, RNase A family, 5 (ANG), mRNA [NM_001145]	0.8	10.9	4.2	0.0002	0.0318
A_23_P390172	RNASEL	NM_021133	Homo sapiens ribonuclease L (2',5'-oligoisoadenylate synthetase-dependent) (RNASEL), mRNA [NM_021133]	0.4	9.5	4.2	0.0002	0.0318
A_23_P72117	SMPDL3A	NM_006714	Homo sapiens sphingomyelin phosphodiesterase, acid-like 3A (SMPDL3A), mRNA [NM_006714]	0.6	8.4	4.2	0.0002	0.0321
A_23_P404259	LOC493869	NM_001008397	Homo sapiens similar to RIKEN cDNA 2310016C16 (LOC493869), mRNA [NM_001008397]	0.6	8.0	4.2	0.0002	0.0321
A_23_P87545	IFITM3	NM_021034	Homo sapiens interferon induced transmembrane protein 3 (1-8U) (IFITM3), mRNA [NM_021034]	1.6	8.5	4.2	0.0002	0.0321
A_23_P41145	FAM3D	NM_138805	Homo sapiens family with sequence similarity 3, member D (FAM3D), mRNA [NM_138805]	1.2	6.8	4.2	0.0002	0.0321
A_32_P87531	C1orf67	BC042869	Homo sapiens cDNA clone IMAGE:5270407. [BC042869]	-0.7	7.8	-4.2	0.0002	0.0322
A_23_P92222	MUC20	NM_152673	Homo sapiens mucin 20, cell surface associated (MUC20), mRNA [NM_152673]	1.1	10.1	4.2	0.0002	0.0324
A_23_P74299	ADORA1	NM_000674	Homo sapiens adenosine A1 receptor (ADORA1), transcript variant 1, mRNA [NM_000674]	-1.0	5.9	-4.2	0.0002	0.0324
A_32_P113462	A_32_P113462	A_32_P113462	Unknown	0.7	9.3	4.2	0.0002	0.0324
A_24_P248053	TOP1MT	NM_052963	Homo sapiens topoisomerase (DNA) I, mitochondrial (TOP1MT), nuclear gene encoding mitochondrial protein, mRNA [NM_052963]	-0.6	14.0	-4.2	0.0003	0.0327
A_23_P76538	TESC	NM_017899	Homo sapiens tescalcin (TESC), mRNA [NM_017899]	1.1	9.0	4.2	0.0003	0.0327
A_23_P37205	NDRG2	NM_201535	Homo sapiens NDRG family member 2 (NDRG2), transcript variant 1, mRNA [NM_201535]	1.0	11.2	4.2	0.0003	0.0327
A_32_P116813	LYRM4	NM_020408	Homo sapiens LYR motif containing 4 (LYRM4), mRNA [NM_020408]	-0.4	10.2	-4.2	0.0003	0.0327
A_23_P72737	IFITM1	NM_003641	Homo sapiens interferon induced transmembrane protein 1 (9-27) (IFITM1), mRNA [NM_003641]	2.0	9.5	4.2	0.0003	0.0327
A_23_P24960	NARS2	NM_024678	Homo sapiens asparaginyl-tRNA synthetase 2, mitochondrial (putative) (NARS2), mRNA [NM_024678]	-0.5	9.5	-4.2	0.0003	0.0328
A_32_P101313	PTPLAD2	NM_001010915	Homo sapiens protein tyrosine phosphatase-like A domain containing 2 (PTPLAD2), mRNA [NM_001010915]	0.9	10.3	4.2	0.0003	0.0329
A_23_P208579	THC2517558	THC2517558	BC026690 CD97 antigen, isoform 2 precursor {Homo sapiens} (exp=-1; wgp=0; cg=0), partial (15%) [THC2517558]	-0.6	7.2	-4.2	0.0003	0.0330
A_32_P137336	BC013077	BC013077	Homo sapiens, clone IMAGE:3459334, mRNA. [BC013077]	-0.7	7.2	-4.2	0.0003	0.0330
A_32_P180435	WBSCR19	NM_175064	Homo sapiens Williams Beuren syndrome chromosome region 19 (WBSCR19), mRNA [NM_175064]	-0.5	6.5	-4.2	0.0003	0.0330
A_23_P312920	POU2AF1	NM_006235	Homo sapiens POU domain, class 2, associating factor 1 (POU2AF1), mRNA [NM_006235]	-0.8	8.3	-4.2	0.0003	0.0331
A_23_P257111	FBP1	NM_000507	Homo sapiens fructose-1,6-bisphosphatase 1 (FBP1), mRNA [NM_000507]	1.0	12.7	4.2	0.0003	0.0334
A_32_P30345	THC2615857	THC2615857	Unknown	0.9	8.1	4.2	0.0003	0.0335
A_23_P161428	ANKRD22	NM_144590	Homo sapiens ankyrin repeat domain 22 (ANKRD22), mRNA [NM_144590]	1.5	8.3	4.2	0.0003	0.0338
A_24_P418044	BC089454	BC089454	Homo sapiens cDNA clone MGC:105145 IMAGE:30563285, complete cds. [BC089454]	0.7	12.2	4.2	0.0003	0.0338
A_23_P85640	INPP5B	NM_005540	Homo sapiens inositol polyphosphate-5-phosphatase, 75kDa (INPP5B), nuclear gene encoding mitochondrial protein, mRNA [NM_005540]	-0.7	6.7	-4.1	0.0003	0.0339
A_23_P216225	EGR3	NM_004430	Homo sapiens early growth response 3 (EGR3), mRNA [NM_004430]	-1.1	5.8	-4.1	0.0003	0.0339
A_24_P911545	THC2523411	THC2523411	AB011399 AF-6 {Homo sapiens} (exp=-1; wgp=0; cg=0), partial (4%) [THC2572409]	-0.6	5.3	-4.1	0.0003	0.0339
A_23_P59069	HIST1H2BO	NM_003527	Homo sapiens histone cluster 1, H2bo (HIST1H2BO), mRNA [NM_003527]	0.5	12.8	4.1	0.0003	0.0339
A_23_P150979	SBEM	NM_058173	Homo sapiens small breast epithelial mucin (SBEM), mRNA [NM_058173]	0.8	7.1	4.1	0.0003	0.0339
A_32_P151570	A_32_P151570	A_32_P151570	Unknown	-0.7	5.6	-4.1	0.0003	0.0339
A_24_P236091	ENO2	NM_001975	Homo sapiens enolase 2 (gamma, neuronal) (ENO2), mRNA [NM_001975]	-0.6	10.4	-4.1	0.0003	0.0339
A_24_P193011	CCND1	NM_053056	Homo sapiens cyclin D1 (CCND1), mRNA [NM_053056]	-0.8	10.0	-4.1	0.0003	0.0346
A_24_P287043	IFITM2	NM_006435	Homo sapiens interferon induced transmembrane protein 2 (1-8D) (IFITM2), mRNA [NM_006435]	1.1	7.4	4.1	0.0003	0.0346
A_32_P69930	CCDC96	NM_153376	Homo sapiens coiled-coil domain containing 96 (CCDC96), mRNA [NM_153376]	0.7	9.2	4.1	0.0003	0.0346
A_23_P94795	TEAD4	NM_003213	Homo sapiens TEA domain family member 4 (TEAD4), transcript variant 1, mRNA [NM_003213]	-0.5	11.0	-4.1	0.0003	0.0346
A_24_P203696	ENST00000301171	ENST00000301171	Cytochrome P450 2A6 (EC 1.14.14.1) (CYPIIA6) (Coumarin 7-hydroxylase) (P450 IIA3) (CYP2A3) (P450(I)). [Source:Uniprot/SWISSPROT;Acc:P11509] [ENST00000301171]	-0.7	6.6	-4.1	0.0003	0.0346
A_32_P5251	RARA	NM_000964	Homo sapiens retinoic acid receptor, alpha (RARA), transcript variant 1, mRNA [NM_000964]	-1.0	11.3	-4.1	0.0003	0.0346
A_24_P940166	PAPSS2	NM_001015880	Homo sapiens 3'-phosphoadenosine 5'-phosphosulfate synthase 2 (PAPSS2), transcript variant 2, mRNA [NM_001015880]	1.0	9.1	4.1	0.0003	0.0346
A_23_P310483	C8orf58	NM_001013842	Homo sapiens chromosome 8 open reading frame 58 (C8orf58), mRNA [NM_001013842]	-0.8	7.3	-4.1	0.0003	0.0346
A_23_P14458	C14orf149	NM_144581	Homo sapiens chromosome 14 open reading frame 149 (C14orf149), mRNA [NM_144581]	-0.5	6.7	-4.1	0.0003	0.0348
A_23_P15603	MRM1	NM_024864	Homo sapiens mitochondrial rRNA methyltransferase 1 homolog (S. cerevisiae) (MRM1), mRNA [NM_024864]	-0.7	9.0	-4.1	0.0003	0.0348
A_24_P256512	STH	NM_001007532	Homo sapiens saitohin (STH), mRNA [NM_001007532]	-0.8	5.7	-4.1	0.0003	0.0348
A_32_P157391	PSMAL	NM_153696	Homo sapiens growth-inhibiting protein 26 (PSMAL), mRNA [NM_153696]	0.7	6.5	4.1	0.0003	0.0348
A_23_P137248	PRKY	NM_002760	Homo sapiens protein kinase, Y-linked (PRKY), mRNA [NM_002760]	-0.5	7.5	-4.1	0.0003	0.0351
A_24_P64442	AK026517	AK026517	Homo sapiens cDNA: FLJ22864 fis, clone KAT02164. [AK026517]	1.1	11.5	4.1	0.0003	0.0351
A_23_P162668	CPM	NM_001874	Homo sapiens carboxypeptidase M (CPM), transcript variant 1, mRNA [NM_001874]	1.3	8.8	4.1	0.0003	0.0357
A_23_P251795	GPC2	NM_152742	Homo sapiens glypican 2 (GPC2), mRNA [NM_152742]	-0.6	9.1	-4.1	0.0003	0.0360
A_32_P155563	BU153693	BU153693	AGENCOURT_7782569 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6137082 5', mRNA sequence [BU153693]	-0.6	7.2	-4.1	0.0003	0.0360

ProbeName	GeneName	SystematicName	Description	logFC	AveExpr	t	P.Value	adj.P.Val
A_23_P168882	TP53INP1	NM_033285	Homo sapiens tumor protein p53 inducible nuclear protein 1 (TP53INP1), mRNA [NM_033285]	0.8	11.7	4.1	0.0003	0.0360
A_32_P53234	AK123333	AK123333	Homo sapiens cDNA FLJ41339 fis, clone BRASW1000053. [AK123333]	-0.5	5.2	-4.1	0.0003	0.0360
A_32_P185317	LOC645249	BC033117	Homo sapiens hypothetical protein LOC645249, mRNA (cDNA clone IMAGE:3357292), partial cds. [BC033117]	-0.6	10.7	-4.1	0.0003	0.0360
A_24_P261052	MTMR9	NM_015458	Homo sapiens myotubularin related protein 9 (MTMR9), mRNA [NM_015458]	-0.5	7.0	-4.1	0.0003	0.0364
A_23_P93180	HIST1H2BC	NM_003526	Homo sapiens histone cluster 1, H2bc (HIST1H2BC), mRNA [NM_003526]	0.6	10.6	4.1	0.0003	0.0365
A_23_P155425	THC2665061	THC2665061	Q5TVY2_ANOGA (Q5TVY2) ENSANGP00000027655 (Fragment), partial (7%) [THC2665061]	-0.7	7.2	-4.1	0.0003	0.0365
A_23_P214244	ENPP5	NM_021572	Homo sapiens ectonucleotide pyrophosphatase/phosphodiesterase 5 (putative function) (ENPP5), mRNA [NM_021572]	0.8	8.1	4.1	0.0003	0.0366
A_23_P360193	CCDC35	AK074178	Homo sapiens mRNA for FLJ00251 protein. [AK074178]	-0.5	6.1	-4.1	0.0003	0.0368
A_32_P163125	SGMS1	NM_147156	Homo sapiens transmembrane protein 23 (TMEM23), mRNA [NM_147156]	0.6	10.4	4.1	0.0004	0.0375
A_23_P211727	FGF12	NM_004113	Homo sapiens fibroblast growth factor 12 (FGF12), transcript variant 2, mRNA [NM_004113]	0.8	7.9	4.1	0.0004	0.0375
A_32_P118532	BC036625	BC036625	Homo sapiens cDNA clone IMAGE:5288145. [BC036625]	-0.7	7.4	-4.1	0.0004	0.0375
A_23_P79794	TGIF2	NM_021809	Homo sapiens TGFB-induced factor homeobox 2 (TGIF2), mRNA [NM_021809]	-0.7	6.8	-4.1	0.0004	0.0377
A_23_P82823	PINX1	NM_017884	Homo sapiens PIN2-interacting protein 1 (PINX1), mRNA [NM_017884]	-0.5	7.5	-4.0	0.0004	0.0387
A_23_P62115	TIMP1	NM_003254	Homo sapiens TIMP metallopeptidase inhibitor 1 (TIMP1), mRNA [NM_003254]	0.7	7.3	4.0	0.0004	0.0387
A_24_P343929	OAS2	NM_016817	Homo sapiens 2'-5'-oligoadenylate synthetase 2, 69/71kDa (OAS2), transcript variant 1, mRNA [NM_016817]	1.4	5.5	4.0	0.0004	0.0387
A_23_P359174	BC069659	BC069659	Homo sapiens cDNA clone IMAGE:7262526, with apparent retained intron. [BC069659]	-0.9	8.8	-4.0	0.0004	0.0387
A_24_P263767	ENST00000376793	ENST00000376793	MHC class I HLA-J antigen (Fragment). [Source:Uniprot/SPTREMBL;Acc:Q30175] [ENST00000376793]	0.7	9.4	4.0	0.0004	0.0387
A_24_P158298	AK025116	AK025116	Homo sapiens cDNA: FLJ21463 fis, clone COL04765. [AK025116]	-0.4	4.8	-4.0	0.0004	0.0387
A_24_P324986	ERCC8	NM_001007234	Homo sapiens excision repair cross-complementing rodent repair deficiency, complementation group 8 (ERCC8), transcript variant 3, mRNA [NM_001007234]	-0.5	7.0	-4.0	0.0004	0.0387
A_24_P922606	NUP160	NM_015231	Homo sapiens nucleoporin 160kDa (NUP160), mRNA [NM_015231]	-0.5	8.1	-4.0	0.0004	0.0387
A_23_P99967	BBS4	NM_033028	Homo sapiens Bardet-Biedl syndrome 4 (BBS4), mRNA [NM_033028]	0.8	11.0	4.0	0.0004	0.0390
A_24_P332647	ENST00000326470	ENST00000326470	Protein phosphatase Slingshot homolog 1 (EC 3.1.3.48) (EC 3.1.3.16) (SSH-1L) (hSSH-1L). [Source:Uniprot/SWISSPROT;Acc:Q8WYL5] [ENST00000326470]	-0.6	7.1	-4.0	0.0004	0.0390
A_24_P797455	AK056910	AK056910	Homo sapiens cDNA FLJ32348 fis, clone PROST2007200. [AK056910]	-0.5	9.0	-4.0	0.0004	0.0390
A_23_P83838	CA8	NM_004056	Homo sapiens carbonic anhydrase VIII (CA8), mRNA [NM_004056]	-1.2	9.1	-4.0	0.0004	0.0394
A_23_P390665	LOC146177	AY360464	Homo sapiens hypothetical protein LOC146177-like protein mRNA, complete cds; alternatively spliced. [AY360464]	0.5	5.5	4.0	0.0004	0.0394
A_23_P145096	PLA2G7	NM_005084	Homo sapiens phospholipase A2, group VII (platelet-activating factor acetylhydrolase, plasma) (PLA2G7), mRNA [NM_005084]	1.0	7.3	4.0	0.0004	0.0403
A_24_P912136	AK094603	AK094603	Homo sapiens cDNA FLJ37284 fis, clone BRAMY2013590. [AK094603]	0.7	5.5	4.0	0.0004	0.0403
A_23_P92983	BHMT2	NM_017614	Homo sapiens betaine-homocysteine methyltransferase 2 (BHMT2), mRNA [NM_017614]	0.8	6.9	4.0	0.0004	0.0403
A_32_P164637	A_32_P164637	A_32_P164637	Unknown	0.8	6.5	4.0	0.0004	0.0403
A_24_P392496	ENST00000315293	ENST00000315293	BA251O17.4 (Similar to methylenetetrahydrofolate dehydrogenase(NADP+ dependent),methenyltetrahydrofolate cyclohydrolase,formyltetrahydrofolate synthetase (MTHFD1)) (Fragment). [Source:Uniprot/SPTREMBL;Acc:Q8WX68] [ENST00000315293]	-0.7	8.4	-4.0	0.0004	0.0403
A_23_P57709	PCOLCE2	NM_013363	Homo sapiens procollagen C-endopeptidase enhancer 2 (PCOLCE2), mRNA [NM_013363]	-1.0	6.5	-4.0	0.0004	0.0403
A_24_P56462	DAD1	NM_001344	Homo sapiens defender against cell death 1 (DAD1), mRNA [NM_001344]	0.4	13.7	4.0	0.0004	0.0403
A_32_P174110	THC2557387	THC2557387	AF308801 vacuolar protein sorting protein 16 (Homo sapiens) (exp=-1; wgp=0; cg=0), partial (6%) [THC2557387]	-0.6	5.1	-4.0	0.0004	0.0403
A_23_P36928	POLR1D	NM_015972	Homo sapiens polymerase (RNA) I polypeptide D, 16kDa (POLR1D), transcript variant 1, mRNA [NM_015972]	-0.5	11.2	-4.0	0.0004	0.0403
A_32_P24651	AK095707	AK095707	Homo sapiens cDNA FLJ38388 fis, clone FEBRA2004485. [AK095707]	-0.5	8.4	-4.0	0.0004	0.0403
A_24_P253251	SLC7A1	NM_003045	Homo sapiens solute carrier family 7 (cationic amino acid transporter, y+ system), member 1 (SLC7A1), mRNA [NM_003045]	-0.7	8.8	-4.0	0.0004	0.0403
A_23_P100539	ABCC6	NM_001171	Homo sapiens ATP-binding cassette, sub-family C (CFTR/MRP), member 6 (ABCC6), transcript variant 1, mRNA [NM_001171]	0.9	6.7	4.0	0.0004	0.0403
A_24_P621023	GMEB1	ENST00000373816	Glucocorticoid modulatory element-binding protein 1 (GMEB-1) (Parvovirus initiation factor p96) (PIF p96) (DNA-binding protein p96PIF). [Source:Uniprot/SWISSPROT;Acc:Q9Y692] [ENST00000373816]	-0.6	9.2	-4.0	0.0004	0.0403
A_32_P2766	THC2628739	THC2628739	Q5T9M3_HUMAN (Q5T9M3) Protein tyrosine phosphatase, receptor type, C, partial (55%) [THC2628739]	-0.8	6.4	-4.0	0.0004	0.0403
A_23_P132378	CELSR1	NM_014246	Homo sapiens cadherin, EGF LAG seven-pass G-type receptor 1 (flamingo homolog, Drosophila) (CELSR1), mRNA [NM_014246]	-0.6	9.2	-4.0	0.0004	0.0404
A_23_P29939	SNCA	NM_007308	Homo sapiens synuclein, alpha (non A4 component of amyloid precursor) (SNCA), transcript variant NACP112, mRNA [NM_007308]	-0.5	5.3	-4.0	0.0004	0.0410
A_23_P337422	SERBP1	NM_001018067	Homo sapiens SERPINE1 mRNA binding protein 1 (SERBP1), transcript variant 1, mRNA [NM_001018067]	-1.0	9.7	-4.0	0.0004	0.0410
A_32_P99690	NLN	ENST00000380985	Neurolysin, mitochondrial precursor (EC 3.4.24.16) (Neurotensin endopeptidase) (Mitochondrial oligopeptidase M) (Microsomal endopeptidase) (MEP). [Source:Uniprot/SWISSPROT;Acc:Q9BYT8] [ENST00000380985]	-0.5	6.7	-4.0	0.0004	0.0410
A_24_P252130	PPARD	NM_006238	Homo sapiens peroxisome proliferator-activated receptor delta (PPARD), mRNA [NM_006238]	0.8	11.4	4.0	0.0004	0.0410
A_32_P219942	ZIC5	NM_033132	Homo sapiens Zic family member 5 (odd-paired homolog, Drosophila) (ZIC5), mRNA [NM_033132]	-0.6	6.7	-4.0	0.0004	0.0410
A_32_P16323	ENST00000355077	ENST00000355077	protein expressed in prostate, ovary, testis, and placenta 14 [Source:RefSeq_peptide;Acc:NP_001005356] [ENST00000355077]	0.9	8.2	4.0	0.0004	0.0410
A_23_P362228	C1orf213	NM_138479	Homo sapiens chromosome 1 open reading frame 213 (C1orf213), transcript variant 1, mRNA [NM_138479]	-0.5	8.6	-4.0	0.0004	0.0410
A_23_P354908	NMNAT2	NM_015039	Homo sapiens nicotinamide nucleotide adenylyltransferase 2 (NMNAT2), transcript variant 1, mRNA [NM_015039]	0.7	7.6	4.0	0.0004	0.0411
A_24_P778607	AK090897	AK090897	Homo sapiens cDNA FLJ33578 fis, clone BRAMY2011639. [AK090897]	0.8	7.7	4.0	0.0004	0.0413
A_32_P137391	MGC33846	NM_175885	Homo sapiens hypothetical protein MGC33846 (MGC33846), mRNA [NM_175885]	0.8	6.9	4.0	0.0004	0.0413
A_23_P41470	FLJ20035	NM_017631	Homo sapiens hypothetical protein FLJ20035 (FLJ20035), mRNA [NM_017631]	1.4	7.9	4.0	0.0004	0.0413
A_32_P159023	BX114329	BX114329	BX114329 BX114329 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGp998G064741 ; IMAGE:1932317, mRNA sequence [BX114329]	0.6	7.5	4.0	0.0004	0.0413
A_23_P119778	SLC39A10	NM_020342	Homo sapiens solute carrier family 39 (zinc transporter), member 10 (SLC39A10), mRNA [NM_020342]	-1.2	8.0	-4.0	0.0004	0.0413
A_32_P135601	THC2643320	THC2643320	Q306F7_HUMAN (Q306F7) Down syndrome encephalopathy related protein 1, partial (11%) [THC2643320]	-0.6	6.9	-4.0	0.0005	0.0415
A_23_P313828	PRR6	NM_181716	Homo sapiens proline rich 6 (PRR6), mRNA [NM_181716]	-0.6	7.4	-4.0	0.0005	0.0421
A_23_P203920	SSPN	NM_005086	Homo sapiens sarcospan (Kras oncogene-associated gene) (SSPN), mRNA [NM_005086]	0.9	9.4	4.0	0.0005	0.0421
A_23_P145238	HIST1H2BK	NM_080593	Homo sapiens histone cluster 1, H2bk (HIST1H2BK), mRNA [NM_080593]	0.6	13.3	4.0	0.0005	0.0421
A_23_P59375	ID4	NM_001546	Homo sapiens inhibitor of DNA binding 4, dominant negative helix-loop-helix protein (ID4), mRNA [NM_001546]	0.5	5.0	4.0	0.0005	0.0421
A_32_P49516	THC2626957	THC2626957	Unknown	0.7	5.6	4.0	0.0005	0.0421
A_23_P15692	GPR172B	NM_017986	Homo sapiens G protein-coupled receptor 172B (GPR172B), mRNA [NM_017986]	0.9	8.3	4.0	0.0005	0.0421
A_23_P256641	KCNE1L	NM_012282	Homo sapiens KCNE1-like (KCNE1L), mRNA [NM_012282]	0.5	6.3	4.0	0.0005	0.0421
A_23_P131653	PIGF	NM_002643	Homo sapiens phosphatidylinositol glycan anchor biosynthesis, class F (PIGF), transcript variant 1, mRNA [NM_002643]	-0.6	9.9	-4.0	0.0005	0.0421

ProbeName	GeneName	SystematicName	Description	logFC	AveExpr	t	P.Value	adj.P.Val
A_23_P428184	HIST1H2AD	NM_021065	Homo sapiens histone cluster 1, H2ad (HIST1H2AD), mRNA [NM_021065]	0.6	13.4	4.0	0.0005	0.0422
A_23_P366216	HIST1H2BH	NM_003524	Homo sapiens histone cluster 1, H2bh (HIST1H2BH), mRNA [NM_003524]	0.6	13.1	3.9	0.0005	0.0425
A_24_P113674	HLA-B	NM_005514	Homo sapiens major histocompatibility complex, class I, B (HLA-B), mRNA [NM_005514]	0.6	13.6	3.9	0.0005	0.0425
A_23_P46141	CTSS	NM_004079	Homo sapiens cathepsin S (CTSS), mRNA [NM_004079]	0.9	6.4	3.9	0.0005	0.0425
A_23_P26905	POLG2	NM_007215	Homo sapiens polymerase (DNA directed), gamma 2, accessory subunit (POLG2), mRNA [NM_007215]	-0.5	11.4	-3.9	0.0005	0.0425
A_24_P16124	IFITM4P	NR_001590	Homo sapiens interferon induced transmembrane protein 4 pseudogene (IFITM4P) on chromosome 6 [NR_001590]	1.4	7.3	3.9	0.0005	0.0425
A_32_P221429	BM928667	BM928667	AGENCOURT_6726860 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:5798808 5', mRNA sequence [BM928667]	-0.5	11.3	-3.9	0.0005	0.0425
A_23_P213141	TRIM2	NM_015271	Homo sapiens tripartite motif-containing 2 (TRIM2), mRNA [NM_015271]	1.1	9.1	3.9	0.0005	0.0425
A_23_P332713	KIAA1086	NM_015174	Homo sapiens KIAA1086 (KIAA1086), mRNA [NM_015174]	0.9	8.2	3.9	0.0005	0.0426
A_23_P143845	TIPARP	NM_015508	Homo sapiens TCDD-inducible poly(ADP-ribose) polymerase (TIPARP), mRNA [NM_015508]	0.7	11.9	3.9	0.0005	0.0426
A_23_P41344	EREG	NM_001432	Homo sapiens epiregulin (EREG), mRNA [NM_001432]	-2.7	7.6	-3.9	0.0005	0.0431
A_23_P29953	IL15	NM_172174	Homo sapiens interleukin 15 (IL15), transcript variant 1, mRNA [NM_172174]	0.6	9.2	3.9	0.0005	0.0433
A_32_P104432	AK123302	AK123302	Homo sapiens cDNA FLJ41308 fis, clone BRAMY2042612. [AK123302]	0.8	7.7	3.9	0.0005	0.0433
A_24_P263826	AK055694	AK055694	Homo sapiens cDNA FLJ31132 fis, clone IMR3220000953. [AK055694]	0.6	5.8	3.9	0.0005	0.0434
A_23_P207850	TNS4	NM_032865	Homo sapiens tensin 4 (TNS4), mRNA [NM_032865]	-1.1	6.3	-3.9	0.0005	0.0435
A_23_P386942	DIRAS1	NM_145173	Homo sapiens DIRAS family, GTP-binding RAS-like 1 (DIRAS1), mRNA [NM_145173]	-1.0	8.2	-3.9	0.0005	0.0435
A_32_P121751	THC2678624	THC2678624	Q5SLK5_THET8 (Q5SLK5) 2-oxoglutarate dehydrogenase E2 component (Dihydrolipoamide succinyltransferase), partial (5%) [THC2765213]	-0.5	5.6	-3.9	0.0005	0.0439
A_23_P251695	NXT1	NM_013248	Homo sapiens NTF2-like export factor 1 (NXT1), mRNA [NM_013248]	-0.5	11.7	-3.9	0.0005	0.0440
A_24_P344087	REC8	NM_005132	Homo sapiens REC8-like 1 (yeast) (REC8L1), transcript variant 1, mRNA [NM_005132]	1.2	8.5	3.9	0.0005	0.0443
A_23_P214459	PRPH2	NM_000322	Homo sapiens peripherin 2 (retinal degeneration, slow) (PRPH2), mRNA [NM_000322]	0.7	9.1	3.9	0.0005	0.0443
A_23_P217447	LAMP2	NM_002294	Homo sapiens lysosomal-associated membrane protein 2 (LAMP2), transcript variant LAMP2A, mRNA [NM_002294]	0.7	9.6	3.9	0.0005	0.0444
A_23_P36700	TAPBPL	NM_018009	Homo sapiens TAP binding protein-like (TAPBPL), mRNA [NM_018009]	0.6	9.1	3.9	0.0005	0.0446
A_24_P467449	CA438802	CA438802	UI-H-DH0-aus-b-12-0-UI.s1 NCI_CGAP_DH0 Homo sapiens cDNA clone UI-H-DH0-aus-b-12-0-UI 3', mRNA sequence [CA438802]	-0.6	7.0	-3.9	0.0005	0.0446
A_23_P323563	PLEKHG2	NM_022835	Homo sapiens pleckstrin homology domain containing, family G (with RhoGef domain) member 2 (PLEKHG2), mRNA [NM_022835]	-0.8	6.5	-3.9	0.0005	0.0446
A_23_P66854	KRT20	NM_019010	Homo sapiens keratin 20 (KRT20), mRNA [NM_019010]	0.5	5.5	3.9	0.0005	0.0446
A_23_P251293	SNCG	NM_003087	Homo sapiens synuclein, gamma (breast cancer-specific protein 1) (SNCG), mRNA [NM_003087]	0.8	7.4	3.9	0.0005	0.0446
A_32_P60716	FLJ25770	NM_178555	Homo sapiens hypothetical protein FLJ25770 (FLJ25770), transcript variant 2, mRNA [NM_178555]	0.5	6.6	3.9	0.0005	0.0448
A_23_P141345	MPP3	NM_001932	Homo sapiens membrane protein, palmitoylated 3 (MAGUK p55 subfamily member 3) (MPP3), mRNA [NM_001932]	-0.5	10.5	-3.9	0.0005	0.0448
A_24_P716162	A_24_P716162	A_24_P716162	Unknown	-0.5	8.4	-3.9	0.0006	0.0453
A_23_P57417	MMP11	NM_005940	Homo sapiens matrix metallopeptidase 11 (stromelysin 3) (MMP11), mRNA [NM_005940]	-0.7	6.4	-3.9	0.0006	0.0453
A_24_P326739	GLS2	NM_013267	Homo sapiens glutaminase 2 (liver, mitochondrial) (GLS2), nuclear gene encoding mitochondrial protein, mRNA [NM_013267]	-0.7	10.0	-3.9	0.0006	0.0454
A_24_P799680	A_24_P799680	A_24_P799680	Unknown	0.6	5.9	3.9	0.0006	0.0456
A_23_P42178	HIST1H2BF	NM_003522	Homo sapiens histone cluster 1, H2bf (HIST1H2BF), mRNA [NM_003522]	0.5	13.4	3.9	0.0006	0.0456
A_23_P93109	ENST00000366800	ENST00000366800	Uncharacterized protein C6orf54 (Protein HGC6.1). [Source:Uniprot/SWISSPROT;Acc:Q9Y6Z4] [ENST00000366800]	0.3	5.0	3.9	0.0006	0.0457
A_23_P157736	PPAPDC3	NM_032728	Homo sapiens phosphatidic acid phosphatase type 2 domain containing 3 (PPAPDC3), mRNA [NM_032728]	0.8	6.1	3.9	0.0006	0.0457
A_23_P334308	KIAA0774	NM_001033602	Homo sapiens KIAA0774 (KIAA0774), transcript variant 1, mRNA [NM_001033602]	0.5	5.9	3.9	0.0006	0.0459
A_24_P382187	IGFBP4	NM_001552	Homo sapiens insulin-like growth factor binding protein 4 (IGFBP4), mRNA [NM_001552]	-0.9	7.4	-3.9	0.0006	0.0459
A_23_P56933	RTN4	NM_020532	Homo sapiens reticulon 4 (RTN4), transcript variant 1, mRNA [NM_020532]	0.6	12.1	3.9	0.0006	0.0467
A_24_P551067	BC032409	BC032409	Homo sapiens cDNA clone IMAGE:5217021, with apparent retained intron. [BC032409]	0.5	7.1	3.9	0.0006	0.0467
A_32_P163891	THC2672319	THC2672319	Unknown	1.2	6.6	3.9	0.0006	0.0467
A_23_P380881	ANKRD13B	NM_152345	Homo sapiens ankyrin repeat domain 13B (ANKRD13B), mRNA [NM_152345]	-0.8	6.5	-3.9	0.0006	0.0467
A_23_P45999	FBXO2	NM_012168	Homo sapiens F-box protein 2 (FBXO2), mRNA [NM_012168]	1.1	9.2	3.9	0.0006	0.0474
A_23_P212160	C3orf31	AK126587	Homo sapiens cDNA FLJ44624 fis, clone BRACE2017410. [AK126587]	-0.5	11.7	-3.9	0.0006	0.0474
A_23_P101737	BC032803	BC032803	Homo sapiens, clone IMAGE:3865769, mRNA, partial cds. [BC032803]	-0.7	8.2	-3.9	0.0006	0.0474
A_23_P819	ISG15	NM_005101	Homo sapiens ISG15 ubiquitin-like modifier (ISG15), mRNA [NM_005101]	1.4	11.9	3.9	0.0006	0.0474
A_23_P125748	ZMAT1	NM_032441	Homo sapiens zinc finger, matrin type 1 (ZMAT1), transcript variant 3, mRNA [NM_032441]	1.1	8.4	3.9	0.0006	0.0477
A_23_P25060	FLJ13769	AK023831	Homo sapiens cDNA FLJ13769 fis, clone PLACE4000222. [AK023831]	-0.6	5.6	-3.9	0.0006	0.0477
A_24_P111031	LOC414300	BC032033	Homo sapiens serine PI Kazal type 5-like 1, mRNA (cDNA clone IMAGE:4826141). [BC032033]	-0.9	5.0	-3.8	0.0006	0.0480
A_24_P201381	FGFBP3	NM_152429	Homo sapiens fibroblast growth factor binding protein 3 (FGFBP3), mRNA [NM_152429]	-0.7	7.4	-3.8	0.0006	0.0480
A_23_P166109	FLRT3	NM_198391	Homo sapiens fibronectin leucine rich transmembrane protein 3 (FLRT3), transcript variant 2, mRNA [NM_198391]	1.4	8.5	3.8	0.0006	0.0480
A_23_P23048	S100A9	NM_002965	Homo sapiens S100 calcium binding protein A9 (S100A9), mRNA [NM_002965]	0.6	5.2	3.8	0.0006	0.0480
A_24_P370702	GBP3	NM_018284	Homo sapiens guanylate binding protein 3 (GBP3), mRNA [NM_018284]	1.4	7.1	3.8	0.0006	0.0482
A_23_P211985	SNRK	NM_017719	Homo sapiens SNF related kinase (SNRK), mRNA [NM_017719]	-0.6	7.1	-3.8	0.0006	0.0482
A_32_P212578	AW579245	AW579245	AW579245 PM2-DT0042-120100-001-f07 DT0042 Homo sapiens cDNA, mRNA sequence [AW579245]	-0.4	6.3	-3.8	0.0006	0.0482
A_23_P56978	PTK6	NM_005975	Homo sapiens PTK6 protein tyrosine kinase 6 (PTK6), mRNA [NM_005975]	0.7	10.8	3.8	0.0006	0.0482
A_24_P479793	THC2707492	THC2707492	Unknown	-0.5	6.1	-3.8	0.0006	0.0482
A_23_P487	UCK2	NM_012474	Homo sapiens uridine-cytidine kinase 2 (UCK2), mRNA [NM_012474]	-0.6	12.6	-3.8	0.0006	0.0482
A_23_P30776	HIST1H2BE	NM_003523	Homo sapiens histone cluster 1, H2be (HIST1H2BE), mRNA [NM_003523]	0.5	12.9	3.8	0.0006	0.0482
A_32_P130536	THC2650668	THC2650668	Q4RGI5_TETNG (Q4RGI5) Chromosome undetermined SCAF15099, whole genome shotgun sequence. (Fragment), partial (6%) [THC2650668]	-0.5	6.4	-3.8	0.0006	0.0482
A_32_P126410	AK000901	AK000901	Homo sapiens cDNA FLJ10039 fis, clone HEMBA1000975. [AK000901]	-0.5	6.2	-3.8	0.0006	0.0482
A_32_P29140	THC2724046	THC2724046	AA344632 EST50565 Gall bladder I Homo sapiens cDNA 5' end, mRNA sequence [AA344632]	-0.7	6.1	-3.8	0.0006	0.0484
A_23_P103588	HMGCS2	NM_005518	Homo sapiens 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2 (mitochondrial) (HMGCS2), mRNA [NM_005518]	2.7	8.3	3.8	0.0007	0.0486
A_23_P91891	COPB2	NM_004766	Homo sapiens coatomer protein complex, subunit beta 2 (beta prime) (COPB2), mRNA [NM_004766]	0.4	14.4	3.8	0.0007	0.0489
A_23_P311144	CCDC138	NM_144978	Homo sapiens hypothetical protein FLJ32745 (FLJ32745), mRNA [NM_144978]	-0.5	6.4	-3.8	0.0007	0.0489
A_23_P413952	AK098579	AK098579	Homo sapiens cDNA FLJ25713 fis, clone TST05089. [AK098579]	-0.8	5.8	-3.8	0.0007	0.0492