CANCER TESTIS GENE EXPRESSION AS A BIOMARKER OF ONE-CARBON METABOLIC ACTIVITY, DRUG SENSITIVITY AND PHENOTYPIC HETEROGENEITY IN NON-SMALL CELL LUNG CANCER AND MALIGNANT MELANOMA

A DISSERTATION SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

MOLECULAR BIOLOGY AND GENETICS

By

Kerem Mert Şenses

September 2016

CANCER TESTIS GENE EXPRESSION AS A BIOMARKER OF ONE-CARBON METABOLIC ACTIVITY, DRUG SENSITIVITY AND PHENOTYPIC HETEROGENEITY IN NON-SMALL CELL LUNG CANCER AND MALIGNANT MELANOMA

By Kerem Mert Şenses

September 2016

We certify that we have read this dissertation and that in our opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

_____

Ali Osmay Güre (Advisor)

_____

Özlen Konu

_____

Zeynep Işıl Kalaylıoğlu

_____

Sreeparna Banerjee

_____

Serkan İsmail Göktuna

Approved for the Graduate School of Engineering and Science:

_____

Levent Onural

Director of the Graduate School

ABSTRACT

CANCER TESTIS GENE EXPRESSION AS A BIOMARKER OF ONE-CARBON METABOLIC ACTIVITY, DRUG SENSITIVITY AND PHENOTYPIC HETEROGENEITY IN NON-SMALL CELL LUNG CANCER AND MALIGNANT MELANOMA

Kerem Mert ŞENSES

Ph.D. in Molecular Biology and Genetics

Supervisor: Ali O. Güre

September, 2016

Expression of cancer-testis (CT) genes on X chromosome (CT-X) is restricted to tumors, with very low or no expression in normal adult tissues. CT-X gene expression is one of the factors contributing to tumor heterogeneity and is variably observed in various types of cancers including non-small cell lung cancer (NSCLC) and malignant melanoma. Regulation of CT-X gene expression has been strongly linked to DNA methylation of promoter regions, however, mechanisms leading to re-expression of these genes in tumors, driven by promoter hypomethylation, is remained unsolved. Although tumors expressing CT-X genes are shown to be associated with higher tumor stage, larger tumors and aggressiveness, differential drug sensitivity of CT-X positive and negative tumors have not been investigated. In this thesis, we aimed to find the association between one-carbon pathway polymorphisms and CT-X gene expression. Moreover, we tested various tools and methods to find effective drugs for tumors in different phenotypic subgroups determined by CT-X gene expression or by other factors.

Keywords: cancer-testis genes, DNA damage, one-carbon pathway, drug sensitivity, NSCLC, melanoma, microarray.

ÖZET

KÜÇÜK HÜCRE DIŞI AKCİĞER KANSERİ VE MALİN MELANOM'DA TEK-KARBON METABOLİK YOLAĞI AKTİVİTESİ, İLAÇ HASSASİYETİ VE FENOTİPİK HETEROJENİTE'NİN BİYO-BELİRTECİ OLARAK KANSER-TESTİS GEN İFADESİ

Kerem Mert ŞENSES

Moleküler Biyoloji ve Genetik, Doktora

Danışman: Ali O. Güre

Eylül, 2016

X kromozomu üzerinde bulunan kanser-testis (KT-X) genlerinin ifadesi tümörlerle sınırlı olup, normal yetişkin dokularında çok düşüktür veya hiç görülmez. KT-X gen ifadesi tümör heterojenitesine katkıda bulunan faktörlerden biridir ve küçük hücre dışı akciğer kanseri (KHDAK) ve malin melanom gibi birçok farklı kanser türünde değişen oranlarda görülmektedir. KT-X gene ifadesinin regülasyonu promotor bölge metilasyonu ile kuvvetlice ilişkilendirilmiş, buna karşın, promotor hipometilasyonunun neden olduğu tümörlerde sonradan ortaya çıkan KT-X geni ifade artışının altında yatan mekanizmalar henüz aydınlatılamamıştır. KT-X geni ifadesi gösteren tümörler, yüksek tümör evresi, büyük tümörler ve agresiflik ile ilişkilendirilmekle birlikte, KT-X pozitif ve KT-X negatif tümörlerdeki değişken ilaç hassasiyeti araştırılmamıştır. Bu tez çalışmasında, tek-karbon yolağı enzim polimorfizmlerinin KT-X geni ifadesi ile ilişkisi araştırılmıştır. Ayrıca, tümörlerde KT-X geni ifadesi ile birlikte farklı faktörlerin belirlediği değişik fenotipik alt gruplara karşı etkili olabilecek ilaçların belirlenebilmesi için birçok araç ve yöntem test edilmiştir.

Anahtar sözcükler: Kanser-testis genleri, DNA hasarı, tek-karbon yolağı, ilaç hassasiyeti, KHDAK, malin melanom, microarray.

Bu tezi, bana her zaman destek olan sevgili babam Kazım Şenses'e, öğretici ve yol gösterici tavsiyeleriyle güç ve yön bulduğum ve aynı zamanda çalıştığım alana yönelmemde en büyük katkıya sahip olan ağabeyim Kemal Yasin Şenses'e, sevgisi ve şefkatiyle bugünlere gelmemi sağlayan annem Semra Şenses'e, son olarak hayatıma varlığıyla anlam katan sevgili eşim Fatma Şenses'e adıyorum.

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my advisor Dr. Ali Osmay Güre for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Özlen Konu, Dr. Zeynep Kalaylıoğlu, Dr. Sreeparna Banerjee, and Dr. Serkan Göktuna for their insightful comments and encouragement, but also for the hard questions which incented me to widen my research from various perspectives.

My sincere thanks also goes to Murat İşbilen who helped me on analyses by writing R-based scripts used in PCA-based grouping of tumors.

I thank Akif Burak Tosun for writing the object-oriented Java application which helped us identify drugs from the DTP/NCI60 database.

I thank my fellow labmates for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had during the time we spent in Bilkent University. Also I thank my friends in the Department of Molecular Biology and Genetics in Bilkent University.

Last but not the least, I would like to thank my parents and my wife for supporting me spiritually throughout writing this thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Cancer-Testis Genes

1.1.1 Expression and regulation of cancer-testis genes

Expression of cancer-testis (CT) or cancer-germline genes is detectable only in testis, fetal ovaries and placenta among normal tissues. However, upon tumorigenic transformation, CT genes are reactivated in almost all tumor types with varying levels of expression. First CT gene, melanoma antigen-1 (MAGE-1) was identified in 1990 by T-cell epitope cloning [1,2]. The list of CT genes have expanded by discoveries through serological expression cloning (SEREX) and differential gene expression analyses [3]. To date, more than 140 CT genes in 70 different gene families are grouped as CT-X and non-X-CT genes where CT-X genes constituting about the half of all CT genes [4]. Although expression patterns are similar, structural and functional differences separates two groups of genes. Located exclusively on X chromosome, CT-X genes are arranged as direct or inverted repeats in multigene families. Non-X CT genes, located on autosomes, however, are mostly arranged as single-copy genes. Sequence conservation is high among members within a CT-X gene family, however, there is no homology between different CT-X gene families [5]. Even though there is no conservation across families, reactivation of almost all CT-X genes in cancer, occur simultaneously with genome-wide hypomethylation which also affects repeat sequence activation [6,7]. In support of this, treatment of cell lines with DNA methyltransferase inhibitors such as 5-

aza-2'-deoxycytidine (5-aza-CdR), have shown to result in upregulation of all CT-X genes examined in various studies [8-12], suggesting a common epigenetic mechanism behind CT-X gene regulation. Moreover, histone acetylation has also been shown to facilitate CT-X gene expression when it occurs in parallel with DNA demethylation [7]. Treatment with Trichostatin A (TSA) - a histone deacetylase inhibitor have resulted in an increase in CT-X gene expression, although to a lesser extent when compared to treatment of DNA methyltransferase inhibitor [10,11]. On the contrary, there are exceptions for non-X CT gene upregulation by DNA methyltransferase and histone deacetylase inhibitor treatment, where 5 out of 7 non-X CT genes were failed to be upregulated by 5-aza-CdR and TSA treatment in a recently published study [12]. Other differences between CT-X and non-X CT genes include differential expression in spermatogonia cells during normal testicular germ cell development. CT-X gene expression decreases when these cells enter meiosis [13] with the help of genomic methylation [14,15], loss of Suv39h2 expression and H4 hyperacetylation [16,17]. Non-X CT genes, on the other hand, are continuously expressed during meiosis of gametes [18-20]. Furthermore, CT-X genes are found to be more antigenic than non-X CT genes and are subjected to a number of anti-tumor vaccination studies [21]. Due to these differences between two groups of CT genes, CT-X genes stands out as distinct biomarkers of regulation of gene expression, cellular differentiation and transformation.

CT-X gene expression is considerably varied across different types of tumors as well as individual samples. RT-PCR based measurements of CT-X genes revealed bladder, breast, hepatic, lung, melanoma and ovarian tumors as highly positive whereas colon, gastric, renal and blood-derived malignancies as weakly positive for CT-X gene

expression [22,23]. Moreover, CT-X expression was shown to be varied greatly among histological subtypes of the same tumor where number of patients with squamous subtype was significantly higher than the number of patients with adenocarcinoma of NSCLC when assessed for CT-X expression [24]. In another recent study, triple negative breast cancer (TNBC) samples had the highest level of CT-X expression among other histological subtypes of the same tumor [25].

Variable CT-X gene expression is also observable at intra-tumor or intra-tissue level such that different cells of the same tumor or tissue was shown to have different levels of CT-X gene expression. Intra-tumor heterogeneity was well documented using immunohistochemistry where MAGE-1 and NY-ESO-1 protein expression was shown to be heterogeneously distributed within NSCLC [13] and ovarian cancer [26] section photos, respectively. Even though the level of CT-X expression is greatly varied across tumors, it is well established by our group [24] and others [27] that a tumor sample positive for a CT-X gene will highly likely to be positive for more than one CT-X gene. Coordinate expression of CT-X genes might be an evidence for existence of a common mechanism behind CT-X gene regulation.

Recently, our group has been able to demonstrate dynamic regulation of CT-X genes in Caco-2 spontaneous differentiation model. This is the first report indicating that the mechanism underlying CT gene expression can be dynamically regulated and indicates which pathways might be relevant to this. During differentiation of Caco-2 cells from mesenchymal to an epithelial state (MET), expression of SPANX-B and PAGE-2B increase together with Ten-eleven translocation-2 (TET2) and 5-hydroxymethylated (5hmC) DNA. This correlates with detachment of EZ2H and HP-1

from the promoters of these genes. When MET is reversed to epithelial-to-mesenchymal transition (EMT) in this model, downregulation of both CT-X genes was observed, together with the reversal of all molecular mechanisms involved [28]. Although this model indicates a link between EMT and CT gene expression, it can not fully explain why some highly epithelial tumors remain CT-negative; suggesting the presence of other yet unidentified mechanisms involved in CT gene expression regulation.

1.1.2 One-carbon metabolism and its possible impact on CT-X gene expression

Despite the presence of common epigenetic mechanisms controlling CT-X gene expression, there is no consensus on why some tumors or tumor cells express CT-X genes and some do not. We hypothesized that a reason for this discrepancy could be the differential S-adenosylmethionine (SAM) production rates of cells, which is the sole methyl donor for all cellular methylation reactions. One of the main research aims of this thesis was to address this question by subgrouping NSCLC tumor samples based on genetic markers of one-carbon pathway enzymes that are known to affect SAM production rates, which in turn could possibly have impact on CT-X gene expression.

Folate, an essential vitamin supplemented especially in vegetables, fruits and grains, is a one-carbon source for DNA, RNA and protein synthesis as well as methylation and maintenance of various biomolecules. One-carbon molecules of folate is utilized by the enzymes of one-carbon metabolism for the *de novo* synthesis of purines and pyrimidines as well as the synthesis of SAM [29]. Biological methylation reactions of phospholipids, neurotransmitters, proteins, DNA, RNA, and other small

molecules are catalyzed by the transfer of a methyl group from SAM, the only known methyl donor in the cell. Therefore, maintaining intracellular SAM levels is crucial for the tightly controlled cellular processes such as control of gene expression, DNA methylation and DNA repair in normal physiology. SAM levels were shown to be affected by inadequate intake of folate or the inefficient catalytic capacity of polymorphic enzyme variants of the one-carbon metabolism. A recent study has been shown that methyl-deficient diet resulted in significantly reduced levels of SAM and significantly decreased ratio of SAM/S-adenosyl homocysteine (SAH; the more stable by-product of SAM) where both events were concurrent with increased global DNA hypomethylation in rats [30]. Independent from folate absorption, SAM levels have been shown to be affected directly or indirectly by fluctuations in activities of certain enzymes which are responsible for the production of one-carbon derivatives [31-43]. Since CT-X gene expression occurs in parallel with global DNA hypomethylation and promoter-specific demethylation, inadequate SAM production may facilitate the coordinate upregulation of many CT-X genes in cancer.

Synthesis of SAM starts with intracellular transport of folate controlled by three different proteins; folate receptor (FR), reduced folate carrier (RFC) and proton-coupled folate transporter (PCFT). RFC and PCFT are members of solute carrier superfamily, which are distinguished from FR by the ability to transport folate by ion exchange gradient. FR, on the other hand, internalizes and transports folate in a receptor-mediated endocytic process [44]. RFC-1, the gene encoding RFC transmembrane protein, harbors a common non-synonymous polymorphism (RFC 80G>A, dbSNP ID: rs1051266) which results in the change of arginine to histidine at the 27^{th} amino acid position. This

polymorphism have been associated with moderately increased plasma folate concentration [32], suggesting an impaired intake of folate from serum. Intracellular transport of folate is followed by its rapid reduction to dihydrofolate (DHF) and subsequently to tetrahydrofolate (THF) by dihydrofolate reductase (DHRF) which catalyzes the entrance of folate into 1-carbon pools. THF is important for the synthesis of THF polyglutamate derivatives catalyzed by folylpolyglutamate synthetase (FPGS). These folate polyglutamate molecules serve as one-carbon donors in one-carbon transfer reactions when they are associated with methyl, formyl or methylene one-carbon moieties at the N^5 and/or N^{10} positions [44]. Alternatively, vitamin B6-dependent serine hydroxymethyltransferase (SHMT) directly converts THF to N^5-N^{10}-methylene-THF [45] which is subsequently catalyzed by thymidylate synthase (TS) by transferring a one-carbon group from N^5-N^{10}-methylene-THF to deoxyuridylate for thymidylate synthesis. TS transfers the one-carbon group rapidly and irreversibly, resulting in oxidation of THF to DHF. Thymidylate synthesis has the potential to deplete cellular N^5-N^{10}-methylene-THF pools, however, DHFR rapidly reduces DHF to THF to maintain the flow of one-carbon moieties [44]. Therefore, maintaining level of N^5-N^{10}-methylene-THF is critical for thymidylate synthesis in the cell. N^5-N^{10}-methylene-THF levels are also critical for the production of N^5-methyl-THF, the predominant form of folate in blood used in vitamin B-12 dependent remethylation of homocysteine (Hcy) to methionine [45]. N^5-methyl-THF is the only one-carbon source for methionine and SAM biosynthesis, thus concentration of N^5-methyl-THF is thought to be the limiting factor for the synthesis for both metabolites [46]. One of the major factors affecting N^5-methyl-THF concentration is the activity of N^5-N^{10}-methylene-THF reductase

(MTHFR) enzyme which catalyzes the reduction of N^5-N^{10}-methylene-THF to N^5-methyl-THF. Importance of MTHFR enzyme for one-carbon metabolism was emphasized in two separate studies where transgenic mice lacking exon 3 of mouse orthologue MTHFR (designated as MTHFR-deficient) was evaluated for plasma total Hcy, N^5-methyl-THF, SAM, SAH and DNA hypomethylation levels. One of these studies have shown that mice heterozygous or homozygous for MTHFR deficiency had significantly higher total plasma Hcy and lower N^5-methyl-THF levels compared to wildtype littermates where the latter metabolite was evaluated for both brain and liver tissues. Furthermore, MTHFR-deficient homozygous mice have been reported to die within the first five weeks of life [47]. In both studies, MTHFR-deficient mice exhibited significantly reduced levels of SAM, significantly increased levels of SAH and significantly increased level of [^{3}H]dCTP incorporation, indicative of increased global DNA hypomethylation [47,48].

Two common non-synonymous polymorphisms in the human MTHFR gene (MTHFR 677 C>T and MTFHR 1298 A>C) have been shown to result in comparable phenotypes with MTHFR-deficient mice. MTHFR 677 C>T polymorphism (dbSNP ID: rs1801133), located within the N-terminal catalytic domain, changes the protein amino acid sequence from the 222nd alanine to a valine [40]. This polymorphism have been associated with increased thermolability [35], reduced enzymatic activity [35,49,50], increased plasma Hcy [32,35,43,49-54], increased DNA hypomethylation [43,54,55], and decreased plasma folate [32,43,49,52-54,56] levels in numerous studies. The second most common polymorphism in MTHFR gene, MTFHR 1298 A>C (dbSNP ID: rs1801131) changes the protein sequence from 429th glutamate (E, Glu) to alanine (A,

Ala). When compared with MTFHR 677 C>T polymorphism, the effect of MTFHR 1298 A>C is limited on phenotype, where no significant difference was detected in plasma Hcy [50,53], plasma folate [43,53] and DNA hypomethylation [43,55] levels between individuals with MTFHR 1298 A/A versus C/C genotype in several studies. Only in one study, plasma Hcy levels were significantly increased in MTFHR 1298 C/C genotype (at the $p < 0.05$ level) compared with MTFHR 1298 A/A individuals [43]. Although the effect of MTHFR 1298 A>C polymorphism is limited on phenotype when assessed alone, combined heterozygosity for both MTHFR 677 C>T and MTHFR 1298 A>C polymorphisms have been shown to express similar phenotype with that of MTHFR 677 TT homozygote individuals [53,57].

The flow of one-carbon moieties in one-carbon metabolism continue as methionine synthase (MS), a methylcobalamin (vitamin B-12)-dependent enzyme encoded by the MTR gene, converts N^5-methyl-THF into THF by transferring a methyl group to Hcy which results in the formation of methionine. Subsequently, methionine adenosyltransferase (MAT) transfers the adenosyl portion of ATP to methionine for SAM synthesis (Figure 1.1). Upon remethylation of Hcy to methionine, the active cobalamin (I) MS [or Cbl (I) MTR] complex, is oxidized to inactive form, Cbl (II) MS. Methionine synthase reductase (MSR), encoded by the MTRR gene, reactivates Cbl (II) MS complex by a methyl group transfer from SAM, a process known as reductive methylation [58].

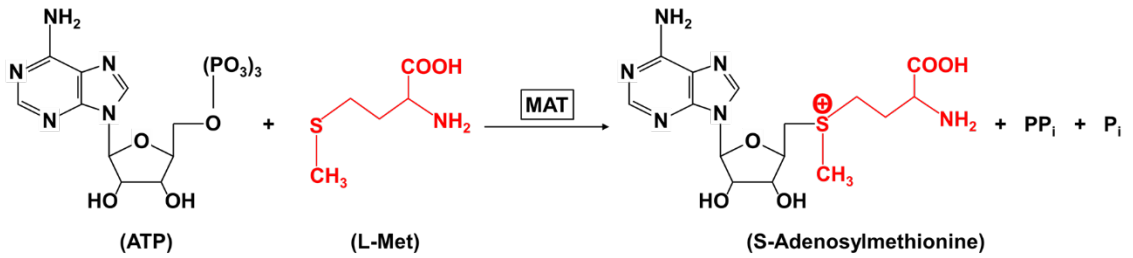


Figure 1.1: Reaction scheme of SAM synthesis. Synthesis of SAM (also referred as AdoMet) is catalyzed by methionine adenosyltransferase (MAT or AdoMet synthetase) by covalently attaching an adenosyl group from ATP with sulphur atom of methionine generating a sulphonium bond between the 5-carbon atom of the ribose (of ATP) and the sulphur atom of the methionine amino acid (L-Met) resulting in highly energetic SAM molecule [59].

Enzymatic activity of MS and MSR enzymes are influenced by two different non-synonymous polymorphisms; MTR 2756 A>G (dbSNP ID: rs1805087) and MTRR 66 A>G (dbSNP ID: rs1801394) changing 919th aspartic acid (D, Asp) to glycine (G, Gly) of MS and 22nd isoleucine (I, Ile) to methionine (M, Met) of MSR, respectively. Primary tumor samples harboring the less common MTR 2756 G/G genotype have been associated with decreased CpG island methylation compared to MTR 2756 A/A genotype [55] emphasizing the indirect but eminent role of MS on SAM production. Secondly, the variant G allele of MSR enzyme have been associated with 75% reduced in vitro enzymatic activity measured by the molar amount of MSR required for maximal activation of MS [38]. Therefore, it can be speculated that, the variant alleles of both MTR and MTRR genes may lead to impaired access for SAM which is utilized as the only cofactor in DNA methylation reactions.

Even though, diverse one-carbon donors are produced in one-carbon pathway reactions, high level of intermingling between these metabolic reactions ultimately

affect intracellular SAM concentrations directly or indirectly, resulting in a complex interplay between bio-processes such as DNA methylation, gene expression regulation, maintenance of DNA integrity and DNA synthesis. Here, it was hypothesized that tumors with hypoactive one-carbon pathway due to polymorphisms on critical enzymes, will result in expressing CT-X genes due to inefficient utilization of folate for the production of the sole methyl donor SAM.

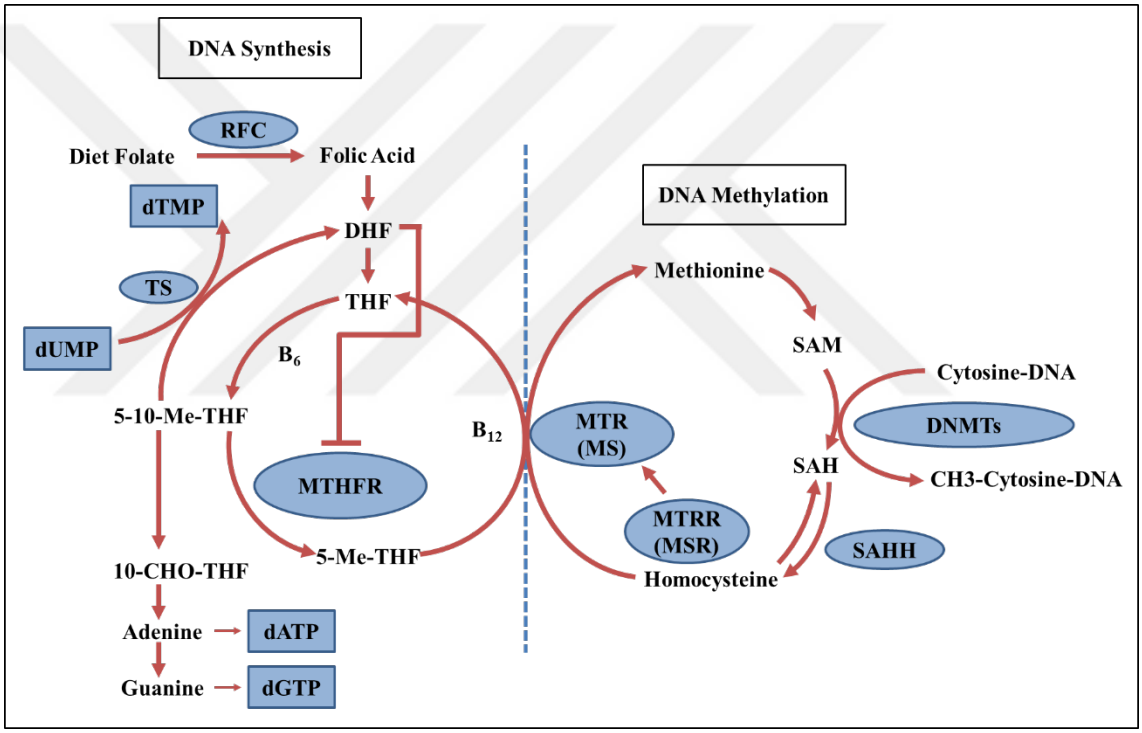


Figure 1.2: Schematic representation of one-carbon metabolism related with DNA synthesis and DNA methylation reactions. Abbreviations: RFC, reduced folate carrier; dUMP, deooxyuridine monophosphate; dTMP, thymidine monophosphate; dATP, deoxyadenosine triphosphate; dGTP, deoxyguasonosine triphosphosphate; TS, thymidylate synthase; DHF, dihydrofolate; THF, tetrahydrofolate; 5-10-Me-THF, N^5-N^{10}-methylene-THF, 5-Me-THF, N^5-methyl-THF; 10-CHO-THF, N^{10}-formyl-THF; MTHFR, methylenetetrahydrofolate reductase; MS, methionine synthase; MSR, methionine synthase reductase; SAM, S adenosylmethionine; SAH, S-adensylhomocysteine; SAHH, SAH Hydrolase; DNMTs, DNA Methyltransferases. Adopted from [60].

1.1.3 One-carbon metabolism and DNA damage

One-carbon metabolism is critical for maintaining physiological levels of purines and pyrimidines. Adequate levels of purines and pyrimidines are required for the synthesis and maintenance of DNA, therefore efficient repair of DNA damage is dependent on the efficient utilization of purines and pyrimidines through DNA damage-repair proteins. Since dTMP is synthesized by the transfer of a methyl group from N^5-N^{10}-methylene-THF to dUMP, deficiency in folate transport or decreased levels of N^5-N^{10}-methylene-THF as a consequence of decreased one-carbon metabolic activity results in dUMP accumulation, leading to misincorporated uracil in DNA [61]. If not removed by DNA glycosylase, misincorporated uracil results in single or double stranded DNA breaks, chromosome breakage and micronuclei formation [62]. Transfer of one-carbon units from folate is also required for the synthesis of purine ring from N^{10}-formyl-THF [63]. Decreased purine biosynthesis is linked to decreased folate availability as well as decreased enzymatic activity of methylenetetrahydrofolate dehydrogenase (MTHFD1) which in turn reduces availability of purines during DNA synthesis and repair [64,65]. Apart from the direct effect of folate deficiency and decreased one-carbon metabolic activity, DNA damage have been shown to be induced by DNA demethylation during primordial germ cell (PGC) development in mice. Mouse PGCs have been shown to accumulate single strand breaks and activate base excision repair (BER) during epigenetic reprogramming events such as changes in nuclear architecture driven by genome-wide active DNA demethylation [66]. Since DNA demethylation, the main control mechanism of CT-X gene expression, leads to increased DNA damage in PGCs,

it can be speculated that upregulation of CT-X gene expression can be an indicative of increased DNA damage in tumor cells.

In this thesis, association between polymorphisms of the one-carbon pathway enzymes (detailed in section 1.1.2) and CT-X gene expression was investigated in NSCLC with the aim of making a link between regulation of CT-X gene expression and SAM availability. Because folate metabolism is an important determinant of efficient DNA synthesis and repair, level of DNA damage in CT-X positive [CT-X (+)] and negative [CT-X (-)] cancer cell lines was also investigated.

Clinical features such as increased aggressiveness, poor prognosis and higher stage, which were previously associated with CT-X expression, might be indicative of differential drug sensitivity in CT-X (+) and (-) tumors. Therefore, using NCI-60 in silico drug screening and gene expression data, effectiveness of chemotherapeutic agents as well as single molecule inhibitors on CT-X (+) and (-) cancer cell lines was investigated.

1.2 Molecular subgroups of melanoma in relation to drug sensitivity

Although increased CT-X gene expression associates with worse outcome in several tumors undergoing conventional treatment, it can be an indicator of better outcome if specific therapeutic strategies, such as autologous vaccination, are to be used [67]. As CT-X expression is also related to EMT, and the fact that its expression is related to differential clinical outcome, determination of molecular markers in addition to CT-X gene expression might be informative for differential drug sensitivity. We chose to use melanoma as a model to test various markers for drug sensitivity as a large percentage of samples from this tumor are known to express CT-X genes, and because we were able to study a precious primary cell line culture obtained from melanoma patients.

Melanoma is the 7th most common malignancy after breast, lung, prostate, colon, lymphoma and bladder cancer with estimated 76,380 new cases in USA by the end of year 2016 [68]. For the early stage of disease, 5-year survival rate is 92% [69], however, median survival for stage IV metastatic melanoma is less than 1 year and 5-year survival rate is less than 10% [70,71].

Melanoma exhibits extensive inter- and intra-individual heterogeneity in which genomic and transcriptomic landscape substantially contributes to therapy resistance [72]. Perturbations in MAPK-ERK pathway is common in melanoma as a result of genetic alterations reported for oncogenes such as BRAF, NRAS and MAP2K1. BRAF V600E, the most predominant mutation in melanoma, was reported in 40% of a primary melanoma cohort, while only 15% of the same cohort was reported to be mutant for NRAS [73]. Intriguingly, a recent study have shown that melanomas wild-type for

BRAF and NRAS had higher mutation loads compared to patients carrying a mutation on either oncogene [74], which emphasizes the complex heterogeneity of melanoma. Furthermore, BRAF mutation is not a prerequisite for melanogenesis where benign nevi have shown to be mutant for BRAF at a high frequency. Additional genetic aberrations such as deletions or loss-of-function (LOF) mutations in genes involved in cell cycle control ($CDKN2A^{p16}$, TP53, BAP1, and PTEN) or copy number gain of transcription factors (MITF, MYC and ETV1) are necessary for melanogenesis (reviewed by Shtivelman et al., 2014 [75]). The complex interplay between different genomic events contributes to inextricable phenotypic heterogeneity of melanoma.

In addition to studies of genomic landscape in melanoma, molecular subgroups based on transcription is subjected to extensive research. Next generation RNA sequencing (NGS) have recently provided considerable advantage over microarray-based approaches, however, most of the transcriptomic signature studies on melanoma have been published using diverse microarray platforms. Studies which have reported gene expression profiles associated with melanoma prognosis have shown limited overlap between each other where only two biomarkers, MHC-II and osteopontin were common in two independent studies among 14 different studies. Hence, a commercial kit for clinical use in melanoma such as the clinically approved Agendia Mammaprint used for the prognostic evaluation of breast carcinoma patients, is still lacking. The reasons for sparsity of prognostic biomarkers in melanoma are listed by Tremante, 2012 [76] as; i) technical difficulties in extracting pure melanoma cells from the site of lesion and low level of availability for tumor material in array-based testing, ii) lack of association between genetic and transcriptional signatures (e.g., BRAFV600E mutation

is not correlated with any gene expression signature studied to date), iii) low level of molecular differences between normal melanocytes and malignant melanoma cells, and iv) lack of consensus for a melanoma stem-cell signature.

Melanoma cells show transcriptome based heterogeneity when studied in vitro, and this is important as it has been shown to correlate with drug sensitivity [77]. Thus, complete cure of melanoma might have to be pursued from two arms, one based on in vivo effects and therefore the microenvironment of the tumor cells and the other based on in vitro defined mechanisms related to drug sensitivity. Switching between phenotypes has been observed for melanoma in vivo, as well as in vitro, and can result in altered drug sensitivity, irrespective of mutational status for melanoma cells [78,79] suggesting that a combination treatment consisting of two drugs affecting each one of the two phenotypes might be ultimately necessary for efficient tumor eradication.

Drug resistance has traditionally been associated with a stem-cell-like phenotype and a more mesenchymal state, as compared to an epithelial one in almost all tumors. The “stemness” of melanoma cells is thought to exist in a dynamic state as melanoma cells with stem-cell markers have been reported to arise from cells lacking these markers, stem-cell like melanoma cells might not be required for tumor initiation, and the stem and non-stem-cell phenotypes can switch dynamically [80,81]. Similar to the plasticity reported for stemness in melanoma, epithelial-to-mesenchymal transition (EMT) has also been implied as a factor driving metastasis and causing drug resistance in melanoma [82,83]. Thirdly, melanoma specific in vitro analyses aiming primarily to explain differing metastatic potential generated gene lists that could robustly classify

melanoma cell lines or tumor cells into either a “proliferative” or an “invasive” character [84-86]; and switching among these phenotypes was shown as well [78,81].

Metastasis, drug resistance and invasion have all been attributed both to the stem cell phenotype, as well as to EMT and have been observed in those classes defined by various researchers for melanoma. In this study we asked if there was overlap between melanoma cells classified by these various signatures and one that would be based on CT-X gene expression, and if such melanoma subtypes would have unique responses to targeted therapy. We also aimed the identification of a minimum number of biomarkers that could help classify melanoma subtypes with differing drug sensitivity profiles.

Chapter 2

Materials and Methods

2.1 Materials

Table 2-1: List of instruments used during experiments

Name of Instrument	Company	Catalogue of part number	Assay Operated
Applied Biosystems 7500 Real-Time PCR System	Thermo Fisher Scientific (Waltham, MA USA)	4351106	Quantitative measurement of mRNA expression
Applied Biosystems GeneAmp PCR System 9700 End Point PCR Machine	Thermo Fisher Scientific (Waltham, MA USA)	N805-0200	RFLP-PCR, cDNA production, restriction digestion
Stratagene Mx3005P qPCR System	Agilent (Santa Clara, CA, USA)	401514	TaqMan SNP Genotyping experiments
Qubit™ fluorometer 1.0	Thermo Fisher Scientific (Waltham, MA USA)	Q32857	Quantification of genomic DNA and mRNA samples
NanoDrop 2000 Spectrophotometer	Thermo Fisher Scientific (Waltham, MA USA)	ND-2000	Quantification of genomic DNA and mRNA samples

Centrifuge 5810 R	Eppendorf (Hamburg, GERMANY)	5810R	RNA, DNA extractions, cell pelleting
XCell SureLock™ Mini-Cell Electrophoresis System	Life Sciences (CA, USA)	EI0001	RFLP-PCR PAGE
AxioCam MRc5 image capture device	Carl Zeiss (Oberkochen, GERMANY)		Visualization of comet slides
Ultraviolet Crosslinker	UVP, LLC (Upland, CA, USA)	CL-1000	Inducing DNA double strand breaks for H2AX phosphorylation assay

Table 2-2: List of chemicals, enzymes and kits.

Name	Catalog or part number	Company	Used in Experiment
TRIzol reagent	15596018	Ambion by Life Sciences (CA, USA)	Total RNA extraction
Chloroform	372978	Sigma-Aldrich Chemie GmbH (Taufkirchen Germany)	Genomic DNA, total RNA extraction
Phenol: Chloroform: IAA, 25:24:1, pH 6.6	AM9730	Ambion by Life Sciences (CA, USA)	Genomic DNA extraction
Proteinase K	P2308	Sigma Aldrich (St. Louis, USA)	Genomic DNA extraction

Nuclease free water	AM9930	Ambion by Life Sciences (CA, USA)	Various experiments
DNA-free™ Kit DNase Treatment and Removal Reagents	AM1906	Ambion by Life Sciences (CA, USA)	Removal of DNA from total RNA extracts
High-Capacity cDNA Reverse Transcription Kits	4368814	Thermo Fisher Scientific (Waltham, MA USA)	cDNA production
DyNAzyme II DNA Polymerase	# F-501S	Thermo Fisher Scientific (Waltham, MA USA)	End point PCR
OneTaq Hot Start DNA Polymerase	M0481S	New England BioLabs Inc.	End point PCR
SYBR® Green PCR Master Mix	4309155	Applied Biosystems by Life Sciences (CA, USA)	Quantitative measurement of mRNA expression
TaqMan® Genotyping Master Mix	4371355	Applied Biosystems by Life Sciences (CA, USA)	TaqMan SNP genotyping experiments
TaqMan® SNP Genotyping Assays, MTHFR 677 C>T	4351376 (assay ID: C___1202883_20)	Applied Biosystems by Life Sciences (CA, USA)	TaqMan SNP genotyping experiments
TaqMan® SNP Genotyping Assays, MTHFR 1298 A>C	4351376 (assay ID: C____850486_20)	Applied Biosystems by Life Sciences (CA, USA)	TaqMan SNP genotyping experiments
TaqMan® SNP Genotyping Assays, MTR 2756 A>G	4351376 (assay ID: C__12005959_10)	Applied Biosystems by Life Sciences (CA, USA)	TaqMan SNP genotyping experiments
TaqMan® SNP Genotyping Assays, MTRR 66 A>G	4351376 (assay ID: C___3068176_10)	Applied Biosystems by Life Sciences (CA, USA)	TaqMan SNP genotyping experiments

HinP1I	R0124S	New England Biolabs, Hertfordshire, UK	PCR- RFLP of RFC G80A polymorphism
HinF1I	R0155S	New England Biolabs, Hertfordshire, UK	PCR- RFLP of MTHFR C677T polymorphism
MboII	R0148S	New England Biolabs, Hertfordshire, UK	PCR- RFLP of MTHFR A1298C polymorphism
Nde1	R0111S	New England Biolabs, Hertfordshire, UK	PCR- RFLP of MTRR G66A polymorphism
HaeII	R0108S	New England Biolabs, Hertfordshire, UK	PCR- RFLP of MTR A2756G polymorphism
Novex® TBE Gels, 10%, 15 well	EC62755BOX	Thermo Fisher Scientific (Waltham, MA USA)	Separation of PCR-RFLP fragments
AGAROSE GEL	HS-8000	PRONA BASICA LE	Agarose gel electrophoresis
Quant-iT™ dsDNA HS Assay Kits	Q33120	Thermo Fisher Scientific (Waltham, MA USA)	Quantification of genomic DNA
Quant-iT™ RNA Assay Kit	Q33140	Thermo Fisher Scientific (Waltham, MA USA)	Quantification of total RNA
CellTiter-Glo® Luminescent Cell Viability Assay	G7572	Promega (Madison WI, USA)	Quantification of cytotoxicity and proliferation
CyQUANT® Cell Proliferation Assay Kit	C7026	Thermo Fisher Scientific (Waltham, MA USA)	Quantification of cytotoxicity and proliferation
Saracatinib (AZD0530)	S1006	Selleck Chemicals (Houston, TX, USA)	In vitro drug screening

Selumetinib (AZD6244)	S1008	Selleck Chemicals (Houston, TX, USA)	In vitro drug screening
17-AAG (Tanespimycin)	S1141	Selleck Chemicals (Houston, TX, USA)	In vitro drug screening
DMSO (Dimethyl sulfoxide)	A3672,0100	AppliChem MO, USA	Drug solvent, cryopreservation of cell lines
H2A.X Phosphorylation Assay Kit (Flow Cytometry)	17-344	MILLIPORE (Temecula, CA)	Measurement of DNA double strand breaks
Recombinant Human/Mouse/Rat Activin A	338-AC-010	R & D SYSTEMS (Minneapolis, MN, USA)	INHBA treatment experiments
RPMI 1640 medium	F 1215	Biochrom AG (Berlin, Germany)	Tissue culture applications
RPMI 1640 medium, no folic acid	27016021	GIBCO- Thermo Fisher Scientific (Waltham, MA, USA)	Culturing of cancer cell lines for comet assay
Dialyzed Fetal Bovine Serum	26400044	GIBCO- Thermo Fisher Scientific (Waltham, MA, USA)	Culturing of cancer cell lines for comet assay
Folic acid	F7876-1G	Sigma-Aldrich Chemie GmbH (Taufkirchen Germany)	Culturing of cancer cell lines for comet assay
FBS, ORIGIN: AUSTRALIA	S 0415	Biochrom AG (Berlin, Germany)	Tissue culture applications
DMEM medium	FG 0415	Biochrom AG (Berlin, Germany)	Tissue culture applications
Trypsin-EDTA	CC-5012	Lonza (Basel, Switzerland)	Tissue culture applications
L-Glutamine	17-605E	Lonza (Basel, Switzerland)	Tissue culture applications
Penicillin/	17-602E	Lonza (Basel,	Tissue culture

Streptomycin		Switzerland)	applications

Table 2-3: Primers used for q-RT-PCR analyses of MLANA and INHBA

Primer Name	Sequence	Product Length (bp*)	Tm (^{0}C)
GAPDH F	5’-GGAGCGAGATCCCTCCAAAAT-3’	101	64
GAPDH R	5’-GGCTGTTGTCATACTTCTCATGG -3’		64
MLANA F	5’-CAGCCGTGGTGTAAGAGTGG -3’	108	60
MLANA R	5’-ATTAAGGAAGGTGTCCTGTGCC -3’		61
INHBA F	5’-GAAGAGCCAGACTTCTGCACG -3’	100	62
INHBA R	5’-GTTTGCCGAGTCAGGAACAGC -3’		61

*bp, base pairs

Table 2-4: Primers used for genotyping RFC 80 G>A with PCR-RFLP

Primer Name	Sequence	Primer Length (bp*)	Tm (^{0}C)
RFCG80A_A1	5'-CTCCCGCGTGAAGTTCTT-3'	18	57
RFCG80A_A2	5'-AGCCGTAGAAGCAAAGGTAGC-3'	21	60
RFCG80A_B1	5'-TGCATTCGTCTCCAGGGTG-3'	19	60
RFCG80A_B2	5'-AGCGTCACCTTCGTCCCCTC-3'	20	64

*bp, base pairs

2.2 Methods

2.2.1 Patients, Tumor Samples and Cell Lines

2.2.1.1 NSCLC Patients and Tumor Samples

Tumor samples (n = 763) were obtained from primary NSCLC patients undergoing curative surgical resection between 1991 to July 2005 at the Department of Cardio-Thoracic Surgery, Weill Medical College of Cornell University. Informed consent which was approved by the Institutional Review Board of Weill Medical College of Cornell University, was obtained from all patients. Tumors for genotyping studies were selected by semi-quantitative PCR and the criteria was based on evaluation for the presence of transcripts from 9 different CT-X genes (MAGE-A1, MAGE-A3, MAGE-A4, MAGE-A10, CT-7, NY-ESO-1, LAGE-1, SSX2, and SSX4). Tumor samples having expression in at least 4 CT-X genes out of 9 or with strong expression in one CT-X gene was considered as CT-X (+) (n = 21). Tumors with no expression in 9 of the CT-X genes profiled were considered as CT-X (-) (n = 29, see Appendix B-Table 7-1 for details). CT-X gene positivity was determined in three levels; strong (+++), intermediate (++), weak (+ or +/-), or none (-), as described previously [24]. 12% of patients had no clinical data.

2.2.1.2 Cancer Cell Lines

Primary melanoma cell lines (PrMCLs) were generated as described in a previous study [67] and were a kind gift from Prof. Michal Lotem, Hadassah Medical Center,

Jerusalem, Israel. All cell lines (primary and commercial) were incubated in humidified incubators supplied with 5% CO2 and cultured in RPMI1640 medium supplemented with 10% FBS, 1% 10K/10K Penicillin-Streptomycin (P/S Solution, Lonza, Basel, Switzerland) and 1% 200mM LGlutamine (L-Glu, Lonza, Basel, Switzerland).

2.2.1.3 Other Cancer Cell Lines

Cancer cell lines used throughout the thesis were: SK-LC-17, COLO-205, SW-620, K-562 and SK-BR-3.

2.2.2 Tissue Culture Procedures and Experiments

2.2.2.1 Growing Conditions of Cell Lines

All commercial and primary cell lines were grown in complete RPMI-1640 (Lonza, Basel, Switzerland) medium consisted of 10% Fetal Bovine Serum (FBS) (GIBCO, Thermo Fisher Scientific, Waltham, MA, USA), 1% 10K/10K Penicillin-Streptomycin (P/S, Lonza, Basel, Switzerland) and 1% 200mM L-Glutamine (L-Glu, Lonza, Basel, Switzerland). All reagents were sterile filtered before supplementing into RPMI1640 bottles. Reagents for tissue culturing were stored based on manufacturers' recommendations. All cell lines and reagents were routinely tested for mycoplasma contamination using MycoAlert™ Mycoplasma Detection Kit (Lonza, Basel, Switzerland). Cultures positive for mycoplasma contamination were immediately discarded if not, treated with MycoZap™ Reagent (Lonza, Basel, Switzerland) according to manufacturer's instructions.

2.2.2.2 Harvesting Cell Lines

For adherent cell lines, growth media was removed from tissue culture flasks. Flask surface was washed with 5 mL 1 X PBS (phosphate buffered saline). Cells were detached using 1 mL of Trypsin-EDTA® (Lonza, Basel, Switzerland). Trypsin-EDTA solution was quenched by adding 3 mL of FBS containing RPMI-1640 media. Cell suspension was resuspended in 15 ml falcon and pelleted at 1000 rpm for 3 minutes. For non-adherent cell lines, cells were harvested directly from tissue culture flasks into 15 mL falcons. Cells were then pelleted at 1000 rpm for 3 minutes.

2.2.2.3 Freezing Cell Lines

Cell pellet was gently mixed with 90% FBS + 10% DMSO (Dimethyl sulfoxide, AppliChem MO, USA) solution and transferred into cryovials (Greiner Bio One, Kremsmünster, Austria). Cryovials were gradually frozen in -20^{0}C for 1h, then in -80^{0}C for 24 hours followed by indefinite storage in liquid nitrogen (-170^{0}C). All cryovials were stored in cryogenic storage boxes (USA Scientific, Ocala, FL, USA) which were stored in liquid nitrogen tank (Thermo Fisher Scientific, Waltham, MA, USA).

2.2.2.4 In vitro scratch assay

In vitro migration of melanoma cell lines were measured by in vitro scratch assay as previously described [87]. Cell lines were grown in six-well plates in normal growing conditions (see section 2.2.2.1) until the day of scrape (wound) formation with a 1000µL pipette tip. After scratch is created on cell layer, old complete media was removed and cells were gently washed with 1X PBS for the removal of remaining

serum. Two parallel lines are drawn ~2-3mm apart each other on the outer surface of every well of six-well plate which crosses the scratch with 90 degrees for generating a reference point for photographing the wound in different time intervals. In order to reduce the effect of proliferation on wound closure, cells were cultured in 1% FBS (instead of 10%) containing RPMI-1640 supplemented with 1% P/S and 1% L-Glu for the following days of the assay. Three different wounds in three different wells were photographed for all cell lines (except M187; only two photos were captured due to technical problems) and time intervals (0h, 24h and 48h). Remaining scratch area was measured for all images by the GIMP v2.8 Image Processing software. Briefly, all the scratch areas in all time intervals were measured by manually selecting scratch area with the software's "free select tool" and then following "Windows → Dockable Dialogs → Histogram" menu to make the software calculate pixel count for the selected area. Pixel count numeric value was considered as the area of that corresponding scratch.

2.2.3 Extraction and Quantification of DNA and RNA from Cell Lines

2.2.3.1 Genomic DNA (gDNA) Extraction from Cancer Cell Lines

Genomic DNA extraction was performed after harvesting cancer cell lines from culture flasks or dishes. Cell pellets were kept in ice cold PBS until the extraction procedure to prevent any DNA degradation due to DNAse enzymes. Briefly, cell pellets were washed with 0.3 ml of digestion buffer per 3 × 107 cells in an overnight long period at 50^0 C. For cell concentrations higher than 108, 1 ml of digestion buffer were used per

cell line. After overnight incubation, cell lysates were treated with equal volume of phenol/chloroform/isoamyl alcohol. Eppendorf tubes were then centrifuged 10 min at 1700 x g. The aqueous top layer were transferred into a new eppendorf and 0.15 ml (1/2 volume) of 7.5 M ammonium acetate and 0.3 ml (original amount of phenol/chloroform/isoamyl and cell lysate mix) of 100% ethanol were added. The amount of ammonium acetate and 100% ethanol should be adjusted if higher amount of digestion solution is used in the beginning. The stringy precipitate is recovered by centrifugation at 1700 x g for 2 minutes. Pellet was then rinsed with 70% ethanol following decanting and air drying. Extracted DNA was resuspended in ~1ml of TE buffer kept at 65^0 C for several hours for effective solubilization.

2.2.3.2 Total RNA Extraction from Adherent Cell Lines

Total RNA was extracted directly from cell culture flasks or dishes using TRIzol® reagent (Ambion® Thermo Fisher Scientific, Waltham, MA USA) according to manufacturer's instructions. Briefly, growth media was removed from culture flasks or dishes. Cells were lysed and homogenized by adding 1ml of TRIzol reagent per 10cm2 culturing area. Homogenized samples were transferred into 1.5ml eppendorf tubes and kept for 5 minutes at room temperature. Tubes were vigorously shaken after adding 0.2ml of chloroform per 1ml of TRIzol reagent and incubated 2-3 minutes at room temperature. Tubes were centrifuged at 12,000 x g for 15 minutes at 4^0C. Upper aqueous phase was removed into a new eppendorf since the extracted RNA remains exclusively in aqueous phase. To precipitate RNA in aqueous phase, 0.5ml of isopropanol was added and followed by 10 minutes of incubation at room temperature.

Eppendorf tubes were then centrifuged at 12,000 x g for 10 minutes at 4^0C. Supernatant was removed and RNA pellet was washed with 1ml of 75% ethanol and centrifuged at 7500 x g for 5 minutes at 4^0C. Ethanol containing wash solution was discarded and the remaining RNA pellet was air dried for 5 min at room temperature. RNA was reconstituted in nuclease-free ddH2O (Ambion® Thermo Fisher Scientific, Waltham, MA USA).

2.2.3.3 Quantification of gDNA and RNA Samples

Genomic DNA samples obtained from tumors and cell line pellets were quantified using NanoDrop™ (Thermo Fisher Scientific, Waltham, MA USA) and QuBit® fluorometer (Thermo Fisher Scientific, Waltham, MA USA) according to manufacturer's instructions. Briefly, NanoDrop measurements was achieved by loading 1µL of gDNA or RNA sample on the lower measurement pedestal. gDNA and RNA samples were considered as pure if the 260nm/280nm ratio was ~1.8 and ~2.0, respectively.

Quantities of gDNA and RNA was measured by QuBit fluorometer with Quant-iT™ dsDNA High-Sensitivity (HS) Assay Kit and Quant-iT™ RNA Broad-Range BR Assay Kit kits (see Table 2-2), according to manufacturer's instructions.

2.2.4 SNP Genotyping

2.2.4.1 SNP Genotyping with Quantitative Real Time PCR

SNPs with dbSNP IDs rs1801133, rs1801131, rs1805087, and rs1801394 (also known as MTHFR 677 C>T, MTHFR 1298 A>C, MTR 2756 A>G and MTRR 66 A>G,

respectively) were genotyped using TaqMan® SNP Genotyping assays (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA USA). Assay IDs were C___1202883_20, C___850486_20, C___12005959_10, and C___3068176_10, respectively for each SNP. The assays were run on Stratagene Mx3005P q-PCR machine (Agilent Technologies). TaqMan® Genotyping Master Mix (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) including AmpliTaq Gold® DNA Polymerase (Ultra Pure), dNTP mix (with dUTP), Passive Reference 1 (ROX), Uracil-D-glycosylase (UDG) and reaction buffer was used in amplification reactions. Thermal profile used in TaqMan® SNP Genotyping reactions were used according to manufacturer's instructions. ROX passive reference dye was used to minimize the effect of pipetting errors during reaction setup. UDG was used to prevent carry over contamination during pre-PCR. All samples were run as triplicates in each genotyping reaction to achieve maximum reproducibility. Each sample was genotyped at least twice in separate genotyping reactions. Genotype frequencies were in Hardy-Weinberg equilibrium (Table 2-5).

2.2.4.2 PCR-Restriction Fragment Length Polymorphism (RFLP)

RFC 80 G>A (rs1051266) polymorphism was genotyped using PCR-RFLP method. PCR primers are given in Table 2-4. Primer sequences were previously published [32]. PCR amplicons were digested with HinP1I restriction enzyme. Digested PCR products were run on Novex® TBE PAGE Gels, (10%, 15 well) or agarose gel electrophoresis (Table 2-2). All samples were genotyped at least twice. Genotype frequencies were in Hardy-Weinberg equilibrium (Table 2-5).

Table 2-5: Hardy-Weinberg Distributions of Single Nucleotide Polymorphisms in NSCLC Patients.

Polymorphism	Genotype	Observed	Observed/ Expected	$\chi 2$	P*
MTHFR 677 C>T (rs1801133)	CC	20	1.1	1.389	0.24
	CT	20	0.8		
	TT	10	1.3		
MTHFR 1298 A>C (rs1801131)	AA	27	1	0.078	0.78
	AC	20	1		
	CC	3	0.9		
MTR2756 A>G (rs1805087)	AA	36	1	1.325	0.25
	AG	14	1.2		
	GG	0	-		
MTRR66 A>G (rs1801394)	AA	15	1.1	0.703	0.4
	GA	22	0.9		
	GG	13	1.1		
RFC80 G>A (rs1051266)	GG	18	1.1	0.469	0.49
	GA	22	0.9		
	AA	10	1.1		

* Chi-Square test (one degree of freedom) (from Senses et al. BMC Med Genet, 2013 [143]).

2.2.5 Gene Expression Analyses

2.2.5.1 Complementary DNA (cDNA) Synthesis

Total RNA were converted to cDNA using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) according to manufacturer's instructions. All the cDNA reaction tubes were stored at -80 ^{0}C before use.

2.2.5.2 Quantitative Real Time PCR Analyses for gene expression

Gene expression was measured by SYBR® Green PCR Master Mix (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) with PCR primers given in Table 2-3. PCR was run in 20 µL of reaction volume and thermal profile was 95^0C for 10 min (AmpliTaq Gold® Polymerase Activation) followed by 40 cycles of denaturation (95^0C for 15 sec) and annealing/extension (60^0C for 1 min) steps. Melt curve was generated after 40 cycles to check for nonspecific product formation (amplification). Reaction was run on 7500 Real-Time PCR System (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA, see Table 2-1 for details).

2.2.6 Quantification of DNA Damage in Cancer Cell Lines

2.2.6.1 Single Cell Gel Electrophoresis (SCGE) Assay-Comet Assay

Experimental setup was adopted from a previous study (Figure 2.1) [88]. CT-X (+) SK-LC-17 and CT-X (-) COLO205 cancer cell lines were cultured in 12nM folic acid (FA) containing RPMI-1640 medium (Appendix B-Table 7-3). CT-X expression of both cell lines were measured in a previous study [89]. Unmodified RPMI-1640 and FBS contains supra-physiological levels of folic acid, therefore, 12nM FA containing RPMI-1640 was prepared by supplementing FA externally into FA-free RPMI-1640 and dialyzed FBS (see Table 2-2). DNA damage of cell lines was measured on 3 day intervals till day 9 to measure the effect of prolonged folic acid deficiency. Assay controls were Comet assay protocol was adopted from Dhawan et al. [90]. Comet assay was run on alkaline conditions. The separated DNA was visualized by ethidium

bromide staining under fluorescent microscope (AxioCam MRc5 image capture device Carl Zeiss (Oberkochen, GERMANY). Comet images are then analyzed by CASPLab Comet image analyses software [91] according to manufacturer's instructions. Buffers and chemicals used in comet assay protocol are given in Table 2-6. Detailed explanation of preparation of these substances can be find in [99]. This protocol is composed of three main parts. In the first part (steps a-to-g) include preparation of slides for electrophoresis. Steps "h-to-k" include the electrophoresis of slides in alkaline conditions. Steps "l-to-n" include visualization of comets and analyses of tail parameters with CASPLab software. This steps are detailed as follows:

a. Slides were cleaned by dipping in methanol and then burning over a blue flame. This step is required to remove the oil and dust on slides.
b. One-third of slides is dipped in melted normal melting agarose (NMA). Undersides of slides were wiped with towels to remove excess agarose and left in a flat surface for the NMA to dry.
c. Adherent cultures were detached by scraping to avoid extra DNA damage caused by enzymatic detachment methods (e.g., trypsination). Cell viability assay by trypan blue staining was performed on every assay. Cell viability below 80% is not acceptable, hence this will result in biased DNA damage measurements.
d. 75 µL of low melting point agarose (LMPA, 0.5%; 37ºC) was mixed gently with 10 µL of cell suspension (concentration: 1,000 cells/µL) in an eppendorf tube.

e. The LMPA + cell suspension was pipetted on the surface of NMA coated slides [ADD LINK TO PREVIOUS STEP FOR CLARITY] and covered with a coverslip. Slides were moved on ice packs for the LMPA to dry (for 5-10 min).

f. Coverslip were gently removed from slides and the third layer of low-melting agarose (80 µL LMPA) was added onto the slide. Slide was covered with a new coverslip and moved again on ice packs for the LMPA to dry (for 5-10 min).

g. Coverslips were removed very gently and slowly. Then, slides were dipped into cold, freshly made lysing solution followed by incubation in dark at 4ºC for 2 hours.

h. Slides were gently removed from lysing solution and placed side by side on horizontal gel box near one end. Slides were placed together as close as possible. Electrophoresis tank was filled with freshly made electrophoresis buffer (pH>13) until the liquid level completely covers the slides (bubbles were avoided over the agarose, if existed, removed by micropipette tips).

i. Slides were left in alkaline buffer for 20 minutes in dark to allow unwinding of the DNA and to show the expression of alkali-labile single strand DNA damage.

j. 0.74 V/cm fixed voltage was used and current was adjusted to 300 milliamperes by fine tuning the level of buffer in electrophoresis tank. Electrophoresis reaction was ran for 30 min. Electrophoresis tank was kept in cold room during the reaction to avoid melting of LMPA.

k. Slides were removed from the tank and tilted dropwise into a drain tray. Tray was filled with neutralization buffer and slides were kept there for 5 minutes.

After 5 min. neutralization buffer was discarded. This procedure was repeated with fresh neutralization buffer two more times.

l. Slides were stained with 80µL of 1X ethidium bromide and left for 5 min. and then dipped in chilled distilled water to remove excess stain. Slides were covered with coverslips and images were taken under fluorescent microscope. At least 30 comet images from every slide was taken to increase statistical power.

m. Slides which were to be evaluated later on were dried for archival purposes. Those slides were dipped in cold 100% ethanol (or cold 100% methanol) for 20 min. Then, slides were air dried and incubated at 50^0C for 30 min. Slides prepared with this protocol can be stored in low humidified area and can be stored for years. Evaluation of archived slides were achieved by rehydrating with chilled distilled water for 30 min. and staining with 80µL 1X Ethidium Bromide, as in step "l".

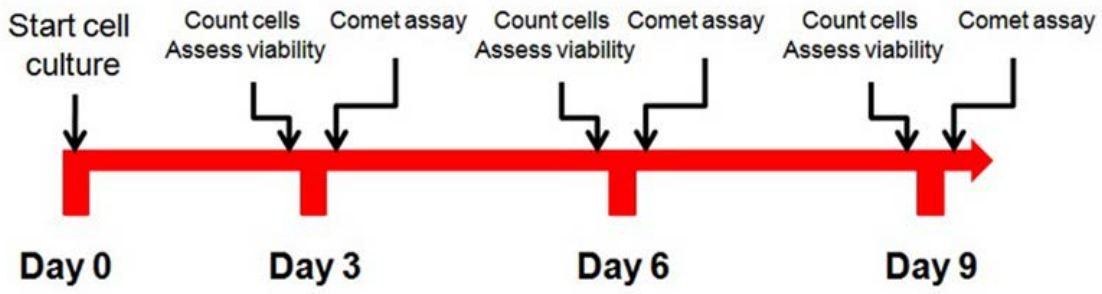


Figure 2.1: Comet assay experimental setup. SK-LC-17 and COLO-205 cancer cell lines were cultured in 75 cm2 flasks. In 3 day intervals till day 9, cells were detached and counted by trypan blue staining, viability was assessed. Cell supernatants were used for comet assay protocol.

Table 2-6: Chemicals or buffers used in comet assay.

Name of Chemical or Buffer	Ingredients
Normal Melting Agarose (NMA)	
Low Melting Point Agarose (LMPA)	
Absolute ethanol or methanol	
Coverslips (No. 1, 24 x 60 mm)	
PBS (Ca++, Mg++ free):	**137 mmol/l NaCl**
	2.7 mmol/l KCl
	10 mmol/l Na2HPO4
	1.8 mmol/l KH2PO4
	Adjust pH to 7.4
Lysing Solution:	**Ingredients per 1000 mL**
	2.5 M NaCl 146.1 g
	100 mM EDTA 37.2 g
	10 mM Trizma base 1.2 g
	1% Triton X-100
	10% DMSO
Electrophoresis Buffer:	**300 mM NaOH**
	1 mM EDTA
Neutralization Buffer:	**0.4 M Tris**
	pH adjusted to 7.5 with concentrated (>10 M) HCl
Staining Solution:	**1X EtBr Solution**
	(Prepared from stock solution: EtBr; 10X Stock - 20 µg/mL)

2.2.6.2 Histone H2A.X Phosphorylation Assay by Flow Cytometry

DNA double strand breaks were quantified using H2A.X Phosphorylation Assay Flow Cytometry Kit (MILLIPORE, Temecula, CA) according to manufacturer's instructions. This kit detects Histone H2A.X phosphorylation at serine 139. Cell lines were grown either in FA-free RPMI-1640 or normal RPMI-1640 culture media conditions in 60 mm culture dishes (see section 2.2.2.1 for growing conditions). Phosphorylated H2A.X was measured on two setups. In the first setup, cell lines were grown in tissue culture for 5 days and H2AX phosphorylation was directly measured by flow cytometry. In the second setup, after 5 days of incubation, DNA double strand breaks were induced by 100j/m2 UV-C irradiation using UV crosslinker and cells were incubated for 2 hours in tissue culture conditions (see section 2.2.2.1) for recovery before H2A.X measurement. Till the day of H2A.X measurement cell lines were kept in sub-confluent conditions to reduce inherent DNA damage. List of cell lines used in H2A.X measurement analyses are given in Appendix B-Table 7-3.

2.2.7 In vitro drug screening analyses

2.2.7.1 Treatment of cancer cell lines with single molecule inhibitors

AZD0530, 17AAG and AZD6244 were purchased from Selleck Chemicals. 10a was developed in a previous study and a gift from Prof. Botta [CITATION]. Research Drugs were screened on transparent 96-well-plates in which cancer cell lines were seeded as 1000 cell/well 96 hours before cytotoxicity measurement. Drugs were screened using 10 different doses ranging from 0.001 µM to 50 µM except 10a for which 6 different doses

ranging from 0.625µM to 20µM were tested. All stock solutions were prepared in 0.1% DMSO. IC50 values were calculated as previously described [92].

2.2.7.2 Cytotoxicity measurements and cell number quantification methods

CellTiter-Glo® Luminescent Cell Viability Assay (Promega, Madison WI, USA) was used for drug cytotoxicity measurement according to manufacturer's instructions. CyQUANT® Cell Proliferation Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA) was used for quantification of cell proliferation according to manufacturer's instructions.

2.2.8 In silico drug screening analyses

2.2.8.1 In silico drug screening analyses using NCI-60 data

NCI-60 microarray gene expression dataset (GSE5720) was normalized as detailed in section 2.2.9.1. NCI60 cell line panel were classified into 8 groups based on CT-X expression, with R-based script (detailed in section 2.2.9.2). CT-X expression of 11 cell lines from this panel was validated for MAGE-A3, NY-ESO-1, GAGE and SSX4-4B expression by q-RT-PCR. Developmental Therapeutics Programs' drug screening data was then downloaded and searched for compounds that had statistically distinct affects on CT-X (+) and CT-X (-) cells. This task was accomplished by an object-oriented Java application developed ("DrugScreen"), which compares the GI50, TGI and LC50 values of the ~49,000 drugs for CT (+) and CT (-) cell lines utilizing the Students' T-Test with Welch's correction. Cells were initially divided into two groups, the first excluding

colorectal cancer and melanoma cell lines known to be CT-X (-) and CT-X (+), respectively. Drugs that showed a preferential affect on CT-X (+) or CT-X (-) were then validated, in silico, against the second cell line group.

2.2.8.2 In silico drug screening analyses in melanoma

Normalized activity area (ActArea_norm) of 24 different commercial drugs were downloaded from CCLE database (https://portals.broadinstitute.org/ccle/home) [92] for in silico drug screening analyses.

2.2.9 In Silico Analyses of Gene Expression

2.2.9.1 Microarray Gene Expression Analyses

Microarray gene expression analyses were conducted using publically available datasets and dataset related information downloaded from Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo) [93] and/or ArrayExpress (www.ebi.ac.uk/arrayexpress) databases. All the normalization, gene selection, sample grouping, group comparisons and statistical analyses were done using BRB-ArrayTools [94] and GeneSpring12.5 (Agilent, Santa Clara, CA, USA). Publically available Affymetrix GeneChip arrays were normalized using the RMA method [95] without background correction and no other gene filtering criteria was used unless specified in an alternative method. Hierarchical clustering analyses were done using internal clustering engine of GeneSpring12.5 or Cluster3.0 [96]. Clustered data files were visualized with Java Treeview [97]. Principal component analyses (PCA) and volcano

plot visualization was conducted using custom made R-scripts [98]. Further formatting of heatmaps and data filtering were done using Microsoft® Office Excel 2013™, if necessary.

Table 2-7: Microarray datasets used to observe CT-X gene expression across tumors.

#	Tumor Type	Dataset Accession No	Microarray Platform	Reference
1	Melanoma	GSE15605	HGU-133 Plus2[$]	[99]
2	NSCLC*	GSE33532	HGU-133 Plus2	[100]
3	Breast Cancer	GSE65194	HGU-133 Plus2	[101-103]
4	Ovarian Cancer	GSE18520	HGU-133 Plus2	[104]
5	Colorectal Cancer	GSE22598	HGU-133 Plus2	[105]
6	CLL[£]	GSE31048	HGU-133 Plus2	[106]
7	Pancreas Cancer	GSE15471	HGU-133 Plus2	[107]
8	Pancreas Cancer	GSE16515	HGU-133 Plus2	[108]

*NSCLC: non-small cell lung cancer, [$]Affymetrix GeneChip Platform, [£]CLL: chronic lymphocytic leukemia

2.2.9.2 In silico association analyses of AML

Microarray gene expression dataset (GSE14471) and SNP genotyping dataset (GSE15714) of an AML cohort was used. Using a custom R script and normalized gene expression data (see section 2.2.9.1 for normalization details), firstly, Principal Component Analyses (PCA) was performed to calculate the 1st PC of CT-X probeset expression values of nine CT-X gene families in AML samples (probesets are given in Appendix A-Table 6-1). 1st PC values were explaining the 48% of the variation in the data set. Then, using 1st PC values, AML samples were clustered using K-means

clustering algorithm. AML samples were divided into five K-means clusters in which the optimum number of clusters were determined based on Elbow criterion as previously described [109].

2.2.9.3 Generation of a heuristic fibroblast detection algorithm

These analyses were conducted in two rounds. In the first round of analyses, GSE19234 was normalized as described in section 2.2.9.1. Exporting the normalized gene expression data of GSE19234 into Heuristic Online Phenotype Prediction (HOPP) tool (http://www.jurmo.ch/hopp) [85] Pro (r) and Inv (r) values of these samples were calculated. Melanoma samples with the highest 25% Inv (r) value were selected for the second round analyses, as the "invasive" samples are more likely to be misclassified with fibroblast cells than "proliferative" samples. In the second round of analyses, GSE17032 (skin fibroblast dataset) and most invasive melanoma samples were normalized together. The most differentially expressed genes in fibroblasts vs. invasive melanoma samples were determined according to a defined p-value and fold change (FC) cutoff criteria (p-value < 0.0001 and absolute FC ≥ 4). 122 different probesets fulfilled the criteria (list of probesets are given in Appendix C-Table 8-1). Matching probesets with the Luminex1000 dataset of Primary melanoma cell lines (PrMCLs) was determined and the Pearson's correlation coefficient (r) was calculated between the median values for the 122 probesets of fibroblast samples (designated as Fib(r)) and Luminex1000 values. This calculation was repeated between invasive melanoma samples (designated as Mel(r)) and Luminex1000 values. Pearson's r value was calculated for other skin fibroblast datasets (GSE11919, GSE16715, GSE21648,

GSE34309, and GSE35551), cancer-associated fibroblasts (CAFs, GSE46824), commercial melanoma cell lines (GSE36133) and other published primary melanoma datasets (GSE4840, GSE4841 and GSE4843).

2.2.9.4 EMT, CSC and PRO/INT/INV score calculation

EMT and CSC scores were calculated based on an R script 18(http://www.R-project.org/) which runs a principal component analyses (PCA) with corresponding probeset expression values. For EMT scoring, PCA was run on probeset expression values of ECAD (201131_s_at) and VIM (201426_s_at). For CSC scoring, PCA was run on CD271 (205858_at), CD20 (228599_at, 228592_at, 217418_x_at, 210356_x_at), CD133 (204304_s_at) and ALDH1 (212224_at) which were previously published melanoma stem-cell markers [110]. CCLE and primary melanoma cell lines were classified as PRO, INT and INV according to Pro (r) and Inv (r) values generated using the online HOPP tool [85]. CCLE cell lines with a Pro (r) value of more than 0.652 and an Inv (r) value of less than 0.171 were designated as “PRO”, those with a Pro (r) value between 0.041 to 0.630 and an Inv (r) value between 0.254 to 0.600 were designated as “INT” and cells with a Pro (r) value of less than -0.200 and an Inv (r) value of more than 0.445 were designated as “INV” cells. For primary cell lines, designations were as follows: PRO: Pro (r) > 0.533 and Inv (r) < 0.162; INT: Pro (r): 0.186 to 0.588 and Inv (r): 0.221 to 0.581; INV: Pro (r) < 0.03 and Inv (r) > 0.554

Chapter 3

Results

3.1 CT-X gene expression in different types of tumors according to in silico data

In an effort to show the level of CT-X gene expression in silico and in different types of tumors, microarray analyses were conducted using publically available mRNA gene expression datasets of melanoma, NSCLC, breast cancer, ovarian cancer, colon cancer, leukemia, and pancreas cancer (see Table 2-7 and section 2.2.9.1 for detailed explanation of analyses) were normalized and tumor samples were hierarchically clustered based on expression values of CT-X probesets [89] (Appendix A-Table 6-1). In concordance with literature, hierarchical clustering results of tumor samples revealed melanoma, NSCLC, breast and ovarian cancer as tumors with highly positive CT-X gene expression compared to leukemia, colon and pancreas cancer in which tumor samples had similar CT-X expression compared to their normal counterparts (see Appendix A- Figure 6.1 for melanoma, Figure 6.2 for NSCLC, Figure 6.3, for breast cancer, Figure 6.4 for ovarian cancer, Figure 6.5 for CLL, Figure 6.6 for CRC, Figure 6.7 and Figure 6.8 for pancreas cancer). Normal tissues of all tumors analyzed had lower, if not least expression of CT-X genes compared to their malignant counterparts which is one of the hallmarks of CT-X gene expression. Although, healthy samples in GSE16515 pancreas cancer dataset had slightly increased CT-X expression compared tumors (Figure 6.8), this was not observed in the other pancreas cancer dataset analyzed

(Figure 6.7). According to hierarchical clustering analyses of breast cancer samples, triple negative breast cancer (TNBC) had higher NY-ESO-1 (CTAG1A), MAGEA2, MAGEA3, MAGEA6 and MAGEA12 expression when compared to all other breast cancer histological subtypes which was in concordance with previously published data [25].

These in silico analyses justifies the fact that NSCLC and melanoma are good candidates for sample stratification according CT-X gene expression. Therefore NSCLC and melanoma were selected for investigating the role of CT-X genes on tumor biology and differential drug sensitivity.

3.2 CT-X gene expression is associated with MTHFR 677 C>T polymorphism in NSCLC

Relationship between one-carbon pathway metabolic activity and CT-X gene expression was studied in NSCLC tumor samples, stratified for CT-X expression. Even though enzymes of the one-carbon pathway have different functions in metabolizing folate as one-carbon donors, it was thought that cumulative effect of variant alleles which cause reduction in enzyme activity (detailed in section 1.1.2), may lead to increased CT-X expression due to inefficient SAM availability in NSCLC tumors. To test this hypothesis, variant alleles was genotyped by TaqMan SNP genotyping and PCR-RFLP methods. Results of these genotyping analyses are discussed below.

Statistical evaluation of demographics and clinical characteristics of NSCLC patients was done using patient data given in Appendix B-Table 7-1. CT-X (-) tumors were associated with non-squamous histology and earlier tumor (T-1 and 2) stage, however, CT-X (+) tumors were associated with squamous histology and higher (T-3 and 4) stage tumors (Table 3-1). In other words, earlier stage tumors with non-squamous histology were statistically likely to be CT-X (-) and higher stage tumors with squamous histology were CT-X (+). This association results was in concordance with previously reported results [24,111].

Table 3-1: Demographics and Clinical Characteristics of NSCLC Patients.

		CT (+) Patients (n=21)	CT (-) Patients (n=29)	*P**
Age	>60	12	19	0.74
	<=60	6	7	
	Unknown[$]	3	3	
Sex	Male	10	12	0.76
	Female	8	14	
	Unknown	3	3	
Smoking history	No	1	5	0.37
	Yes	15	18	
	Unknown	5	6	
Histology	SQCC[#]	8	2	0.002
	non-SQCC	7	24	
	Unknown	6	3	
Pathological Tumor size	>3cm	10	8	0.21
	<=3cm	8	17	
	Unknown	3	4	
T stage	1	5	14	0.007
	2	8	12	
	3	3	0	
	4	2	0	
	Unknown	3	3	
TNM Stage (Pathologic stage of primary tumor)	I	9	18	0.37
	II	5	3	
	III	4	5	
	IV	0	0	
	Unknown	3	3	

*Chi-square (Fisher's exact test, two sided) or chi-square test for trend; # SQCC: squamous cell carcinoma (from Senses et al. BMC Med Genet, 2013 [143]).

Individual genotypic distributions of CT-X (+) and (-) tumors are shown in Table 3-2. The normo-active genotype of MTHFR, 677CC, was significantly associated with CT-X (+) tumors (P = 0.03) which was opposite to the initial hypothesis that CT-X

expression might be associated with hypo-active enzymes. There was no significant association between other polymorphisms and CT-X gene expression. Raw genotyping results are given in Appendix B-Table 7-2.

Table 3-2: Individual genotypic distribution of CT-X (+) and (-) tumors.

Polymorphism	***Genotype***	***CT-X (+) Tumors, n (%)***	***CT-X (-) Tumors, n (%)***	***χ^2***	***P****
MTHFR 677 C>T (rs1801133)	CC	13 (62%)	7 (24%)	7.30	0.03
	CT	5 (24%)	15 (52%)		
	TT	3 (14%)	7 (24%)		
MTHFR 1298 A>C (rs1801131)	AA	13 (62%)	14 (48%)	0.91	0.63
	AC	7 (33%)	13 (45%)		
	CC	1 (5%)	2 (7%)		
MTR 2756 A>G (rs1805087)	AA	14 (67%)	22 (76%)	0.51	0.53**
	AG	7 (33%)	7 (24%)		
	GG	0 (0%)	0 (0%)		
MTRR 66 A>G (rs1801394)	AA	4 (19%)	9 (31%)	0.92	0.63
	AG	10 (48%)	12 (41%)		
	GG	7 (33%)	8 (28%)		
RFC 80 G>A (rs1051266)	GG	6 (29%)	12 (41%)	1.90	0.39
	GA	9 (43%)	13 (45%)		
	AA	6 (29%)	4 (14%)		

* Chi-square (2 degrees of freedom); ** Fisher's exact test, two sided (from Senses et al. BMC Med Genet, 2013 [143]).

To understand whether MTHFR 677CC genotype is independently associated with CT-X expression, multivariate logistic regression analyses (MVA) which included MTHFR 677 genotype, histology, age, sex and T-stage parameters were performed. MVA revealed MTHFR 677 genotype and histology as significant predictors of CT-X gene expression (Table 3-3). However, significance of MTHFR 677 genotype was

higher compared to histology, therefore, MTHFR 677 genotype was an independent predictor of CT-X gene expression in this NSCLC cohort.

Table 3-3: Multivariate Analysis of MTHFR 677 genotypes with CT-X Gene Expression.

Parameter		***OR***	***95% CI***	***P****
MTHFR 677 C>T (rs1801133)	CC	32.33	2.42-431.52	0.003
	CT/TT	1**		
Age[$]		1.03	0.94-1.12	0.69
Sex (female vs. male)		0.28[‡]	0.02-3.86	0.52
Histology	SQCC[#]	18.46	1.19-284.49	0.04
	non-SQCC	1**		
T Stage	T1	0.31	0.03-3.56	0.53
	T2, 3, 4	1**		

* Computed from a logistic regression model using the EXACT option of PROC LOGISTIC in SAS to account for the small data set; $continuous variable. # SQCC: squamous cell carcinoma; ‡reference = male; **reference group (from Senses et al. BMC Med Genet, 2013 [143]).

When the association between CT-X gene expression and all the polymorphisms genotyped in this study analyzed by MVA on a per allele basis, controlling for age, sex, histology and T stage, only MTHFR 677 C>T polymorphism was significantly associated with CT-X gene expression (Table 3-4).

Table 3-4: Multivariate Logistic Regression Modeling of Association between CT-X Gene Expression and One-Carbon Enzyme Alleles

Polymorphism	***OR***[$]	***95% CI***	***P****
MTHFR 677 (rs1801133)	13.18	1.96-88.5	0.004
MTHFR 1298 (rs1801131)	0.56	0.15-2.09	0.53
MTR 2756 (rs1805087)	0.81	0.14-4.75	1
MTRR 66 (rs1801394)	0.67	0.22-2.00	0.52
RFC 80 (rs1051266)	0.7	0.24-2.07	0.62

[$] Based on number of major alleles, adjusted for age, sex, histology and T stage; * computed using the EXACT option of PROC LOGISTIC in SAS to account for the small data set (from Senses et al. BMC Med Genet, 2013 [143]).

Next, we aimed to validate our results on a different type of cancer. In silico data of an Acute Myeloid Leukemia (AML) cohort was used where both microarray gene expression and SNP genotyping data were available for matched AML samples (see section 2.2.9.2 for details) [112]. Out of 111 AML samples, clusters 1 and 2 were designated as CT-X (-) (n=45), cluster 3 were CT-X intermediate (n=33) and clusters 4 and 5 were CT-X high (n=31) (Figure 3.1). Two samples were excluded from the analyses as genotype data for those samples was not available. These analyses, however, did not reveal a significant association between MTHFR 677 C>T genotype and CT-X gene expression in this AML cohort (Table 3-5).

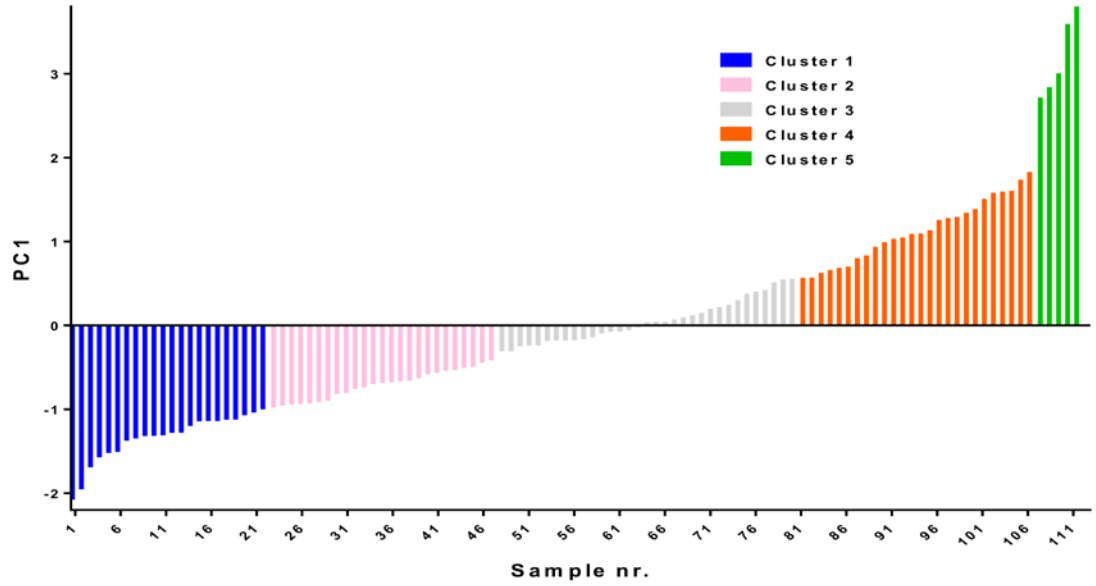


Figure 3.1: Principal component analysis based in silico clustering of AML. Tumor samples are shown ordered from the lowest to the highest first principal component (PC1) value. The 5 clusters generated by K-means clustering are indicated. Tumors with low, intermediate and high CT gene expression correspond to clusters 1-2, 3, and 4-5, respectively (from Senses et al. BMC Med Genet, 2013 [143]).

Table 3-5: Association Analyses of CT-X Gene Expression with MTHFR 677 Genotypes in Acute Myeloid Leukemia.

CT gene expression clusters	*MTHFR* 677 C>T (rs1801133)			*P* (chi-square)
	CC	CT/TT	Untyped*	
High	15	16	0	0.17
Intermediate	20	13	0	
Low	29	16	2	

* Untyped samples were not included in analysis (from Senses et al. BMC Med Genet, 2013 [143]).

3.3 DNA damage measurements in CT-X (+) and (-) cancer cell lines

Prior to genotyping NSCLC tumor samples for one-carbon pathway enzymes, CT-X expression was thought to be a result of DNA hypomethylation due to folate underutilization and insufficient SAM production which is in turn due to presence of hypo-active one-carbon pathway enzymes. However, association analyses have shown that CT-X (+) NSCLC tumors were associated with normo-active form of MTHFR (677 CC genotype of C allele). Since CT-X (+) tumors have several distinctive features compared to (-) tumors in various types of cancers such as, (i) difference in aggressiveness and prognosis (CT-X (+) tumors are more aggressive than (-) tumors) [111,113-119] and (ii) CT-X (+) tumors show genome-wide hypomethylation [22,24], it can be speculated that increased aggressiveness and increased DNA hypomethylation may lead to increased DNA damage and repair in CT-X (+) tumors. Therefore, as normal cells transform to malignant cells, cells with hypo-active MTHFR will not survive to the point of ultimate aggressiveness, therefore, will remain CT-X (-) and less aggressive, however, cells only with normo-active MTHFR will be able to cope with increased DNA damage, continue to grow and ultimately express CT-X genes.

3.3.1 Measurement of alkali labile DNA damage in CT-X (+) and (-) cancer cell lines

To test whether CT-X (+) cells have increased DNA damage compared to (-) cells, level of DNA damage was measured in CT-X (+) SK-LC-17 lung cancer cell line and CT-X (-) COLO-205 colon cancer cell line using comet assay (assay conditions are detailed in section 2.2.6.1). Cell lines were cultured in sub-physiological folic acid concentrations

(12 nM) to induce DNA damage as it known that FA deficiency causes DNA damage in mammalian cultures [61].

DNA damage was higher in SK-LC-17 cell line after 3 day incubation when compared to COLO-205 ($p<0.0001$), however, there was no significant difference after 6 or 9 day incubation between two cell lines (Figure 3.2). DNA damage was not significantly increased after 9 days of incubation in sub-physiological FA conditions in SK-LC-17 cell line, however, DNA damage in COLO-205 was significantly increased after 6 and 9 days of incubation when compared to day 3 measurements. Furthermore, there was no significant difference in tail DNA percentage when SK-LC-17 and COLO-205 were compared in day 6 and 9 measurements.

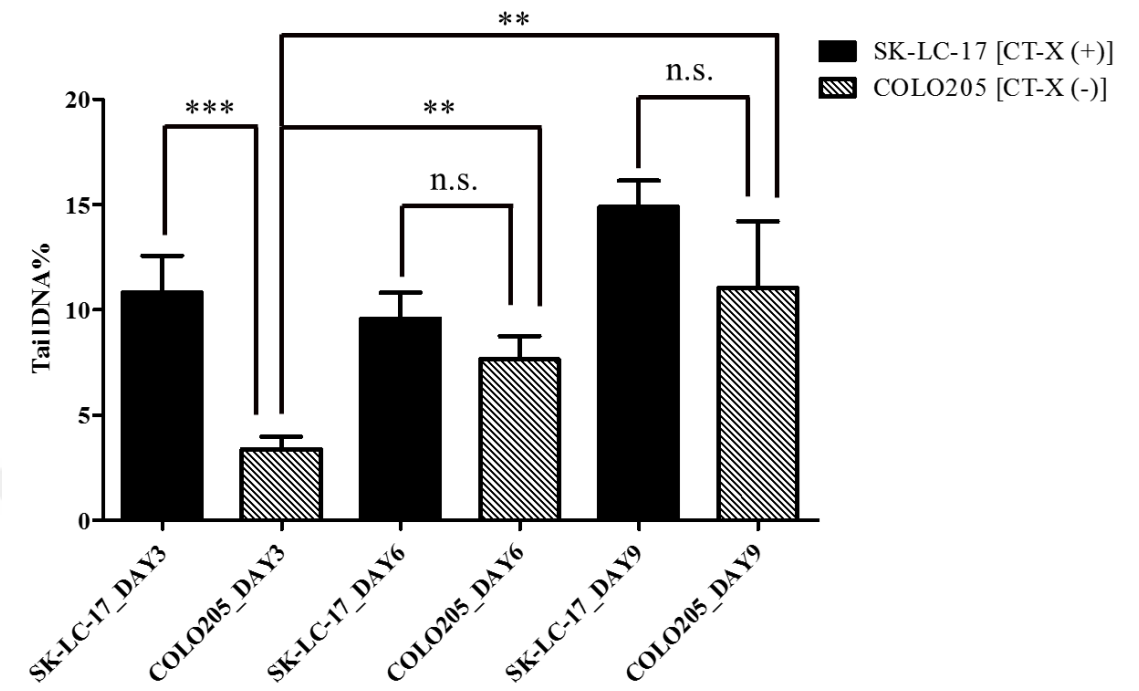


Figure 3.2: Level of DNA damage measured in sub-physiological FA conditions in CT-X (+) SK-LC-17 and CT-X (-) COLO205 cell lines using comet assay. DNA damage is shown as tail DNA percentage in 3, 6 and 9 day incubation periods. ***$p<0.0001$ (SK-LC-17-DAY3 vs. COLO205-DAY3), **$p<0.01$ (COLO205-DAY3 vs. COLO205-DAY6 and COLO205-DAY3 vs. COLO205-DAY9), n.s. (not significant, SK-LC-17-DAY6 vs. COLO205-DAY6, SK-LC-17-DAY9 vs. COLO205-DAY9), SK-LC-17-DAY3 vs. SK-LC-17-DAY6 and SK-LC-17-DAY3 vs. SK-LC-17-DAY9 was n.s. ($p = 0.55$ and 0.059, respectively).

3.3.2 Measurement of Histone H2A.X phosphorylation in CT-X (+) and (-) cancer cell lines

Since phosphorylated H2A.X (pH2A.X) is increased in the nucleus upon DNA damage which occurs in both strands [120-122], pH2A.X levels in CT-X (+) and (-) cell lines was measured indirectly by a flow cytometry-based assay (detailed in section 2.2.6.2).

Geometric mean (GM) of pH2A.X fluorescence intensity was higher in CT-X (+) SK-BR-3 and K-562 cell lines compared to CT-X (-) COLO-205 and SW-620 cell lines cultured in normal medium conditions as the measurement was done after 2 hours of recovery from UVC radiation (Figure 3.3). Folic acid concentration was influential on pH2A.X levels where cell lines cultured in FA-free conditions had higher pH2A.X compared to cell lines cultured in normal RPMI-1640 in matched UVC treatment groups (Figure 3.3) and this was in concordance with the literature [61]. Increase of pH2A.X in CT-X (-) cell lines was higher than CT-X (+) cells when cells were treated with UVC radiation in the presence of FA-free culture media compared to normal culture media conditions (Figure 3.4, 59% and 81% increase for COLO205 and SW620 compared to 13% and 39% of increase for K562 and SK-BR-3, respectively, Table 3-6). This result and the comet assay result from previous section reveals that CT-X (-) cells might be more susceptible to low folic acid concentration compared to CT-X (+) cells in terms of DNA damage and repair. Although the sample size is small, MTHFR 677 genotypes of cell lines selected for H2A.X measurement were C/T, T/T, C/C and C/C for COLO205, SW620, K562 and SK-BR-3, respectively. Therefore, one reason for increased susceptibility of CT-X (-) cells to low folic acid concentration might be due to

variant allele of MTHFR. There was no difference between CT-X (+) and (-) cell line groups in H2A.X phosphorylation measured directly (without UVC treatment) after 5 days of incubation either in FA-free or normal FA containing conditions (Figure 3.5).

These results, although inconclusively, demonstrates that CT-X gene expression might be a biomarker of increased DNA damage. Therefore, between CT-X (+) and (-) tumors, there might be differential sensitivity to certain chemotherapeutic drugs which effect DNA synthesis or cause DNA damage. Another conclusion is that CT-X (-) cells might be more susceptible to FA deficiency when these cells are evaluated by DNA damage and repair capacity.

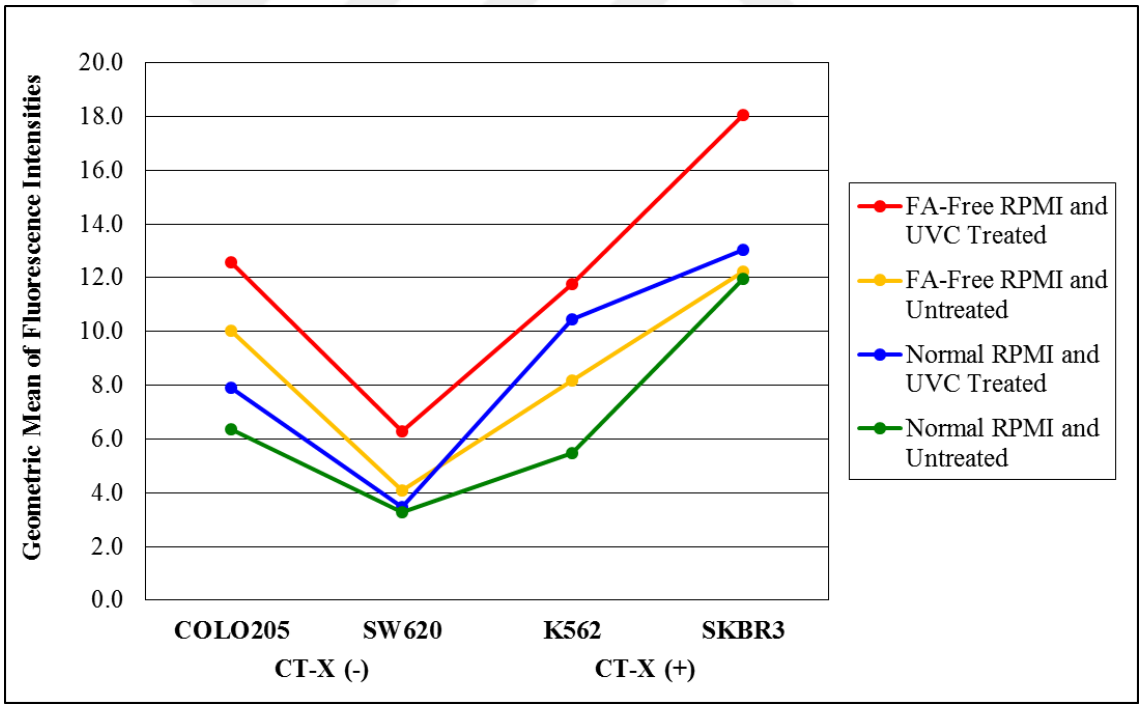


Figure 3.3: Geometric mean of fluorescence intensity of pH2AX in CT-X (-) COLO205 and SW620 and CT-X (+) K562 and SK-BR-3 cell lines.

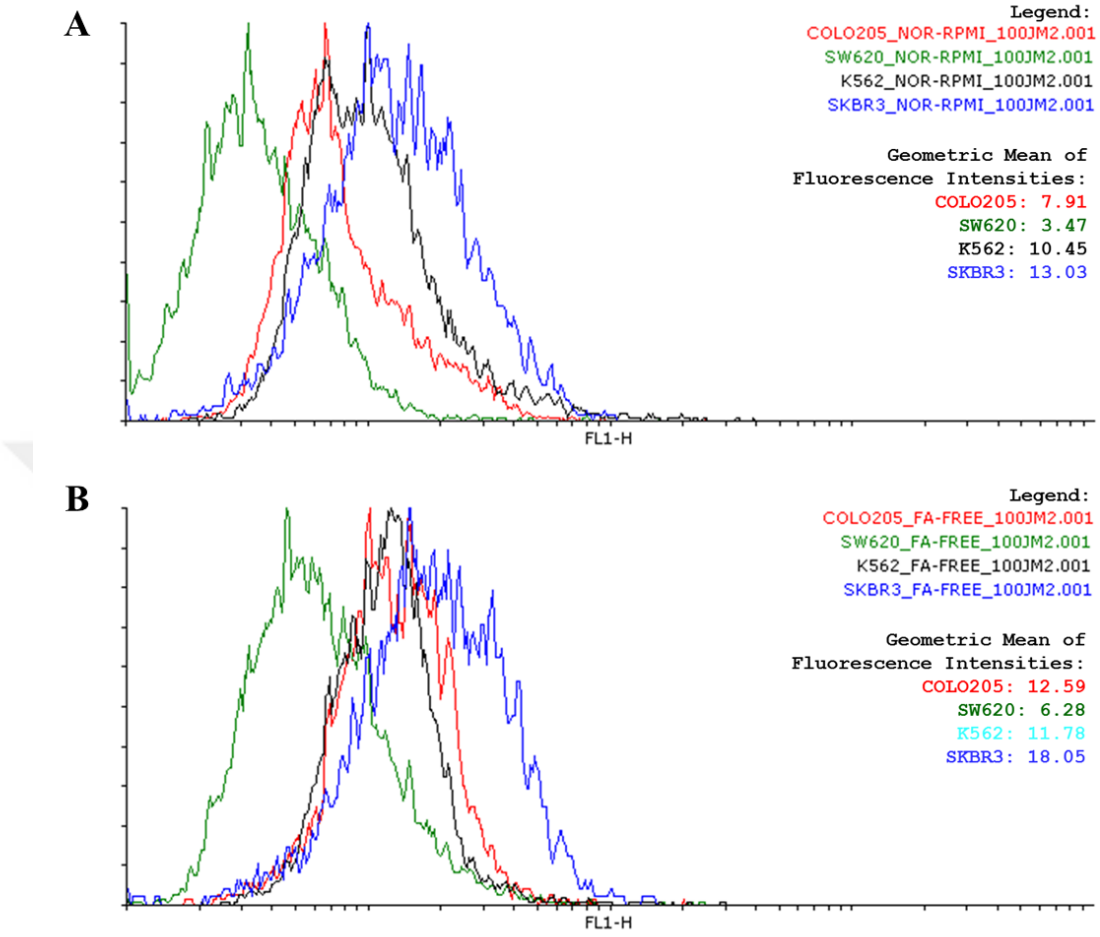


Figure 3.4: Level of H2AX phosphorylation in CT-X (+) K562 and SK-BR-3 and CT-X (-) COLO205 and SW620 cell lines cultured in normal (A) and FA-free (B) RPMI-1640 culture media with UVC radiation treatment. 59%, 81%, 13% and 39% of increase was calculated for, COLO205, SW620, K562 and SK-BR-3, respectively (Table 3-6).

Table 3-6: Increase of pH2A.X in CT-X (-) and (+) cells upon FA deficiency and UVC treatment.

#	Cell Line	*Treatment #1	$Treatment #2	% Difference (‡Tr. #2/Tr. #1)
1	**COLO205**	†7.9	†12.6	59.1
2	**SW620**	†3.5	†6.3	80.9
3	**K562**	†10.4	†11.8	12.8
4	**SKBR3**	†13.0	†18.0	38.5

*Normal RPMI and UVC Treated; $FA-Free RPMI and UVC Treated; †Geometric mean intensity of pH2A.X level measured by flow cytometry; ‡treatment.

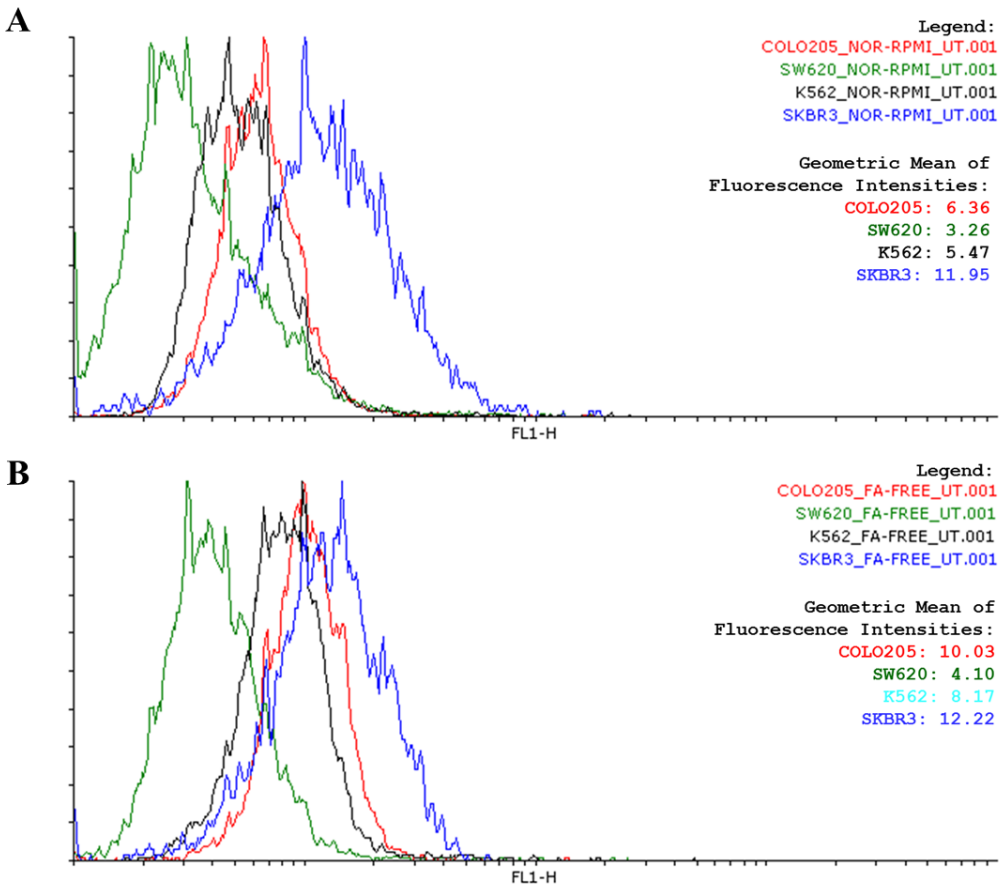


Figure 3.5: Level of H2AX phosphorylation in CT-X (+) K562 and SK-BR-3 and CT-X (-) COLO205 and SW620 cell lines cultured in normal (A) and FA-free (B) RPMI-1640 culture media without UVC treatment.

3.4 In silico drug screening analyses in CT-X (+) and (-) NCI-60 cancer cell lines

In order to find drugs which were differentially effective on CT-X (+) and (-) tumors, in silico data of one of the first known in silico drug screening data bases was used [123]. 53 cancer cell lines from the NCI-60 panel grouped according to CT-X gene expression as CT-X (+) and (-). Common drugs in test set and validation set which were significantly differentially effective on both CT-X (+) and (-) cell lines were determined (see section 2.2.8.1). Only two FDA-unapproved chemical compounds were determined with this approach (Figure 3.6).

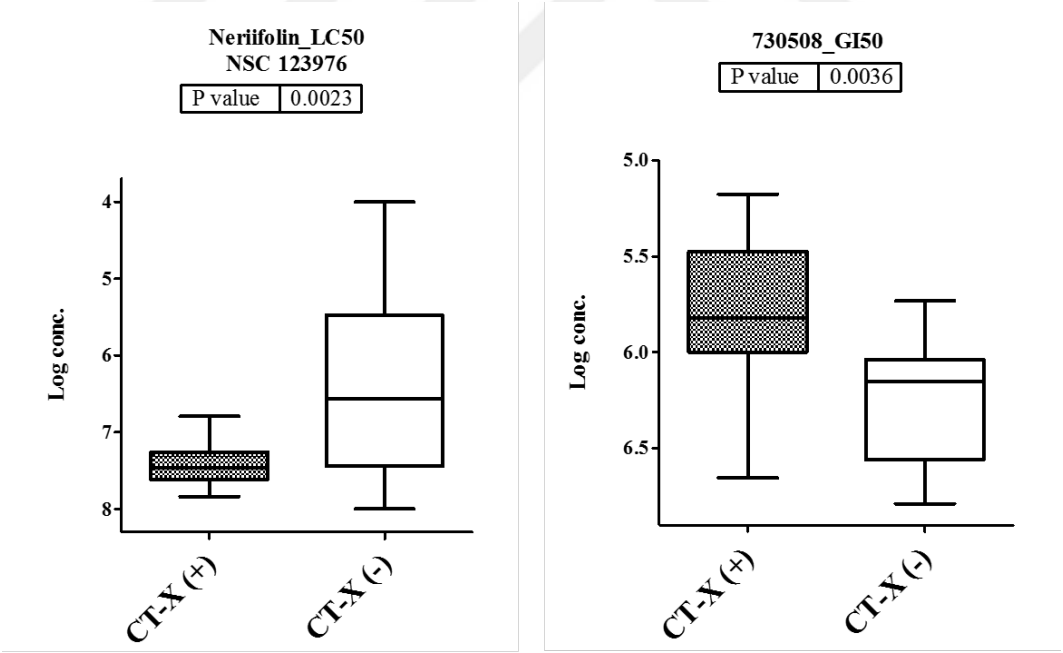


Figure 3.6: Significant drugs determined through analyses NCI-60 in silico data. Neriifolin was more effective on CT-X (+), whereas chemical compound with unique NSC id "730508" was more effective on CT-X (-) cancer cell lines.

In silico drug screening analyses results obtained by the NCI-60 data were limited and fragmented making hard to conclude on the type of chemotherapeutics or single molecule inhibitors which were effective on CT-X (+) and (-) cancer cell lines. Therefore, tools and methods used for in silico drug identification was changed and will be discussed in the upcoming chapter.

3.5 An approach to identify and validate effectiveness of drugs on different subgroups of malignant melanoma

In this section, in silico and in vitro experiments are described which were designed in the aim of identifying differentially effective drugs on different subgroups of melanoma cell lines, regardless of CT-X gene expression profiles. These experiments were performed on a unique primary melanoma cell line panel using the panel's Luminex1000 microarray transcriptome profiling data generated previously as part of the C-MAP project [67].

3.5.1 Identification of a melanoma gene expression signature

3.5.1.1 Heuristic algorithm discriminates melanoma cells from fibroblasts

Primary melanoma cell lines (PrMCLs) which were previously tested for melanoma markers S100, MART-1 and gp100 to differentiate them from non-melanoma cells [67], were further compared with a user-defined fibroblast and melanoma gene expression signature which was generated based on previously defined algorithm (see section 2.2.9.3) [85]. This algorithm was effective for separating melanoma from fibroblast cells as shown in Figure 3.7. There were only 4 melanoma cells out of 36 which have Mel (r) value lower than "0" and Fib (r) value higher than "0.5". All of the fibroblast cells compared with PrCMLs were negative for Mel (r) value and most of those cells' Fib (r) values were higher than "0.5". The 4 melanoma cell lines with high Fib (r) value were clearly melanoma cell lines as judged by morphology and expression for melanoma markers previously measured [67]. PrMCLs other from 36, which had no

expression of S100, MART-1 or gp100, and which lacked a melanoma-like morphology had Mel (r) value less than “0” and Fib (r) value higher than “0.6” were excluded from further studies (data not shown). Most of the samples of the two primary melanoma cell line datasets out of three were highly melanoma like according to our analyses, however for the third dataset (GSE4840) almost half of the samples were fibroblast-like (Figure 3.8). It was consistently observed that fibroblast-like cells from our cohort grow much slower, compared to those with high Mel (r) scores. This might explain the sparsity of commercially available established melanoma cell lines with high Fib (r) value since these cells are grown in culture for years (Figure 3.9).

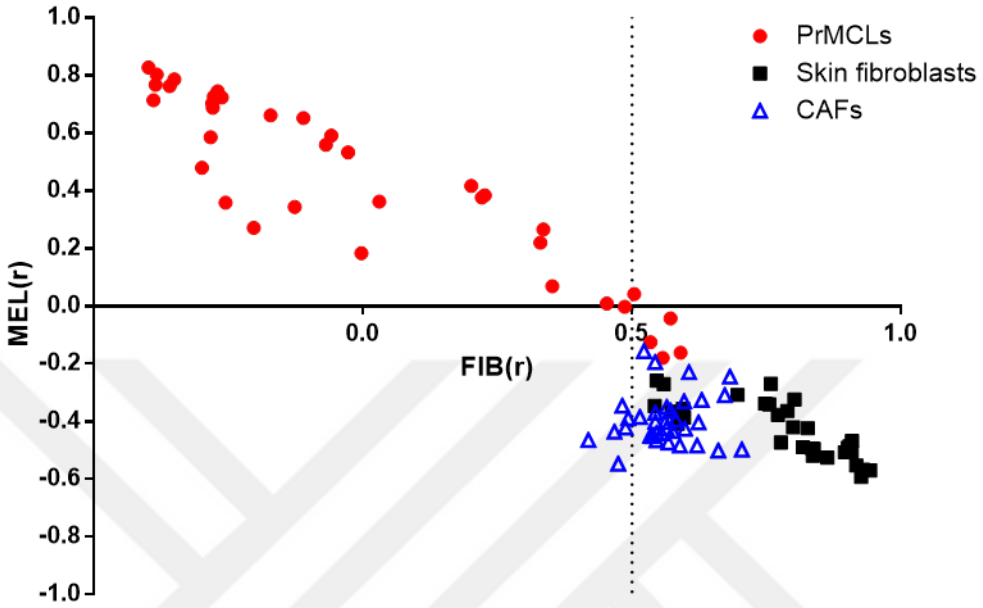


Figure 3.7: Correlation plot for fibroblast detection using gene expression data of PrMCLs, skin fibroblasts and cancer associated fibroblasts (CAFs).

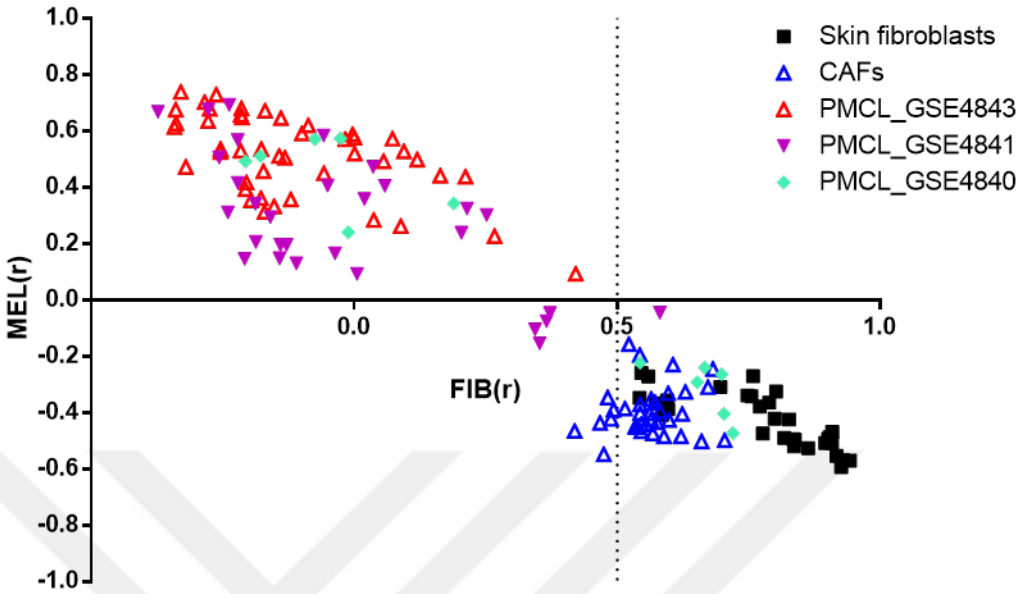


Figure 3.8: Correlation plot for fibroblast detection using gene expression data of skin fibroblasts, cancer associated fibroblasts (CAFs) and primary melanoma cell lines from different cohorts.

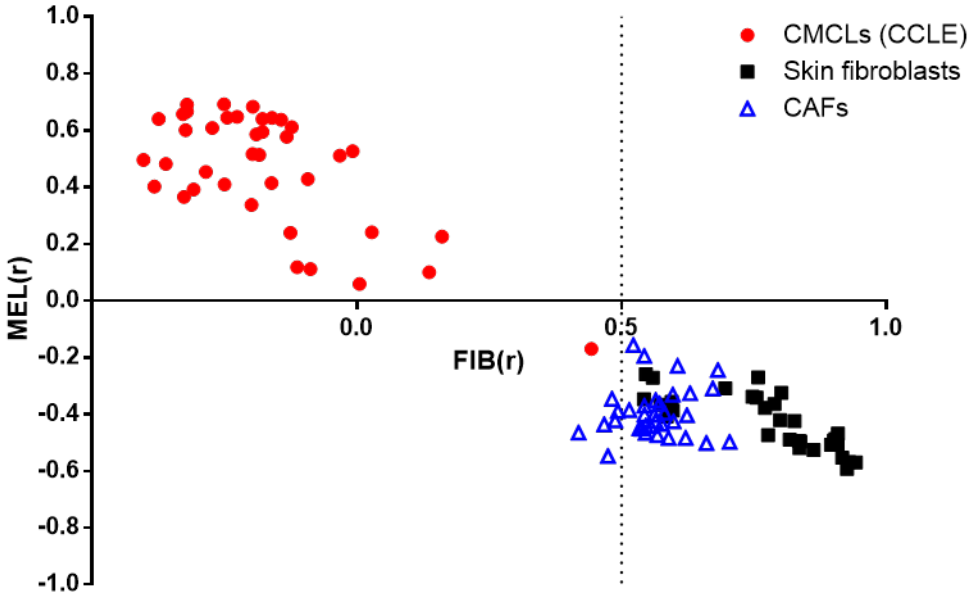


Figure 3.9: Correlation plot for fibroblast detection using gene expression data of skin fibroblasts, cancer associated fibroblasts (CAFs) and commercially available melanoma cells lines (CMCLs).

3.5.1.2 Molecular subgrouping of melanoma

To observe the heterogeneity of the unique PrMCL cohort, a gene list was created which harbors melanoma differentiation factors as well as factors which are not associated with melanoma biology. Gene expression data of CCLE [92] and CGP [124] databases which comprise high number of established cancer cell lines were used. In the first step of analyses, CCLE and CGP microarray gene expression data (GSE36133 and E-MTAB-783) were analyzed independently by comparing melanoma cell lines (group A) with cell lines from diverse cancer lineages such as lung, colorectal, breast, hepatocellular, pancreatic and gastric cancers (group B) (Figure 3.10). Probesets with Bonferroni corrected p-value less than 0.0001 and fold change value higher than 4 in either direction were considered as significant. In the second step, common significant probesets from both comparisons were selected. Established cancer cells lines was thought to provide an established signature of tissue specific gene expression, since these cell lines are grown in culture for years. Established expression signature was thought to facilitate finding best candidate genes which will discriminate melanoma-like from non-melanoma-like cells from our PrMCL cohort.

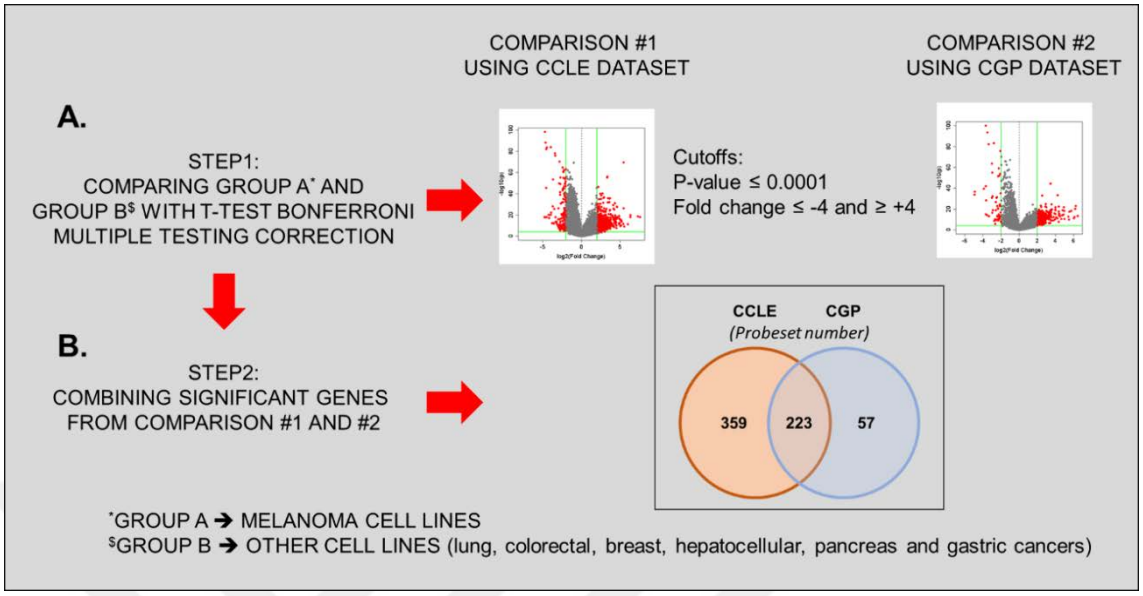


Figure 3.10: Schematic representation of the workflow for generating a gene list with discriminatory power of melanoma and non-melanoma cell lines. (A) Differentially expressed genes were determined in both CCLE and CGP datasets. (B) Common significant genes found in comparison #1 and #2 were determined. Number of common probesets were 223, matching to 171 different gene symbols. Of this 171 genes, 149 were upregulated and 22 were downregulated in CCLE and CGP melanoma cell lines compared with cell lines in group B.

The number of probesets fulfilling significance criteria were 582 (matching with 401 different gene symbols) and 280 (matching with 209 different gene symbols) in CCLE and CGP datasets, respectively. Among significant probesets, 223 were common in both datasets matching to 171 gene symbols. Gene list was composed of two types of genes where, 149 were upregulated and 22 were downregulated in established melanoma cell lines compared to other group of cell lines. The reason of biased distribution towards upregulated genes was probably the heterogeneous composition of group B cell lines; consensus in gene expression among different cancer lineages was low. However, melanoma group of cell lines were more homogenous in gene expression and this resulted in higher number of melanoma-specific upregulated gene expression.

The full list of probesets are given in Appendix C-Table 8-2. Intriguingly, in this probeset list, there were 3 different CT-X probesets belonging to MAGE-A3, MAGE-A6 and MAGE-A12 and a non-X CT gene probeset belonging to PRAME (Cancer/Testis Antigen 130) which were upregulated in melanoma cell lines compared to group B cell lines.

3.5.1.3 Clustering of PrMCLs using the gene list generated through CCLE and CGP database analyses

After defining the putative melanoma gene expression signature, PrMCLs were hierarchically clustered based on the gene list given in Table 8-2. Although Luminex1000 platform shares significant amount of Affymetrix HGU133A probeset annotation information, we retrieved 171 gene symbols' gene expression values instead of 223 unique probeset ID from Luminex1000 dataset. This approach was chosen to maximize the number of genes represented from the gene list. The number of genes represented from the list in Luminex1000 platform was 165 matching to 321 probesets. These probesets divided PrMCLs into two major clusters (cluster A and B, Figure 3.11). Samples in cluster A was characterized with the upregulation of genes associated with non-melanoma biology (genes downregulated in melanoma cell lines given in Table 8-2) and samples in cluster B were characterized with upregulation of genes associated with melanoma biology (genes upregulated in melanoma cell lines given in Table 8-2). Therefore, this PrMCL cohort might be classified as differentiated (melanoma-like) and de-differentiated (non-melanoma like).

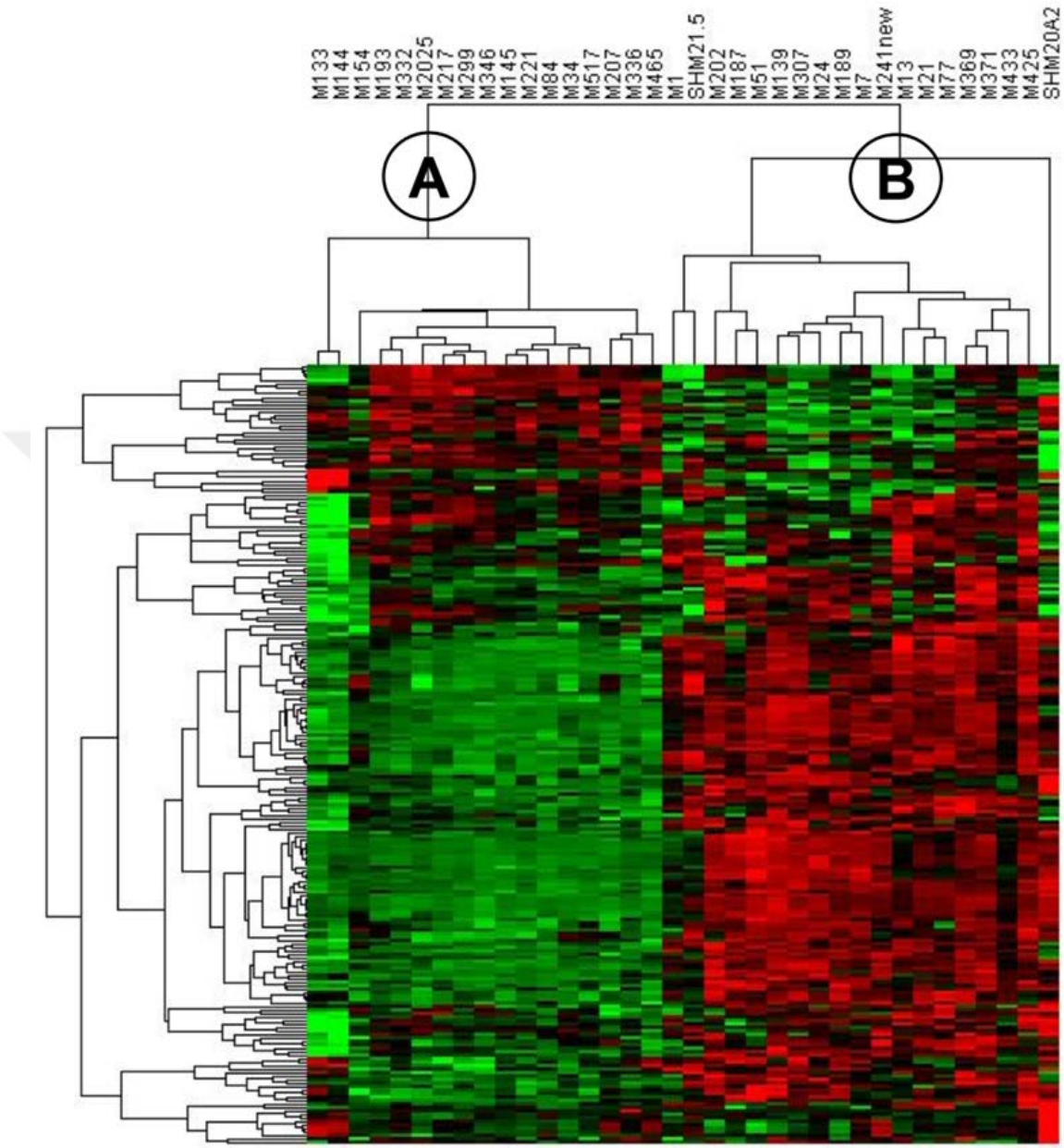


Figure 3.11: Hierarchical clustering analysis of PrMCLs. HCA based on 223 melanoma specific probesets using Luminex1000 data identifies two major groups (A and B).

3.5.1.4 Clustering PrMCLs using published alternative approaches and gene lists

Clustering results which were obtained with the gene list generated for molecular subgrouping of melanoma was compared with other published methods. One of these methods was Heuristic Online Phenotype Prediction (HOPP) which was generated for the aim of classifying melanoma cell lines and tumors into proliferative and invasive phenotypic groups [85]. Another one, developed through a class comparison analyses have resulted in generating a gene list with two groups of genes where one group was co-expressed with MITF and the other group was inversely correlated with MITF expression [86].

Using other published gene lists and methods, PrMCLs were clustered to make a comparison with the clustering result obtained in section 3.5.1.3. Even though gene lists and methods were different in other studies, PrMCLs samples were clustered almost identically in all three clustering methods. Only three of the samples (M1, SHM21.5 and M433) were clustered in different clusters.

This tool generates two different Pearson's correlation coefficient (r) values for a given melanoma sample by comparison of predefined gene expression values and the samples' own expression values of 97 different genes in a linear regression model. The correlation coefficient values are abbreviated as Pro(r) and Inv(r) for defining proliferative and invasive phenotypes, respectively. Samples with high Pro(r) value are designated as more proliferative, whereas samples with high Inv(r) value are designated as more invasive.

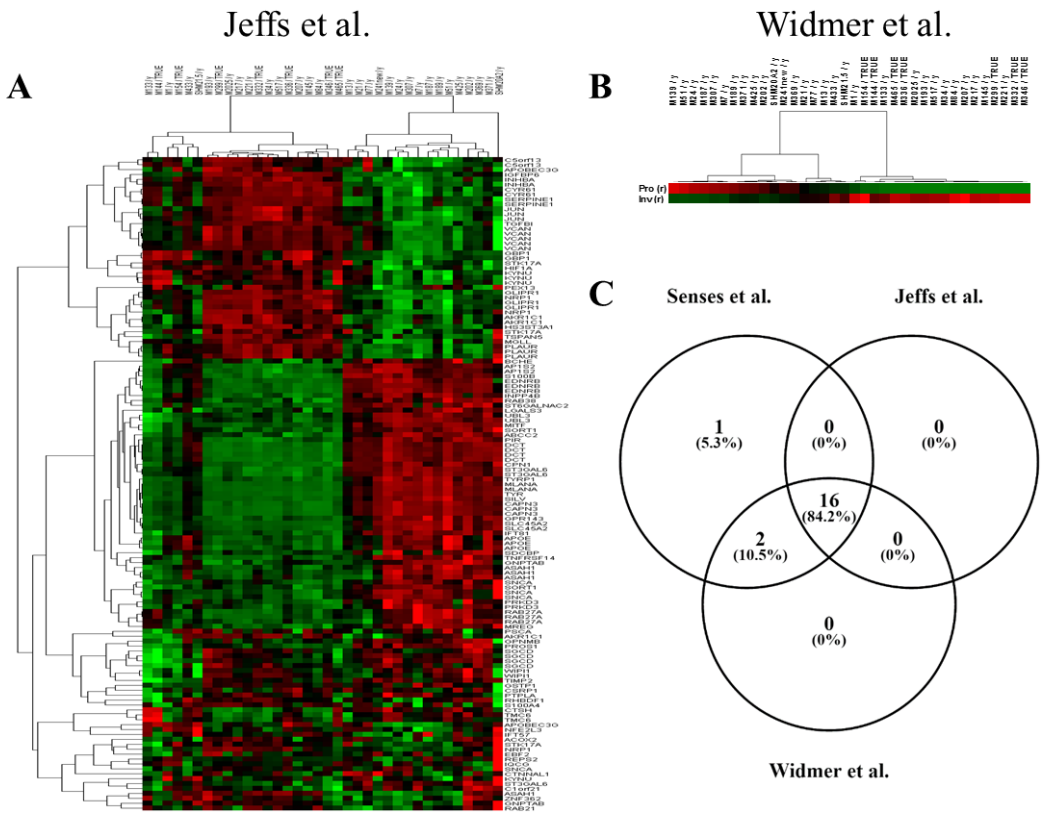
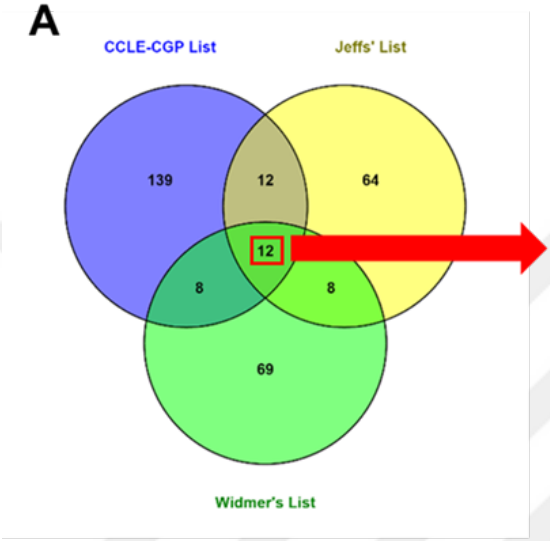


Figure 3.12: Comparison of CCLE and CGP derived gene list-based clustering with two different published clustering methods. (A) Primary melanoma cell lines were clustered using 96 gene signature predictive of invasive potential [86]. This clustering method generated two major clusters, similar to that in Figure 3.11. (B) Pro(r) and Inv(r) values of primary melanoma cell lines were calculated with HOPP tool (http://jurmo.ch/hopp/) [85] followed by clustering based on Pro(r) and Inv(r) values with Cluster 3.0 software. Similar to previous clustering results, this clustering method generated two major clusters. (C) Table was generated to show the overlapping samples in cluster B of Figure 3.11. 84.2% of samples were clustered in identical clusters.

12 genes were common in three of the gene lists used for clustering PrMCLs, all of which were melanoma-specific (Figure 3.13). Therefore, there was a lack of commonality between three clustering methods in distinguishing non-melanoma-like (or invasive) from melanoma-like samples.

Clustering PrMCLs with 12 common genes using Luminex1000 gene expression data generated 3 different clusters (Figure 3.13-B). 94% of the samples were in same cluster as in Figure 3.11.



#	Gene Symbol	Regulation in CCLE and CGP Melanoma Cell Lines
1	CAPN3	UP
2	GPR143	UP
3	INPP4B	UP
4	DCT	UP
5	MITF	UP
6	MLANA	UP
7	TYR	UP
8	TYRP1	UP
9	RAB38	UP
10	SLC45A2	UP
11	ST3GAL6	UP
12	TNFRSF14	UP

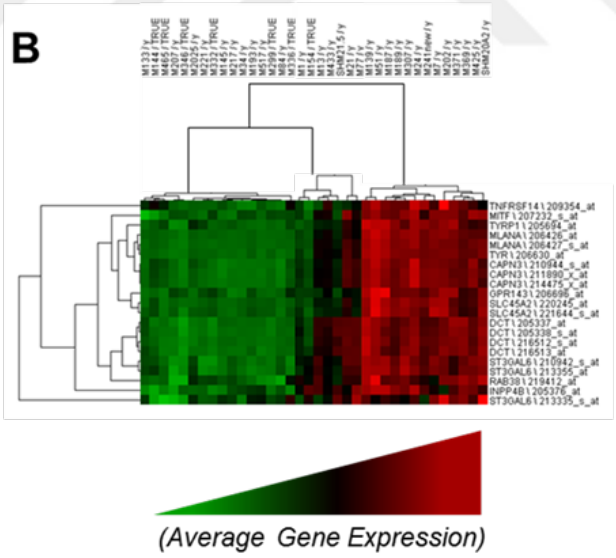


Figure 3.13: (A) Venn diagram shows distribution of numbers of genes in three gene lists (VENNY "Oliveros, J.C. (2007-15) An interactive tool for comparing lists with Venn Diagrams. (http://bioinfogp.cnb.csic.es/tools/venny/index.html)". 12 of the common gene symbols are shown. All of these gene were upregulated in CCLE and CGP melanoma cell lines (right panel). (B) Unique primary melanoma cell lines were clustered using 12 common genes. These genes generated 3 major clusters with average expression levels from low to high.

3.5.1.5 Clustering primary melanoma cell lines using GSEA

Gene Set Enrichment Analyses (GSEA) was conducted on Luminex1000 dataset to test gene signatures of two clusters in primary melanoma cell lines [125]. Clusters which were generated in Figure 3.11 were used as phenotype labels. Input phenotype labels for GSEA software are given in Table 3-7. Default parameters were used with the exception of selecting “gene_set” instead of “phenotype” from the permutation type menu. The most significantly enriched gene sets in 1st cluster samples were associated with formation of extracellular matrix and immune system (NES > 2.0 and FWER p-value ≤ 0.05) (Figure 3.14-A). The most significantly enriched gene sets in 2nd cluster samples, on the other hand, were associated with proliferation and pigmentation (Figure 3.14-B). Consistent with previous reports [78,79,86], GSEA revealed that our primary melanoma cell line panel majorly exists in two transcriptomically distinguished phenotypic states. Melanoma cells in cluster A which were enriched with extracellular matrix and immune related gene sets suggests a phenotypic state with a very active relationship with the surrounding environment, thus having high motility and being highly sensitive to extracellular stimuli resembling a mesenchymal-like phenotype. On the other hand, melanoma cells in the cluster B with enrichment of proliferation and pigmentation associated gene sets are transcriptomically more differentiated and in a more epithelial-like phenotypic state.

Table 3-7: Input phenotype labels for the GSEA and cluster information of primary cell lines.

#	Cell Line Name	GSEA Phenotype Label
1	M133 / y	Cluster A
2	M144 / TRUE	Cluster A
3	M207 / y	Cluster A
4	M84 / y	Cluster A
5	M465 / TRUE	Cluster A
6	M154 / TRUE	Cluster A
7	M2025 / y	Cluster A
8	M193 / y	Cluster A
9	M299 / TRUE	Cluster A
10	M145 / y	Cluster A
11	M217 / y	Cluster A
12	M346 / TRUE	Cluster A
13	M34 / y	Cluster A
14	M336 / TRUE	Cluster A
15	M332 / TRUE	Cluster A
16	M221 / y	Cluster A
17	M517 / y	Cluster A
18	M1 / y	Cluster B
19	SHM21.5 / y	Cluster B
20	M202 / y	Cluster B
21	M139 / y	Cluster B
22	M24 / y	Cluster B
23	M189 / y	Cluster B
24	M307 / y	Cluster B
25	M241new / y	Cluster B
26	M7 / y	Cluster B
27	M187 / y	Cluster B
28	M51 / y	Cluster B
29	M13 / y	Cluster B
30	M21 / y	Cluster B
31	M77 / y	Cluster B
32	M369 / y	Cluster B
33	M371 / y	Cluster B
34	M433 / y	Cluster B
35	M425 / y	Cluster B
36	SHM20A2 / y	Cluster B

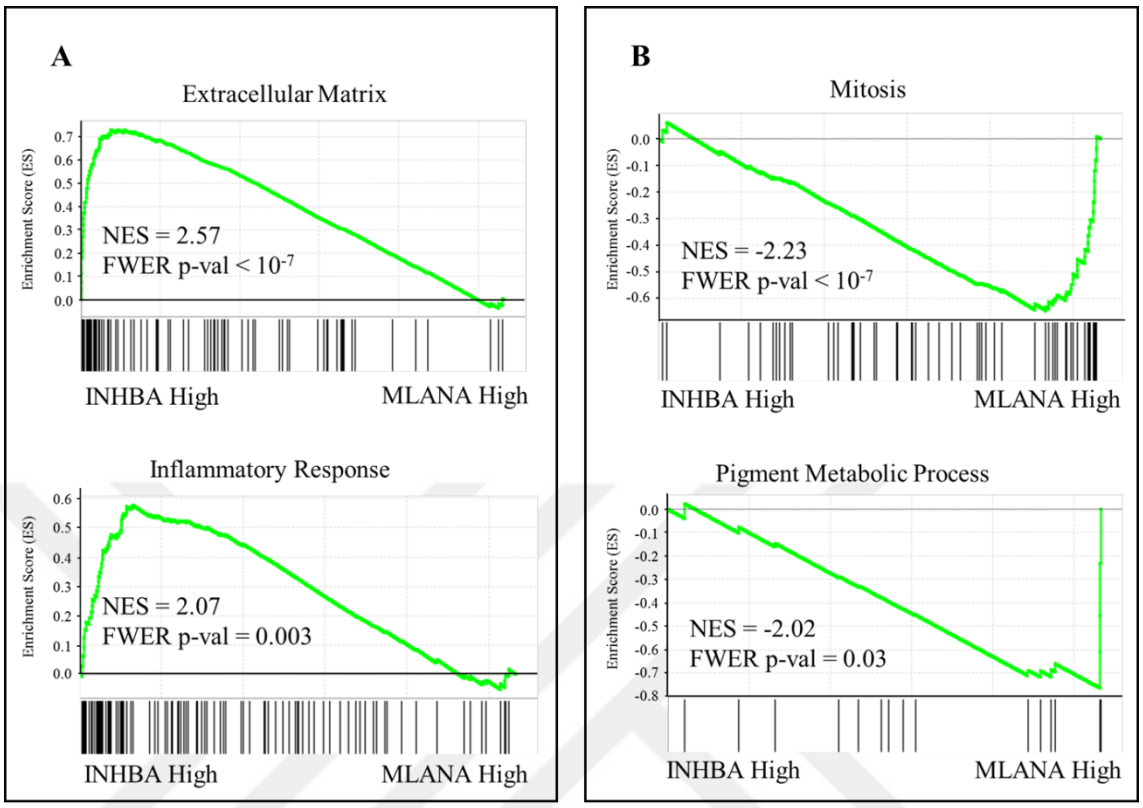


Figure 3.14: GSEA plots of PrMCLs from cluster A and B. Two of the significantly enriched gene sets are shown. Green line represents the enrichment profile, vertical black lines shows the rank of individual genes in a given gene set. (NES: normalized enrichment score, FWER: family-wise error rate)

3.5.1.6 Classification of PrCMLs based on the combination of EMT, CSC and PRO-INT-INV scores

Each PrMCL, as well as CMCLs were evaluated for EMT, CSC and PRO-INT-INV signatures *in silico*, to determine similarities between these signatures in clustering melanoma, and to understand whether if one or a combination of these signatures could be used to better explain sensitivity to a molecularly targeted drug. To generate these signatures, cells were first classified into proliferative (PRO), invasive (INV) or

intermediate (INT) categories using Pearson's correlation coefficients generated by an online tool (HOPP URL: http://jurmo.ch/hopp/) [85]. EMT and CSC scores were then assigned on each cell (detailed in 2.2.9.4) and plotted together with PRO-INT-INV classification information. Only one of the CMCL, and 3 PrCMLs were both classified as epithelial and CSC-like (Q1), as these two classification measures are expected to show mutual exclusivity (Figure 3.15**)**. Similarly, most INV cells had higher mesenchymal score, whereas PRO cells had high epithelial score. However, while the INT cells among CMCLs were mostly CSC-like and mesenchymal-like (Q3, Figure 3.15, right) the INT group among primary melanoma cell lines were more non-CSC-like and epithelial-like features (Q4, Figure 3.15, left).

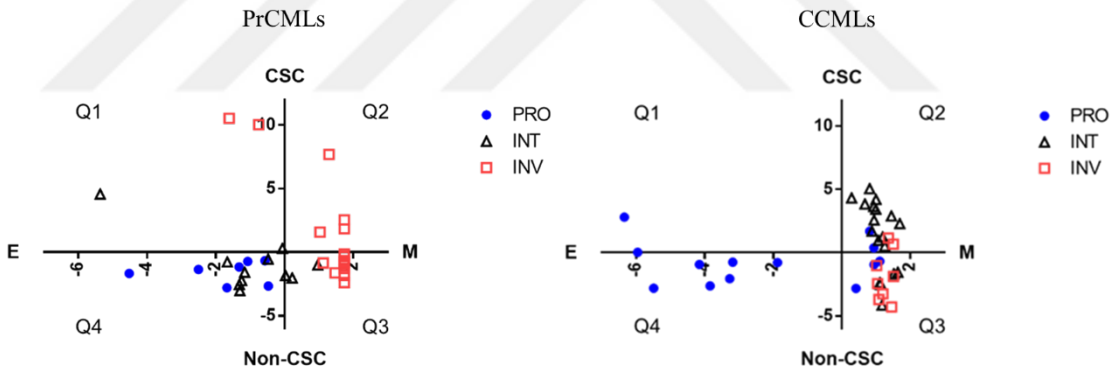


Figure 3.15: Clustering melanoma cell lines based on three molecular signatures. Distribution of CCMLs categorized into PRO, INV or INT groups according to an EMT score (X axis) and CSC score (Y axis). E: epithelial, M: mesenchymal phenotype. CSC: cancer stem-cell-like phenotype, Non-CSC: non-cancer-stem-cell-like phenotype. Left: PrMCLs, right: CCMLs.

3.5.2 Validation of transcriptomic subgroups through in vitro assays

In silico analyses using Luminex1000 gene expression data of primary melanoma cell line panel revealed two major phenotypes which are transcriptomically distinguishable by diverse tools and methods (e.g., gene lists and GSEA). Whether this transcriptomic status leads to phenotypic change in cellular behavior, such as, change in motility or change in rate of proliferation was yet to be validated by in vitro experiments. Experiments which were performed to measure gene expression, proliferation and motility are detailed in following sections.

3.5.2.1 Measurement of INHBA and MLANA by quantitative RT-PCR

First step of validation experiments was the validation of Luminex1000 gene expression data by measuring INHBA and MLANA gene expression. Total RNA samples extracted from 17 of the primary cell lines was used in q-RT-PCR experiments. Candidate genes for the validation of the clusters A an B (from Figure 3.11) were selected as INHBA and MLANA, respectively. There were three reasons for selecting INHBA and MLANA as candidate genes; (i) they were among the most significant genes separating two clusters (Figure 3.16) and (ii) MLANA is a skin specific gene which will easily distinguish melanoma-like from non-melanoma-like cell lines, iii) MLANA was among the common genes with other published gene lists (see section 3.5.1.4).

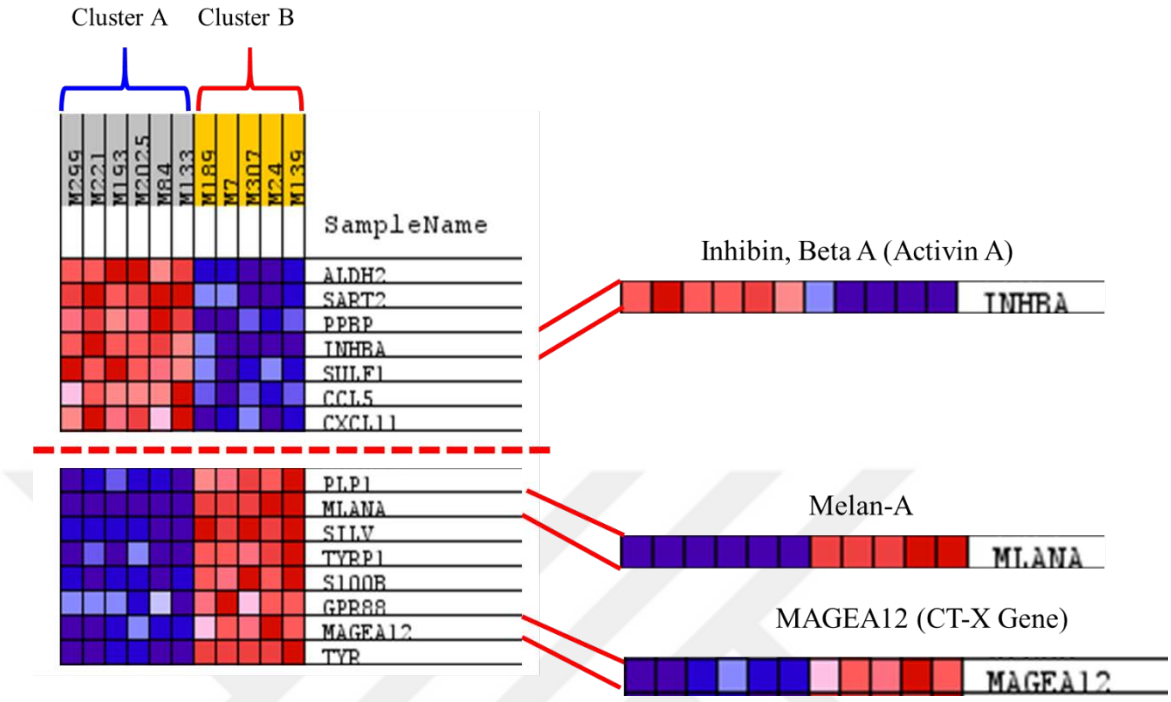


Figure 3.16: Heatmap generated by GSEA showing microarray gene expression level of INHBA and MLANA. High and low expression is shown by red and blue color, respectively. One of the CT-X genes (MAGEA12) is shown to be upregulated in cluster B PrCMLs.

Gene expression levels of INHBA and MLANA measured with q-RT-PCR using 17 of the primary melanoma cell lines were compared with Luminex1000 data (Figure 3.17). INHBA expression was consistent between PCR and Luminex1000 data in 15 out of 17 cell lines (Figure 3.17, left). This number was 13 for MLANA expression (Figure 3.17, right). Inconsistent cell lines for both INHBA and MLANA expression were M221 and M133, where completely opposite pattern of expression was measured in q-RT-PCR and Luminex1000 microarray platforms for INHBA and MLANA. Cell lines M77 and M21 were inconsistent for only MLANA expression when two platforms were compared. Inconsistency of gene expression have changed the clustering information of inconsistent cell lines, where M221 and M133 switched to 2nd cluster and M77 and M21

switched to 1st cluster (original cluster information generated with Luminex1000 data are given in Table 3-7).

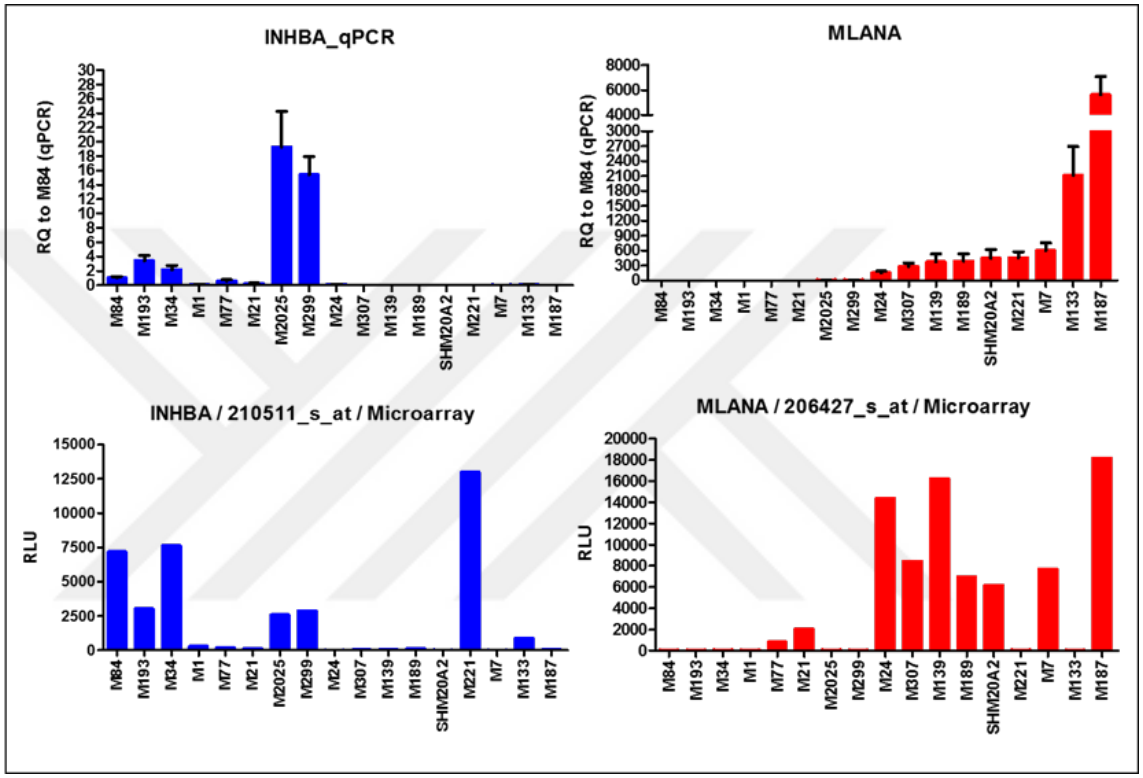


Figure 3.17: Comparison of q-RT-PCR and Luminex1000 data for INHBA and MLANA gene expression. Relative quantity (RQ) levels were measured by DDCt method [126]. M84 and GAPDH were reference sample and gene for the analyses, respectively. Percentage of consistent samples between q-RT-PCR and Luminex1000 expression was (A) 88% (15/17) for INHBA and (B) 76% (13/17) for MLANA. Inconsistent samples for INHBA were M221 and M133 and for MLANA were M221, M133, M77 and M21.

In a second round of q-RT-PCR analyses, INHBA and MLANA was measured using RNA samples from PrCMLs as well as CCMLs which were deposited on Bilkent University, Department of Molecular Biology and Genetics. These results were in concordance with first round measurement results.

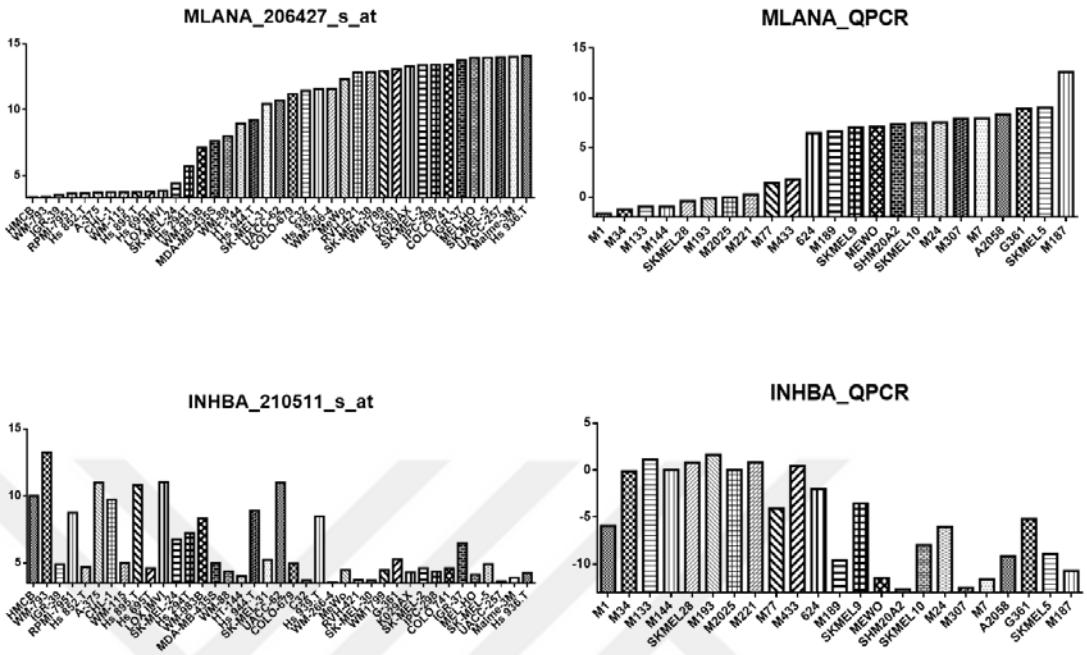


Figure 3.18: Melan-A (MLANA) and Inhibin beta-A (INHBA) expression in melanoma cell lines. In silico expression of MLANA and INHBA in CMCLs (using dataset GSE36133, left panel); and q-RT-PCR based MLANA and INHBA expression of PrMCLs and CMCLs (right panel). Expression of these two genes are inversely correlated. Pearson's r and p values for left panel (in silico data of CMCLs) and right panel (q-RT-PCR data of PrMCLs and CMCLs) are -0.6 ($p<0.0001$) and -0.79 ($p<0.0001$), respectively.

3.5.2.2 Measurement of proliferation in primary melanoma cell lines

Proliferation experiments were performed using 13 primary melanoma cell lines by culturing up to 96 hours and measuring growth rates at 24th, 48th and 96th hours, using CyQuant Assay Kit (see section 2.2.7.2). Proliferation ratios were calculated by dividing the 48th and 96th hour raw fluorescence unit (RLU) values to 24 hour RLU values. This calculation yielded two time point growth ratios (for 48 and 96 hour measurements). Growth rate differences between 48th and 96th hour measurements were compared across cluster 1 and 2 cell lines. Cluster information was based on INHBA

and MLANA gene expression values measured by q-RT-PCR (section 3.5.2.1 and Figure 3.17). Three different comparisons were made to show the difference of proliferation rates between cluster A and B; (i) comparison of average growth rates between 48th and 96th hour measurements, (ii) comparison of average slope of graphs shown in Figure 3.19 and (iii) comparison of doubling times calculated by exponential growth equation. There were no significant results obtained in these comparisons, however, cell lines M307 and M7 which were in cluster B, had the highest proliferation rates among other cell lines tested.

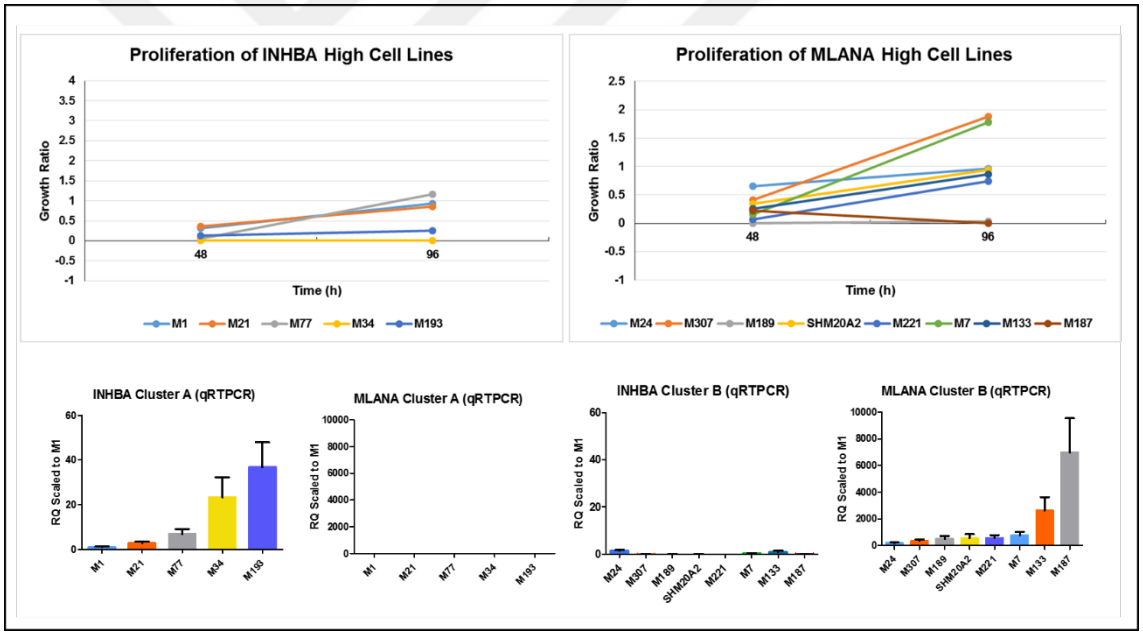


Figure 3.19: Proliferation measurements of primary melanoma cell lines. (A) Cell lines from cluster 1 and (B) cluster 2 were cultured for 96 hours in 5% CO2 supplied 37^{0}C incubator. Growth rates were measured as a function of DNA content using CyQuant assay kit. INHBA and MLANA expression values were calculated by using cell line M1 as reference and GAPDH gene as endogenous control. Apparent from the bar graphs, cell lines in 1st cluster had high INHBA and low MLANA, cell lines in cluster B had low INHBA and high MLANA according to q-RT-PCR measurements.

3.5.2.3 Migration capacity of PrMCLs compared to their gene expression profiles

In vitro migration of melanoma cell lines from the 1st cluster (M1, M77 and M193) were compared with cell lines from the 2nd cluster (M7, M24, M133, M187 and M221) using in vitro scratch assay (described in section 2.2.2.4). Mean remaining scratch area percentage of cluster A and B cell lines were compared using three different images for all cell lines and all time intervals (except M187 in which two images were captured due to technical difficulties) (Figure 3.20-A, for sample images). Although cell lines in cluster A were slightly faster (Figure 3.20-B), there was no significant difference between two clusters at 0.05 significance level.

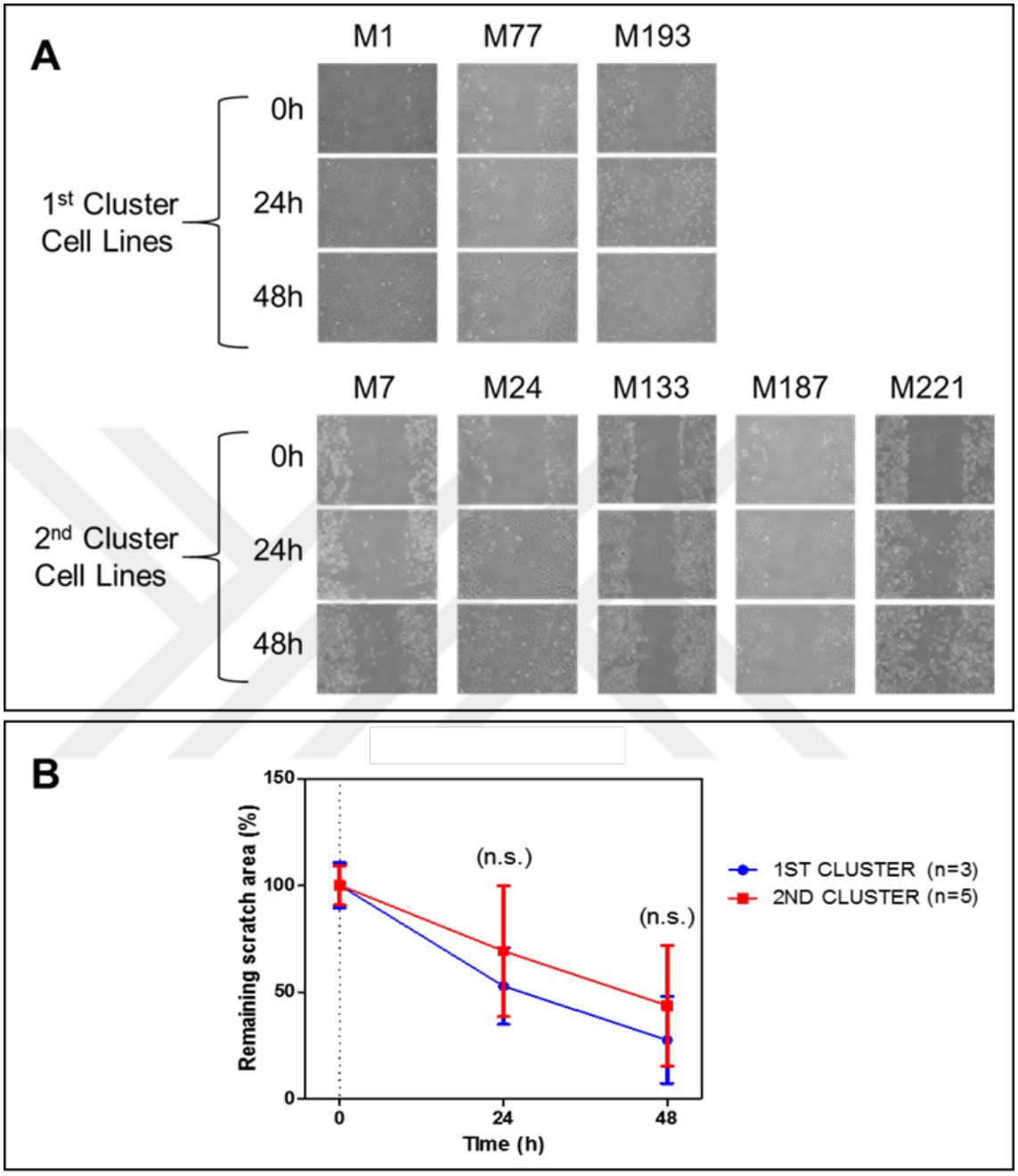


Figure 3.20: In vitro scratch assay of melanoma cell lines. (A) Images of 3 cell lines from 1st cluster (cluster A) and 5 cell lines from 2nd cluster (cluster B) were captured with 100X magnification up to 48 hours in 24 hour time intervals. (B) Three different images were captured for all cell lines for all time intervals (except M187; two images were captured for this cell line) for the comparison of mean 2D motility of 1st and 2nd cluster cell lines. 1st cluster cell lines were slightly faster in motility, however, the difference between two clusters was not significant for 24 and 48 hour intervals.

3.5.3 In silico drug identification using CCLE drug screening database

To identify effective drugs on different molecular subgroups of melanoma, correlation between INV (r) and PRO (r) values and drug sensitivity data (Activity Area) were obtained for melanoma cell lines which have cytotoxicity data in the CCLE database (see, 2.2.8.2) [92]. There was a strong positive correlation between cytotoxicity data of two MEK inhibitors, AZD6244 and PD0325901 and PRO (r) values. Furthermore, there was a simultaneous negative correlation between INV (r) values and cytotoxicity data of AZD6244 and PD0325901. Three other drugs, the HSP90 inhibitor 17-AAG, the SRC-inhibitor AZD0530, and a c-MET inhibitor PF-2341066 have shown the exact opposite pattern (Table 3-8). Because it has been previously shown that HSP90 and MEK inhibitors affect opposite phenotypes [127], and similar observations exist for SRC and MEK inhibitors [77], we decided to perform further analyses on 17AAG, AZD6244 and AZD0530. First, combination of EMT, CSC and "PRO/INV/INT" classifications were tested if drug sensitivity was predicted better than "PRO/INV/INT" classification alone. Cell classified ad "PRO" were more sensitive to AZD6244 which was in concordance with previously published data [128]. However, INV cells were more sensitive to 17AAG and AZD0530. We further classified CCMLs into quadrants based on Figure 3.15 (based on EMT, CSC and PRO-INT-INV classification detailed in section 3.5.1.6). We failed to identify a molecular subgroup that was more sensitive to any of the three drugs with in silico cytotoxicity data when compared to PRO-INT-INV classification (Figure 3.21). This indicates that EMT, CSC and PRO-INT-INV combined classification is able to identify distinct cell populations that have different chemosensitivity profiles but does not lead to identify cells that have increased

sensitivity to a given drug, compared to the classification generated by the PRO-INT-INV score alone.

Table 3-8: Linear correlation measures of drugs with PRO and INV values of CCLs

DRUG	TARGET	PROr - AA(norm) - Pearson's r	INVr - AA(norm) - Pearson's r	PROr-RANK	INVr - RANK	RANK-SUM
AZD0530	SRC	-0.488	0.401	1	1	2
PF-2341066	c-MET	-0.390	0.307	2	3	5
17-AAG	HSP90	-0.274	0.327	4	2	6
TAE684	ALK	-0.335	0.259	3	4	7
LBW242	XIAP	-0.130	0.200	7	5	12
Topotecan	TOP1	-0.203	0.092	6	7	13
Lapatinib	EGFR	-0.130	0.099	8	6	14
Irinotecan	TOP1	-0.258	0.069	5	9	14
Nilotinib	ABL	-0.093	0.078	11	8	19
ZD-6474	EGFR	-0.111	0.048	10	12	22
Paclitaxel	TUBB1	-0.125	0.044	9	13	22
PHA-665752	c-MET	-0.044	0.063	13	10	23
PD-0332991	CDK4	0.012	0.055	15	11	26
TKI258	FGFR	-0.082	0.017	12	14	26
Erlotinib	EGFR	-0.044	-0.052	14	17	31
L-685458	GS	0.037	0.011	17	15	32
AEW541	IGF1R	0.027	-0.105	16	18	34
Nutlin-3	MDM2	0.135	0.010	19	16	35
RAF265	RAF	0.210	-0.178	20	19	39
Panobinostat	HDAC	0.094	-0.288	18	21	39
PLX4720	RAF	0.303	-0.256	22	20	42
Sorafenib	RTK	0.268	-0.394	21	22	43
PD0325901	MEK	0.535	-0.429	24	23	47
AZD6244	MEK	0.526	-0.445	23	24	47

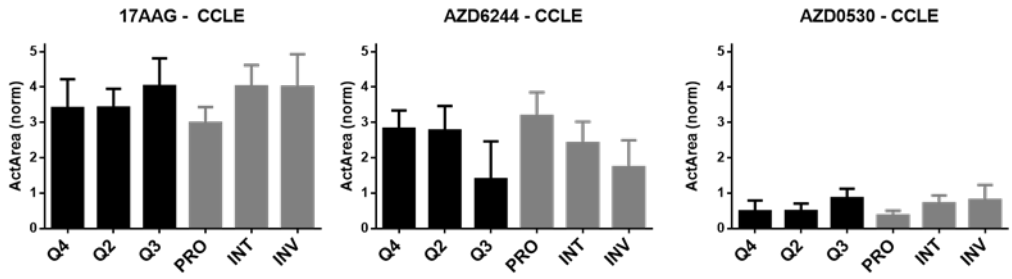


Figure 3.21: In silico cytotoxicity analyses based on EMT, CSC and PRO-INT-INV on CMCLs. Clustering information based on quadrants in Figure 3.15 as well as according to the PRO-INT-INV classification alone were analyzed for sensitivity against 17AAG, AZD6244 and AZD0530. Cytotoxicity is shown as normalize activity area (Y axis).

Next, we aimed to validate the *in silico* cytotoxicity correlations by *in vitro* screening experiments using PrMCLs and their corresponding in silico data. Different from CMCLs, PRO, INT and INV groups of PrMCLs were not separated clearly into quadrants (Figure 3.22). Most PRO and INT cells were scattered in Q4, while INV cells were scattered in quadrants 1, 2 and 3, which is a different from CMCL classification. However, cytotoxicity profiles of PrMCLs, were quite different between quadrant-based and PRO-INT-INV based classification (Figure 3.22). INV cells were most sensitive to 17-AAG, but INT cells were generally more resistant than PRO cells. For the MEK inhibitor AZD6244, PRO and cells classified as Q4 were relatively sensitive. Sensitivity profile of PrCMLs to AZD0530 was similar to the *in silico* data of CCMLs, in which it did not show much variation among groups. A novel SRC inhibitor 10a was also tested, which has a specific SRC inhibitor activity [129]. Intriguingly, 10a was cytotoxic to completely opposite clusters of cells when compared to AZD0530, the other SRC inhibitor tested. AZD0530 is a BCR-ABL/SRC inhibitor, however, 10a is shown to be

relatively specific for SRC [129]. These in vitro profiling analyses reveals that cytotoxic profile of 10a is almost identical to that of AZD6244. As SRC and MAPK functions in similar pathways, PRO cells might have more an active SRC and MAPK proteins compared to INV cells. Although AZD0530 and 10a are opposite in terms of cytotoxicity, AZD0530 is weakly cytotoxic to both CMCLs and PrMCLs (Figure 3.21 and Figure 3.22) suggesting that AZD0530 might also be active on a non-SRC-kinase.

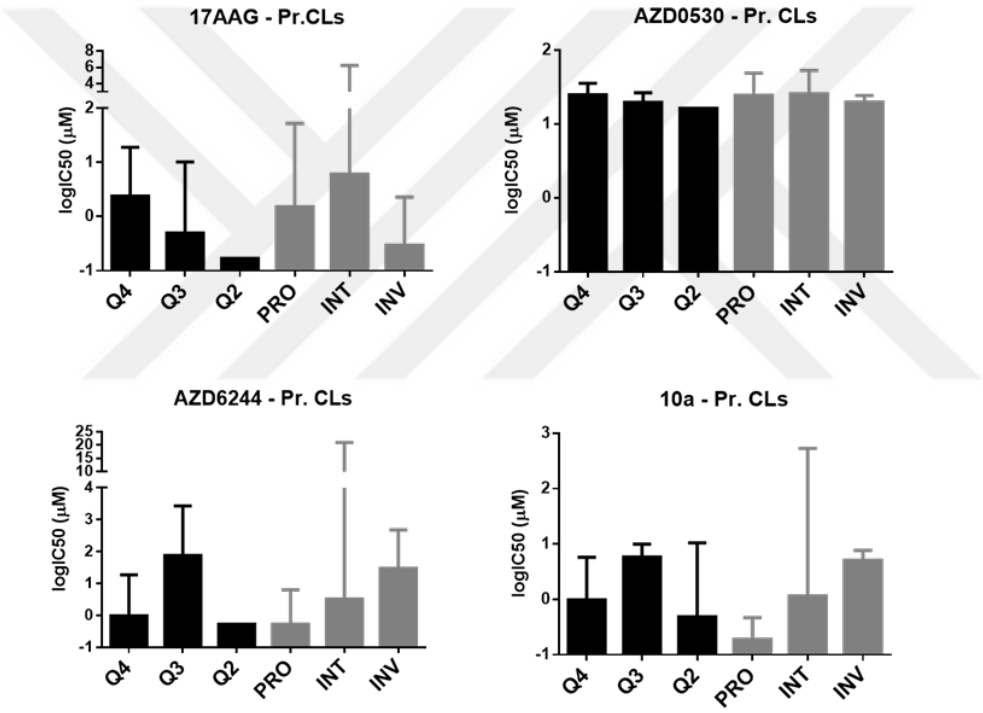


Figure 3.22: In vitro cytotoxicity analyses based on EMT, CSC and PRO-INT-INV on PrCMLs. Clustering information based on quadrants in Figure 3.15 as well as according to the PRO-INT-INV classification alone were analyzed for sensitivity against 17AAG, AZD6244 and AZD0530. Cytotoxicity is shown as logIC50 (µm, Y axis).

3.5.4 MLANA and INHBA as biomarkers of drug sensitivity in melanoma

Gene expression based classifications of melanoma which are compared with drug sensitivity profiles so far were based on a large number of genes. However, as the EMT, CSC and PRO-INT-INV based classifications generated clusters with similar drug sensitivity profiles, it was argued that gene expression profiles generated with a much smaller number of genes might be as effective as the ones generated with large number of genes. As genes which classify PrCMLs into clusters A and B (Figure 3.11) are related to differentiation markers of melanoma, we aimed to identify genes whose differential expression could separate melanoma-like from non-melanoma-like PrCMLs simply by comparing gene expression profiles of cluster A and B. These class comparison analyses (detailed in section 3.5.1.2 and 2.2.9.1) generated a gene list which classified PrCMLs into two distinct clusters (Figure 3.23).

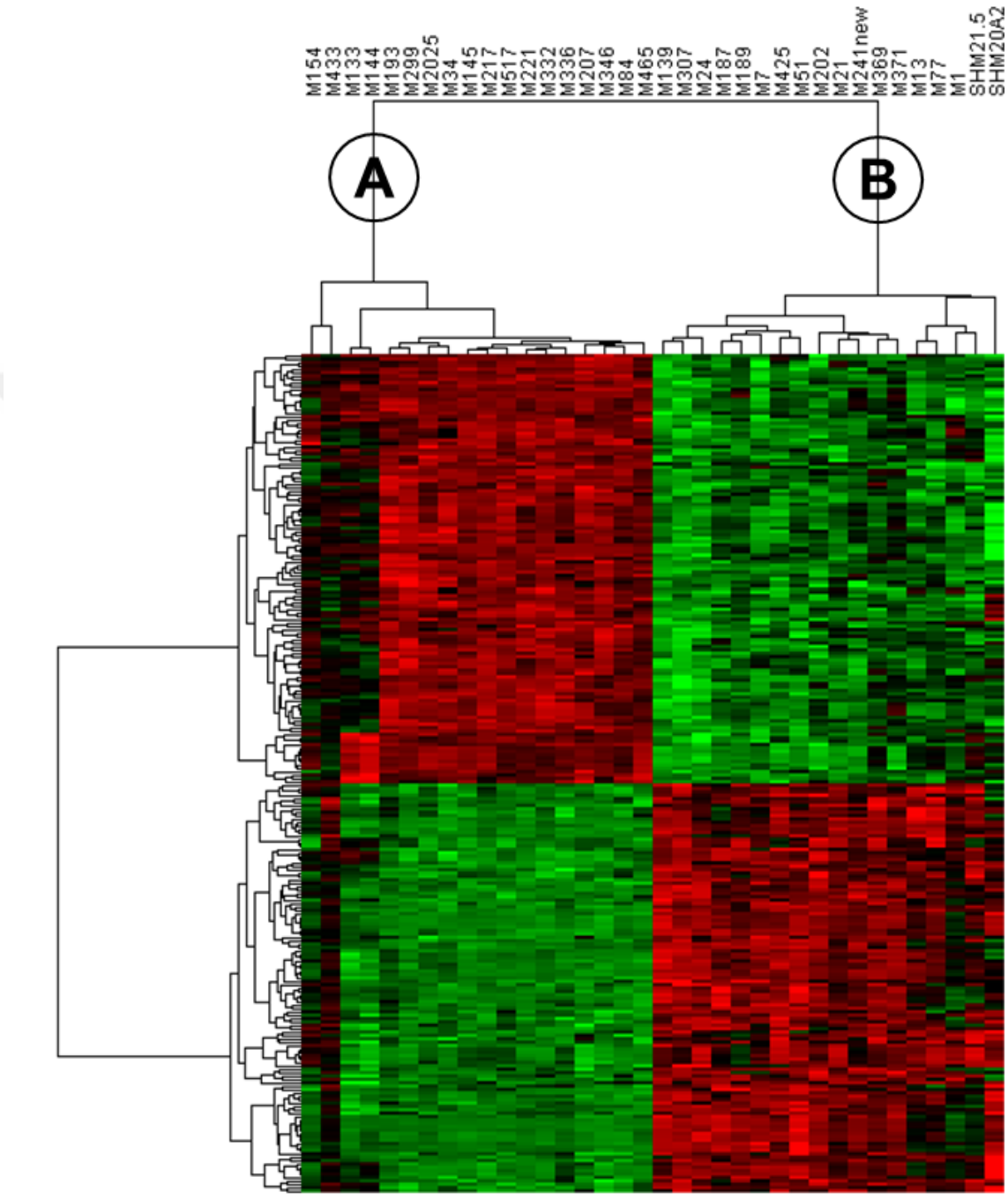


Figure 3.23 Hierarchical clustering analyses of PrCMLs using the most differentially expressed genes between clusters A and B. Heatmap generated with the probesets in Appendix C-Table 8-3.

MLANA and INHBA were among the genes with highest significance as they both separated melanoma cells into two distinct clusters by both in silico and in vitro approaches. Expression of these two genes are inversely correlated in both commercial as well as in PrMCLs (Figure 3.18). Although there is a clear mutual exclusivity in expression of two genes (quadrants 1 and 3, Figure 3.24), few number of cells express both or have little expression of both genes. Quadrants which separated PrCMLs on MLANA and INHBA basis are shown in Figure 3.24 and were used for discriminating cells based on differential sensitivity to HSP90, MEK and SRC inhibitors. Classifying melanoma cells with this method was successful in revealing drug sensitivity differences for both CMCLs and PrMCLs. Cells in Q2, 3 and 4 were combined as their sensitivity profiles were similar.

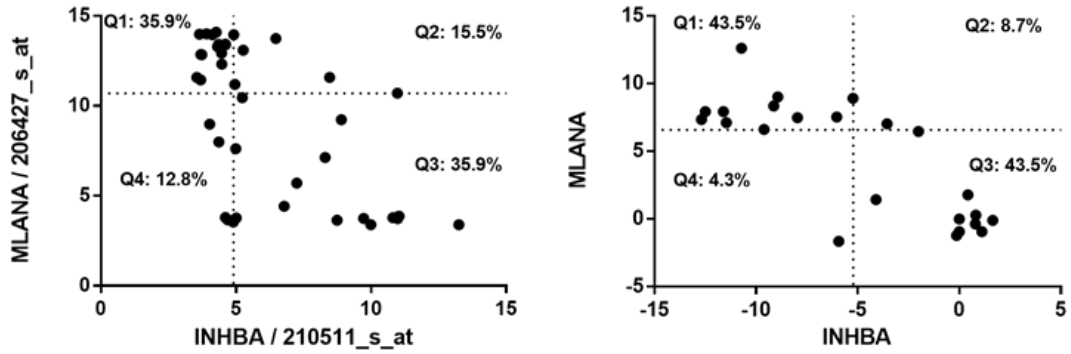


Figure 3.24: MLANA and INHBA expression plot of CCMLs and PrMCLs. Cell lines divided into quadrants based on expression of MLANA and INHBA. Thresholds shown as dashed lines correspond to median expression. Most cells are clustered in Q1 and 3 in in silico analysis of CCMLs (left) and in vitro q-RT-PCR based analysis of PrMCLs (right), as expected.

Sensitivity of both CCMLs and PrCMLs were compared between Q1 versus Q2, 3 and 4 groups (Figure 3.25 and Figure 3.26). This classification was similar to that of classification shown in Figure 3.21 and Figure 3.22. According to *in silico* analysis, Q1 cells (MLANA-high, INHB-low) were more sensitive to AZD6244, while they were more resistant to 17AAG and AZD0530, in comparison to Q2, 3, 4 cells (Figure 3.25). Validation experiments of these findings have included screening results for 10a as well. As shown in Figure 3.26, AZD6244 and 10a were both more effective on Q1 cells compared to Q2, 3, 4, whereas 17AAG and AZD0530 were more effective on Q2, 3, 4 cells.AZD6244 was significantly effective on Q1 cells ($p < 0.01$). There was no significance detected for other drugs tested. We therefore conclude that MLANA and INHBA based classification of melanoma cells is able to identify cells which are sensitive to MEK inhibitor AZD6244. Although not significant, two-gene based classification is also predictive on 10a and 17AAG sensitivity in melanoma.

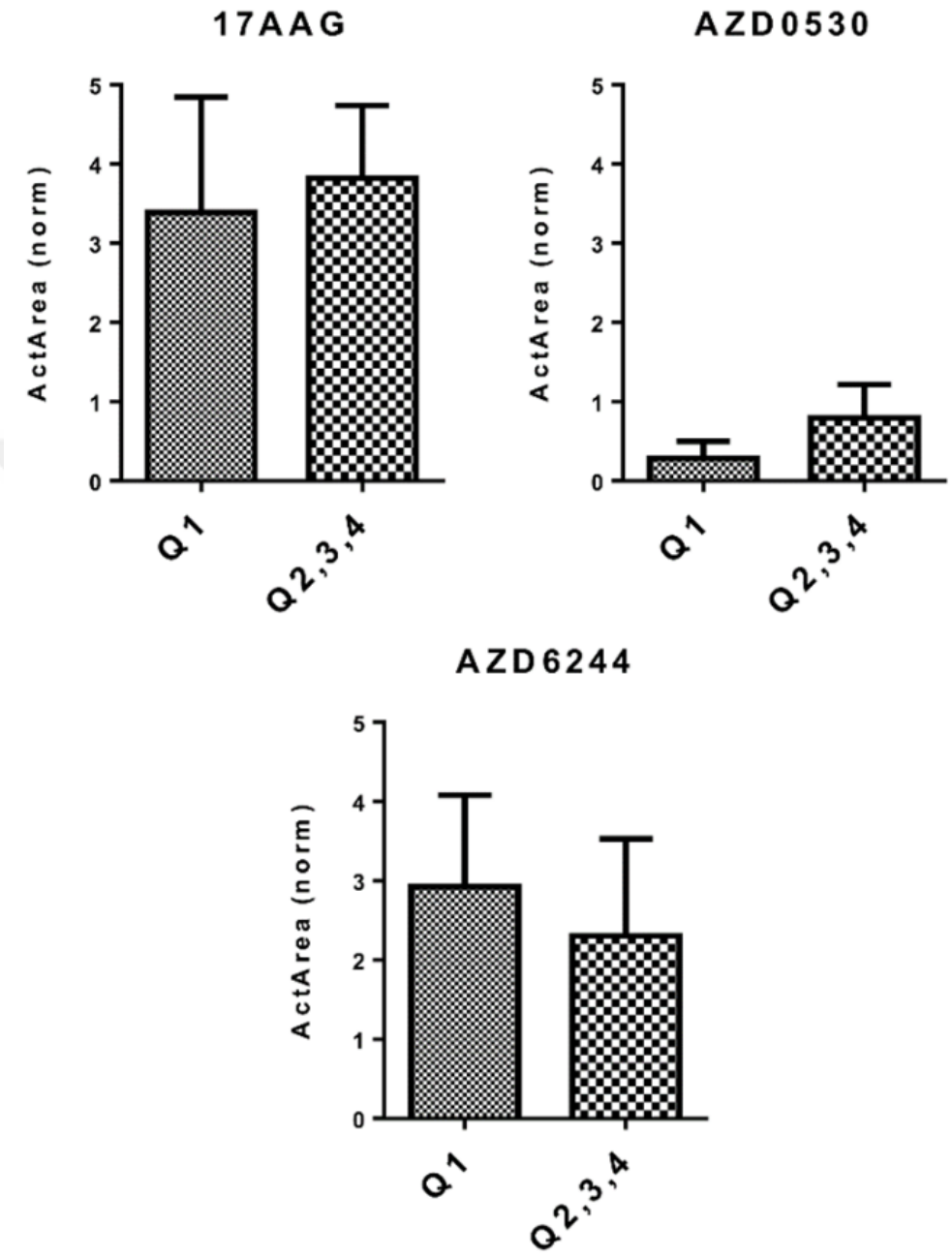


Figure 3.25: MLANA and INHBA based in silico drug sensitivity analysis of CMCLs. Cells classified in quadrants identified in Figure 3.24 (left) are used in this analyses. P values (t-test) for 17AAG, AZD0530 and AZD6244 are 0.26, 0.0002 and 0.13, respectively.

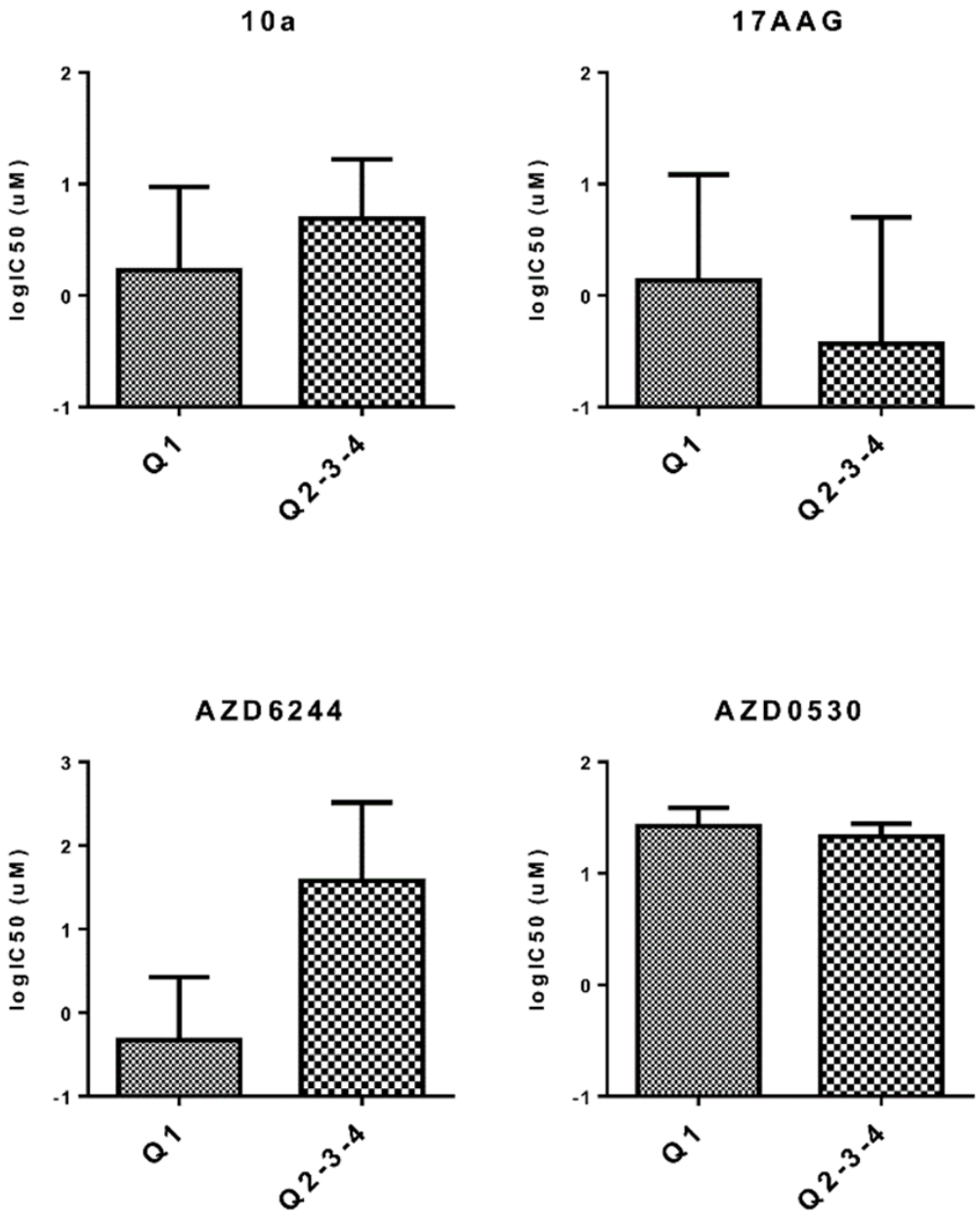


Figure 3.26: MLANA and INHBA based in vitro drug sensitivity analysis of PrCMLs. Cells classified in quadrants identified in Figure 3.24 (right) are used in this analyses. P values (t-test) for 10a, 17AAG, AZD6244, and AZD0530 and are 0.1, 0.43, 0.009 and 0.32, respectively.

Chapter 4

Discussion

In this thesis study, a significant association was found between CT-X gene expression and MTHFR 677 CC genotype together with the MTHFR 677 C allele. Other four markers analyzed were not associated with CT-X gene expression. This result was in contrast with a previously published study where the MTHFR 677 T allele was associated with decreased MAGE-A1 gene expression as well as decreased DNA methylation [130]. One reason for this could be the variance in genotype frequencies between NSCLC cohort genotyped in this study and the general population. However, distribution of one-carbon pathway genotypes is similar between other cancer cohorts, normal samples and NSCLC samples genotyped in this study (Appendix B-Table 7-4, Table 7-5, Table 7-6, Table 7-7 and Table 7-8).

Previously, squamous carcinoma of NSCLC have been reported to show stronger CT-X gene expression compared to adenocarcinoma of same type of tumor [24,111], raising the possibility that histology is a confounding factor for CT-X gene expression. However, multivariate logistic regression analyses have shown that MTHFR 677 CC genotype or MTHFR 677 C allele was associated independently from histology with CT-X expression. Tumor type is among the factors affecting CT-X gene expression, as hematological malignancies and kidney cancers have low level of CT-X gene expression [22,23]. Accordingly, NSCLC association analyses results were failed to be validated in AML and this may stem from the fact that PCA-based AML classification

is artificial due to negligible CT-X expression in this type of tumor. Future validation studies using in silico data derived from NSCLC might reveal associations not identified in this study. One such data may come from The Cancer Genome Atlas (TCGA) where both whole genome sequencing and RNA sequencing data is available for a diverse array of tumors [131-134]. Detailed analyses of TCGA data can be conducted to validate the association analyses described for NSCLC in this thesis as well as other tumors with high CT-X gene expression.

Power analyses were performed to calculate the sample size which is required to detect a significant association between four other polymorphisms genotyped and CT-X gene expression, except MTHFR 677 C>T. These analyses have revealed that at least 250 patients would be required for another significant association result (not shown). Therefore, larger cohorts are required to reveal additional associations as well as other compound effects of one-carbon pathway polymorphisms on CT-X gene expression.

Although DNA demethylation is expected to be modulated by decreased SAM concentrations, gene expression regulation controlled by a defined SAM concentration threshold might be affected by various other factors not tested in this study. One of these factors might be thymidylate synthase (TS) which is known to be expressed in varying levels in cancer with inhibitory activity against MTHFR [135]. Since CT-X (+) tumors are known to be more aggressive and associated with poor clinical outcome compared to CT-X (-) tumors, increased TS activity due to increased demand for thymidylate synthesis in CT-X (+) tumors will increase MTHFR inhibition due to potent inhibitory effect of DHF on MTHFR. To cope with increased MTHFR inhibition, tumors carrying normo-active form of MTHFR (which is encoded by the 677 C allele)

might be the reason of significant association between CT-X gene expression and MTHFR 677 C allele. In this line, DNA damage of CT-X (+) cell lines were found to be higher than (-) cell lines suggesting a link between CT-X expression, DNA synthesis and DNA damage. Moreover, CT-X (-) cell lines were found to be more DNA damage prone when cultured in FA-free conditions compared to CT-X (+) cells. Association between genotypes of one-carbon pathway enzymes and DNA damage cannot be conducted due to limited number of cell lines studied. Future experiments may provide insights on the level of association between CT-X gene expression DNA synthesis, DNA damage and polymorphisms of relevant genes of the one-carbon pathway.

The role of MTHFR function on CT-X gene expression could be better understood if MTHFR expression is reduced or abolished by siRNA knockdown. Decreasing or abolishing MTHFR expression could, however, have off-target effects on other one-carbon derivatives which will further complicate the scenario. Although transgenic mice lacking MTHFR exon 3 have been associated with a severe phenotype [47], role of MTHFR polymorphisms on cancer progression is subjected to long term debate [136-139].

Heterogeneity of melanoma is eminent from a global gene expression point of view together with CT-X gene expression. Although melanoma is among tumors with high CT-X gene expression (Appendix A-Figure 6.1), major transcriptomic clusters identified in this study have high expression of MAGE and PRAME which are CT-X and non-X-CT genes, respectively. Other CT genes were not included in the list of highly differentially expressed genes among clusters A and B (Figure 3.11, Figure 3.23, Appendix C-Table 8-3). Therefore, CT genes other than MAGE and PRAME have

expression in both melanoma subgroups which makes it difficult to subgroup melanoma solely based on CT gene expression. Accordingly, instead of characterizing melanoma subgroups using CT-X genes, we switched to an unsupervised approach and focused on all transcriptome based analyses. This revealed two major molecular subgroups which had differential sensitivity to MEK inhibitor AZD6244. This result was in concordance with a previously published data [128], however, to best of our knowledge, sensitivity to AZD6244 was first related to MLANA and INHBA expression in this study. Although there was a lack of significance with other drugs tested, sensitivity profiles of melanoma cells from different clusters were varied.

For melanoma related drug sensitivity analyses, we hypothesized that different melanoma cells with different phenotypic profiles would have variably active signaling pathways which are sensitive to one of the recent single molecule inhibitors. Furthermore, we hypothesized that the distinct phenotypes defined for melanoma could have overlapping features. Although the number of cell lines tested in this study were limited, we generated data which supports our second hypothesis, as we have shown that cells molecularly subgrouped by multiple gene lists are not necessarily better classified in terms of drug sensitivity, compared to subgrouping with 2 gene-based approaches (MLANA and INHBA). If melanoma cells switch between different phenotypes in vivo [78], utilization of different pathways during this switching process will result in altered sensitivity to single molecular inhibitor. Investigation of altered drug sensitivity during this switching process could have interesting insights on melanoma therapy.

We were able to identify two genes with high mutual exclusivity (MLANA and INHBA) however, transcriptomic subgroups defined by these genes and other multigene lists was not able to identify a distinct group sensitive to a specific targeted drug. One explanation for this fact could be that mRNA expression based subgrouping, however they may be useful in identifying cells with homogenous transcriptomic phenotypes, do not necessarily identify cells that utilize the same signaling pathways. This discrepancy could stem from the fact that simultaneous but diverse pathways could be active in a given cell. Despite the lack of statistical significance, similar sensitivity profiles observed for 17AAG and AZD0530, which target two non-overlapping signaling pathways, support this line of thought. If a cell is able to use various pathways at the same time, then the cell will use one out of the several ones when possible and dynamically switch to another. .

Our results have shown that melanoma cells are possibly utilizing multiple survival pathways due to phenotypic homogeneity. The best personalized treatment approach for melanoma will be to first identify the signaling pathway which is critical for the survival in each cell, followed by the utilization of a drug or a combination of drugs that can target all of these pathways simultaneously. Our results demonstrated that MLANA and INHBA based profiling of melanoma cells is capable of classifying in vitro responses to molecularly targeted drugs. Larger validation experiments both in vitro and in vivo which include tests of drug synergy is needed to confirm these observations.

Bibliography

1. Traversari C, van der Bruggen P, Van den Eynde B, Hainaut P, Lemoine C, et al. (1992) Transfection and expression of a gene coding for a human melanoma antigen recognized by autologous cytolytic T lymphocytes. Immunogenetics 35: 145-152.

2. van der Bruggen P, Traversari C, Chomez P, Lurquin C, De Plaen E, et al. (1991) A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. Science 254: 1643-1647.

3. Scanlan MJ, Gure AO, Jungbluth AA, Old LJ, Chen YT (2002) Cancer/testis antigens: an expanding family of targets for cancer immunotherapy. Immunol Rev 188: 22-32.

4. Fratta E, Coral S, Covre A, Parisi G, Colizzi F, et al. (2011) The biology of cancer testis antigens: putative function, regulation and therapeutic potential. Mol Oncol 5: 164-182.

5. Caballero OL, Chen YT (2009) Cancer/testis (CT) antigens: potential targets for immunotherapy. Cancer Sci 100: 2014-2021.

6. Ehrlich M (2009) DNA hypomethylation in cancer cells. Epigenomics 1: 239-259.

7. Loriot A, Reister S, Parvizi GK, Lysy PA, De Smet C (2009) DNA methylation-associated repression of cancer-germline genes in human embryonic and adult stem cells. Stem Cells 27: 822-824.

8. Weber J, Salgaller M, Samid D, Johnson B, Herlyn M, et al. (1994) Expression of the MAGE-1 tumor antigen is up-regulated by the demethylating agent 5-aza-2'-deoxycytidine. Cancer Res 54: 1766-1771.

9. Coral S, Sigalotti L, Altomonte M, Engelsberg A, Colizzi F, et al. (2002) 5-aza-2'-deoxycytidine-induced expression of functional cancer testis antigens in human renal cell carcinoma: immunotherapeutic implications. Clin Cancer Res 8: 2690-2695.

10. Güre AO, Wei IJ, Old LJ, Chen YT (2002) The SSX gene family: characterization of 9 complete genes. Int J Cancer 101: 448-453.

11. Gjerstorff M, Burns JS, Nielsen O, Kassem M, Ditzel H (2009) Epigenetic modulation of cancer-germline antigen gene expression in tumorigenic human mesenchymal stem cells: implications for cancer therapy. Am J Pathol 175: 314-323.

12. Almatrafi A, Feichtinger J, Vernon EG, Escobar NG, Wakeman JA, et al. (2014) Identification of a class of human cancer germline genes with transcriptional silencing refractory to the hypomethylating drug 5-aza-2'-deoxycytidine. Oncoscience 1: 745-750.

13. Jungbluth AA, Stockert E, Chen YT, Kolb D, Iversen K, et al. (2000) Monoclonal antibody MA454 reveals a heterogeneous expression pattern of MAGE-1 antigen in formalin-fixed paraffin embedded lung tumours. Br J Cancer 83: 493-497.

14. Marchal R, Chicheportiche A, Dutrillaux B, Bernardino-Sgherri J (2004) DNA methylation in mouse gametogenesis. Cytogenet Genome Res 105: 316-324.

15. Reik W, Dean W, Walter J (2001) Epigenetic reprogramming in mammalian development. Science 293: 1089-1093.

16. O'Carroll D, Scherthan H, Peters AH, Opravil S, Haynes AR, et al. (2000) Isolation and characterization of Suv39h2, a second histone H3 methyltransferase gene that displays testis-specific expression. Mol Cell Biol 20: 9423-9433.

17. Sonnack V, Failing K, Bergmann M, Steger K (2002) Expression of hyperacetylated histone H4 during normal and impaired human spermatogenesis. Andrologia 34: 384-390.

18. Inoue N, Hess KD, Moreadith RW, Richardson LL, Handel MA, et al. (1999) New gene family defined by MORC, a nuclear protein required for mouse spermatogenesis. Hum Mol Genet 8: 1201-1207.

19. Kim E, Nishimura H, Baba T (2003) Differential localization of ADAM1a and ADAM1b in the endoplasmic reticulum of testicular germ cells and on the surface of epididymal sperm. Biochem Biophys Res Commun 304: 313-319.

20. Sage J, Martin L, Meuwissen R, Heyting C, Cuzin F, et al. (1999) Temporal and spatial control of the Sycp1 gene transcription in the mouse meiosis: regulatory elements active in the male are not sufficient for expression in the female gonad. Mech Dev 80: 29-39.

21. Whitehurst AW (2014) Cause and consequence of cancer/testis antigen activation in cancer. Annu Rev Pharmacol Toxicol 54: 251-272.

22. Simpson AJ, Caballero OL, Jungbluth A, Chen YT, Old LJ (2005) Cancer/testis antigens, gametogenesis and cancer. Nat Rev Cancer 5: 615-625.

23. Old LJ (2001) Cancer/testis (CT) antigens - a new link between gametogenesis and cancer. Cancer Immun 1: 1.

24. Gure AO, Chua R, Williamson B, Gonen M, Ferrera CA, et al. (2005) Cancer-testis genes are coordinately expressed and are markers of poor outcome in non-small cell lung cancer. Clin Cancer Res 11: 8055-8062.

25. Chen YT, Ross DS, Chiu R, Zhou XK, Chen YY, et al. (2011) Multiple cancer/testis antigens are preferentially expressed in hormone-receptor negative and high-grade breast cancers. PLoS One 6: e17876.

26. Odunsi K, Jungbluth AA, Stockert E, Qian F, Gnjatic S, et al. (2003) NY-ESO-1 and LAGE-1 cancer-testis antigens are potential targets for immunotherapy in epithelial ovarian cancer. Cancer Res 63: 6076-6083.

27. Mashino K, Sadanaga N, Tanaka F, Yamaguchi H, Nagashima H, et al. (2001) Expression of multiple cancer-testis antigen genes in gastrointestinal and breast carcinomas. Br J Cancer 85: 713-720.

28. Yilmaz-Ozcan S, Sade A, Kucukkaraduman B, Kaygusuz Y, Senses KM, et al. (2014) Epigenetic mechanisms underlying the dynamic expression of cancer-

testis genes, PAGE2, -2B and SPANX-B, during mesenchymal-to-epithelial transition. PLoS One 9: e107905.

29. Stover PJ (2009) One-carbon metabolism-genome interactions in folate-associated pathologies. J Nutr 139: 2402-2405.

30. Pogribny IP, Ross SA, Wise C, Pogribna M, Jones EA, et al. (2006) Irreversible global DNA hypomethylation as a key step in hepatocarcinogenesis induced by dietary methyl deficiency. Mutat Res 593: 80-87.

31. Botto N, Andreassi MG, Manfredi S, Masetti S, Cocci F, et al. (2003) Genetic polymorphisms in folate and homocysteine metabolism as risk factors for DNA damage. Eur J Hum Genet 11: 671-678.

32. Chango A, Emery-Fillon N, de Courcy GP, Lambert D, Pfister M, et al. (2000) A polymorphism (80G->A) in the reduced folate carrier gene and its associations with folate status and homocysteinemia. Mol Genet Metab 70: 310-315.

33. Christensen B, Frosst P, Lussier-Cacan S, Selhub J, Goyette P, et al. (1997) Correlation of a common mutation in the methylenetetrahydrofolate reductase gene with plasma homocysteine in patients with premature coronary artery disease. Arterioscler Thromb Vasc Biol 17: 569-573.

34. Friso S, Girelli D, Trabetti E, Olivieri O, Guarini P, et al. (2005) The MTHFR 1298A>C polymorphism and genomic DNA methylation in human lymphocytes. Cancer Epidemiol Biomarkers Prev 14: 938-943.

35. Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, et al. (1995) A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. Nat Genet 10: 111-113.

36. Heijmans BT, Boer JM, Suchiman HE, Cornelisse CJ, Westendorp RG, et al. (2003) A common variant of the methylenetetrahydrofolate reductase gene (1p36) is associated with an increased risk of cancer. Cancer Res 63: 1249-1253.

37. Leclerc D, Campeau E, Goyette P, Adjalla CE, Christensen B, et al. (1996) Human methionine synthase: cDNA cloning and identification of mutations in patients of the cblG complementation group of folate/cobalamin disorders. Hum Mol Genet 5: 1867-1874.

38. Olteanu H, Munson T, Banerjee R (2002) Differences in the efficiency of reductive activation of methionine synthase and exogenous electron acceptors between the common polymorphic variants of human methionine synthase reductase. Biochemistry 41: 13378-13385.

39. Relton CL, Wilding CS, Pearce MS, Laffling AJ, Jonas PA, et al. (2004) Gene-gene interaction in folate-related genes and risk of neural tube defects in a UK population. J Med Genet 41: 256-260.

40. Sharp L, Little J (2004) Polymorphisms in genes involved in folate metabolism and colorectal neoplasia: a HuGE review. Am J Epidemiol 159: 423-443.

41. Stern LL, Mason JB, Selhub J, Choi SW (2000) Genomic DNA hypomethylation, a characteristic of most cancers, is present in peripheral leukocytes of individuals who are homozygous for the C677T polymorphism in the

methylenetetrahydrofolate reductase gene. Cancer Epidemiol Biomarkers Prev 9: 849-853.

42. Whetstine JR, Gifford AJ, Witt T, Liu XY, Flatley RM, et al. (2001) Single nucleotide polymorphisms in the human reduced folate carrier: characterization of a high-frequency G/A variant at position 80 and transport properties of the His(27) and Arg(27) carriers. Clin Cancer Res 7: 3416-3422.

43. Castro R, Rivera I, Ravasco P, Camilo ME, Jakobs C, et al. (2004) 5,10-methylenetetrahydrofolate reductase (MTHFR) 677C-->T and 1298A-->C mutations are associated with DNA hypomethylation. J Med Genet 41: 454-458.

44. Zhao R, Matherly LH, Goldman ID (2009) Membrane transporters and folate homeostasis: intestinal absorption and transport into systemic compartments and tissues. Expert Rev Mol Med 11: e4.

45. Crider KS, Yang TP, Berry RJ, Bailey LB (2012) Folate and DNA methylation: a review of molecular mechanisms and the evidence for folate's role. Adv Nutr 3: 21-38.

46. Anguera MC, Field MS, Perry C, Ghandour H, Chiang EP, et al. (2006) Regulation of folate-mediated one-carbon metabolism by 10-formyltetrahydrofolate dehydrogenase. J Biol Chem 281: 18335-18342.

47. Chen Z, Karaplis AC, Ackerman SL, Pogribny IP, Melnyk S, et al. (2001) Mice deficient in methylenetetrahydrofolate reductase exhibit

hyperhomocysteinemia and decreased methylation capacity, with neuropathology and aortic lipid deposition. Hum Mol Genet 10: 433-443.

48. Devlin AM, Arning E, Bottiglieri T, Faraci FM, Rozen R, et al. (2004) Effect of Mthfr genotype on diet-induced hyperhomocysteinemia and vascular function in mice. Blood 103: 2624-2629.

49. van der Put NM, Steegers-Theunissen RP, Frosst P, Trijbels FJ, Eskes TK, et al. (1995) Mutated methylenetetrahydrofolate reductase as a risk factor for spina bifida. Lancet 346: 1070-1071.

50. Lievers KJ, Boers GH, Verhoef P, den Heijer M, Kluijtmans LA, et al. (2001) A second common variant in the methylenetetrahydrofolate reductase (MTHFR) gene and its relationship to MTHFR enzyme activity, homocysteine, and cardiovascular disease risk. J Mol Med (Berl) 79: 522-528.

51. Markus HS, Ali N, Swaminathan R, Sankaralingam A, Molloy J, et al. (1997) A common polymorphism in the methylenetetrahydrofolate reductase gene, homocysteine, and ischemic cerebrovascular disease. Stroke 28: 1739-1743.

52. Harmon DL, Woodside JV, Yarnell JW, McMaster D, Young IS, et al. (1996) The common 'thermolabile' variant of methylene tetrahydrofolate reductase is a major determinant of mild hyperhomocysteinaemia. QJM 89: 571-577.

53. van der Put NM, Gabreëls F, Stevens EM, Smeitink JA, Trijbels FJ, et al. (1998) A second common mutation in the methylenetetrahydrofolate reductase gene: an additional risk factor for neural-tube defects? Am J Hum Genet 62: 1044-1051.

54. Friso S, Choi SW, Girelli D, Mason JB, Dolnikowski GG, et al. (2002) A common mutation in the 5,10-methylenetetrahydrofolate reductase gene affects genomic DNA methylation through an interaction with folate status. Proc Natl Acad Sci U S A 99: 5606-5611.

55. Paz MF, Avila S, Fraga MF, Pollan M, Capella G, et al. (2002) Germ-line variants in methyl-group metabolism genes and susceptibility to DNA methylation in normal tissues and human primary tumors. Cancer Res 62: 4519-4524.

56. Molloy AM, Mills JL, Kirke PN, Whitehead AS, Weir DG, et al. (1998) Whole-blood folate values in subjects with different methylenetetrahydrofolate reductase genotypes: differences between the radioassay and microbiological assays. Clin Chem 44: 186-188.

57. Weisberg I, Tran P, Christensen B, Sibani S, Rozen R (1998) A second genetic polymorphism in methylenetetrahydrofolate reductase (MTHFR) associated with decreased enzyme activity. Mol Genet Metab 64: 169-172.

58. Castro R, Rivera I, Blom HJ, Jakobs C, Tavares de Almeida I (2006) Homocysteine metabolism, hyperhomocysteinaemia and vascular disease: an overview. J Inherit Metab Dis 29: 3-20.

59. Chiang PK, Gordon RK, Tal J, Zeng GC, Doctor BP, et al. (1996) S-Adenosylmethionine and methylation. FASEB J 10: 471-480.

60. James SJ, Pogribny IP, Pogribna M, Miller BJ, Jernigan S, et al. (2003) Mechanisms of DNA damage, DNA hypomethylation, and tumor progression in the

folate/methyl-deficient rat model of hepatocarcinogenesis. J Nutr 133: 3740S-3747S.

61. Fenech M (2012) Folate (vitamin B9) and vitamin B12 and their function in the maintenance of nuclear and mitochondrial genome integrity. Mutat Res 733: 21-33.

62. Blount BC, Mack MM, Wehr CM, MacGregor JT, Hiatt RA, et al. (1997) Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage. Proc Natl Acad Sci U S A 94: 3290-3295.

63. Stover PJ (2004) Physiology of folate and vitamin B12 in health and disease. Nutr Rev 62: S3-12; discussion S13.

64. Duthie SJ, Narayanan S, Sharp L, Little J, Basten G, et al. (2004) Folate, DNA stability and colo-rectal neoplasia. Proc Nutr Soc 63: 571-578.

65. Krajinovic M, Lemieux-Blanchard E, Chiasson S, Primeau M, Costea I, et al. (2004) Role of polymorphisms in MTHFR and MTHFD1 genes in the outcome of childhood acute lymphoblastic leukemia. Pharmacogenomics J 4: 66-72.

66. Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP, et al. (2010) Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. Science 329: 78-82.

67. Lotem M, Merims S, Frank S, Hamburger T, Nissan A, et al. (2016) Adjuvant Autologous Melanoma Vaccine for Macroscopic Stage III Disease: Survival,

Biomarkers, and Improved Response to CTLA-4 Blockade. J Immunol Res 2016: 8121985.

68. Siegel RL, Miller KD, Jemal A (2016) Cancer statistics, 2016. CA Cancer J Clin 66: 7-30.

69. Miller KD, Siegel RL, Lin CC, Mariotto AB, Kramer JL, et al. (2016) Cancer treatment and survivorship statistics, 2016. CA Cancer J Clin.

70. Jang S, Atkins MB (2013) Which drug, and when, for patients with BRAF-mutant melanoma? Lancet Oncol 14: e60-69.

71. Balch CM, Gershenwald JE, Soong SJ, Thompson JF, Atkins MB, et al. (2009) Final version of 2009 AJCC melanoma staging and classification. J Clin Oncol 27: 6199-6206.

72. Roesch A (2015) Tumor heterogeneity and plasticity as elusive drivers for resistance to MAPK pathway inhibition in melanoma. Oncogene 34: 2951-2957.

73. Colombino M, Capone M, Lissia A, Cossu A, Rubino C, et al. (2012) BRAF/NRAS mutation frequencies among primary tumors and metastases in patients with melanoma. J Clin Oncol 30: 2522-2529.

74. Mar VJ, Wong SQ, Li J, Scolyer RA, McLean C, et al. (2013) BRAF/NRAS wild-type melanomas have a high mutation load correlating with histologic and molecular signatures of UV damage. Clin Cancer Res 19: 4589-4598.

75. Shtivelman E, Davies MQ, Hwu P, Yang J, Lotem M, et al. (2014) Pathways and therapeutic targets in melanoma. Oncotarget 5: 1701-1752.

76. Tremante E, Ginebri A, Lo Monaco E, Frascione P, Di Filippo F, et al. (2012) Melanoma molecular classes and prognosis in the postgenomic era. Lancet Oncol 13: e205-211.

77. Ferguson J, Arozarena I, Ehrhardt M, Wellbrock C (2013) Combination of MEK and SRC inhibition suppresses melanoma cell growth and invasion. Oncogene 32: 86-96.

78. Hoek KS, Eichhoff OM, Schlegel NC, Döbbeling U, Kobert N, et al. (2008) In vivo switching of human melanoma cells between proliferative and invasive states. Cancer Res 68: 650-656.

79. Sáez-Ayala M, Montenegro MF, Sánchez-Del-Campo L, Fernández-Pérez MP, Chazarra S, et al. (2013) Directed phenotype switching as an effective antimelanoma strategy. Cancer Cell 24: 105-119.

80. Roesch A, Fukunaga-Kalabis M, Schmidt EC, Zabierowski SE, Brafford PA, et al. (2010) A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. Cell 141: 583-594.

81. Hoek KS, Goding CR (2010) Cancer stem cells versus phenotype-switching in melanoma. Pigment Cell Melanoma Res 23: 746-759.

82. Jayachandran A, Anaka M, Prithviraj P, Hudson C, McKeown SJ, et al. (2014) Thrombospondin 1 promotes an aggressive phenotype through epithelial-to-mesenchymal transition in human melanoma. Oncotarget 5: 5782-5797.

83. Kreiseder B, Holper-Schichl YM, Muellauer B, Jacobi N, Pretsch A, et al. (2015) Alpha-catulin contributes to drug-resistance of melanoma by activating NF-κB and AP-1. PLoS One 10: e0119402.

84. Hoek KS, Schlegel NC, Brafford P, Sucker A, Ugurel S, et al. (2006) Metastatic potential of melanomas defined by specific gene expression profiles with no BRAF signature. Pigment Cell Res 19: 290-302.

85. Widmer DS, Cheng PF, Eichhoff OM, Belloni BC, Zipser MC, et al. (2012) Systematic classification of melanoma cells by phenotype-specific gene expression mapping. Pigment Cell Melanoma Res 25: 343-353.

86. Jeffs AR, Glover AC, Slobbe LJ, Wang L, He S, et al. (2009) A gene expression signature of invasive potential in metastatic melanoma cells. PLoS One 4: e8461.

87. Gure AO, Chua R, Williamson B, Gonen M, Ferrera CA, et al. (2005) Cancer-testis genes are coordinately expressed and are markers of poor outcome in non-small cell lung cancer. Clin Cancer Res 11: 8055-8062.

88. Liang CC, Park AY, Guan JL (2007) In vitro scratch assay: a convenient and inexpensive method for analysis of cell migration in vitro. Nat Protoc 2: 329-333.

89. Crott JW, Mashiyama ST, Ames BN, Fenech M (2001) The effect of folic acid deficiency and MTHFR C677T polymorphism on chromosome damage in human lymphocytes in vitro. Cancer Epidemiol Biomarkers Prev 10: 1089-1096.

90. Akbas-Avci D (2009) An SSX4 Knock-In Cell Line Model and In Silico Analysis of Gene Expression Data as Two Approaches for Investigating Mechanisms of Cancer/Testis Gene Expression. Ankara: Bilkent University.

91. Dhawan A, Bajpayee M, Pandey AK, Parmar D ITRC: The SCGE/ Comet Assay Protocol/ Protocol for the Single Cell Gel Electrophoresis / Comet Assay for Rapid Genotoxicity Assessment. In: Division DT, Centre ITR, editors.

92. Końca K, Lankoff A, Banasik A, Lisowska H, Kuszewski T, et al. (2003) A cross-platform public domain PC image-analysis program for the comet assay. Mutat Res 534: 15-20.

93. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, et al. (2012) The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 483: 603-607.

94. Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30: 207-210.

95. Simon R, Lam A, Li MC, Ngan M, Menenzes S, et al. (2007) Analysis of gene expression data using BRB-ArrayTools. Cancer Inform 3: 11-17.

96. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, et al. (2003) Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 31: e15.

97. de Hoon MJ, Imoto S, Nolan J, Miyano S (2004) Open source clustering software. Bioinformatics 20: 1453-1454.

98. Saldanha AJ (2004) Java Treeview--extensible visualization of microarray data. Bioinformatics 20: 3246-3248.

99. Team RC (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. .

100. Raskin L, Fullen DR, Giordano TJ, Thomas DG, Frohm ML, et al. (2013) Transcriptome profiling identifies HMGA2 as a biomarker of melanoma progression and prognosis. J Invest Dermatol 133: 2585-2592.

101. Meister M, Belousov A, Xu E, Schnabel P, Warth A, et al. (2014) Intra-tumor Heterogeneity of Gene Expression Profiles in Early Stage Non-Small Cell Lung Cancer. Journal of Bioinformatics Research Studies 1: 1.

102. Maire V, Némati F, Richardson M, Vincent-Salomon A, Tesson B, et al. (2013) Polo-like kinase 1: a potential therapeutic option in combination with conventional chemotherapy for the management of patients with triple-negative breast cancer. Cancer Res 73: 813-823.

103. Maire V, Baldeyron C, Richardson M, Tesson B, Vincent-Salomon A, et al. (2013) TTK/hMPS1 is an attractive therapeutic target for triple-negative breast cancer. PLoS One 8: e63712.

104. Maubant S, Tesson B, Maire V, Ye M, Rigaill G, et al. (2015) Transcriptome analysis of Wnt3a-treated triple-negative breast cancer cells. PLoS One 10: e0122333.

105. Mok SC, Bonome T, Vathipadiekal V, Bell A, Johnson ME, et al. (2009) A gene signature predictive for outcome in advanced ovarian cancer identifies a survival factor: microfibril-associated glycoprotein 2. Cancer Cell 16: 521-532.

106. Okazaki S, Ishikawa T, Iida S, Ishiguro M, Kobayashi H, et al. (2012) Clinical significance of UNC5B expression in colorectal cancer. Int J Oncol 40: 209-216.

107. Wang L, Shalek AK, Lawrence M, Ding R, Gaublomme JT, et al. (2014) Somatic mutation as a mechanism of Wnt/β-catenin pathway activation in CLL. Blood 124: 1089-1098.

108. Badea L, Herlea V, Dima SO, Dumitrascu T, Popescu I (2008) Combined gene expression analysis of whole-tissue and microdissected pancreatic ductal adenocarcinoma identifies genes specifically overexpressed in tumor epithelia. Hepatogastroenterology 55: 2016-2027.

109. Pei H, Li L, Fridley BL, Jenkins GD, Kalari KR, et al. (2009) FKBP51 affects cancer cell response to chemotherapy by negatively regulating Akt. Cancer Cell 16: 259-266.

110. Tibshirani R, Walther G, Hastie T (2001) Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63: 411-423.

111. Medema JP (2013) Cancer stem cells: the challenges ahead. Nat Cell Biol 15: 338-344.

112. Akbas-Avci D (2009) An SSX4 Knock-In Cell Line Model and In Silico Analysis of Gene Expression Data as Two Approaches for Investigating Mechanisms of Cancer/Testis Gene Expression. Ankara: Bilkent University.

113. Yoshida N, Abe H, Ohkuri T, Wakita D, Sato M, et al. (2006) Expression of the MAGE-A4 and NY-ESO-1 cancer-testis antigens and T cell infiltration in non-small cell lung carcinoma and their prognostic significance. Int J Oncol 28: 1089-1098.

114. Radtke I, Mullighan CG, Ishii M, Su X, Cheng J, et al. (2009) Genomic analysis reveals few genetic alterations in pediatric acute myeloid leukemia. Proc Natl Acad Sci U S A 106: 12944-12949.

115. Atanackovic D, Hildebrandt Y, Jadczak A, Cao Y, Luetkens T, et al. (2010) Cancer-testis antigens MAGE-C1/CT7 and MAGE-A3 promote the survival of multiple myeloma cells. Haematologica 95: 785-793.

116. Kim J, Reber HA, Hines OJ, Kazanjian KK, Tran A, et al. (2006) The clinical significance of MAGEA3 expression in pancreatic cancer. Int J Cancer 118: 2269-2275.

117. Liu W, Cheng S, Asa SL, Ezzat S (2008) The melanoma-associated antigen A3 mediates fibronectin-controlled cancer progression and metastasis. Cancer Res 68: 8104-8112.

118. Shantha Kumara HM, Grieco MJ, Caballero OL, Su T, Ahmed A, et al. (2012) MAGE-A3 is highly expressed in a subset of colorectal cancer patients. Cancer Immun 12: 16.

119. Xu Y, Wang C, Zhang Y, Jia L, Huang J (2015) Overexpression of MAGE-A9 Is Predictive of Poor Prognosis in Epithelial Ovarian Cancer. Sci Rep 5: 12104.

120. Yakirevich E, Sabo E, Lavie O, Mazareb S, Spagnoli GC, et al. (2003) Expression of the MAGE-A4 and NY-ESO-1 cancer-testis antigens in serous ovarian neoplasms. Clin Cancer Res 9: 6453-6460.

121. Zou C, Shen J, Tang Q, Yang Z, Yin J, et al. (2012) Cancer-testis antigens expressed in osteosarcoma identified by gene microarray correlate with a poor patient prognosis. Cancer 118: 1845-1855.

122. Talasz H, Helliger W, Sarg B, Debbage PL, Puschendorf B, et al. (2002) Hyperphosphorylation of histone H2A.X and dephosphorylation of histone H1 subtypes in the course of apoptosis. Cell Death Differ 9: 27-39.

123. Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. J Biol Chem 273: 5858-5868.

124. Rogakou EP, Boon C, Redon C, Bonner WM (1999) Megabase chromatin domains involved in DNA double-strand breaks in vivo. J Cell Biol 146: 905-916.

125. Shoemaker RH (2006) The NCI60 human tumour cell line anticancer drug screen. Nat Rev Cancer 6: 813-823.

126. Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, et al. (2012) Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature 483: 570-575.

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

128. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.

129. Zajac M, Gomez G, Benitez J, Martínez-Delgado B (2010) Molecular signature of response and potential pathways related to resistance to the HSP90 inhibitor, 17AAG, in breast cancer. BMC Med Genomics 3: 44.

130. Zipser MC, Eichhoff OM, Widmer DS, Schlegel NC, Schoenewolf NL, et al. (2011) A proliferative melanoma cell phenotype is responsive to RAF/MEK inhibition independent of BRAF mutation status. Pigment Cell Melanoma Res 24: 326-333.

131. Kruewel T, Schenone S, Radi M, Maga G, Rohrbeck A, et al. (2010) Molecular characterization of c-Abl/c-Src kinase inhibitors targeted against murine tumour progenitor cells that express stem cell markers. PLoS One 5: e14143.

132. Cadieux B, Ching TT, VandenBerg SR, Costello JF (2006) Genome-wide hypomethylation in human glioblastomas associated with specific copy number alteration, methylenetetrahydrofolate reductase allele status, and increased proliferation. Cancer Res 66: 8469-8476.

133. Chin L, Hahn WC, Getz G, Meyerson M (2011) Making sense of cancer genomic data. Genes Dev 25: 534-555.

134. Chin L, Andersen JN, Futreal PA (2011) Cancer genomics: from discovery science to personalized medicine. Nat Med 17: 297-303.

135. (2015) The future of cancer genomics. Nat Med 21: 99.

136. Tomczak K, Czerwinska P, Wiznerowicz M (2015) The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) 19: A68-77.

137. Matthews RG, Daubner SC (1982) Modulation of methylenetetrahydrofolate reductase activity by S-adenosylmethionine and by dihydrofolate and its polyglutamate analogues. Adv Enzyme Regul 20: 123-131.

138. Shi Q, Zhang Z, Li G, Pillow PC, Hernandez LM, et al. (2005) Sex differences in risk of lung cancer associated with methylene-tetrahydrofolate reductase polymorphisms. Cancer Epidemiol Biomarkers Prev 14: 1477-1484.

139. Kotsopoulos J, Zhang WW, Zhang S, McCready D, Trudeau M, et al. (2008) Polymorphisms in folate metabolizing enzymes and transport proteins and the risk of breast cancer. Breast Cancer Res Treat 112: 585-593.

140. Suzuki T, Matsuo K, Hiraki A, Saito T, Sato S, et al. (2007) Impact of one-carbon metabolism-related gene polymorphisms on risk of lung cancer in Japan: a case control study. Carcinogenesis 28: 1718-1725.

141. Shen H, Spitz MR, Wang LE, Hong WK, Wei Q (2001) Polymorphisms of methylene-tetrahydrofolate reductase and risk of lung cancer: a case-control study. Cancer Epidemiol Biomarkers Prev 10: 397-401.

142. Charasson V, Hillaire-Buys D, Solassol I, Laurand-Quancard A, Pinguet F, et al. (2009) Involvement of gene polymorphisms of the folate pathway enzymes in gene expression and anticancer drug sensitivity using the NCI-60 panel as a model. Eur J Cancer 45: 2391-2401.

143. Senses KM, Gonen M, Barutcu AR, Kalaylioglu Z, Isbilen M, et al. (2013) Cancer-testis gene expression is associated with the methylenetetrahydrofolate reductase 677 C>T polymorphism in non-small cell lung carcinoma. BMC Med Genet 14: 97.

144. Shi Q, Zhang Z, Li G, Pillow PC, Hernandez LM, et al. (2005) Polymorphisms of methionine synthase and methionine synthase reductase and risk of lung cancer: a case-control analysis. Pharmacogenet Genomics 15: 547-555.

Appendix A

Figures and Supporting Information for section 3.1

Table 6-1: CT-X probesets used in microarray analyses.

#	CT-X Gene Family	Gene Symbol	Probe Set ID	Microarray Platform[$]
1	NY-ESO-1	CTAG1A	210546_x_at	HGU-133 Plus2
2	NY-ESO-1	CTAG1A	211674_x_at	HGU-133 Plus2
3	NY-ESO-1	CTAG1A	217339_x_at	HGU-133 Plus2
4	GAGE	GAGE1	208283_at	HGU-133 Plus2
5	GAGE	GAGE1	207739_s_at	HGU-133 Plus2
6	GAGE	GAGE1	207086_x_at	HGU-133 Plus2
7	GAGE	GAGE1	208155_x_at	HGU-133 Plus2
8	GAGE	GAGE12B	206640_x_at	HGU-133 Plus2
9	GAGE	GAGE12B	208235_x_at	HGU-133 Plus2
10	MAGEA	MAGEA1	207325_x_at	HGU-133 Plus2
11	MAGEA	MAGEA10	210295_at	HGU-133 Plus2
12	MAGEA	MAGEA10-MAGEA5	1553585_a_at	HGU-133 Plus2
13	MAGEA	MAGEA10-MAGEA5	214642_x_at	HGU-133 Plus2
14	MAGEA	MAGEA11	210503_at	HGU-133 Plus2
15	MAGEA	MAGEA12	210467_x_at	HGU-133 Plus2
16	MAGEA	MAGEA2	1553830_s_at	HGU-133 Plus2
17	MAGEA	MAGEA2	214603_at	HGU-133 Plus2
18	MAGEA	MAGEA3	209942_x_at	HGU-133 Plus2
19	MAGEA	MAGEA4	214254_at	HGU-133 Plus2
20	MAGEA	MAGEA6	214612_x_at	HGU-133 Plus2
21	MAGEA	MAGEA8	210274_at	HGU-133 Plus2

22	**MAGEA**	**MAGEA9**	**210437_at**	**HGU-133 Plus2**
23	**MAGEB**	**MAGEB1**	**207534_at**	**HGU-133 Plus2**
24	**MAGEB**	**MAGEB18**	**1552913_at**	**HGU-133 Plus2**
25	**MAGEB**	**MAGEB2**	**206218_at**	**HGU-133 Plus2**
26	**MAGEB**	**MAGEB3**	**207579_at**	**HGU-133 Plus2**
27	**MAGEB**	**MAGEB4**	**207580_at**	**HGU-133 Plus2**
28	**MAGEB**	**MAGEB4**	**207581_s_at**	**HGU-133 Plus2**
29	**MAGEB**	**MAGEB6**	**1552858_at**	**HGU-133 Plus2**
30	**MAGEC**	**MAGEC1**	**206609_at**	**HGU-133 Plus2**
31	**MAGEC**	**MAGEC2**	**220062_s_at**	**HGU-133 Plus2**
32	**MAGEC**	**MAGEC3**	**216592_at**	**HGU-133 Plus2**
33	**SPANX**	**SPANXA1**	**220922_s_at**	**HGU-133 Plus2**
34	**SPANX**	**SPANXA1**	**224032_x_at**	**HGU-133 Plus2**
35	**SPANX**	**SPANXA1**	**220217_x_at**	**HGU-133 Plus2**
36	**SPANX**	**SPANXB1**	**220921_at**	**HGU-133 Plus2**
37	**SSX**	**SSX1**	**206627_s_at**	**HGU-133 Plus2**
38	**SSX**	**SSX1**	**206626_x_at**	**HGU-133 Plus2**
39	**SSX**	**SSX2**	**215881_x_at**	**HGU-133 Plus2**
40	**SSX**	**SSX2**	**207493_x_at**	**HGU-133 Plus2**
41	**SSX**	**SSX2**	**210497_x_at**	**HGU-133 Plus2**
42	**SSX**	**SSX2**	**216471_x_at**	**HGU-133 Plus2**
43	**SSX**	**SSX3**	**208528_x_at**	**HGU-133 Plus2**
44	**SSX**	**SSX3**	**207666_x_at**	**HGU-133 Plus2**
45	**SSX**	**SSX3**	**211731_x_at**	**HGU-133 Plus2**
46	**SSX**	**SSX3**	**211670_x_at**	**HGU-133 Plus2**
47	**SSX**	**SSX4**	**208586_s_at**	**HGU-133 Plus2**
48	**SSX**	**SSX4**	**210394_x_at**	**HGU-133 Plus2**
49	**SSX**	**SSX4**	**211425_x_at**	**HGU-133 Plus2**

[$]Affymetrix GeneChip Platform

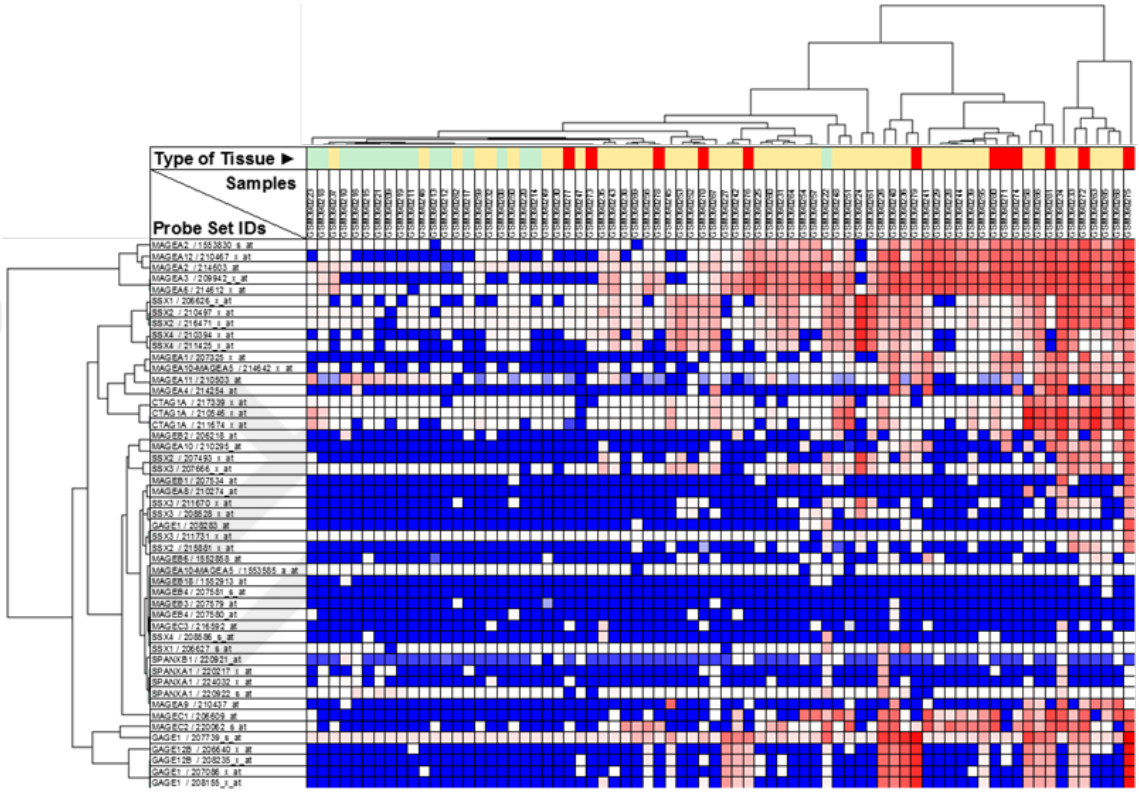


Figure 6.1: Expression of 49 CT-X probesets (Table 6-1) in melanoma. Heatmap generated by hierarchical clustering of dataset with accession number GSE15605 (Table 2-7). Details of clustering algorithm are given in section 2.2.9.1. Type of tissue is color coded where green, yellow and red represents normal skin, primary melanoma and malignant melanoma, respectively.

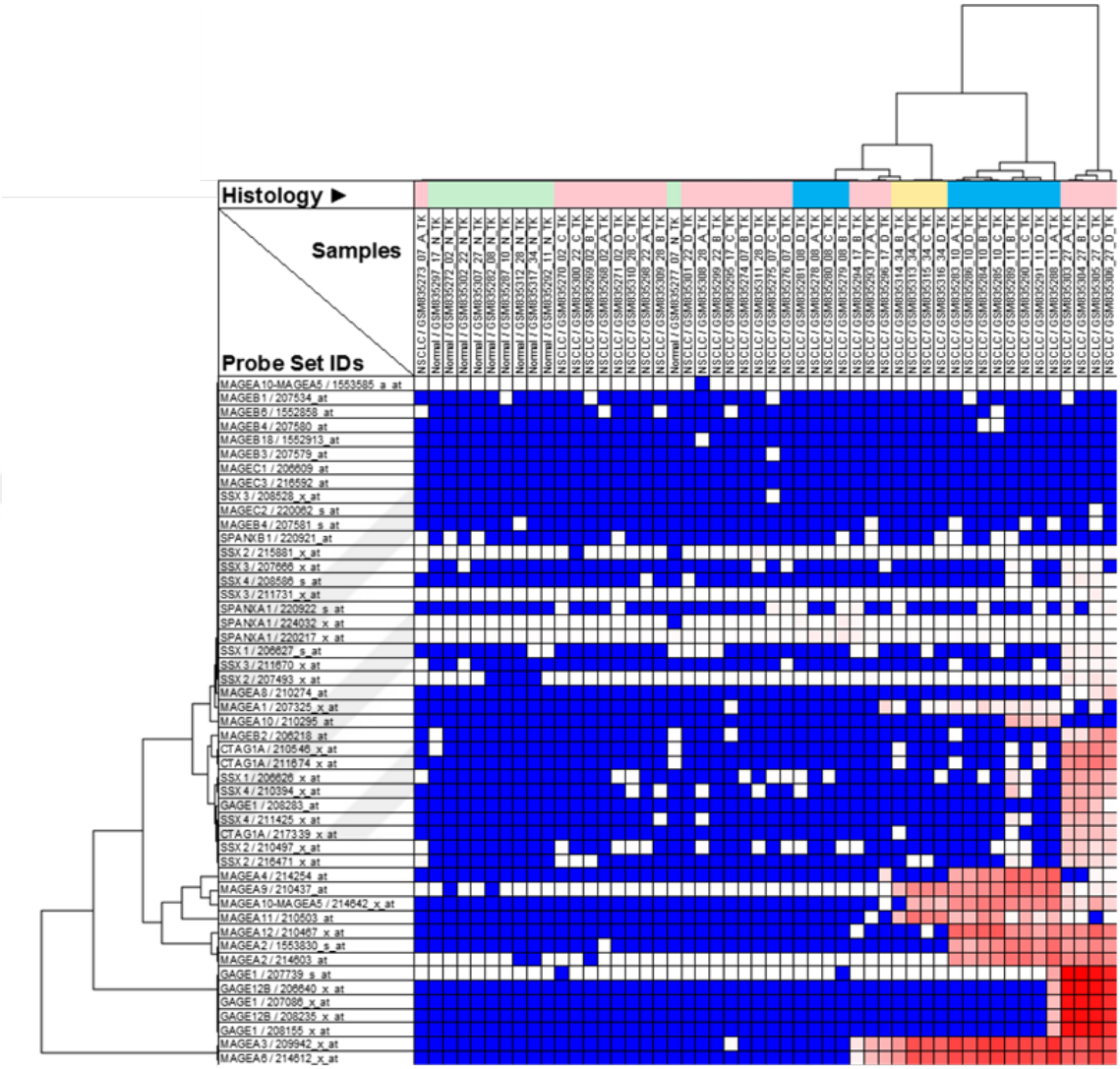


Figure 6.2: Expression of 49 CT-X probesets (Table 6-1) in NSCLC. Heatmap generated by hierarchical clustering of dataset with accession number GSE33532 (see Table 2-7). Details of clustering algorithm are given in section 2.2.9.1. Type of tissue is color coded where green, yellow, pink and blue represents normal lung, squamous carcinoma, adenocarcinoma and mixed histology, respectively.

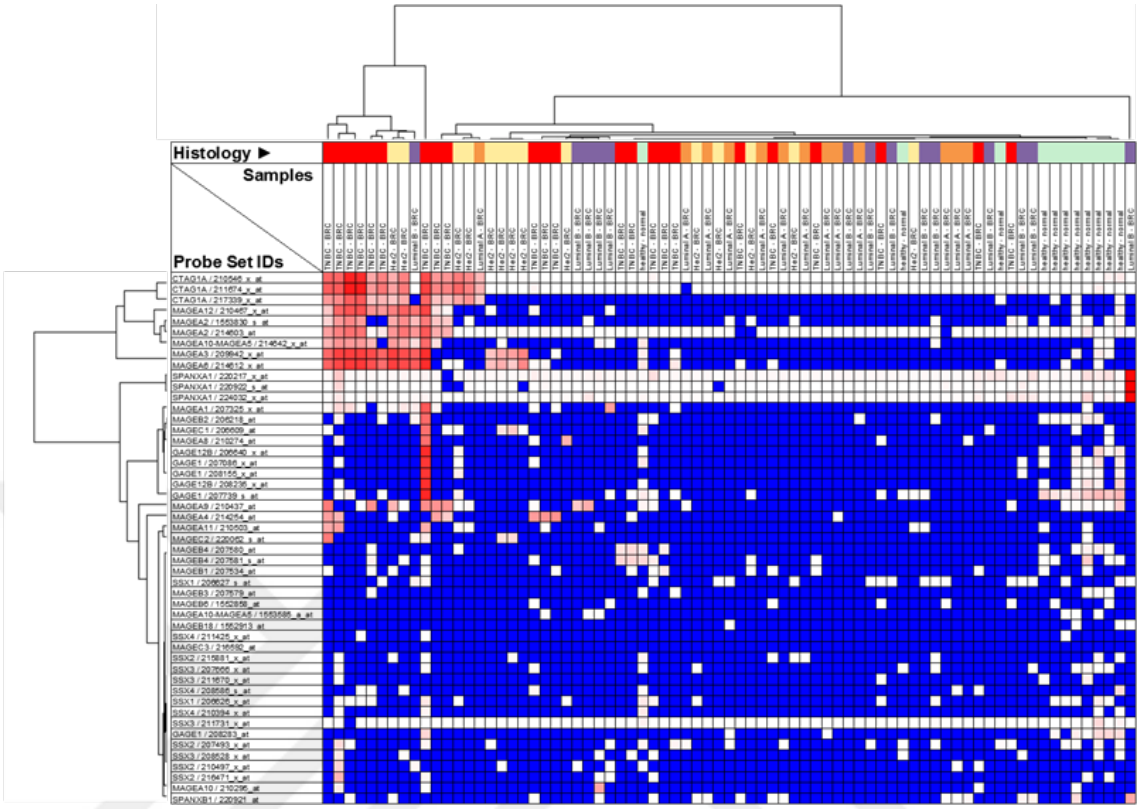


Figure 6.3: Expression of 49 CT-X probesets (Table 6-1) in breast cancer. Heatmap generated by hierarchical clustering of dataset with accession number GSE65194 (see Table 2-7). Details of clustering algorithm are given in section 2.2.9.1. Type of tissue is color coded where green, yellow, red, orange and purple represents normal breast, Her2 positive breast cancer, triple-negative breast cancer, luminal A and luminal B, respectively.

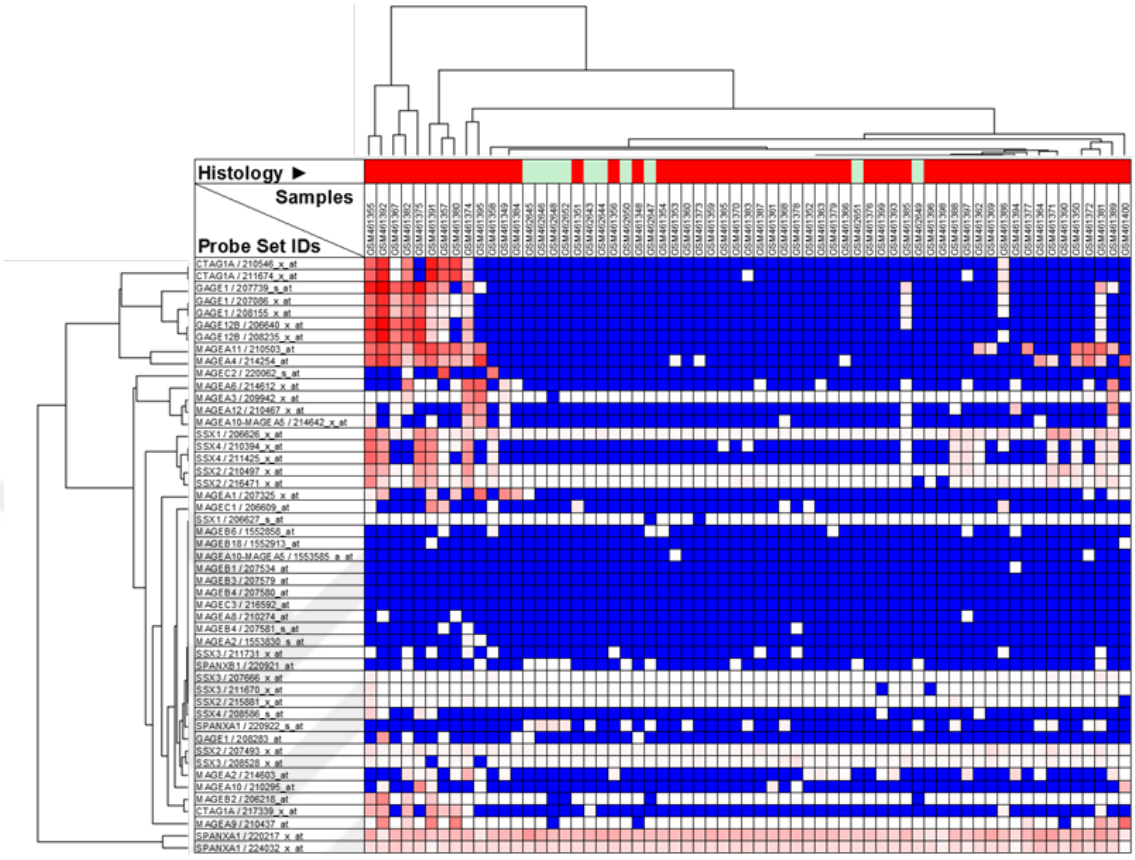


Figure 6.4: Expression of 49 CT-X probesets (Table 6-1) in ovarian cancer. Heatmap generated by hierarchical clustering of dataset with accession number GSE65194 (see Table 2-7). Details of clustering algorithm are given in section 2.2.9.1. Type of tissue is color coded where green and red represents normal ovaries and ovarian cancer, respectively.

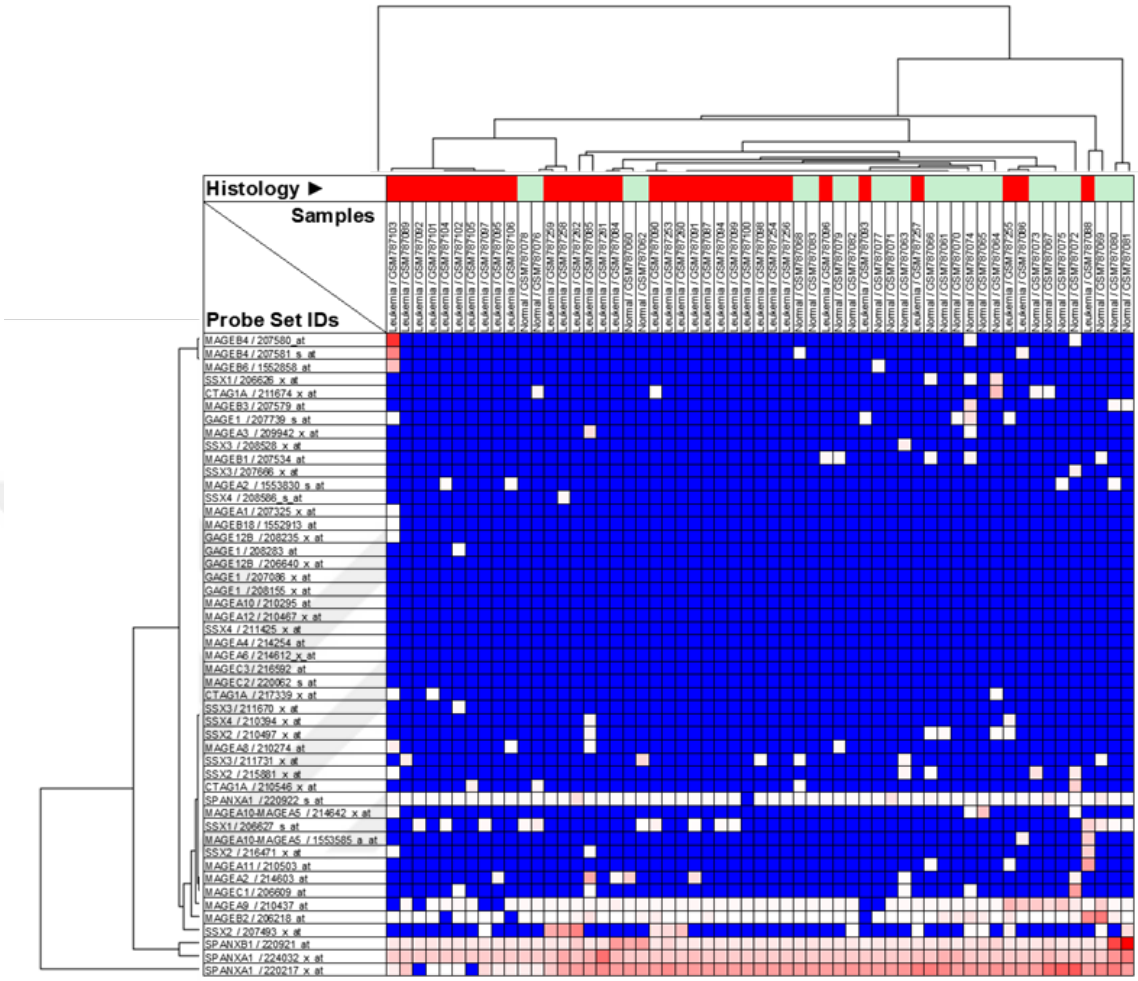


Figure 6.5 Expression of 49 CT-X probesets (Table 6-1) in CLL. Heatmap generated by hierarchical clustering of dataset with accession number GSE31048 (see Table 2-7). Details of clustering algorithm are given in section 2.2.9.1. Type of tissue is color coded where green and red represents healthy samples and CLL, respectively.

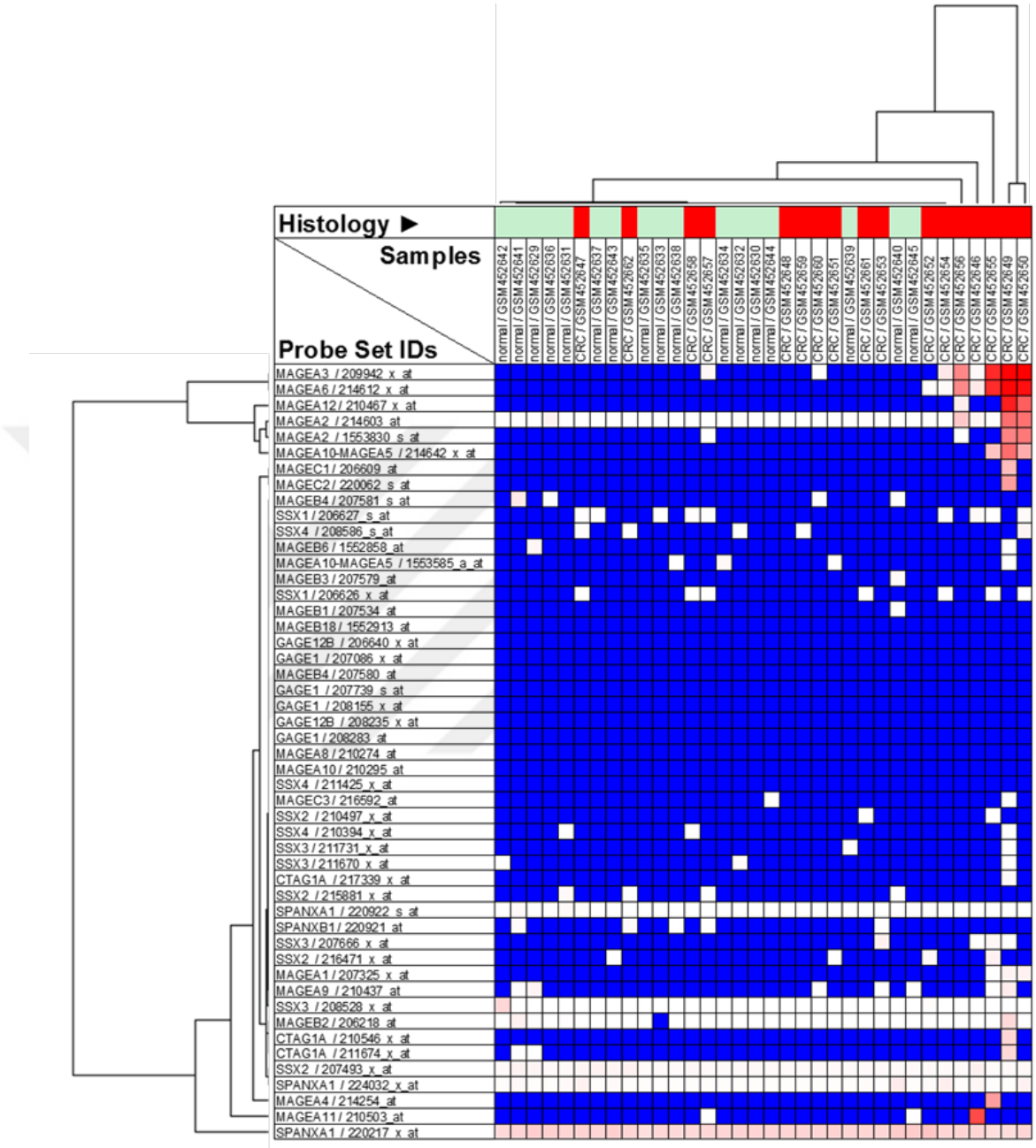


Figure 6.6 Expression of 49 CT-X probesets (Table 6-1) in Colorectal Cancer. Heatmap generated by hierarchical clustering of dataset with accession number GSE22598 (see Table 2-7). Details of clustering algorithm are given in section 2.2.9.1. Type of tissue is color coded where green and red represents healthy samples and tumor samples, respectively.

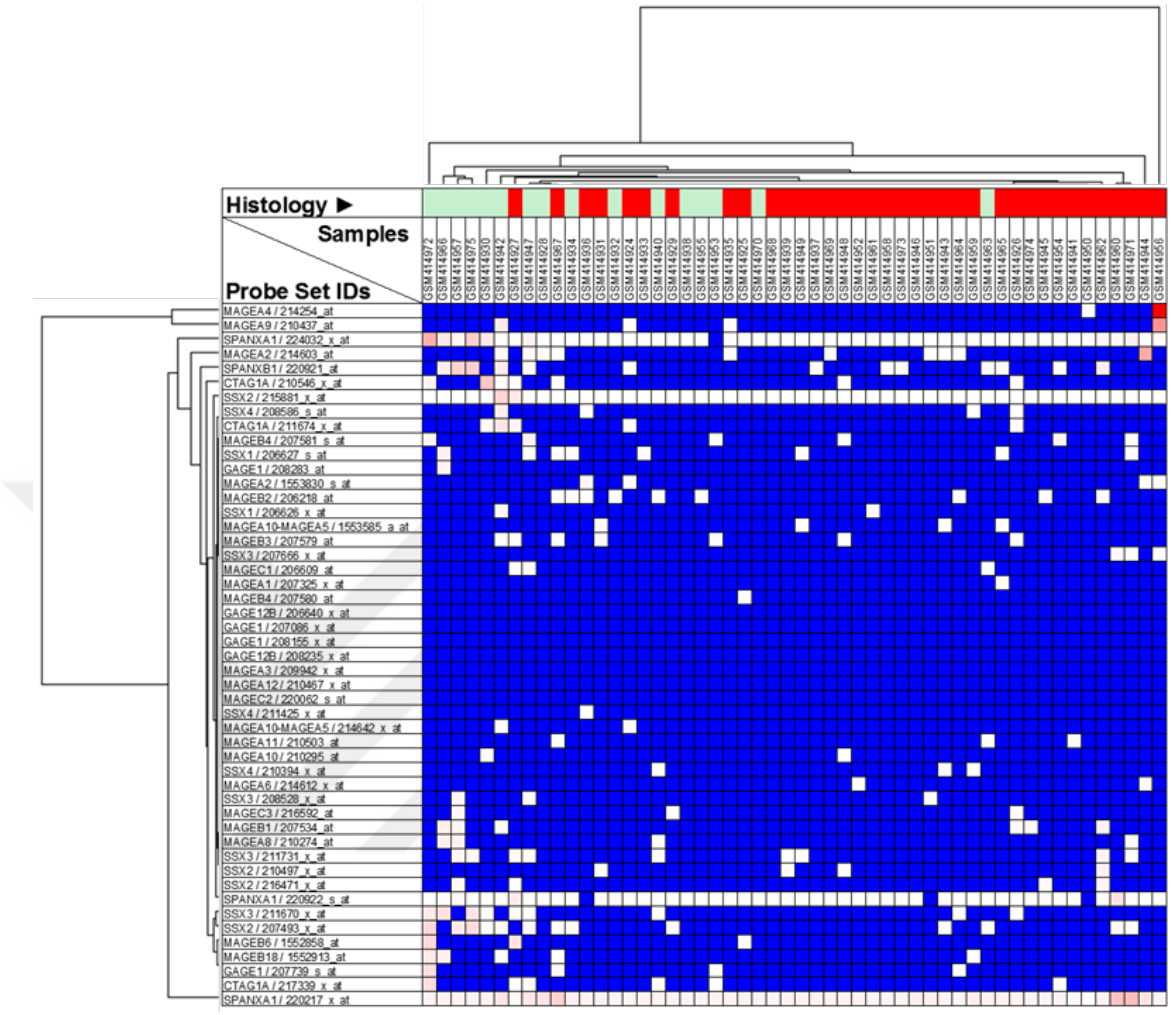


Figure 6.7 Expression of 49 CT-X probesets (Table 6-1) in Pancreas Cancer. Heatmap generated by hierarchical clustering of dataset with accession number GSE16515 (see Table 2-7). Details of clustering algorithm are given in section 2.2.9.1. Type of tissue is color coded where green and red represents healthy samples and pancreas cancer samples, respectively.

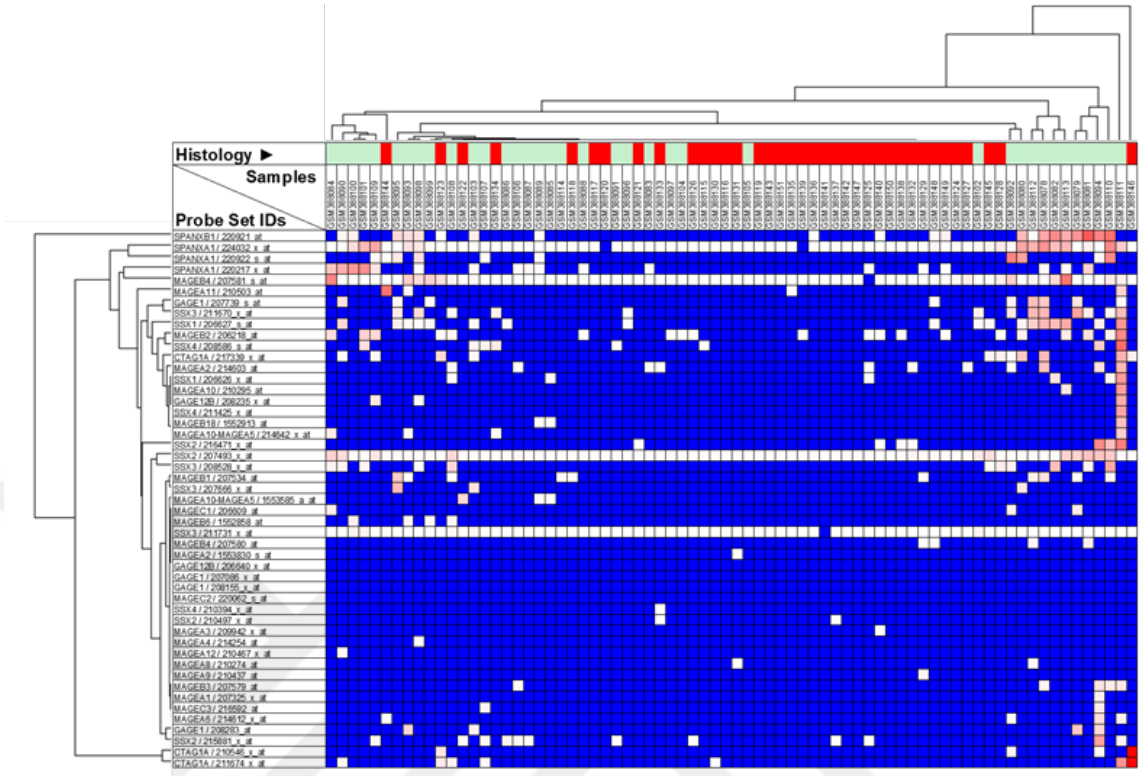


Figure 6.8 Expression of 49 CT-X probesets (Table 6-1) in Pancreas Cancer. Heatmap generated by hierarchical clustering of dataset with accession number GSE16515 (see Table 2-7). Details of clustering algorithm are given in section 2.2.9.1. Type of tissue is color coded where green and red represents healthy samples and pancreas cancer samples, respectively.

Appendix B

Tables for SNP Genotyping Analyses

Table 7-1: CT-X Gene Expression and Clinical Data of NSCLC Patients.

#	Tumor (LU #)	NY-ESO-1	LAGE-1	MAGE-A1	MAGE-A3	MAGE-A4	MAGE-A10	CT-7	SSX-2	SSX-4	Age	Sex	T Stage	Histology
1	31	(+++)	(+/-)	(+)	(+++)	(+++)	N.D.	(+)	N.D.	N.D.	57	M	2	SQCC
2	68	(+++)	N.D.	(+)	(+)	N.D.	N.D.	(-)	N.D.	N.D.	79	M	1	Adeno
3	87	N.D.	(-)	(+++)	(+++)	N.D.	N.D.	(+++)	N.D.	N.D.	65	M	1	Adeno
4	89	(+++)	(+++)	(+++)	(+++)	(+/-)	(+)	(+/-)	N.D.	N.D.	85	F	3	Adeno
5	111	N.D.	(-)	(+++)	(+++)	(+++)	N.D.	(+++)	N.D.	N.D.	55	M	4	Lgcell
6	131	(+++)	(+/-)	(+)	(+++)	(+++)	N.D.	N.D.	N.D.	N.D.	64	M	1	SQCC
7	168	(+++)	(+++)	(+++)	(+++)	(+++)	N.D.	(+)	N.D.	N.D.	70	F	1	SQCC
8	185	(+++)	(+++)	(+++)	(+++)	(+/-)	(-)	(-)	(++)	(-)	56	F	2	NSCLC
9	186	(+++)	(+++)	N.D.	(+++)	N.D.	N.D.	N.D.	N.D.	N.D.	61	F	3	Adeno
10	219	(+)	(+)	(+++)	(+++)	(+++)	N.D.	(-)	(+/-)	(+++)	54	M	2	SQCC
11	223	(+++)	(+++)	N.D.	(+++)	(-)	N.D.	N.D.	N.D.	N.D.	84	M	2	SQCC
12	649	(+++)	(+/-)	(+++)	(+++)	(+/-)	(+++)	(-)	(-)	N.D.	78	M	1	AdenoBAC
13	652	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)	(-)	(-)	N.D.	75	M	2	SQCC
14	658	(+/-)	(+++)	(+++)	(+++)	(+++)	(+/-)	N.D.	(+++)	N.D.	80	F	3	SQCC

15	726	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)	N.D.	(++)	(+++)	56	F	4	Adeno
16	736	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)	N.D.	(+/-)	(-)	62	F	2	Adeno
17	739	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)	N.D.	(++)	(+++)	60	F	2	Lgcell
18	745	(+++)	(+++)	(+++)	(+++)	(+++)	(+/-)	N.D.	(++)	(-)	63	M	2	SQCC
19	752	(+++)	(+++)	(+++)	(+++)	(+++)	(+/-)	N.D.	(-)	(-)	U	U	U	U
20	753	(++)	(+/-)	(+++)	(+++)	(+++)	(+/-)	N.D.	N.D.	(-)	U	U	U	U
21	759	(+++)	(+++)	(+++)	(+++)	(+++)	(+/-)	N.D.	N.D.	(-)	U	U	U	U
22	69	(-)	(-)	(-)	(-)	N.D.	N.D.	(-)	N.D.	N.D.	75	M	2	BAC
23	77	(-)	(-)	(-)	(-)	N.D.	N.D.	(-)	N.D.	N.D.	57	M	2	SQCC
24	88	(-)	(-)	(-)	(-)	N.D.	N.D.	(-)	N.D.	N.D.	78	M	1	Adeno
25	90	(-)	(-)	(-)	(-)	N.D.	N.D.	(-)	N.D.	N.D.	32	M	2	Adeno
26	108	(-)	(-)	(-)	(-)	N.D.	N.D.	(-)	N.D.	N.D.	74	M	1	Adeno
27	110	(-)	(-)	(-)	(-)	N.D.	N.D.	(-)	N.D.	N.D.	60	M	1	Adeno
28	112	(-)	(-)	(-)	(-)	N.D.	N.D.	(-)	N.D.	N.D.	72	M	2	SQCC
29	180	(-)	(-)	(-)	(-)	N.D.	N.D.	(-)	(-)	(-)	74	F	2	Adeno
30	183	(-)	(-)	(-)	(-)	N.D.	N.D.	(-)	(-)	(-)	64	F	1	BAC
31	191	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	N.D.	N.D.	62	F	1	Adeno
32	221	(-)	(-)	(-)	N.D.	(-)	N.D.	(-)	N.D.	N.D.	62	F	2	Adeno
33	225	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	(-)	79	F	2	Adeno
34	639	(-)	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	N.D.	68	F	2	AdenoBAC
35	656	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	N.D.	N.D.	83	F	1	AdenoBAC
36	670	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	N.D.	N.D.	82	F	2	AdenoBAC
37	692	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	56	M	1	AdenoBAC
38	693	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	U	U	U	U
39	694	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	U	U	U	U
40	698	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	63	F	2	Adeno
41	706	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	75	M	1	Adeno
42	707	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	58	F	1	Adeno

43	713	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	48	F	1	AdenoBAC
44	716	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	80	F	1	AdenoBAC
45	718	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	48	M	1	AdenoBAC
46	728	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	72	F	2	AdenoBAC
47	748	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	76	M	1	AdenoBAC
48	749	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	78	M	1	Adeno
49	751	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	70	F	2	AdenoBAC
50	763	(-)	(-)	(-)	(-)	(-)	(-)	N.D.	(-)	(-)	U	U	U	U

N.D.: not determined; U: unknown; M: male; F: female; SQCC: squamous cell carcinoma; Adeno: adenocarcinoma; BAC: bronchoalveolar carcinoma; AdenoBAC: adenocarcinoma with bronchoalveolar featuresl Lgcell: large cell carcinoma; NSCLC: other non-small cell carcinoma.

Table 7-2: Genotypes of NSCLC Patients.

#	Tumor (LU #)	***MTHFR 677 C>T*** **(rs1801133)**	***MTHFR 1298 A>C*** **(rs1801131)**	***MTR2756 A>G*** **(rs1805087)**	***MTRR66 A>G*** **(rs1801394)**	†***RFC80 G>A*** **(rs1051266)**
1	31	C/T	A/C	A/A	A/A	A/A
2	68	C/T	A/A	A/A	G/G	G/A
3	87	C/C	A/C	A/A	A/G	G/G
4	89	C/C	A/C	A/G	A/G	G/A
5	111	C/C	A/C	A/A	A/A	G/G
6	131	C/C	A/A	A/A	G/G	G/G
7	168	C/C	A/A	A/G	A/G	A/A
8	185	C/T	A/A	A/G	A/G	G/A
9	186	C/C	A/C	A/A	A/G	G/G
10	219	C/T	A/A	A/A	G/G	G/G

11	223	C/T	A/A	A/A	G/G	G/A
12	649	C/C	C/C	A/G	G/G	A/A
13	652	C/C	A/A	A/A	A/G	A/A
14	658	C/C	A/C	A/G	A/G	G/A
15	726	C/C	A/C	A/A	A/G	G/A
16	736	C/C	A/A	A/A	A/G	G/A
17	739	T/T	A/A	A/A	G/G	G/A
18	745	T/T	A/A	A/A	G/G	G/G
19	752	C/C	A/A	A/A	A/G	A/A
20	753	C/C	A/A	A/G	A/A	A/A
21	759	T/T	A/A	A/G	A/A	G/A
22	69	C/T	A/C	A/A	G/G	G/A
23	77	T/T	A/A	A/A	A/G	G/G
24	88	C/T	A/A	A/A	A/G	G/G
25	90	C/T	A/A	A/A	A/A	G/A
26	108	C/C	A/C	A/A	A/A	G/A
27	110	C/C	C/C	A/A	G/G	A/A
28	112	C/T	A/C	A/A	G/G	G/A
29	180	T/T	A/A	A/G	A/A	G/A
30	183	C/T	A/C	A/A	A/G	A/A
31	191	C/C	A/C	A/A	A/G	G/A
32	221	T/T	A/A	A/G	G/G	G/G
33	225	C/C	A/C	A/A	A/A	G/A
34	639	C/T	A/C	A/A	A/G	G/G
35	656	T/T	A/A	A/A	A/G	G/G
36	670	C/T	A/A	A/A	A/G	G/A
37	692	C/T	A/A	A/G	A/G	G/G
38	693	T/T	A/A	A/A	A/A	G/A

39	694	C/T	A/C	A/A	A/A	G/G
40	698	C/T	A/C	A/G	A/G	G/G
41	706	T/T	A/A	A/G	A/A	A/A
42	707	C/T	A/C	A/G	A/G	G/A
43	713	C/C	A/C	A/A	G/G	G/G
44	716	C/T	A/C	A/G	G/G	G/G
45	718	C/T	A/A	A/A	A/A	A/A
46	728	C/C	A/C	A/A	A/A	G/G
47	748	C/T	A/A	A/A	G/G	G/A
48	749	C/T	A/A	A/A	A/G	G/A
49	751	T/T	A/A	A/A	G/G	G/A
50	763	C/C	C/C	A/A	A/G	G/G

* Genotyped by TaqMan® SNP Genotyping Assay, †genotyped by PCR-RFLP.

Table 7-3: Genotypes of cell lines used in DNA damage measurement experiments.

Cell Line	CT-X Expression	Type of Measurement	*MTHFR 677 C>T	*MTHFR 1298 A>C	*MTR 2756 A>G	*MTRR 66 A>G	†RFC 80 G>A	Ref.
SK-LC-17	POSITIVE	Comet assay	T/T	A/A	A/A	A/A	G/A	Own typing results
COLO-205	NEGATIVE	Comet assay and H2AX Phosphorylation assay	C/T	A/A	G/G	G/G	G/A	
SW-620	NEGATIVE	H2AX Phosphorylation assay	T/T	A/A	A/A	A/A	G/A	
SK-BR-3	POSITIVE	H2AX Phosphorylation assay	C/C	C/C	A/A	G/G	G/G	
K-562	POSITIVE	H2AX Phosphorylation assay	C/C	A/C	A/G	N.A.	G/G	[140]

*TaqMan SNP genotyping results, †PCR-RFLP,

Table 7-4: MTHFR 677 C>T Genotype Population Distributions

Tissue/Genotype	C	C/T	T	Reference
NSCLC (Q-PCR) n=50	40%	40%	20%	Our Genotyping Results [141]
Lung Cancer n=1051	46%	45%	10%	[136]
Healthy Tissue n=1141	44%	45%	11%	[136]
Breast Cancer n=940	41%	45%	15%	[137]
Healthy Tissue n=782	37%	50%	13%	[137]
Lung Cancer n=550	44%	46%	10%	[139]
Healthy Tissue n=554	44%	45%	10%	[139]

Table 7-5: MTHFR 1298 A>C Genotype Population Distributions

Tissue/Genotype	A	A/C	C	Reference
NSCLC (Q-PCR) n=50	54%	40%	6%	Our Genotyping Results [141]
Lung Cancer n=1051	46%	44%	10%	[136]
Healthy Tissue n=1141	49%	43%	8%	[136]
Breast Cancer n=940	50%	42%	9%	[137]
Healthy Tissue n=782	51%	40%	9%	[137]
Lung Cancer n=550	47%	45%	8%	[139]
Healthy Tissue n=554	48%	45%	7%	[139]

Table 7-6: MTRR 66 A>G Genotype Population Distributions

Tissue/Genotype	A	A/G	G	Reference
NSCLC (Q-PCR) n=50	26%	44%	30%	Our Genotyping Results [141]
Lung Cancer n=1035	16%	48%	35%	[142]
Healthy Tissue n=1148	20%	48%	33%	[142]
Breast Cancer n=940	29%	48%	24%	[137]
Healthy Tissue n=782	31%	46%	23%	[137]

Table 7-7: MTR 2756 A>G Genotype Population Distributions

Tissue/Genotype	A	A/G	G	Reference
NSCLC (Q-PCR) n=50	72%	28%	0%	Our Genotyping Results [141]
Lung Cancer n=1035	72%	26%	2%	[142]
Healthy Tissue n=1148	72%	26%	2%	[142]
Lung Cancer n=515	62%	34%	4%	[138]
Healthy Tissue n=1029	68%	28%	4%	[138]

Table 7-8: RFC 80 G>A Genotype Population Distributions

Tissue/Genotype	G	G/A	A	Reference
NSCLC (PCR-RFLP) n=50	36%	44%	20%	Our Genotyping Results [141]
Breast Cancer n=937	32%	47%	21%	[137]
Healthy Tissue n=764	32%	45%	23%	[137]

Appendix C

Probeset Lists

Table 8-1: Probesets which define fibroblast and melanoma gene expression signature.

#	Probe Set ID	Gene Symbol	*p (Corr)	$FC (abs)	†FIB Vs. INVMEL	‡FIB (AVERAGE)	‡ INVMEL (AVERAGE)	‡FIB (MEDIAN)	‡ INVMEL (MEDIAN)
1	204830_x_at	PSG5	2.04E-13	55.26	up	11.74	5.96	12.14	6.14
2	201107_s_at	THBS1	5.12E-12	37.59	up	11.88	6.65	12.21	6.50
3	204933_s_at	TNFRSF11B	7.22E-07	31.18	up	11.67	6.70	12.11	6.66
4	204932_at	TNFRSF11B	1.74E-05	27.92	up	11.18	6.38	11.61	6.50
5	220794_at	GREM2	1.63E-12	26.40	up	10.03	5.30	10.21	5.02
6	201109_s_at	THBS1	3.96E-06	25.37	up	13.27	8.60	13.47	8.46
7	240509_s_at	GREM2	2.79E-15	25.25	up	9.25	4.59	9.51	4.50
8	203438_at	STC2	9.58E-11	24.06	up	11.44	6.85	11.49	6.91
9	209594_x_at	PSG9	3.87E-08	23.54	up	11.11	6.55	11.31	6.57
10	214701_s_at	FN1	3.08E-06	23.32	up	12.64	8.10	13.07	7.62
11	203439_s_at	STC2	2.81E-07	21.98	up	10.81	6.35	10.79	6.43

12	201108_s_at	THBS1	1.88E-08	21.44	up	13.48	9.05	13.77	8.93
13	209738_x_at	PSG6	3.87E-10	20.98	up	10.05	5.66	10.22	5.58
14	1555471_a_at	FMN2	4.94E-15	16.17	up	8.95	4.93	9.03	4.87
15	239336_at	THBS1	9.64E-13	16.09	up	8.91	4.91	9.10	4.89
16	208106_x_at	PSG6	1.2E-09	14.68	up	9.69	5.82	9.87	5.93
17	210261_at	KCNK2	1.41E-08	14.22	up	9.72	5.89	9.85	5.61
18	204948_s_at	FST	3.42E-07	14.05	up	12.33	8.52	12.63	8.31
19	204422_s_at	FGF2	1.22E-05	13.71	up	9.61	5.83	9.83	5.46
20	239468_at	MKX	1.64E-07	13.55	up	8.79	5.03	9.21	4.91
21	1554685_a_at	KIAA1199	2.64E-12	13.21	up	10.59	6.86	10.89	6.86
22	222899_at	ITGA11	7.02E-10	13.07	up	11.38	7.67	11.58	7.32
23	223734_at	MGARP	1.66E-09	11.29	up	8.30	4.81	8.46	4.71
24	223618_at	FMN2	8.51E-14	11.20	up	9.31	5.82	9.39	5.76
25	235086_at	THBS1	1.73E-09	11.13	up	10.00	6.52	10.24	6.34
26	234951_s_at	COL12A1	1.03E-11	10.68	up	9.27	5.85	9.46	5.90
27	235504_at	GREM2	1.91E-13	10.54	up	9.17	5.77	9.45	5.81
28	213568_at	OSR2	1.65E-05	10.48	up	10.74	7.35	10.96	7.27
29	203399_x_at	PSG3	3.08E-10	10.06	up	9.69	6.36	9.85	6.41
30	241902_at	MKX	2.15E-09	10.03	up	8.12	4.79	8.25	4.72

31	236129_at	GALNT5	4.14E-05	9.97	up	9.28	5.96	9.40	5.82
32	201308_s_at	"SEP11"	2.82E-06	9.81	up	10.56	7.26	10.75	7.33
33	211741_x_at	PSG3	3.45E-06	9.30	up	10.11	6.89	10.26	7.06
34	242722_at	LMO7	4.72E-06	9.15	up	10.35	7.16	10.49	6.76
35	244852_at	DSEL	1.57E-11	9.02	up	8.65	5.48	8.87	5.46
36	1569290_s_at	GRIA3	1.68E-08	8.90	up	8.01	4.85	8.13	4.70
37	210702_s_at	PTGIS	3.06E-09	8.90	up	8.73	5.58	8.90	5.37
38	233947_s_at	TBX5-AS1	7.23E-10	8.78	up	8.45	5.31	8.77	5.19
39	232531_at	EMX2OS	1.86E-08	8.69	up	10.21	7.09	10.40	6.77
40	202434_s_at	CYP1B1	3.15E-06	8.36	up	9.52	6.46	9.56	6.11
41	204421_s_at	FGF2	5.58E-11	8.33	up	9.94	6.88	10.15	6.82
42	219654_at	PTPLA	3.98E-05	7.98	up	10.40	7.40	10.42	7.41
43	231211_s_at	YIF1B	5.2E-07	7.88	up	9.07	6.09	9.11	5.91
44	241986_at	BMPER	4.4E-09	7.88	up	7.78	4.81	7.85	4.85
45	1557181_s_at	C11orf87	2.15E-06	7.87	up	7.56	4.58	7.81	4.54
46	220092_s_at	ANTXR1	6.62E-06	7.79	up	9.35	6.39	9.34	6.38
47	214927_at	ITGBL1	3.88E-05	7.77	up	10.87	7.91	11.08	7.79
48	232825_s_at	DSEL	1.1E-12	6.93	up	9.08	6.28	9.20	6.24
49	236532_at	C11orf87	3.97E-06	6.87	up	7.78	5.00	8.25	5.07

50	239218_at	PDE1C	1.29E-06	6.85	up	8.31	5.54	8.64	5.40
51	217220_at	LOC100287387	4.44E-07	6.66	up	8.34	5.60	8.49	5.33
52	209286_at	CDC42EP3	9.15E-09	6.62	up	11.09	8.36	11.29	8.39
53	206673_at	GPR176	2.3E-05	6.62	up	9.56	6.83	9.55	6.82
54	215821_x_at	PSG3	8.19E-12	5.94	up	9.51	6.94	9.52	6.84
55	1554819_a_at	ITGA11	2.54E-08	5.88	up	10.07	7.52	10.15	7.57
56	213498_at	CREB3L1	9.17E-08	5.84	up	8.88	6.33	9.11	6.23
57	225977_at	PCDH18	8.31E-06	5.84	up	9.58	7.03	9.67	6.98
58	211124_s_at	KITLG	4.76E-07	5.77	up	7.68	5.15	7.62	5.08
59	239919_at	TBX5-AS1	1.63E-08	5.65	up	8.44	5.94	8.67	5.97
60	230071_at	"SEP11"	1.6E-05	5.58	up	9.44	6.96	9.64	6.86
61	211892_s_at	PTGIS	7.21E-08	5.58	up	8.49	6.01	8.71	5.89
62	207345_at	FST	6.56E-07	5.55	up	9.35	6.87	9.42	6.80
63	203895_at	PLCB4	1.59E-05	5.52	up	8.77	6.31	8.83	6.25
64	207733_x_at	PSG9	1.19E-05	5.32	up	10.47	8.05	10.57	8.03
65	206404_at	FGF9	5.87E-07	4.89	up	7.77	5.48	7.88	5.40
66	206814_at	NGF	1.09E-06	4.81	up	8.91	6.64	9.06	6.45
67	233888_s_at	SRGAP1	5.93E-05	4.66	up	9.52	7.30	9.44	7.47
68	1553565_s_at	DDAH1	5.31E-10	4.63	up	8.77	6.56	8.84	6.65

69	213661_at	PAMR1	2.06E-08	4.62	up	10.41	8.20	10.51	8.32
70	230750_at		5.59E-13	4.48	up	7.87	5.71	7.85	5.68
71	1558636_s_at	ADAMTS5	1.02E-05	4.39	up	7.20	5.06	7.43	4.94
72	205818_at	DBC1	3.22E-05	4.38	up	9.04	6.91	9.26	6.95
73	209647_s_at	SOCS5	9.78E-05	4.26	up	10.71	8.62	10.85	8.57
74	213823_at	HOXA11	9.84E-05	4.24	up	8.99	6.90	9.06	6.87
75	205897_at	NFATC4	6.1E-06	4.22	up	9.14	7.06	9.33	6.93
76	230543_at		2.55E-06	4.13	up	7.93	5.88	7.93	6.07
77	201615_x_at	CALD1	2.51E-06	4.09	up	10.66	8.63	10.78	8.58
78	219308_s_at	AK5	5.13E-05	4.07	up	8.93	6.90	9.01	6.75
79	213975_s_at	LYZ	7.92E-05	100.02	down	4.37	11.02	4.38	11.19
80	208894_at	HLA-DRA	2.02E-06	72.07	down	6.18	12.35	6.04	12.29
81	210982_s_at	HLA-DRA	1.94E-05	66.12	down	6.51	12.56	6.42	12.80
82	200795_at	SPARCL1	9.81E-05	65.95	down	4.88	10.92	4.83	11.15
83	211990_at	HLA-DPA1	3.09E-08	54.73	down	6.83	12.61	6.65	12.51
84	209312_x_at	HLA-DRB1	3.72E-07	50.41	down	6.95	12.60	6.63	12.90
85	209619_at	CD74	1.23E-07	49.54	down	7.60	13.23	7.50	13.47
86	215193_x_at	HLA-DRB1	9.28E-05	44.96	down	6.82	12.31	6.58	12.63
87	216834_at	RGS1	1.62E-05	43.03	down	4.04	9.47	4.03	9.61

88	202345_s_at	FABP5	1.29E-07	39.41	down	6.58	11.88	6.55	11.63
89	201721_s_at	LAPTM5	2.46E-06	36.18	down	7.63	12.81	7.52	13.15
90	201859_at	SRGN	5.42E-13	33.14	down	6.99	12.04	6.78	11.97
91	202291_s_at	MGP	3.03E-07	32.95	down	7.34	12.38	6.70	12.79
92	217757_at	A2M	1.5E-05	24.69	down	7.36	11.98	7.17	12.02
93	204122_at	TYROBP	8.15E-05	23.01	down	4.92	9.44	4.95	9.35
94	228155_at	FAM213A	3.75E-10	22.67	down	5.77	10.28	5.63	10.45
95	201690_s_at	TPD52	8.08E-12	21.59	down	4.94	9.37	4.70	9.51
96	202546_at	VAMP8	6.92E-07	21.16	down	5.94	10.35	5.65	10.85
97	221872_at	RARRES1	1.36E-06	17.92	down	4.84	9.01	4.88	9.12
98	224435_at	FAM213A	2.2E-06	17.81	down	6.08	10.24	5.86	10.37
99	219049_at	CSGALNACT1	6.27E-07	13.61	down	6.27	10.04	6.02	10.21
100	218319_at	PELI1	3.14E-06	12.63	down	5.91	9.57	5.84	9.51
101	201689_s_at	TPD52	3.18E-05	11.93	down	3.98	7.55	3.73	7.57
102	203561_at	FCGR2A	6.45E-06	11.91	down	5.61	9.18	5.62	9.19
103	201656_at	ITGA6	2.02E-06	10.86	down	6.66	10.10	6.65	10.10
104	203382_s_at	APOE	3.73E-05	9.09	down	7.40	10.58	7.26	10.49
105	208891_at	DUSP6	8.9E-05	8.89	down	8.55	11.70	8.98	11.65
106	203148_s_at	TRIM14	1.71E-06	8.83	down	6.42	9.57	6.35	9.77

107	202241_at	TRIB1	5.87E-07	8.60	down	7.34	10.45	7.35	10.24
108	221802_s_at	KIAA1598	9.18E-05	8.60	down	5.24	8.34	5.09	8.18
109	208892_s_at	DUSP6	3.53E-05	8.23	down	8.08	11.12	8.43	11.09
110	211340_s_at	MCAM	4.34E-05	7.18	down	7.86	10.70	7.49	10.59
111	203765_at	GCA	4.06E-05	7.01	down	4.77	7.58	4.62	7.79
112	218404_at	SNX10	3.65E-07	6.99	down	5.29	8.09	5.14	8.02
113	1552701_a_at	CARD16	2.16E-06	6.52	down	6.56	9.26	6.48	9.28
114	203455_s_at	SAT1	2.09E-12	5.74	down	9.28	11.80	9.21	11.67
115	209087_x_at	MCAM	5.69E-06	5.64	down	7.80	10.29	7.70	10.37
116	209099_x_at	JAG1	3.07E-05	5.47	down	7.65	10.10	7.47	10.34
117	210592_s_at	SAT1	1.39E-09	5.46	down	9.48	11.93	9.39	11.82
118	228314_at	LRRC8C	3.13E-06	5.26	down	6.00	8.39	6.02	8.40
119	221899_at	N4BP2L2	9.91E-05	5.22	down	6.20	8.58	6.14	8.54
120	225606_at	BCL2L11	5.28E-05	4.24	down	7.44	9.53	7.37	9.50
121	209846_s_at	BTN3A2	2.83E-05	4.17	down	7.36	9.42	7.27	9.32
122	214093_s_at	FUBP1	3.83E-05	4.10	down	6.35	8.38	6.26	8.48

*Bonferroni corrected p-value, [$]FC-fold change (fibroblast vs. invasive melanoma), [†]Up or down regulation based on fibroblast vs. invasive melanoma comparison, [‡]Average of median gene expression value of fibroblast or invasive melanoma samples.

Table 8-2: Gene list generated for subgrouping of melanoma

#	Probeset ID	Gene Symbol	Gene Description	*Regulation
1	200665_s_at	SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	up
2	200696_s_at	GSN	gelsolin	up
3	201141_at	GPNMB	glycoprotein (transmembrane) nmb	up
4	201147_s_at	TIMP3	TIMP metallopeptidase inhibitor 3	up
5	201148_s_at	TIMP3	TIMP metallopeptidase inhibitor 3	up
6	201149_s_at	TIMP3	TIMP metallopeptidase inhibitor 3	up
7	201150_s_at	TIMP3	TIMP metallopeptidase inhibitor 3	up
8	201278_at	DAB2	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	up
9	201279_s_at	DAB2	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	up
10	201280_s_at	DAB2	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	up
11	201525_at	APOD	apolipoprotein D	up
12	201739_at	SGK1	serum/glucocorticoid regulated kinase 1	up
13	201860_s_at	PLAT	plasminogen activator, tissue	up

14	202036_s_at	SFRP1	secreted frizzled-related protein 1	up
15	202037_s_at	SFRP1	secreted frizzled-related protein 1	up
16	202132_at	WWTR1	WW domain containing transcription regulator 1	up
17	202133_at	WWTR1	WW domain containing transcription regulator 1	up
18	202191_s_at	GAS7	growth arrest-specific 7	up
19	202192_s_at	GAS7	growth arrest-specific 7	up
20	202202_s_at	LAMA4	laminin, alpha 4	up
21	202341_s_at	TRIM2	tripartite motif-containing 2	up
22	202342_s_at	TRIM2	tripartite motif-containing 2	up
23	202376_at	SERPINA3	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	up
24	202421_at	IGSF3	immunoglobulin superfamily, member 3	up
25	202450_s_at	CTSK	cathepsin K	up
26	202478_at	TRIB2	tribbles homolog 2 (Drosophila)	up
27	202479_s_at	TRIB2	tribbles homolog 2 (Drosophila)	up
28	202898_at	SDC3	syndecan 3	up
29	202947_s_at	GYPC	glycophorin C (Gerbich blood group)	up
30	203083_at	THBS2	thrombospondin 2	up
31	203234_at	UPP1	uridine phosphorylase 1	up
32	203299_s_at	AP1S2	adaptor-related protein complex 1, sigma 2 subunit	up

33	203300_x_at	AP1S2	adaptor-related protein complex 1, sigma 2 subunit	up
34	203348_s_at	ETV5	ets variant 5	up
35	203349_s_at	ETV5	ets variant 5	up
36	203400_s_at	TF	transferrin	up
37	203518_at	LYST	lysosomal trafficking regulator	up
38	203542_s_at	KLF9	Kruppel-like factor 9	up
39	203543_s_at	KLF9	Kruppel-like factor 9	up
40	203607_at	INPP5F	inositol polyphosphate-5-phosphatase F	up
41	203632_s_at	GPRC5B	G protein-coupled receptor, family C, group 5, member B	up
42	203723_at	ITPKB	inositol 1,4,5-trisphosphate 3-kinase B	up
43	203729_at	EMP3	epithelial membrane protein 3	up
44	203759_at	ST3GAL4	ST3 beta-galactoside alpha-2,3-sialyltransferase 4	up
45	203827_at	WIPI1	WD repeat domain, phosphoinositide interacting 1	up
46	203845_at	KAT2B	K(lysine) acetyltransferase 2B	up
47	203979_at	CYP27A1	cytochrome P450, family 27, subfamily A, polypeptide 1	up
48	204011_at	SPRY2	sprouty homolog 2 (Drosophila)	up
49	204015_s_at	DUSP4	dual specificity phosphatase 4	up
50	204040_at	RNF144A	ring finger protein 144A	up
51	204086_at	PRAME	preferentially expressed antigen in melanoma	up

52	204137_at	GPR137B	G protein-coupled receptor 137B	up
53	204197_s_at	RUNX3	runt-related transcription factor 3	up
54	204198_s_at	RUNX3	runt-related transcription factor 3	up
55	204271_s_at	EDNRB	endothelin receptor type B	up
56	204273_at	EDNRB	endothelin receptor type B	up
57	204457_s_at	GAS1	growth arrest-specific 1	up
58	204467_s_at	SNCA	synuclein, alpha (non A4 component of amyloid precursor)	up
59	204469_at	PTPRZ1	protein tyrosine phosphatase, receptor-type, Z polypeptide 1	up
60	204527_at	MYO5A	myosin VA (heavy chain 12, myoxin)	up
61	204591_at	CHL1	cell adhesion molecule with homology to L1CAM (close homolog of L1)	up
62	204627_s_at	ITGB3	integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)	up
63	204629_at	PARVB	parvin, beta	up
64	204638_at	ACP5	acid phosphatase 5, tartrate resistant	up
65	204724_s_at	COL9A3	collagen, type IX, alpha 3	up
66	204857_at	MAD1L1	MAD1 mitotic arrest deficient-like 1 (yeast)	up
67	204955_at	SRPX	sushi-repeat-containing protein, X-linked	up
68	205030_at	FABP7	fatty acid binding protein 7, brain	up
69	205097_at	SLC26A2	solute carrier family 26 (sulfate transporter), member 2	up
70	205174_s_at	QPCT	glutaminyl-peptide cyclotransferase	up

71	205227_at	IL1RAP	interleukin 1 receptor accessory protein	up
72	205334_at	S100A1	S100 calcium binding protein A1	up
73	205337_at	DCT	dopachrome tautomerase (dopachrome delta-isomerase, tyrosine-related protein 2)	up
74	205338_s_at	DCT	dopachrome tautomerase (dopachrome delta-isomerase, tyrosine-related protein 2)	up
75	205433_at	BCHE	butyrylcholinesterase	up
76	205681_at	BCL2A1	BCL2-related protein A1	up
77	205694_at	TYRP1	tyrosinase-related protein 1	up
78	205862_at	GREB1	growth regulation by estrogen in breast cancer 1	up
79	205954_at	RXRG	retinoid X receptor, gamma	up
80	206256_at	CPN1	carboxypeptidase N, polypeptide 1	up
81	206376_at	SLC6A15	solute carrier family 6 (neutral amino acid transporter), member 15	up
82	206426_at	MLANA	melan-A	up
83	206427_s_at	MLANA	melan-A	up
84	206498_at	OCA2	oculocutaneous albinism II	up
85	206560_s_at	MIA	melanoma inhibitory activity	up
86	206630_at	TYR	tyrosinase (oculocutaneous albinism IA)	up
87	206696_at	GPR143	G protein-coupled receptor 143	up
88	206701_x_at	EDNRB	endothelin receptor type B	up
89	206898_at	CDH19	cadherin 19, type 2	up

90	207038_at	SLC16A6	solute carrier family 16, member 6 (monocarboxylic acid transporter 7)	up
91	207144_s_at	CITED1	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 1	up
92	207233_s_at	DZIP3	microphthalmia-associated transcription factor	up
93	207323_s_at	MBP	myelin basic protein	up
94	207827_x_at	SNCA	synuclein, alpha (non A4 component of amyloid precursor)	up
95	208370_s_at	RCAN1	regulator of calcineurin 1	up
96	209072_at	MBP	myelin basic protein	up
97	209086_x_at	MCAM	melanoma cell adhesion molecule	up
98	209087_x_at	MCAM	melanoma cell adhesion molecule	up
99	209167_at	GPM6B	glycoprotein M6B	up
100	209168_at	GPM6B	glycoprotein M6B	up
101	209169_at	GPM6B	glycoprotein M6B	up
102	209170_s_at	GPM6B	glycoprotein M6B	up
103	209283_at	CRYAB	crystallin, alpha B	up
104	209354_at	TNFRSF14	tumor necrosis factor receptor superfamily, member 14 (herpesvirus entry mediator)	up
105	209365_s_at	ECM1	extracellular matrix protein 1	up
106	209392_at	ENPP2	ectonucleotide pyrophosphatase/phosphodiesterase 2	up
107	209583_s_at	CD200	CD200 molecule	up
108	209610_s_at	SLC1A4	solute carrier family 1 (glutamate/neutral amino acid transporter), member 4	up

109	209686_at	S100B	S100 calcium binding protein B	up
110	209842_at	SOX10	SRY (sex determining region Y)-box 10	up
111	209843_s_at	SOX10	SRY (sex determining region Y)-box 10	up
112	209848_s_at	SILV	silver homolog (mouse)	up
113	209942_x_at	MAGEA3	melanoma antigen family A, 3	up
114	210139_s_at	PMP22	peripheral myelin protein 22	up
115	210198_s_at	PLP1	proteolipid protein 1	up
116	210467_x_at	MAGEA12	melanoma antigen family A, 12	up
117	210495_x_at	FN1	fibronectin 1	up
118	210613_s_at	SYNGR1	synaptogyrin 1	up
119	210757_x_at	DAB2	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	up
120	210869_s_at	MCAM	melanoma cell adhesion molecule	up
121	210942_s_at	ST3GAL6	ST3 beta-galactoside alpha-2,3-sialyltransferase 6	up
122	210943_s_at	LYST	lysosomal trafficking regulator	up
123	210944_s_at	CAPN3	calpain 3, (p94)	up
124	211067_s_at	GAS7	growth arrest-specific 7	up
125	211340_s_at	MCAM	melanoma cell adhesion molecule	up
126	211546_x_at	SNCA	synuclein, alpha (non A4 component of amyloid precursor)	up
127	211890_x_at	CAPN3	calpain 3, (p94)	up

128	212190_at	SERPINE2	serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 2	up
129	212419_at	ZCCHC24	zinc finger, CCHC domain containing 24	up
130	212423_at	ZCCHC24	zinc finger, CCHC domain containing 24	up
131	212667_at	SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	up
132	212730_at	SYNM	synemin, intermediate filament protein	up
133	212793_at	DAAM2	dishevelled associated activator of morphogenesis 2	up
134	212810_s_at	SLC1A4	solute carrier family 1 (glutamate/neutral amino acid transporter), member 4	up
135	213005_s_at	KANK1	KN motif and ankyrin repeat domains 1	up
136	213113_s_at	SLC43A3	solute carrier family 43, member 3	up
137	213139_at	SNAI2	snail homolog 2 (Drosophila)	up
138	213169_at	SEMA5A	sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5A	up
139	213236_at	SASH1	SAM and SH3 domain containing 1	up
140	213245_at	ADCY1	adenylate cyclase 1 (brain)	up
141	213355_at	ST3GAL6	ST3 beta-galactoside alpha-2,3-sialyltransferase 6	up
142	213638_at	PHACTR1	phosphatase and actin regulator 1	up
143	213836_s_at	WIPI1	WD repeat domain, phosphoinositide interacting 1	up
144	214063_s_at	TF	transferrin	up
145	214297_at	CSPG4	chondroitin sulfate proteoglycan 4	up

146	214321_at	NOV	nephroblastoma overexpressed gene	up
147	214475_x_at	CAPN3	calpain 3, (p94)	up
148	214612_x_at	MAGEA6	melanoma antigen family A, 6	up
149	215253_s_at	RCAN1	regulator of calcineurin 1	up
150	216442_x_at	FN1	fibronectin 1	up
151	216512_s_at	DCT	dopachrome tautomerase (dopachrome delta-isomerase, tyrosine-related protein 2)	up
152	216513_at	DCT	dopachrome tautomerase (dopachrome delta-isomerase, tyrosine-related protein 2)	up
153	217553_at	MGC87042	STEAP family protein MGC87042	up
154	217757_at	A2M	alpha-2-macroglobulin	up
155	218196_at	OSTM1	osteopetrosis associated transmembrane protein 1	up
156	218404_at	SNX10	sorting nexin 10	up
157	218678_at	NES	nestin	up
158	218839_at	HEY1	hairy/enhancer-of-split related with YRPW motif 1	up
159	219282_s_at	TRPV2	transient receptor potential cation channel, subfamily V, member 2	up
160	219412_at	RAB38	RAB38, member RAS oncogene family	up
161	219578_s_at	CPEB1	cytoplasmic polyadenylation element binding protein 1	up
162	219584_at	PLA1A	phospholipase A1 member A	up
163	219926_at	POPDC3	popeye domain containing 3	up
164	220178_at	C19orf28	chromosome 19 open reading frame 28	up

165	220245_at	SLC45A2	solute carrier family 45, member 2	up
166	220425_x_at	ROPN1B	ropporin, rhophilin associated protein 1B	up
167	221489_s_at	SPRY4	sprouty homolog 4 (Drosophila)	up
168	221558_s_at	LEF1	lymphoid enhancer-binding factor 1	up
169	221577_x_at	GDF15 /// LOC100292463	growth differentiation factor 15 /// similar to growth differentiation factor 15	up
170	221644_s_at	SLC45A2	solute carrier family 45, member 2	up
171	221911_at	ETV1	ets variant 1	up
172	222116_s_at	TBC1D16	TBC1 domain family, member 16	up
173	222199_s_at	BIN3	bridging integrator 3	up
174	36711_at	MAFF	v-maf musculoaponeurotic fibrosarcoma oncogene homolog F (avian)	up
175	37966_at	PARVB	parvin, beta	up
176	40560_at	TBX2	T-box 2	up
177	44783_s_at	HEY1	hairy/enhancer-of-split related with YRPW motif 1	up
178	47550_at	LZTS1	leucine zipper, putative tumor suppressor 1	up
179	202986_at	ARNT2	aryl-hydrocarbon receptor nuclear translocator 2	up
180	202185_at	PLOD3	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3	up
181	203167_at	TIMP2	TIMP metallopeptidase inhibitor 2	up
182	204653_at	TFAP2A	transcription factor AP-2 alpha (activating enhancer binding protein 2 alpha)	up
183	207808_s_at	PROS1	protein S (alpha)	up

184	202260_s_at	STXBP1	syntaxin binding protein 1	up
185	202454_s_at	ERBB3	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)	up
186	204014_at	DUSP4	dual specificity phosphatase 4	up
187	203304_at	BAMBI	BMP and activin membrane-bound inhibitor homolog (Xenopus laevis)	up
188	204542_at	ST6GALNAC2	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2	up
189	213417_at	TBX2	T-box 2	up
190	217867_x_at	BACE2	beta-site APP-cleaving enzyme 2	up
191	204466_s_at	SNCA	synuclein, alpha (non A4 component of amyloid precursor)	up
192	212070_at	GPR56	G protein-coupled receptor 56	up
193	203217_s_at	ST3GAL5	ST3 beta-galactoside alpha-2,3-sialyltransferase 5	up
194	205376_at	INPP4B	inositol polyphosphate-4-phosphatase, type II, 105kDa	up
195	212811_x_at	SLC1A4	solute carrier family 1 (glutamate/neutral amino acid transporter), member 4	up
196	201015_s_at	JUP	junction plakoglobin	down
197	201425_at	ALDH2	aldehyde dehydrogenase 2 family (mitochondrial)	down
198	201650_at	KRT19	keratin 19	down
199	201839_s_at	EPCAM	epithelial cell adhesion molecule	down
200	203066_at	CHST15	carbohydrate (N-acetylgalactosamine 4-sulfate 6-O) sulfotransferase 15	down
201	203407_at	PPL	periplakin	down
202	207076_s_at	ASS1	argininosuccinate synthase 1	down

203	208650_s_at	CD24	CD24 molecule	down
204	208651_x_at	CD24	CD24 molecule	down
205	208792_s_at	CLU	clusterin	down
206	208998_at	UCP2	uncoupling protein 2 (mitochondrial, proton carrier)	down
207	209008_x_at	KRT8	keratin 8	down
208	209771_x_at	CD24	CD24 molecule	down
209	209772_s_at	CD24	CD24 molecule	down
210	210715_s_at	SPINT2	serine peptidase inhibitor, Kunitz type, 2	down
211	212148_at	PBX1	pre-B-cell leukemia homeobox 1	down
212	212151_at	PBX1	pre-B-cell leukemia homeobox 1	down
213	212242_at	TUBA4A	tubulin, alpha 4a	down
214	216379_x_at	CD24	CD24 molecule	down
215	216438_s_at	TMSB4X /// TMSL3	thymosin beta 4, X-linked /// thymosin-like 3	down
216	218180_s_at	EPS8L2	EPS8-like 2	down
217	218237_s_at	SLC38A1	solute carrier family 38, member 1	down
218	220161_s_at	EPB41L4B	erythrocyte membrane protein band 4.1 like 4B	down
219	266_s_at	CD24	CD24 molecule	down
220	33322_i_at	SFN	stratifin	down
221	33323_r_at	SFN	stratifin	down

222	202016_at	MEST	mesoderm specific transcript homolog (mouse)	down
223	204679_at	KCNK1	potassium channel, subfamily K, member 1	down

* Regulation between group A (melanoma cell line) vs. group B (lung, colorectal, breast, hepatocellular, pancreatic and gastric cancers)

Table 8-3: The most differentially expressed probe sets between cluster B vs. A in melanoma.

#	ProbeSetID	Gene Symbol	*p (Corr)	$FC (abs)	†Regulation
1	209842_at	SOX10	2.19E-17	41.23	up
2	205338_s_at	DCT	2.84E-14	56.57	up
3	220425_x_at	ROPN1B	1.82E-13	14.02	up
4	206560_s_at	MIA	2.97E-13	23.92	up
5	205337_at	DCT	3.60E-13	32.37	up
6	216512_s_at	DCT	7.64E-13	37.24	up

7	204086_at	PRAME	7.77E-13	32.88	up
8	209170_s_at	GPM6B	1.99E-12	59.01	up
9	47550_at	LZTS1	2.58E-12	5.82	up
10	204271_s_at	EDNRB	4.99E-12	40.98	up
11	209686_at	S100B	5.67E-12	70.22	up
12	210198_s_at	PLP1	7.44E-12	48.69	up
13	216513_at	DCT	7.93E-12	6.59	up
14	221215_s_at	RIPK4	2.17E-11	8.72	up
15	209072_at	MBP	2.44E-11	26.61	up
16	202421_at	IGSF3	2.57E-11	10.38	up
17	206701_x_at	EDNRB	5.45E-11	36.84	up
18	209167_at	GPM6B	7.01E-11	39.85	up
19	218931_at	RAB17	7.94E-11	4.48	up
20	210943_s_at	LYST	9.78E-11	6.38	up
21	203299_s_at	AP1S2	1.22E-10	5.26	up
22	206256_at	CPN1	1.65E-10	4.70	up
23	205334_at	S100A1	3.35E-10	18.93	up
24	207183_at	GPR19	7.09E-10	4.11	up
25	206630_at	TYR	8.52E-10	45.86	up
26	210942_s_at	ST3GAL6	1.06E-09	13.57	up
27	211890_x_at	CAPN3	2.23E-09	22.28	up
28	214475_x_at	CAPN3	3.13E-09	36.27	up
29	207469_s_at	PIR	4.28E-09	16.87	up
30	207144_s_at	CITED1	5.15E-09	5.83	up
31	207323_s_at	MBP	5.79E-09	6.40	up
32	202342_s_at	TRIM2	6.27E-09	13.35	up
33	206427_s_at	MLANA	6.77E-09	107.54	up
34	202454_s_at	ERBB3	8.94E-09	32.94	up

35	206837_at	ALX1	9.55E-09	5.04	up
36	220948_s_at	ATP1A1	1.24E-08	6.09	up
37	220178_at	C19orf28	1.65E-08	4.81	up
38	210944_s_at	CAPN3	1.68E-08	26.58	up
39	209168_at	GPM6B	2.03E-08	11.60	up
40	209169_at	GPM6B	2.90E-08	7.84	up
41	205673_s_at	ASB9	3.55E-08	4.49	up
42	207704_s_at	GAS7	4.49E-08	5.26	up
43	203845_at	KAT2B	5.35E-08	4.08	up
44	218404_at	SNX10	7.05E-08	15.31	up
45	213355_at	ST3GAL6	7.72E-08	9.51	up
46	203300_x_at	AP1S2	9.01E-08	8.07	up
47	210613_s_at	SYNGR1	9.08E-08	4.53	up
48	206898_at	CDH19	9.73E-08	11.17	up
49	204273_at	EDNRB	1.70E-07	24.43	up
50	203518_at	LYST	2.16E-07	7.79	up
51	206426_at	MLANA	2.20E-07	23.10	up
52	207329_at	MMP8	2.51E-07	4.44	up
53	212510_at	GPD1L	2.57E-07	5.66	up
54	209942_x_at	MAGEA3	2.86E-07	56.65	up
55	205227_at	IL1RAP	3.05E-07	4.63	up
56	206653_at	POLR3G	3.20E-07	4.48	up
57	213638_at	PHACTR1	3.21E-07	6.55	up
58	208699_x_at	TKT	3.22E-07	6.11	up
59	206233_at	B4GALT6	3.23E-07	5.16	up
60	214612_x_at	MAGEA6	4.04E-07	40.75	up
61	207232_s_at	MITF	4.31E-07	11.02	up
62	218087_s_at	SORBS1	4.32E-07	8.02	up

63	209848_s_at	SILV	4.38E-07	57.80	up
64	210963_s_at	GYG2	4.72E-07	5.37	up
65	210467_x_at	MAGEA12	5.32E-07	20.88	up
66	208700_s_at	TKT	6.59E-07	5.41	up
67	206576_s_at	CEACAM1	8.04E-07	5.42	up
68	218829_s_at	CHD7	9.38E-07	6.51	up
69	221524_s_at	RRAGD	1.07E-06	14.38	up
70	205413_at	MPPED2	1.32E-06	4.63	up
71	201690_s_at	TPD52	1.36E-06	18.71	up
72	204542_at	ST6GALNAC2	1.41E-06	26.67	up
73	203775_at	SLC25A13	1.53E-06	4.94	up
74	218679_s_at	VPS28	1.72E-06	38.99	up
75	211889_x_at	CEACAM1	1.98E-06	5.75	up
76	205030_at	FABP7	2.48E-06	9.30	up
77	201651_s_at	PACSIN2	2.58E-06	4.84	up
78	205029_s_at	FABP7	2.75E-06	6.91	up
79	221523_s_at	RRAGD	3.16E-06	14.19	up
80	219895_at	FAM70A	3.68E-06	6.02	up
81	211883_x_at	CEACAM1	4.09E-06	5.17	up
82	204584_at	L1CAM	4.60E-06	6.10	up
83	204015_s_at	DUSP4	6.57E-06	9.92	up
84	219412_at	RAB38	6.84E-06	5.78	up
85	220245_at	SLC45A2	7.07E-06	9.45	up
86	221558_s_at	LEF1	7.18E-06	13.13	up
87	219121_s_at	ESRP1	7.39E-06	5.11	up
88	202191_s_at	GAS7	8.05E-06	6.43	up
89	215143_at	DPY19L2P2	8.32E-06	5.04	up
90	214603_at	MAGEA2 /// MAGEA2B	8.35E-06	8.57	up

91	207038_at	SLC16A6	1.04E-05	6.81	up
92	212807_s_at	SORT1	1.06E-05	7.90	up
93	201689_s_at	TPD52	1.14E-05	11.98	up
94	206498_at	OCA2	1.16E-05	4.27	up
95	210964_s_at	GYG2	1.26E-05	7.80	up
96	214297_at	CSPG4	1.40E-05	4.28	up
97	205376_at	INPP4B	1.51E-05	7.27	up
98	208002_s_at	ACOT7	1.54E-05	5.84	up
99	202341_s_at	TRIM2	1.66E-05	9.32	up
100	209610_s_at	SLC1A4	1.70E-05	8.66	up
101	209569_x_at	D4S234E /// FOXP1	1.95E-05	7.14	up
102	212747_at	ANKS1A	2.21E-05	4.07	up
103	212730_at	SYNM	2.44E-05	6.60	up
104	205110_s_at	FGF13	2.59E-05	5.06	up
105	221644_s_at	SLC45A2	2.65E-05	5.30	up
106	204633_s_at	RPS6KA5	3.01E-05	6.14	up
107	212811_x_at	SLC1A4	3.14E-05	6.76	up
108	212810_s_at	SLC1A4	3.39E-05	5.03	up
109	220445_s_at	CSAG2 /// CSAG3	3.43E-05	5.66	up
110	209570_s_at	D4S234E /// FOXP1	4.11E-05	5.09	up
111	209498_at	CEACAM1	4.15E-05	11.29	up
112	201688_s_at	TPD52	4.90E-05	10.00	up
113	201656_at	ITGA6	5.64E-05	5.53	up
114	214039_s_at	LAPTM4B	5.82E-05	6.48	up
115	205694_at	TYRP1	6.74E-05	13.28	up
116	211067_s_at	GAS7	6.87E-05	6.55	up
117	209184_s_at	IRS2	7.42E-05	4.82	up
118	204591_at	CHL1	7.88E-05	5.87	up

119	203123_s_at	SLC11A2	8.30E-05	5.24	up
120	201896_s_at	PSRC1	8.64E-05	4.08	up
121	215695_s_at	GYG2	8.99E-05	6.87	up
122	218854_at	DSE	2.88E-12	45.60	down
123	202766_s_at	FBN1	7.17E-11	36.22	down
124	202391_at	BASP1	1.54E-10	54.11	down
125	202375_at	SEC24D	1.02E-09	5.81	down
126	208091_s_at	VOPP1	1.26E-09	12.79	down
127	209946_at	VEGFC	1.40E-09	21.67	down
128	202765_s_at	FBN1	1.54E-09	26.86	down
129	212385_at	TCF4	2.44E-09	7.21	down
130	210511_s_at	INHBA	2.70E-09	40.62	down
131	212382_at	TCF4	4.30E-09	8.23	down
132	212012_at	PXDN	7.74E-09	9.52	down
133	212473_s_at	MICAL2	1.23E-08	23.24	down
134	204222_s_at	GLIPR1	1.62E-08	10.39	down
135	204337_at	RGS4	1.95E-08	11.77	down
136	201387_s_at	UCHL1	1.99E-08	27.71	down
137	212013_at	PXDN	2.03E-08	10.32	down
138	209094_at	DDAH1	2.06E-08	12.99	down
139	213059_at	CREB3L1	2.12E-08	9.24	down
140	211668_s_at	PLAU	2.86E-08	14.35	down
141	212387_at	TCF4	3.06E-08	9.80	down
142	209687_at	CXCL12	5.58E-08	25.10	down
143	210058_at	MAPK13	5.58E-08	6.89	down
144	201667_at	GJA1	6.27E-08	59.61	down
145	222146_s_at	TCF4	6.93E-08	10.58	down
146	205083_at	AOX1	7.97E-08	5.18	down

147	212472_at	MICAL2	8.04E-08	10.06	down
148	213891_s_at	TCF4	8.28E-08	14.01	down
149	203184_at	FBN2	9.11E-08	14.68	down
150	204135_at	FILIP1L	1.23E-07	8.93	down
151	205479_s_at	PLAU	1.51E-07	11.73	down
152	37408_at	MRC2	1.53E-07	10.03	down
153	207992_s_at	AMPD3	1.66E-07	5.41	down
154	201983_s_at	EGFR	2.71E-07	27.13	down
155	221012_s_at	TRIM8	2.88E-07	4.02	down
156	204004_at	PAWR	3.30E-07	4.78	down
157	203102_s_at	MGAT2	3.59E-07	4.35	down
158	203753_at	TCF4	3.63E-07	10.73	down
159	212298_at	NRP1	3.94E-07	15.16	down
160	203325_s_at	COL5A1	4.81E-07	17.02	down
161	209955_s_at	FAP	5.41E-07	18.01	down
162	218468_s_at	GREM1	5.65E-07	70.74	down
163	212488_at	COL5A1	5.78E-07	21.48	down
164	201984_s_at	EGFR	5.79E-07	4.53	down
165	210764_s_at	CYR61	5.86E-07	13.05	down
166	209210_s_at	FERMT2	6.60E-07	5.57	down
167	202628_s_at	SERPINE1	7.16E-07	90.87	down
168	207876_s_at	FLNC	8.29E-07	4.11	down
169	212386_at	TCF4	8.43E-07	15.79	down
170	214446_at	ELL2	8.47E-07	5.58	down
171	203131_at	PDGFRA	9.07E-07	26.58	down
172	209335_at	DCN	9.60E-07	24.33	down
173	209101_at	CTGF	1.84E-06	15.87	down
174	214696_at	C17orf91	1.99E-06	7.25	down

175	212489_at	COL5A1	1.99E-06	13.23	down
176	207002_s_at	PLAGL1	2.33E-06	7.89	down
177	218469_at	GREM1	2.73E-06	66.82	down
178	216840_s_at	LAMA2	3.13E-06	4.62	down
179	212151_at	PBX1	3.17E-06	5.21	down
180	209278_s_at	TFPI2	3.67E-06	35.14	down
181	202311_s_at	COL1A1	3.72E-06	17.65	down
182	205407_at	RECK	3.79E-06	6.84	down
183	212148_at	PBX1	4.44E-06	6.25	down
184	203088_at	FBLN5	4.67E-06	5.44	down
185	210735_s_at	CA12	5.15E-06	9.24	down
186	202627_s_at	SERPINE1	5.28E-06	34.51	down
187	216438_s_at	TMSB4X /// TMSL3	5.58E-06	7.02	down
188	214452_at	BCAT1	5.99E-06	6.05	down
189	209318_x_at	PLAGL1	6.04E-06	11.14	down
190	201107_s_at	THBS1	6.36E-06	6.33	down
191	201289_at	CYR61	7.96E-06	11.44	down
192	214164_x_at	CA12	9.36E-06	21.12	down
193	211813_x_at	DCN	9.54E-06	22.74	down
194	203570_at	LOXL1	1.17E-05	39.27	down
195	221730_at	COL5A2	1.18E-05	19.84	down
196	221898_at	PDPN	1.39E-05	4.53	down
197	211896_s_at	DCN	1.43E-05	29.56	down
198	219134_at	ELTD1	1.45E-05	6.25	down
199	204780_s_at	FAS	1.47E-05	7.46	down
200	204338_s_at	RGS4	1.50E-05	5.38	down
201	203963_at	CA12	1.51E-05	22.46	down
202	203476_at	TPBG	1.52E-05	17.54	down

203	221760_at	MAN1A1	1.59E-05	11.17	down
204	203085_s_at	TGFB1	1.62E-05	5.97	down
205	207943_x_at	PLAGL1	1.65E-05	5.10	down
206	205128_x_at	PTGS1	1.66E-05	10.36	down
207	203988_s_at	FUT8	1.98E-05	6.06	down
208	213069_at	HEG1	2.04E-05	4.47	down
209	215446_s_at	LOX	2.11E-05	37.86	down
210	221029_s_at	WNT5B	2.16E-05	5.96	down
211	203666_at	CXCL12	2.28E-05	7.63	down
212	202890_at	MAP7	2.31E-05	13.29	down
213	219506_at	C1orf54	2.32E-05	5.27	down
214	205105_at	MAN2A1	2.37E-05	4.81	down
215	220014_at	PRR16	2.52E-05	6.85	down
216	219492_at	CHIC2	2.54E-05	5.31	down
217	201893_x_at	DCN	2.63E-05	20.90	down
218	204604_at	CDK14	2.86E-05	4.92	down
219	208851_s_at	THY1	2.96E-05	6.69	down
220	213661_at	PAMR1	3.29E-05	5.63	down
221	209276_s_at	GLRX	3.34E-05	7.25	down
222	206662_at	GLRX	3.40E-05	7.06	down
223	217430_x_at	COL1A1	3.52E-05	8.79	down
224	202948_at	IL1R1	3.64E-05	21.00	down
225	212859_x_at	MT1E	3.88E-05	15.36	down
226	213905_x_at	BGN	4.03E-05	12.49	down
227	215867_x_at	CA12	4.33E-05	19.17	down
228	31845_at	ELF4	4.61E-05	4.03	down
229	1598_g_at	GAS6	4.73E-05	10.87	down
230	214085_x_at	GLIPR1	4.95E-05	5.80	down

231	204359_at	FLRT2	5.15E-05	10.39	down
232	210059_s_at	MAPK13	5.30E-05	4.52	down
233	201108_s_at	THBS1	5.51E-05	10.49	down
234	204790_at	SMAD7	5.86E-05	5.14	down
235	201842_s_at	EFEMP1	5.93E-05	25.63	down
236	215813_s_at	PTGS1	5.98E-05	8.38	down
237	205828_at	MMP3	6.04E-05	6.91	down
238	212646_at	RFTN1	6.77E-05	8.19	down
239	213869_x_at	THY1	6.82E-05	5.89	down
240	221729_at	COL5A2	6.94E-05	19.22	down
241	204745_x_at	MT1G	6.97E-05	10.13	down
242	202177_at	GAS6	7.40E-05	9.62	down
243	213428_s_at	COL6A1	7.56E-05	11.83	down
244	201261_x_at	BGN	8.02E-05	10.11	down
245	201627_s_at	INSIG1	8.44E-05	6.24	down
246	208789_at	PTRF	8.57E-05	5.68	down
247	209277_at	TFPI2	8.68E-05	6.40	down
248	216336_x_at	MT1E	9.24E-05	7.96	down

*Bonferroni corrected p-value; $absolute fold change; †compared cluster B vs. A from Figure 3.11

PUBLICATIONS

Senses et al. BMC Medical Genetics 2013, 14:97
http://www.biomedcentral.com/1471-2350/14/97

BMC Medical Genetics

RESEARCH ARTICLE Open Access

Cancer-testis gene expression is associated with the *methylenetetrahydrofolate reductase* 677 C>T polymorphism in non-small cell lung carcinoma

Kerem M Senses[1], Mithat Gonen[2], Ahmet R Barutcu[1,6], Zeynep Kalaylioglu[3], Murat Isbilen[1], Ozlen Konu[1], Yao T Chen[4], Nasser K Altorki[5] and Ali O Gure[1*]

Abstract

Background: Tumor-specific, coordinate expression of cancer-testis (CT) genes, mapping to the X chromosome, is observed in more than 60% of non-small cell lung cancer (NSCLC) patients. Although CT gene expression has been unequivocally related to DNA demethylation of promoter regions, the underlying mechanism leading to loss of promoter methylation remains elusive. Polymorphisms of enzymes within the 1-carbon pathway have been shown to affect S-adenosyl methionine (SAM) production, which is the sole methyl donor in the cell. Allelic variants of several enzymes within this pathway have been associated with altered SAM levels either directly, or indirectly as reflected by altered levels of SAH and Homocysteine levels, and altered levels of DNA methylation. We, therefore, asked whether the five most commonly occurring polymorphisms in four of the enzymes in the 1-carbon pathway associated with CT gene expression status in patients with NSCLC.

Methods: Fifty patients among a cohort of 763 with NSCLC were selected based on CT gene expression status and typed for five polymorphisms in four genes known to affect SAM generation by allele specific q-PCR and RFLP.

Results: We identified a significant association between CT gene expression and the *MTHFR* 677 CC genotype, as well as the C allele of the SNP, in this cohort of patients. Multivariate analysis revealed that the genotype and allele strongly associate with CT gene expression, independent of potential confounders.

Conclusions: Although CT gene expression is associated with DNA demethylation, in NSCLC, our data suggests this is unlikely to be the result of decreased MTHFR function.

Keywords: Cancer-testis genes, 1-carbon pathway, Adomet, DNA methylation

Background

Cancer-testis (CT), or cancer-germline genes, currently with more than 100 members, are distinctly expressed in cancer, germline and trophoblast cells but not in other normal tissues in the adult. Most CT genes constitute multigene families organized in clusters along the X chromosome. Members within a family are highly homologous, however, no conservation of sequence exists between families [1]. Despite the lack of sequence similarity (including promoters), re-expression of almost all CT genes in tumors correlates with the demethylation of their promoters that occurs in parallel to a genome-wide demethylation event, primarily affecting repeat regions [2]. The mechanisms leading to CT gene promoter demethylation in cancer are unknown. Increased BORIS expression has been associated with upregulated CT gene expression [3,4], but the protein is likely not the sole responsible factor in this event. Histone acetylation has also been shown to facilitate CT gene expression, primarily when it associates with DNA demethylation [5].

As most CT gene products are highly antigenic they have been utilized in clinical trials based on immunotherapeutic approaches targeting these antigens [6]. Since patient eligibility for CT targeting immunotherapy requires that the tumor express CT genes, it is important to know whether CT gene expression can be induced. It is expected

* Correspondence: agure@bilkent.edu.tr
[1]Department of Molecular Biology and Genetics, Bilkent University, Ankara, Turkey
Full list of author information is available at the end of the article

that any approach leading to CT gene expression should also result in the demethylation of their promoters.

Production of the sole methyl donor in the cell, S-adenosylmethionine (SAM), depends on the efficient utilization of folate, by the 1-carbon pathway. Several enzymes in this pathway contain common polymorphic variants that reduce the efficiency of the enzyme and thus, the rate of SAM production. Hypomorphic alleles of four of these enzymes (methylenetetrahydrofolate reductase (*MTHFR*), methionine synthase reductase (*MTRR*), methionine synthase (*MTR*), and reduced folate carrier (*RFC*)), have been associated with cellular under-utilization of folate and homocysteine, increased DNA hypomethylation, and decreased CpG methylation [7-11]. More recently, the hypomorphic 677 T allele of *MTHFR*, has been associataed with the expression of *MAGE-A1*, a CT gene, in glioblastoma multiforme [12]. Others, however, could not reproduce these findings in ovarian carcinoma [13]. In the present study we asked if polymorphisms of the 1-carbon pathway enzymes associate with CT gene expression in non-small cell lung cancer (NSCLC) patients. Our results show a strong association between the *MTHFR*677 CC genotype as well as the *MTHFR* 677 C allele and CT gene expression independent of age, sex, histology, and tumor stage.

Methods

Patients and tumor material

Tumor samples obtained from patients undergoing curative surgical resection for primary NSCLC at the Department of Cardio-Thoracic Surgery, Weill Medical College of Cornell University, from 1991 to July 2005 were analyzed in this study. Informed consent was obtained from all patients. The study was approved by the Institutional Review Board of Weill Medical College of Cornell University. Fifty tumor samples were selected solely based on CT gene expression from 763 samples that had been evaluated for the presence of transcripts from up to 9 CT genes (*NY-ESO-1, LAGE-1, MAGE-A1, MAGE-A3, MAGE-A4, MAGE-A10, CT-7, SSX2,* and *SSX4*), by semi-quantitative PCR, as described previously [14]. Twenty one samples with CT expression in at least 4 of the 9 CT genes tested, with strong expression in at least one gene, constituted the CT (+) group. Twenty-nine samples with no CT expression in any of the CT genes tested (with a minimum of 5 CT genes tested) were selected as CT (-) tumors for this study. CT gene expression was determined as strong (+++), intermediate (++), weak (+ or +/-), or none (-) as previously described [14], and is shown in Additional file 1: Table S1.

DNA analysis

Genomic DNA extracted from tumor tissues were genotyped using pre-designed 5'-nuclease TaqMan SNP genotyping assays (Applied Biosystems, Foster City, CA) using a Stratagene Mx3005P instrument according to the manufacturer's instructions. The SNPs typed and their reference IDs were: *MTHFR* 677 C> T (rs1801133), *MTHFR* 1298 A>C (rs1801131), *MTR* 2756 A>G (rs1805087), and *MTRR* 66 A>G (rs1801394). Nested PCR-RFLP was used to type the *RFC* 80 G>A (rs1051266) polymorphism for which the first round PCR conditions were previously described [10]. Nested PCR primers were: 5'- AGCCGTAGA AGCAAAGGTAGC-3' and 5'-AGCGTCACCTTCGTCC CCTC-3'. PCR was performed using DyNAzyme™ II Hot Start DNA Polymerase (Finnzymes, Keilaranta, Finland). PCR conditions were: 10' activation at 94°C, followed by 35 cycles of 94°C, 62°C and 72°C; 30" each, with a final 72°C, 7' extension. HinP1I (New England Biolabs, Hertfordshire, UK) digested PCR products were analyzed as described previously [10]. All analyses were repeated at least twice.

Genotypes for all polymorphisms were determined successfully in all cases (Additional file 2: Table S2). Genotype distributions did not deviate from Hardy-Weinberg equilibrium (Additional file 3: Table S3). Minor allele frequencies for individual loci were: 40% for *MTHFR* 677 C > T, 26% for *MTHFR* 1298 A > C, 14% for *MTR* 2756 A > G, 54% for *MTRR* 66, and 42% for *RFC* 80 G > A. *MTHFR* genotypes were not independently distributed across the 2 loci. The major 677C allele was in linkage disequilibrium with the minor 1298C allele (D' = 0.99, r^2 = 0.23) [15].

In silico association analysis

Paired datasets, GSE14471 and GSE15714, containing gene expression and SNP genotyping data, respectively, from 111 pediatric acute myeloid leukemia samples (of which 109 were typed successfully), were analyzed for an association between CT gene expression and *MTHFR* 677 genotype distribution [16]. A principal component analysis using 44 probesets corresponding to 9 CT gene families was performed for the expression dataset. The first principal component, explaining 0.48 of variance for CT gene expression was used to generate groups representing samples with low, intermediate, and high CT gene expression by K means clustering using a customized R code [17]. Optimum number of clusters according to Elbow criterion was determined as five. Therefore, five initial cluster centers were placed equally distant from each other where the first and last centers represented the minimum and maximum values of PC1, respectively. Centers were iteratively updated based on the median value of the reassigned cluster members until no change in cluster membership took place. The five clusters were regrouped into three representing low (clusters 1 & 2), intermediate (cluster 3), and high CT gene expression (clusters 4 & 5).

Statistical analysis

To analyze the association between 1-carbon pathway enzyme polymorphisms and CT gene expression, the genotype distributions were compared in CT (+) and CT (-) tumors by Pearson's Chi-Square (2 degrees of freedom) or Fisher's exact tests. Odds ratios (OR) were estimated by multivariate logistic regression. To evaluate whether CT gene expression was related to sex, smoking status, tumor size, and disease stage, Fisher's exact test or Chi-square tests were used. Race information was available for only 29 patients of which 25 were non-Hispanic white, one was a non-Hispanic black, and 3 were of mixed race, and was not included in statistical analyses. All statistical tests were two-sided with a 5% type I error rate, unless indicated otherwise, and were carried out using SAS (version 9.3) software (SAS Institute, Cary, NC). $P < 0.05$ was considered statistically significant.

Table 1 Demographics and clinical characteristics

		CT (+) patients (n=21)	CT (-) patients (n=29)	P*
Age	>60	12	19	0.74
	<=60	6	7	
	Unknown[$]	3	3	
Sex	Male	10	12	0.76
	Female	8	14	
	Unknown	3	3	
Smoking history	No	1	5	0.37
	Yes	15	18	
	Unknown	5	6	
Histology	SQCC[#]	8	2	0.002
	non-SQCC	7	24	
	Unknown	6	3	
Pathological tumor size	>3 cm	10	8	0.21
	<=3 cm	8	17	
	Unknown	3	4	
T stage	1	5	14	0.007
	2	8	12	
	3	3	0	
	4	2	0	
	Unknown	3	3	
TNM stage (Pathologic stage of primary tumor)	I	9	18	0.37
	II	5	3	
	III	4	5	
	IV	0	0	
	Unknown	3	3	

* Chi-square (Fisher's exact test, two sided) or chi-square test for trend; [#] *SQCC* Squamous cell carcinoma; [$]patients with missing clinical data were not included in statistical analyses.

Results

Demographics and clinical characteristics of patients and their distribution within CT (+) and (-) groups are shown in Table 1 and Additional file 1: Table S1. Tumors with non-squamous cell carcinoma histology and earlier tumor stage (T stage) showed lower CT gene expression, similar to what has been reported previously [14]. Distribution of individual genotypes among CT (+) and (-) tumors are shown in Table 2 and Additional file 2: Table S2. A significant association between the *MTHFR* 677CC genotype and CT expression was observed ($P = 0.03$). CT expression was not related to any other genotype tested. A multivariate logistic regression analysis (MVA) of CT gene expression that included the *MTHFR* 677 genotype distribution, age, sex, histology and T stage revealed that the *MTHFR* 677 genotype and histology were independent predictors of CT gene expression in this cohort (Table 3). The *MTHFR* 677 SNP was found to be associated with CT gene expression when analyzed on a per allele basis, controlling for confounding factors, while other markers were not (Table 4). We performed an *in silico* association analysis for CT expression and the *MTHFR* 677 genotype using two datasets derived from childhood acute myeloid leukemia (AML) where both gene expression and SNP genotyping data were available [16]. This analysis, however, did not reveal a statistically significant association between these two parameters (Table 5 and Additional file 4: Figure S1).

Table 2 Distribution of individual genotypes among CT (+) and CT (-) tumors

Polymorphism	*Genotype*	*CT (+) Tumors, n (%)*	*CT (-) Tumors, n (%)*	*χ2*	*P**
MTHFR 677 C>T (rs1801133)	CC	13 (62%)	7 (24%)	7.30	0.03
	CT	5 (24%)	15 (52%)		
	TT	3 (14%)	7 (24%)		
MTHFR 1298 A>C (rs1801131)	AA	13 (62%)	14 (48%)	0.91	0.63
	AC	7 (33%)	13 (45%)		
	CC	1 (5%)	2 (7%)		
MTR 2756 A>G (rs1805087)	AA	14 (67%)	22 (76%)	0.51	0.53**
	AG	7 (33%)	7 (24%)		
	GG	0 (0%)	0 (0%)		
MTRR 66 A>G (rs1801394)	AA	4 (19%)	9 (31%)	0.92	0.63
	AG	10 (48%)	12 (41%)		
	GG	7 (33%)	8 (28%)		
RFC 80 G>A (rs1051266)	GG	6 (29%)	12 (41%)	1.90	0.39
	GA	9 (43%)	13 (45%)		
	AA	6 (29%)	4 (14%)		

* Chi-square (2 degrees of freedom); ** Fisher's exact test, two sided.

Table 3 Multivariate analysis of CT gene expression with MTHFR 677 genotypes

Parameter		*OR*	*95% CI*	*P**
MTHFR 677 C>T (rs1801133)	CC	32.33	2.42-431.52	0.003
	CT/TT	1**		
Age[$]		1.03	0.94-1.12	0.69
Sex(female vs. male)		0.28[‡]	0.02-3.86	0.52
Histology	SQCC[#]	18.46	1.19-284.49	0.04
	non-SQCC	1**		
T Stage	T1	0.31	0.03-3.56	0.53
	T2, 3, 4	1**		

* Computed from a logistic regression model using the EXACT option of PROC LOGISTIC in SAS to account for the small data set; [$]continuous variable.
[#] *SQCC* Squamous cell carcinoma; [‡]reference = male; **reference group.

Discussion

Among the five markers analyzed in this study, we find a strong association between the major *MTHFR* 677 CC genotype, as well as the *MTHFR* 677 C allele and CT gene expression in lung cancer. This contrasts with earlier studies where the minor allele of this SNP was associated with decreased SAM production, decreased methylation levels and decreased *MAGE-A1* expression [12]. Although our analysis included only 7% of patients within a large cohort with the highest and lowest amount of CT gene expression, we don't think this is a reason for bias, as the distribution of the 1-carbon pathway genotypes of our samples are similar to those where much larger lung cancer patient cohorts were evaluated [18-20]. Tumors of squamous cell histology were previously identified as showing more frequent and stronger CT gene expression; however, MVA shows that the association between *MTHFR* 677 CC genotype or the C allele of the same polymorphism and CT gene expression is independent of histology. On the other hand, tumor type is known to affect CT gene expression rates, as some blood-derived tumors and cancers originating from the kidney rarely express CT genes [21]. In this line, one reason for our inability to replicate our q-PCR based results *in silico* might be related to the fact that AML is not a tumor with strong CT expression and thus, the K-means based classification of this tumor is somewhat artificial. Therefore, a similar analysis with datasets ideally derived from lung cancer might reveal associations not identified in this study.

Table 4 Multivariate logistic regression modeling of association between 1-carbon enzyme alleles and CT gene expression

Polymorphism	*OR*[$]	*95% CI*	*P**
MTHFR 677 (rs1801133)	13.18	1.96-88.5	0.004
MTHFR 1298 (rs1801131)	0.56	0.15-2.09	0.53
MTR 2756 (rs1805087)	0.81	0.14-4.75	1
MTRR 66 (rs1801394)	0.67	0.22-2.00	0.52
RFC 80 (rs1051266)	0.7	0.24-2.07	0.62

[$] Based on number of major alleles, adjusted for age, sex, histology and t stage; * computed using the EXACT option of PROC LOGISTIC in SAS to account for the small data set.

Table 5 *In silico* correlation of CT gene expression with *MTHFR* 677 genotypes in acute Myeloid Leukemia

CT gene expression clusters	*MTHFR* 677 C>T (rs1801133)			*P* (chi-square)
	CC	CT/TT	Untyped*	
High	15	16	0	0.17
Intermediate	20	13	0	
Low	29	16	2	

* Untyped samples were not included in analysis.

We calculated the sample size that would give us 80% power to detect a significant association between polymorphisms other than *MTHFR* 677 and CT gene expression using the observed effect sizes in this study as true values. We found that at least 250 patients would be required to find one more polymorphism significant. Therefore, analysis of larger cohorts might reveal additional associations as well as compound effects of SNPs within the 1-carbon pathway enzymes on CT gene expression. Models to test for such effects were not computed in this study due to the limited sample size.

Although decreased SAM levels might be expected to result in DNA demethylation, the exact SAM concentration threshold required for gene re-expression might be affected by various other parameters not tested in this study. A candidate is thymidylate synthase (TS) whose levels are known to fluctuate widely in cancer and which can inhibit MTHFR activity [22]. CT gene expression is associated with larger tumors and advanced stage [14]. If this is to be taken as a sign of increased proliferation, it would imply increased TS activity, and thus, possibly suppressed MTHFR, which in turn could affect CT gene expression. On the other hand, increased SAM production might indirectly inhibit methylation reactions via methylthioadenosine (MTA), a nucleoside produced from SAM through the polyamine biosynthetic pathway. MTA can strongly inhibit H3K4 methylation, possibly by inhibiting Set1 methyltransferase, which could in turn result in repressed CT gene expression [23-25]. Future studies are necessary to explain which of these primarily affect methylation rates and thus CT gene expression in cancer.

Conclusion

Why some NSCLC cells express CT genes when others don't, remains an interesting and unanswered question. We show a strong association between the normoactive allele of *MTHFR* 677 and CT gene expression in this study. This argues against the hypothesis of low level MTHFR activity leading to DNA hypomethylation, which in turn could lead to genome-wide hypomethylation and

CT gene expression. However, due to the limited power of this study, we might have missed individual or cumulative effects of SNPs within other enzymes of the 1-carbon pathway on CT gene expression. SAM/SAH ratios for the tissues analyzed here were also unknown. Hence, we only contributed to, but did not resolve this interesting story, and hope future studies reveal the intricacies of the relation between CT gene expression and genetic variants of the 1-carbon pathway enzyme genes.

Additional files

Additional file 1: Table S1. CT Gene Expression and Distribution of Clinical Parameters within NSCLC Patients.

Additional file 2: Table S2. Genotypes of NSCLC Patients.

Additional file 3: Table S3. Hardy-Weinberg Distributions of Single Nucleotide Polymorphisms in NSCLC Patients.

Additional file 4: Figure S1. Principal component analysis based *in silico* clustering of AML. Tumor samples are shown ordered from the lowest to the highest first principal component (PC1) value. The 5 clusters generated by K-means clustering are indicated. Tumors with low, intermediate and high CT gene expression correspond to clusters 1-2, 3, and 4-5, respectively.

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

Authors' contributions
Overall study design: KMS, MG, ZK, YTC, AOG; patient recruitment and sample collection: NKA and YTC; genotyping experiments: KMS and ARB; principal component analysis and other *in silico* analyses: MI, OK and KMS; statistical analyses: MG, ZK, OK and KMS. All authors read and approved the final manuscript.

Acknowledgements
We thank Lloyd J. Old, for years of continued support and encouragement; David Kelly and Jay Lee for excellent technical assistance.

Author details
[1]Department of Molecular Biology and Genetics, Bilkent University, Ankara, Turkey. [2]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [3]Department of Statistics, Middle East Technical University, Ankara, Turkey. [4]Department of Pathology, Weill Medical College of Cornell University, New York, NY, USA. [5]Department of Cardiothoracic Surgery, Weill Medical College of Cornell University, New York, NY, USA. [6]Department of Cell and Developmental Biology, University of Massachusetts Medical School, Worcester, MA, USA.

Received: 17 April 2013 Accepted: 20 September 2013
Published: 24 September 2013

References
1. Caballero OL, Chen YT: **Cancer/testis (CT) antigens: potential targets for immunotherapy.** *Cancer science* 2009, **100**(11):2014–2021.
2. Loriot A, Reister S, Parvizi GK, Lysy PA, De Smet C: **DNA methylation-associated repression of cancer-germline genes in human embryonic and adult stem cells.** *Stem cells (Dayton, Ohio)* 2009, **27**(4):822–824.
3. Vatolin S, Abdullaev Z, Pack SD, Flanagan PT, Custer M, Loukinov DI, Pugacheva E, Hong JA, Morse H 3rd, Schrump DS, *et al*: **Conditional expression of the CTCF-paralogous transcriptional factor BORIS in normal cells results in demethylation and derepression of MAGE-A1 and reactivation of other cancer-testis genes.** *Cancer Res* 2005, **65**(17):7751–7762.
4. Hong JA, Kang Y, Abdullaev Z, Flanagan PT, Pack SD, Fischette MR, Adnani MT, Loukinov DI, Vatolin S, Risinger JI, *et al*: **Reciprocal binding of CTCF and BORIS to the NY-ESO-1 promoter coincides with derepression of this cancer-testis gene in lung cancer cells.** *Cancer Res* 2005, **65**(17):7763–7774.
5. Gure AO, Wei IJ, Old LJ, Chen YT: **The SSX gene family: characterization of 9 complete genes.** *International journal of cancer* 2002, **101**(5):448–453.
6. Jager E, Knuth A: **The discovery of cancer/testis antigens by autologous typing with T cell clones and the evolution of cancer vaccines.** *Cancer Immun* 2012, **12**:6.
7. Castro R, Rivera I, Ravasco P, Camilo ME, Jakobs C, Blom HJ, de Almeida IT: **5,10-methylenetetrahydrofolate reductase (MTHFR) 677C–>T and 1298A–>C mutations are associated with DNA hypomethylation.** *J Med Genet* 2004, **41**(6):454–458.
8. Friso S, Girelli D, Trabetti E, Olivieri O, Guarini P, Pignatti PF, Corrocher R, Choi SW: **The MTHFR 1298A>C polymorphism and genomic DNA methylation in human lymphocytes.** *Cancer Epidemiol Biomarkers Prev* 2005, **14**(4):938–943.
9. Olteanu H, Munson T, Banerjee R: **Differences in the efficiency of reductive activation of methionine synthase and exogenous electron acceptors between the common polymorphic variants of human methionine synthase reductase.** *Biochemistry* 2002, **41**(45):13378–13385.
10. Chango A, Emery-Fillon N, de Courcy GP, Lambert D, Pfister M, Rosenblatt DS, Nicolas JP: **A polymorphism (80G->A) in the reduced folate carrier gene and its associations with folate status and homocysteinemia.** *Molecular genetics and metabolism* 2000, **70**(4):310–315.
11. Paz MF, Avila S, Fraga MF, Pollan M, Capella G, Peinado MA, Sanchez-Cespedes M, Herman JG, Esteller M: **Germ-line variants in methyl-group metabolism genes and susceptibility to DNA methylation in normal tissues and human primary tumors.** *Cancer Res* 2002, **62**(15):4519–4524.
12. Cadieux B, Ching TT, Vandenberg SR, Costello JF: **Genome-wide hypomethylation in human glioblastomas associated with specific copy number alteration, methylenetetrahydrofolate reductase allele status, and increased proliferation.** *Cancer Res* 2006, **66**(17):8469–8476.
13. Woloszynska-Read A, Zhang W, Yu J, Link PA, Mhawech-Fauceglia P, Collamat G, Akers SN, Ostler KR, Godley LA, Odunsi K, *et al*: **Coordinated cancer germline antigen promoter and global DNA hypomethylation in ovarian cancer: association with the BORIS/CTCF expression ratio and advanced stage.** *Clin Cancer Res* 2011, **17**(8):2170–2180.
14. Gure AO, Chua R, Williamson B, Gonen M, Ferrera CA, Gnjatic S, Ritter G, Simpson AJ, Chen YT, Old LJ, *et al*: **Cancer-testis genes are coordinately expressed and are markers of poor outcome in non-small cell lung cancer.** *Clin Cancer Res* 2005, **11**(22):8055–8062.
15. Gaunt TR, Rodriguez S, Day IN: **Cubic exact solutions for the estimation of pairwise haplotype frequencies: implications for linkage disequilibrium analyses and a web tool 'CubeX'.** *BMC bioinformatics* 2007, **8**:428.
16. Radtke I, Mullighan CG, Ishii M, Su X, Cheng J, Ma J, Ganti R, Cai Z, Goorha S, Pounds SB, *et al*: **Genomic analysis reveals few genetic alterations in pediatric acute myeloid leukemia.** *Proc Natl Acad Sci USA* 2009, **106**(31):12944–12949.
17. R Core Team: *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing; 2013. http://www.R-project.org/. ISBN 3-900051-07-0.
18. Shen H, Spitz MR, Wang LE, Hong WK, Wei Q: **Polymorphisms of methylene-tetrahydrofolate reductase and risk of lung cancer: a case-control study.** *Cancer Epidemiol Biomarkers Prev* 2001, **10**(4):397–401.
19. Shi Q, Zhang Z, Li G, Pillow PC, Hernandez LM, Spitz MR, Wei Q: **Polymorphisms of methionine synthase and methionine synthase reductase and risk of lung cancer: a case-control analysis.** *Pharmacogenetics and genomics* 2005, **15**(8):547–555.
20. Suzuki T, Matsuo K, Hiraki A, Saito T, Sato S, Yatabe Y, Mitsudomi T, Hida T, Ueda R, Tajima K: **Impact of one-carbon metabolism-related gene polymorphisms on risk of lung cancer in Japan: a case control study.** *Carcinogenesis* 2007, **28**(8):1718–1725.
21. Hofmann O, Caballero OL, Stevenson BJ, Chen YT, Cohen T, Chua R, Maher CA, Panji S, Schaefer U, Kruger A, *et al*: **Genome-wide analysis of cancer/testis gene expression.** *Proc Natl Acad Sci USA* 2008, **105**(51):20422–20427.
22. Matthews RG, Daubner SC: **Modulation of methylenetetrahydrofolate reductase activity by S-adenosylmethionine and by dihydrofolate and its polyglutamate analogues.** *Advances in enzyme regulation* 1982, **20**:123–131.

23. Song MR, Ghosh A: **FGF2-induced chromatin remodeling regulates CNTF-mediated gene expression and astrocyte differentiation.** *Nat Neurosci* 2004, **7**(3):229–235.
24. Chau CM, Lieberman PM: **Dynamic chromatin boundaries delineate a latency control region of Epstein-Barr virus.** *J Virol* 2004, **78**(22):12308–12319.
25. Huang J, Kent JR, Placek B, Whelan KA, Hollow CM, Zeng PY, Fraser NW, Berger SL: **Trimethylation of histone H3 lysine 4 by Set1 in the lytic infection of human herpes simplex virus 1.** *J Virol* 2006, **80**(12):5740–5746.

doi:10.1186/1471-2350-14-97
Cite this article as: Senses *et al.*: Cancer-testis gene expression is associated with the *methylenetetrahydrofolate reductase* 677 C>T polymorphism in non-small cell lung carcinoma. *BMC Medical Genetics* 2013 **14**:97.

Send Orders for Reprints to reprints@benthamscience.net

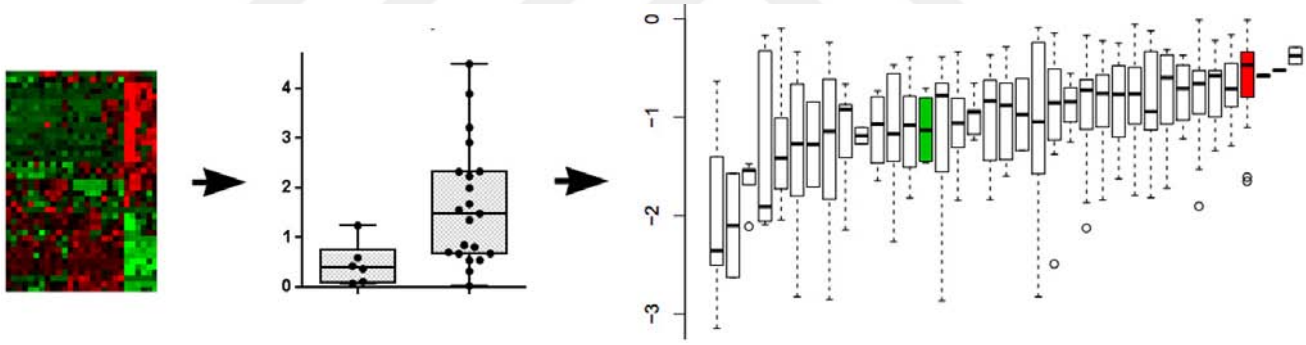
Predicting Chemotherapy Sensitivity Profiles for Breast Cancer Cell Lines with and Without Stem Cell-Like Features

Murat Isbilen, Kerem Mert Senses and Ali Osmay Gure*

Bilkent University, Department of Molecular Biology and Genetics, Ankara, Turkey 06800

Abstract: Our current understanding of cancer-stem cells (CSCs) is that they are slow growing, generally mesenchymal-like cells capable of generating tumors. Convincing evidence for the existence of such cells comes from recent lineage tracing experiments. CSCs have been reported as being resistant to conventional drug treatment and have been considered as being responsible for failure of chemotherapy. Recently, several databases aiming the genetic characterization of a large number of cancer cell lines have been made publicly available. In addition to gene expression data, these databases contain cytotoxicity information for all cell lines for a number of drugs as well. It is possible to classify known cell lines derived from a given tumor, based on how similar they are to CSCs, or in other words, to define their stem-ness, using gene-lists that define such cells. Using two such, independently generated, gene lists we found that breast cancer cell lines could be categorized into two distinct groups which we designate CSC-like and non-CSC-like. We then identified drugs to which the two groups were most sensitive to. We also generated sensitivity profiles for all drugs, within one such database, to identify chemotherapeutics with preferential action on breast cancer. We believe this is a straight-forward approach for swiftly identifying drugs that would selectively target a subpopulation of cells for any given tumor type.

Graphical Abstract



Keywords: Cancer, cancer stem cells, chemotherapy, cytotoxicity, database mining, lapatinib, LBW242, TKI258.

INTRODUCTION

Despite decades of work, cancer remains mostly an incurable disease. One reason for drug resistance in cancer has been proposed to stem from the fact that tumors contain cancer stem cells (CSCs) that can resist therapy. Recently, several papers have clearly documented the presence of CSCs using lineage tracing [1]. However, markers by which such cells can be reproducibly identified in tumors has been difficult due to the fact that tumors show vast heterogeneity [2]. Nevertheless, many ongoing studies aim to identify agents that can target cancer stem cells, as it is agreed that such cells are difficult to kill using conventional drugs. A recent COST action (CM1106) made its priority to characterize such drugs and many other studies have been or are being conducted with the same purpose. The logic behind these is that cells defined by either a marker or a behavior should be tested for drug sensitivity. However, as we show in this study, in cases where a gene expression profile for a CSC-like cell is available, it is possible to use this signature to classify commonly used cancer cell lines into CSC-like and non-CSC-like populations. This information can then be used to identify drugs that affect either population better than the other, and thus predict drugs with preferential action on both CSC-like and non-CSC-like cells. We believe annotating drugs for both populations is critical in that there is building evidence that either group can cause cancer mortality, due to the phenotypic plasticity allowing transition between phenotypes. An extremely useful source for this purpose is the drug cytotoxicity data provided within the Cancer Cell Line Encyclopedia (CCLE) project [3]. We here show that it is possible to identify drugs that can possibly be selected for validation studies, even when their effects on purified CSCs are not yet available.

*Address correspondence to this author at the Bilkent University, Department of Molecular Biology and Genetics, Ankara, Turkey 06800; Tel: +90-312-290-2507; Facsimile: +90-312-266-509; E-mail: agure@bilkent.edu.tr

2212-389X/13 $58.00+.00

MATERIALS AND METHODS

In Silico Classification of Breast Cancer Cell Lines and Identification of Drugs with Differential Effects

The GSE36133 dataset containing expression data from all cell lines with drug screening data in the CCLE project was RMA normalized using GeneSpring v.12.5. Breast cancer cell lines were hierarchically clustered based on the probeset lists reported by Gupta *et al*. [4] or Kao *et al.* [5]; with Euclidian distance measures for both genes and arrays, and complete linkage, using Cluster 3.0 software. Drug toxicity data was obtained from the CCLE project (http://www.broadinstitute.org/ccle) [3]. Groups identified as CSC-like and non-CSC-like were compared using t-test (2-sided, equal variance) for all parameters (activity area, Amax, IC50, and EC50). $p \leq 0.05$ was considered statistically significant.

ANNOTATING GLOBAL EFFECTS OF DRUGS ON CANCER CELL LINES

Drug toxicity (activity area) data was obtained for all drugs contained within the CCLE project database [3], and was plotted for each cell line type following mean-sorting, using the "boxplot" function of the graphics package of R [6]. Cell line annotations are those provided by CCLE, except for the two breast cancer cell types which were identified in this study.

RESULTS

In Silico Classification of Breast Cancer Cell Lines into CSC-Like and Non-CSC-Like Populations

Several recent studies have generated gene lists that define CSC-like populations. One such list generated by Gupta *et al*, consists of genes differentially expressed in cells that resist paclitaxel but are sensitive to salinomycin, drugs known to effect non-CSCs or CSCs, respectively. The list was validated based on its ability to correctly identify breast CSCs as defined by other methods [4]. We asked if this signature could be used to differentiate commonly used breast cancer cell lines into distinct populations. Indeed, hierarchical clustering of the 27 breast cancer cell lines with screening data in the CCLE project revealed two distinct clusters, one that contained 6 cell lines which up-regulated most CSC-related genes and down-regulated all genes associated with non-CSCs, clearly categorizing them as CSC-like. The second cluster behaved in exactly the opposite manner, indicating that they were, or at least contained, primarily non-CSC or epithelial cells (Fig. **1**). In another study aiming to classify breast cancer cell lines into subgrups, Kao *et al*. revealed the presence of 3 distinct subtypes which they coined basal-A, basal-B and luminal [5]. When we used this gene list to cluster all CCLE breast cancer cell lines, we observed that all CSC-like cells as defined by Gupta *et al*.'s

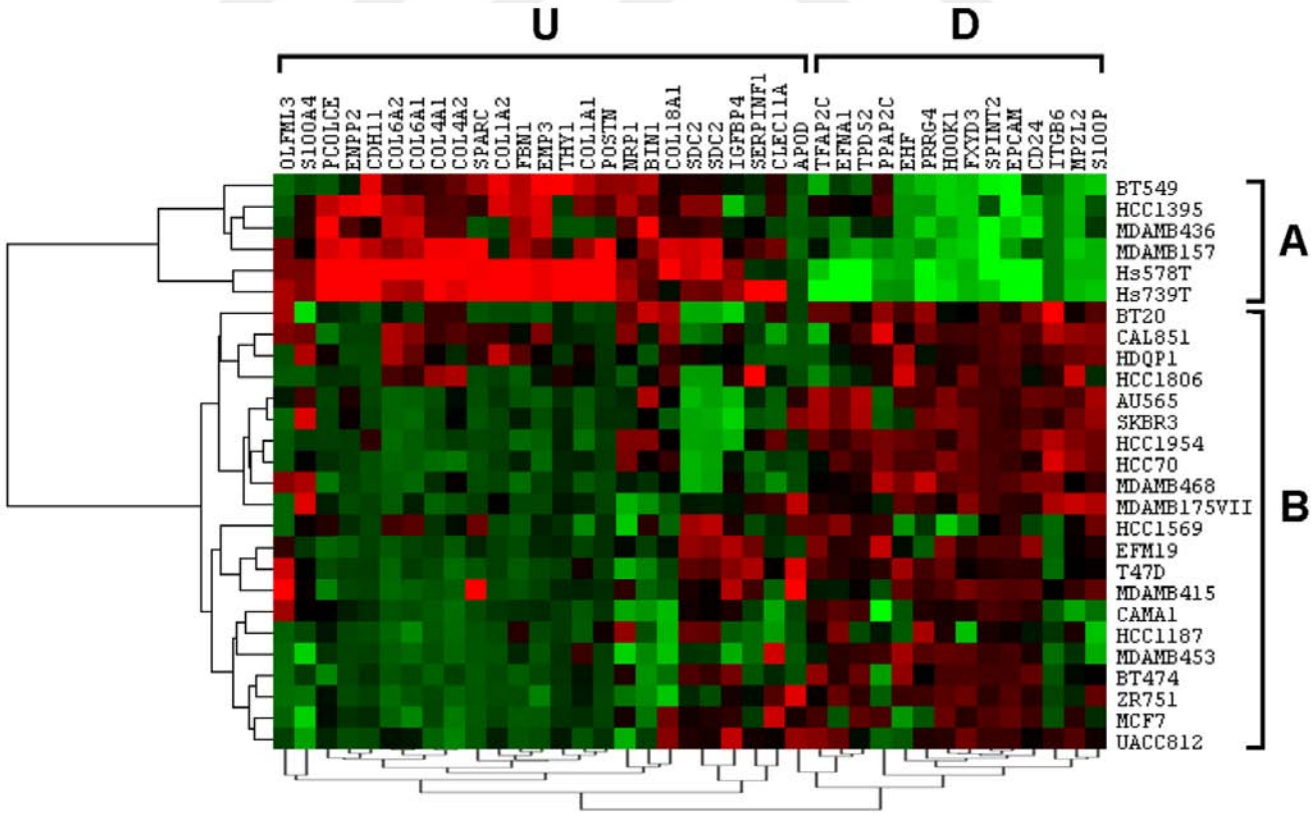


Fig. (1). Hierarchical clustering of the 27 breast cancer cell lines used in the CCLE project with a breast CSC signature (ref. 4). The list used for clustering contains 14 genes that are down-regulated (D), and 25 that are upregulated (U) in salinomycin-sensitive cells (CSCs). Clustering of the CCLE cell lines reveals two distinct groups, one which shows the same expression pattern as CSCs, which we therefore, denote "CSC-like"(A). The second cluster, behaves in the opposite fashion, and are therefore, coined "non-CSC-like" or epithelial cells (B). Colors are standardized to represent maximum and minimum expression for each gene. Red: up-, green: down-regulation.

list corresponded to basal-B cells, while non-CSC-like cells were basal-A or luminal type cell lines (Supplementary Fig. **1**). In our analysis, basal-B cells formed an independent cluster, while basal-A and luminal cells co-clustered. Thus, two independent gene lists classified existing breast cancer cell lines into the same two major groups.

IDENTIFYING DRUGS TARGETING CSC-LIKE BREAST CANCER CELL LINES

The CCLE database contains cytotoxicity data reported as 4 different parameters (IC50, EC50, Amax and activity area) for 24 drugs. Using this data, we determined drugs which showed statistically different effects on the CSC-like

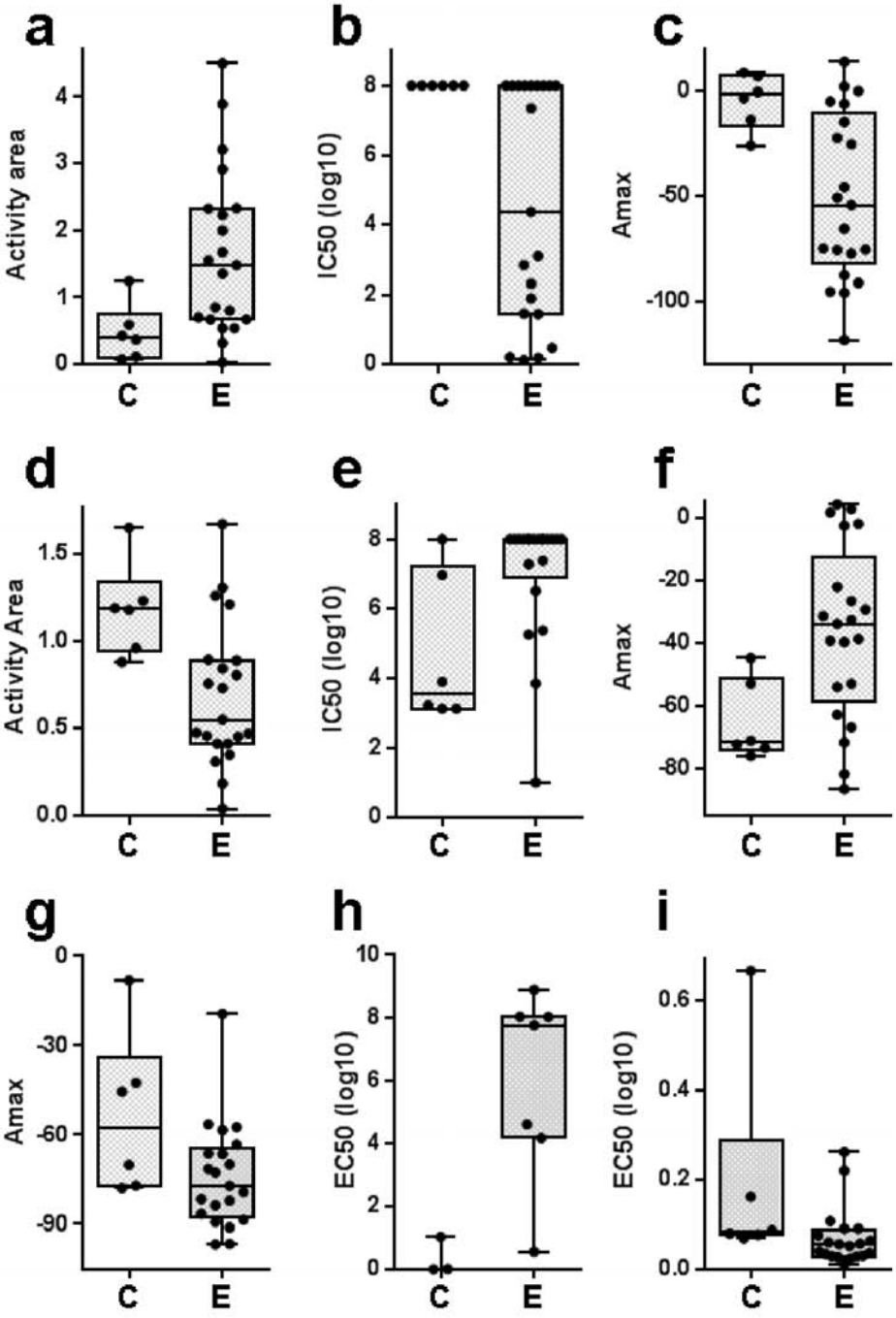


Fig. (2). Drugs with differential effects on CSC-like and non-CSC-like cells. Inhibitory effects that are significantly different among CSC-like (C) and non-CSC-like (E) cell lines are shown. Drugs: Lapatinib (a-c), TKI258 (d-f), 17AAG (g), L685458 (h), and Panobinostat (i). The measure of cytotoxicity for which significance was observed is indicated on the y-axis. Increased activity area, or decreased Amax, IC50 or EC50 values indicate increased drug sensitivity. Box includes values between the 25th and 75th percentiles, with the median indicated. Whiskers are from the minimum to the maximum value. *P* values obtained by the two-tailed t-test are: a: 0.03, b: 0.02, c: 0.01, d: 0.01, e: 0.01, f: 0.02, g: 0.04, h: 0.01, i: 0.05.

and non-CSC-like cell lines. We thus identified TKI258 (dovitinib, a FGFR, VEGFR and PDGFR inhibitor) and L-685458 (a γ -secretase inhibitor) as agents that were preferentially active on CSC-like cells. TKI258 showed significantly better toxicity on the CSC-like cells when analyzed using 3 of the 4 parameters of cell viability (Fig. **2**). Non-CSC cells were relatively sensitive to Lapatinib (a dual HER2/neu and EFGR inhibitor), 17-AAG (a HSP90 inhibitor), and Panobinostat (histone deacetylase inhibitor). Significant sensitivity to Lapatinib was evident from analysis performed with 3 out of 4 measures of cytotoxicity (Fig. **2**).

GLOBAL DRUG EFFECTS ON CANCER CELL LINES BASED ON TISSUE OF ORIGIN

Comparing two groups of cells, albeit useful, has a caveat in that a drug with excellent toxicity for both groups will not be identified if it affects both groups similarly, or alternatively if the drug's active concentration range is narrow. To overcome this caveat we wanted to view the "overall" effect of all CCLE drugs on all cancer cell lines including the two breast cancer cell line groups we identified. We chose "activity area" as the reference measure as it generates a value for each cell/treatment condition as opposed to IC50 or EC50. When tumor cell lines, grouped by tissue of origin, were sorted based on the mean value for activity area from lowest (most active) to the highest (least active) for TKI258, we observed that the two breast cancer subpopulations were indeed quite distant from each other (Fig. **3**). However, the fact that the non-CSC-like breast cancer cell lines were among the most resistant to this drug, obviously contributed to the significant difference observed between the two cell groups. Thus, it needs to be determined whether the doses of this drug at which CSC-like cells are affected can be tolerated *in vivo*. If most CSC-like cells are within this range, then the drug would be a suitable agent for treating CSC-like breast cancer cells. Similar analyses

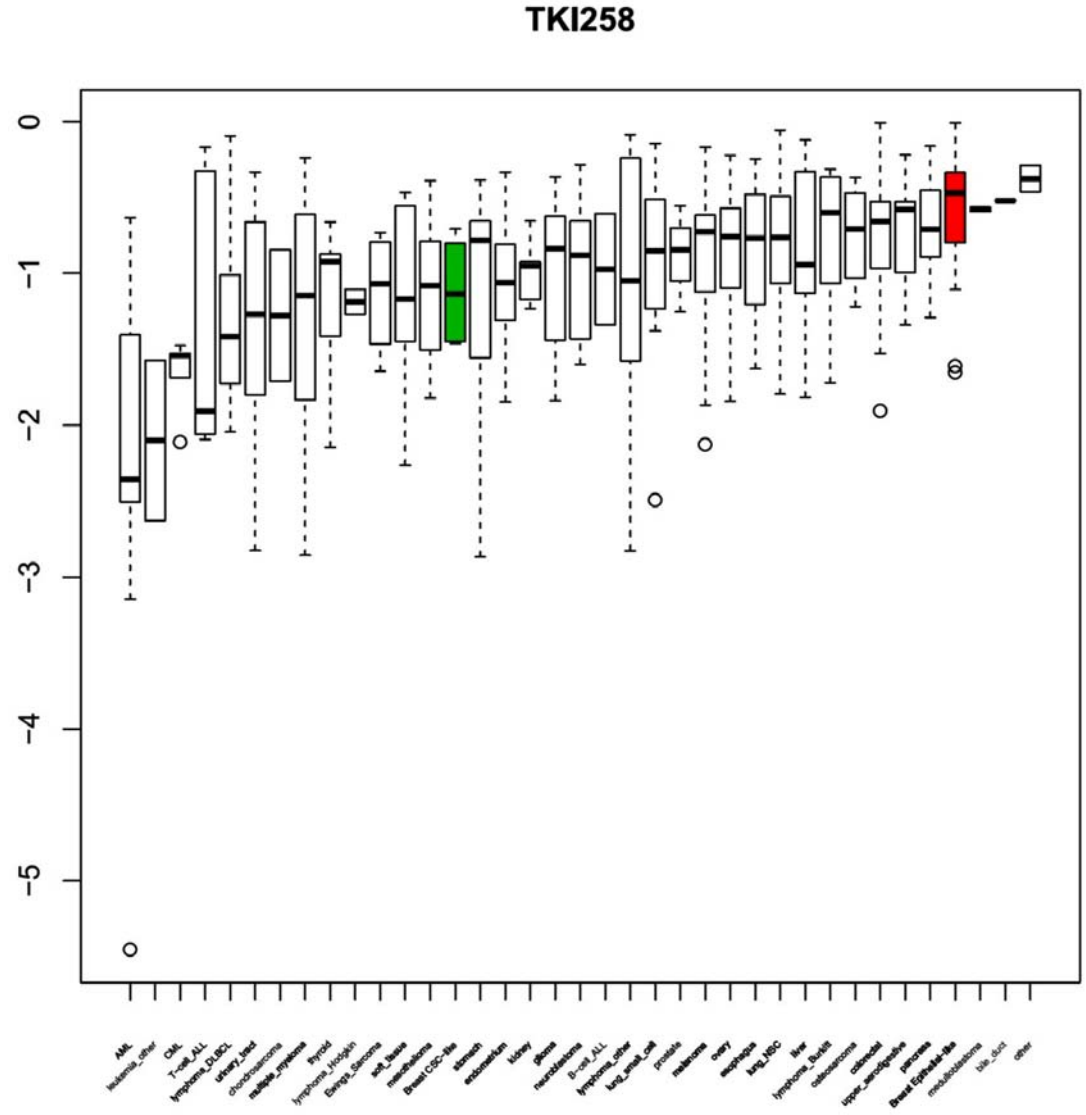


Fig. (3). TKI258 sensitivity of tumor cell lines. TKI258 activity on cell lines (activity area) grouped by tissue of origin is shown. Box represents 25th to 75th percentile, and whiskers the 2 standard deviations above and below the mean value of the activity area values obtained for each cell line type. Outliers are represented by circles. Line indicates the median. Cell lines are mean-sorted from the most sensitive (left) to the least. CSC-like (green) and non-CSC-like (red) breast cancer cell lines are indicated.

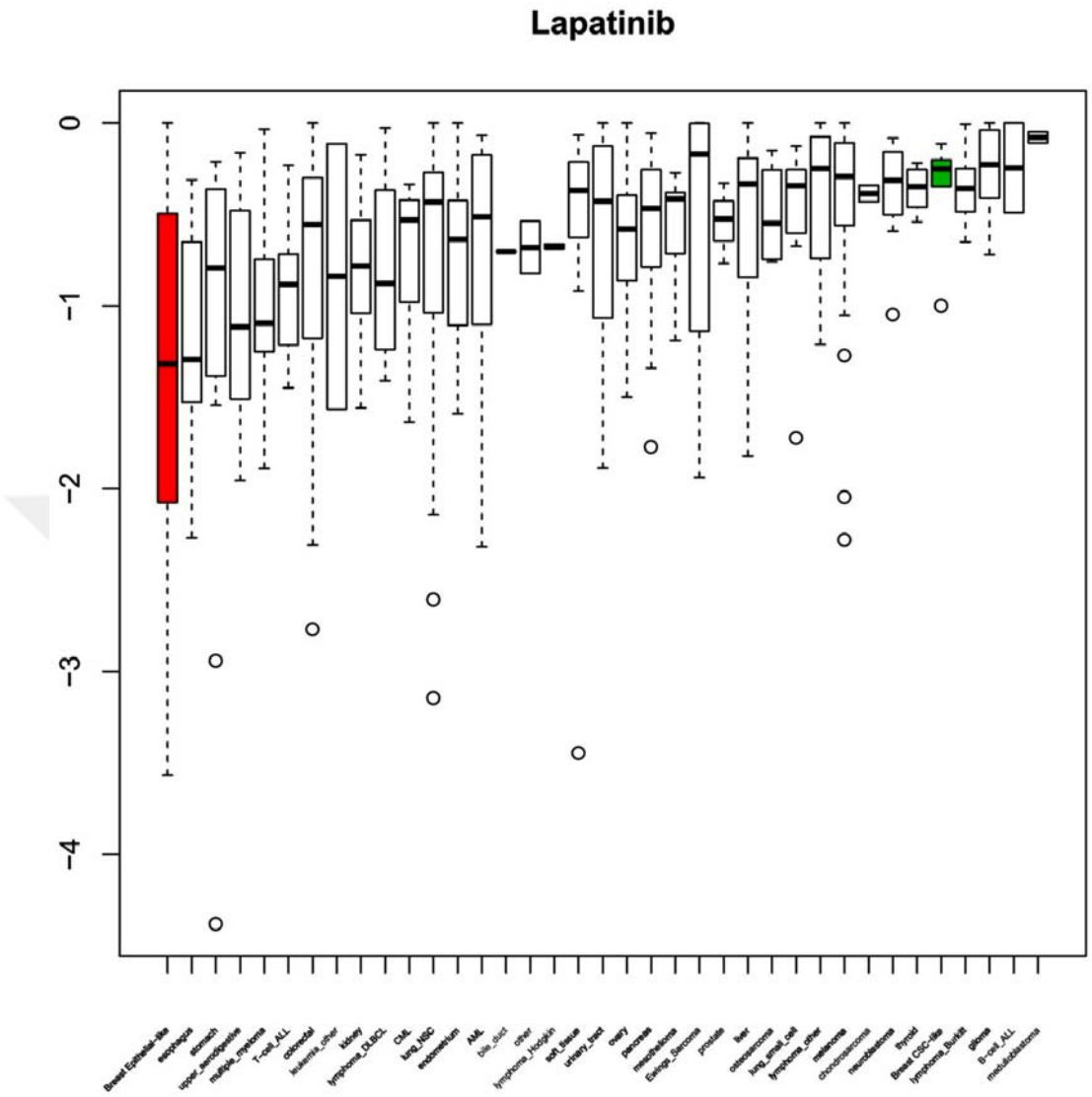


Fig. (4). Lapatinib sensitivity of tumor cell lines. Lapatinib activity on cell lines (activity area) grouped by tissue of origin is shown. Box represents 25th to 75th percentile, and whiskers the 2 standard deviations above and below the mean value of the activity area values obtained for each cell line type. Outliers are represented by circles. Line indicates the median. Cell lines are mean-sorted from the most sensitive (left) to the least. CSC-like (green) and non-CSC-like (red) breast cancer cell lines are indicated.

revealed that the B-Raf and VEGFR2 inhibitor Raf265, although not significant by t-test analysis, was in fact a better differentiator of the two breast cancer cell line populations based on how far apart the two populations were in this graphic (Supplementary Fig. **2**). Again, the value of this observation strongly depends on the effective dose range of the drug *in vivo*. The non-CSC-like breast cancer cells, on the other hand, were found to be the most sensitive cells, compared to all other cancers when evaluated for lapatinib toxicity. In contrast, CSC-like breast cancer cells were very resistant to lapatinib (Fig. **4**). This is an interesting finding as lapatinib has been recently approved for hormone-positive and HER2 positive breast cancer treatment. Even if most non-CSC-like breast cancer cells are sensitive to this agent, our data suggests that lapatinib treatment alone will likely not affect CSC-like breast cancer cells. In contrast, the IAP (inhibitor of apoptosis) inhibitor LBW242 was the drug to which the CSC-like breast cancer cells were most sensitive, when compared to all other drugs (Supplementary Fig. **3**). The drug is only moderately toxic for non-CSC-like breast cancer. Therefore, if tumors are found to contain a mixture of CSC-like and non-CSC-like populations, this analysis suggests that lapatinib/LBW242 combination therapy could be worth further validation.

DISCUSSION

We herein describe two straight-forward approaches by which large-scale gene expression and drug sensitivity data can be utilized to identify suitable chemotherapeutics for a given cell type. Our first approach is based on defining subsets of cell lines based on published gene lists. It is interesting to observe how two such independent lists generate the same clusters for breast cancer cell lines. Gene lists have been used to reproducibly cluster breast cancer into "intrinsic subtypes" that consist primarily of luminal and basal types [7, 8]. It is, therefore, not surprising that

signatures consisting of the most differentially expressed genes among cell groups overlap with the intrinsic subtypes. On the other hand, that such a classification reflects a stem-cell character is in our opinion very valuable. This not only implies that such cell lines can represent CSCs, but it also suggests that larger numbers of cells within a tumor might take upon themselves a stem-like character. In fact, the mechanism behind epithelial to mesencymal transition (EMT) or its reverse (MET), is exactly this [9]. Mesenchymal cells are slow growing, less differentiated, invasive and tumorigenic cells, resembling CSCs in many ways. In this line, for melanoma, colorectal cancer and others, EMT or a similar profile of "switchable" phenotypes has been associated with the tumorigenesis process or the variation of cells within the tumor. Therefore, even if cell line subtypes do not consist of CSCs exclusively, they still can represent behavior of such cells. Nevertheless, analyses described within this study do not eliminate the need for validation studies where CSC-like cells would be tested for qualities of colony formation, proliferation, invasion and such to demonstrate a functional similarity to CSCs. As the drug effects predicted in this study are very likely to be reproduced if the same cells are tested *in vitro*, it would be important to perform clustering and cell toxicity experiments in cells that are not included in the CCLE cell line cohort. In this study, paclitaxel activity on CSC-like and non-CSC-like cells was not found to differ significantly. As Gupta *et al*. showed differential toxicity of this drug against the two populations, the two cell populations identified here in in Gupta *et al*.'s study are not identical [4]. However, as different subtypes of breast cancer (represented by the cell subtypes identified in this study) are known to benefit from different chemotherapeutics, it could be argued that even if cell line sub-populations are not CSCs, they do represent a distinct phenotype that resembles CSCs in certain aspects. The validity of the approach described in this study will be strengthened if drugs that are found to affect CSC-like cells are also able to kill phenotypically similar populations of CSCs.

We believe, a global view of drug sensitivity, showing sensitivity profiles of all tumor cells for a given drug, is helpful in determining chemotherapeutics better suited for clinical use for a given tumor type. A most striking finding is that among all cell lines, non-CSC-like breast cancer cells are most sensitive to lapatinib, a drug currently approved for breast cancer treatment. Equally revealing, is the finding that CSC-like cells are relatively resistant to the same drug, when they are sensitive to agents like LBW242 or Raf265. Thus, our analysis reveals the presence of two distinct breast cancer cell types for which different treatment regimens might be appropriate and those drugs that are candidates for combination therapy. All drugs within the CCLE database are FDA-approved, and have a limited effective concentration range. Generating the *in vitro* counterpart to the maximum tolerable *in vivo* dose for cell lines, combined with this global analysis, would be very helpful in identifying tumor types for which a given drug is suitable *in vivo*.

Received: November 19, 2013 Revised: January 21, 2014 Accepted: January 29, 2014

CONCLUSION

We show that the classification of breast cancer cell lines into phenotypically distinct groups facilitates the identification of chemotherapeutics that might individually target CSC-like as well as non-CSC-like cells. The overall drug sensitivity profile of cell lines, on the other hand, reveals that drugs like Raf265 or LBW242 might affect CSC-like breast cancer cell lines at concentrations that are possibly well tolerated.

CONFLICT OF INTEREST

The author(s) confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

MI: R based bioinformatics analyses, re-annotation of cell lines, cluster analyses and CCLE drug profiling experiments. KMS: Clustering and drug identification based on the CSC-list (Gupta *et al*.). AOG: Overall design of the study.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's web site along with the published article.

LIST OF ABBREVIATIONS

CSC	=	cancer stem cell
CCLE	=	cancer cell line encyclopedia
COST	=	intergovernmental framework for European Cooperation in Science and Technology
EMT	=	epithelial to mesenchymal transition
FDA	=	food and drug administration
RMA	=	robust multi-array average

REFERENCES

[1] Gilbertson, R.J. and T.A. Graham, *Cancer: Resolving the stem-cell debate*. Nature, 2012. **488**(7412): p. 462-3.

[2] Magee, J.A., E. Piskounova, and S.J. Morrison, *Cancer stem cells: impact, heterogeneity, and uncertainty*. Cancer Cell, 2012. **21**(3): p. 283-96.

[3] Barretina, J., *et al*., *The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity*. Nature, 2012. **483**(7391): p. 603-7.

[4] Gupta, P.B., *et al*., *Identification of selective inhibitors of cancer stem cells by high-throughput screening*. Cell, 2009. **138**(4): p. 645-59.

[5] Kao, J., *et al*., *Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery*. PLoS ONE, 2009. **4**(7): p. e6146.

[6] The R Development Core Team, *R: A Language and Environment for Statistical Computing*. 2010.

[7] Perou, C.M., *et al*., *Molecular portraits of human breast tumours*. Nature, 2000. **406**(6797): p. 747-52.

[8] Sorlie, T., *et al*., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications*. Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10869-74.

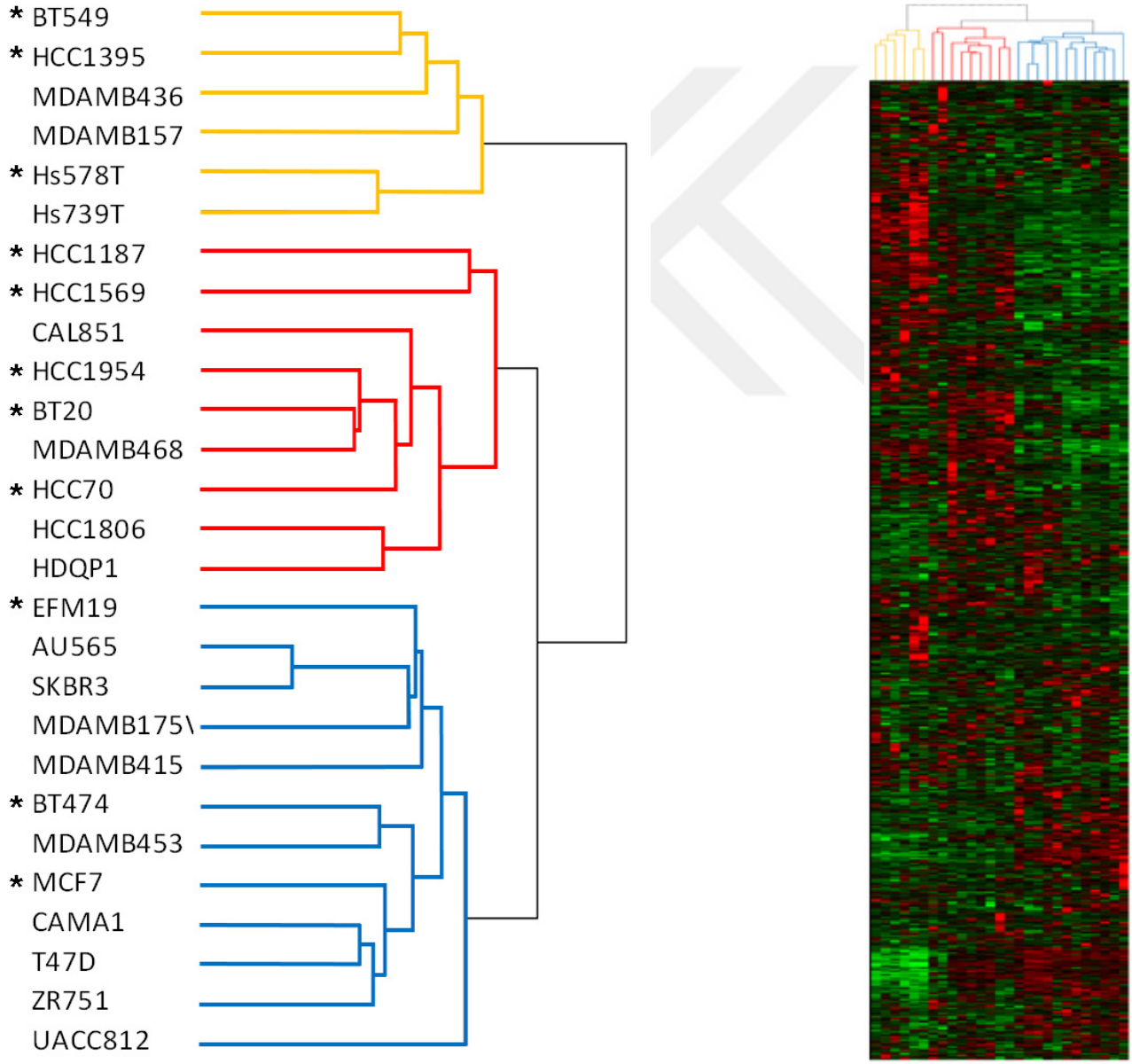
[9] Mani, S.A., *et al*., *The epithelial-mesenchymal transition generates cells with properties of stem cells*. Cell, 2008. **133**(4): p. 704-715.

SUPPLEMENTARY MATERIAL

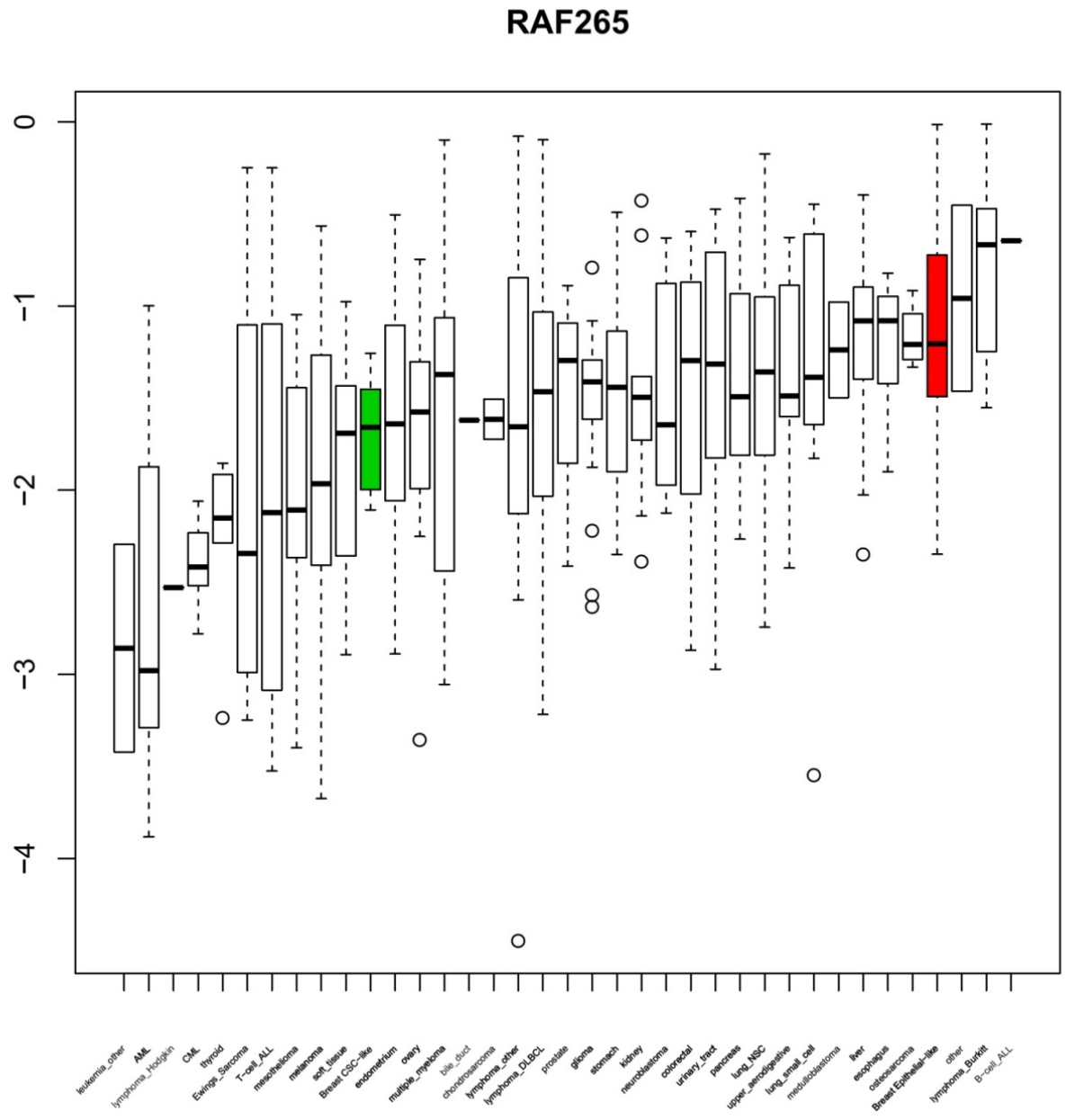
Predicting Chemotherapy Sensitivity Profiles for Breast Cancer Cell Lines with and Without Stem Cell-Like Features

Murat Isbilen, Kerem Mert Senses and Ali Osmay Gure*

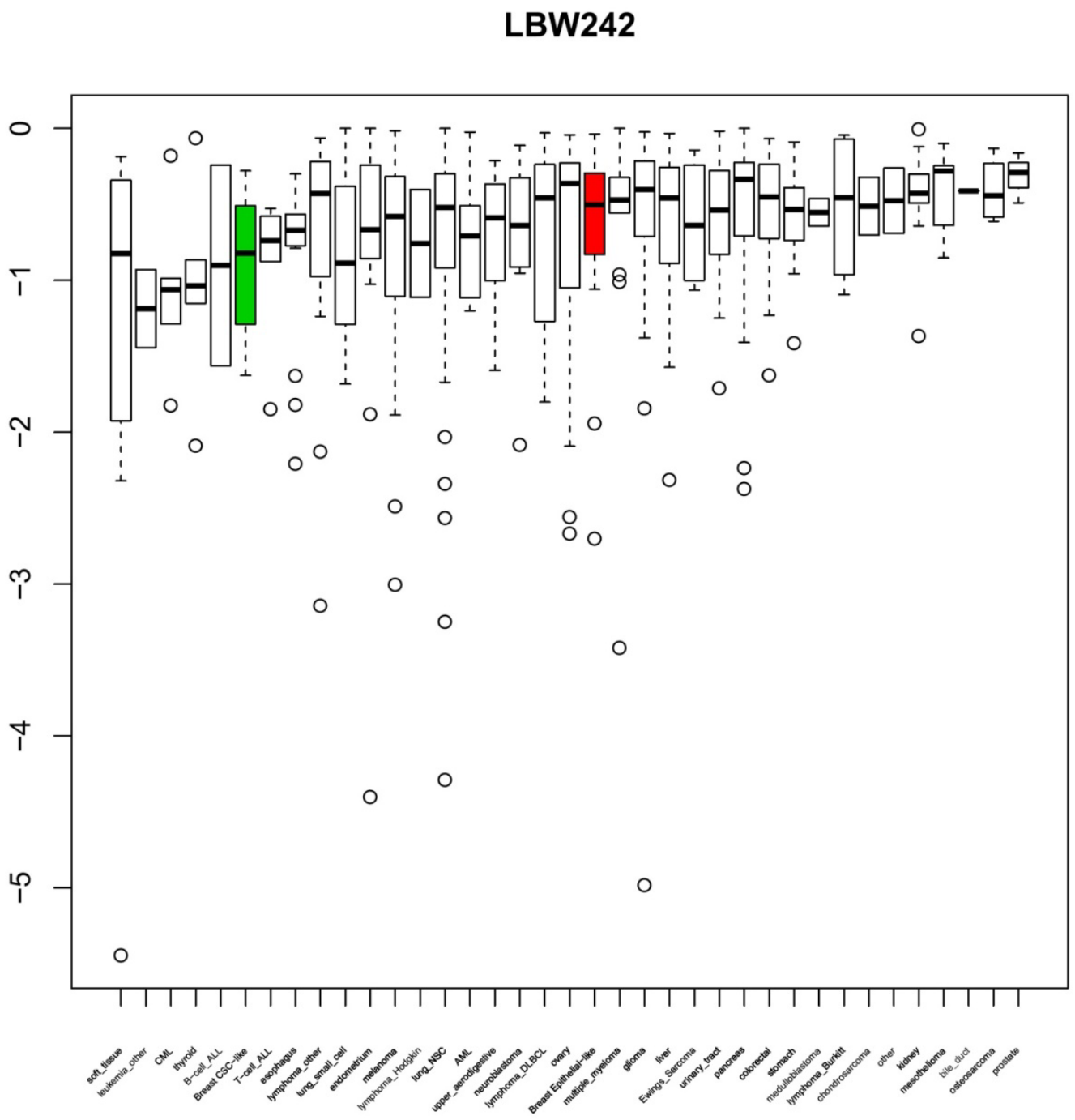
Bilkent University, Department of Molecular Biology and Genetics, Ankara, Turkey 06800



Supplementary Fig. (1). Hierarchical clustering analysis of CCLE breast cancer cell lines with Kao *et al.*'s gene list (ref. 5). Cluster analysis of all breast cancer cell lines, including those that were previously analyzed by Kao *et al*. (ref. 5), reveal three groups: Luminal (blue), Basal A (red) and Basal B (yellow). The heatmap and details of the cluster are shown. The 6 CSC-like cell lines are all of the "Basal B" type. The gene list used for clustering included 12628 probesets corresponding to 5771 of the 8750 genes used for the original analysis, as only these genes had corresponding annotations for the HGU 133 Plus 2.0 platform.



Supplementary Fig. (2). RAF265 sensitivity of tumor cell lines. RAF265 activity on cell lines (activity area) grouped by tissue of origin is shown. Box represents 25th to 75th percentile, and whiskers the 2 standard deviations above and below the mean value of the activity area values obtained for each cell line type. Outliers are represented by circles. Line indicates the median. Cell lines are mean-sorted from the most sensitive (left) to the least. CSC-like (green) and non-CSC-like (red) breast cancer cell lines are indicated.



Supplementary Fig. (3). LBW242 sensitivity of tumor cell lines. LBW242 activity on cell lines (activity area) grouped by tissue of origin is shown. Box represents 25th to 75th percentile, and whiskers the 2 standard deviations above and below the mean value of the activity area values obtained for each cell line type. Outliers are represented by circles. Line indicates the median. Cell lines are mean-sorted from the most sensitive (left) to the least. CSC-like (green) and non-CSC-like (red) breast cancer cell lines are indicated.

OPEN ACCESS Freely available online

PLOS ONE

Epigenetic Mechanisms Underlying the Dynamic Expression of Cancer-Testis Genes, *PAGE2, -2B* and *SPANX-B*, during Mesenchymal-to-Epithelial Transition

Sinem Yilmaz-Ozcan[1], Asli Sade[2], Baris Kucukkaraduman[1], Yasemin Kaygusuz[1], Kerem Mert Senses[1], Sreeparna Banerjee[2], Ali Osmay Gure[1]*

1 Department of Molecular Biology and Genetics, Bilkent University, Ankara, Turkey, **2** Department of Biological Sciences, Middle East Technical University, Ankara, Turkey

Abstract

Cancer-testis (CT) genes are expressed in various cancers but not in normal tissues other than in cells of the germline. Although DNA demethylation of promoter-proximal CpGs of CT genes is linked to their expression in cancer, the mechanisms leading to demethylation are unknown. To elucidate such mechanisms we chose to study the Caco-2 colorectal cancer cell line during the course of its spontaneous differentiation *in vitro*, as we found CT genes, in particular *PAGE2, -2B* and *SPANX-B*, to be up-regulated during this process. Differentiation of these cells resulted in a mesenchymal-to-epithelial transition (MET) as evidenced by the gain of epithelial markers CDX2, Claudin-4 and E-cadherin, and a concomitant loss of mesenchymal markers Vimentin, Fibronectin-1 and Transgelin. PAGE2 and SPAN-X up-regulation was accompanied by an increase in Ten-eleven translocation-2 (TET2) expression and cytosine 5-hydroxymethylation as well as the disassociation of heterochromatin protein 1 and the polycomb repressive complex 2 protein EZH2 from promoter-proximal regions of these genes. Reversal of differentiation resulted in down-regulation of *PAGE2, -2B* and *SPANX-B*, and induction of epithelial-to-mesenchymal transition (EMT) markers, demonstrating the dynamic nature of CT gene regulation in this model.

Citation: Yilmaz-Ozcan S, Sade A, Kucukkaraduman B, Kaygusuz Y, Senses KM, et al. (2014) Epigenetic Mechanisms Underlying the Dynamic Expression of Cancer-Testis Genes, *PAGE2, -2B* and *SPANX-B*, during Mesenchymal-to-Epithelial Transition. PLoS ONE 9(9): e107905. doi:10.1371/journal.pone.0107905

Editor: Keping Xie, The University of Texas MD Anderson Cancer Center, United States of America

Received May 17, 2014; **Accepted** August 22, 2014; **Published** September 17, 2014

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by grant nr. 112S023 from The Scientific and Technological Research Council of Turkey (TUBITAK) to AOG, a Young Investigator Award from The Turkish Academy of Sciences to SB, and by TUBITAK-BIDEP fellowships to SYO and BK. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: agure@bilkent.edu.tr

Introduction

Cancer-testis (CT) or cancer-germline genes are expressed in tumors originating from various tissues, as well as in normal germline and trophoblast cells, but are generally silent in other normal tissues of the adult [1–3]. More than 100 different CT genes can be grouped according to homology into families [2]. Despite the lack of sequence similarity between CT genes from different families, re-expression of all CT genes in tumors has been associated with demethylation of CpG residues within their promoter-proximal regions [4]. This shared mechanism of expression regulation is most likely the reason for their coordinate expression in cancer [5–7]. However, the exact mechanism by which DNA demethylation occurs at CT gene promoter-proximal regions in cancers is unknown. CT genes show mostly a heterogeneous expression pattern in tumors [8–10]; in contrast to their expression in testis, which is demarcated and orderly [11]. A study in which stem-like and non-stem like cells of breast cancer were selectively killed, revealed that CT gene expression is generally a feature of more differentiated, non-stem cells [12]. Similarly, in melanoma, a subgroup of cells with more epithelial features express CT genes, when cells with mesenchymal features don't [13]. Interestingly, melanoma cells can switch between these two classes *in vivo*, suggesting that tumor heterogeneity, as defined by CT gene expression, might represent a transitional phase similar to the switch between epithelial and mesenchymal phenotypes. Indeed, mesenchymal-to-epithelial transition (MET) is associated with the induction of CT gene expression [14]. To define mechanisms involved in regulating CT gene expression in cancer and during MET, we chose to study the Caco-2 spontaneous differentiation model which demonstrates features of MET and EMT during differentiation and de-differentiation, respectively. Our data reveal that the dynamic regulation of the two CT genes, *PAGE* and *SPANX-B* in this model system, involves alterations of polycomb repressive complex 2 (PRC2) and heterochromatin protein 1 (HP1) occupancy within their promoter-proximal regions, with concordant changes in TET expression and cytosine hydroxymethylation (hmC) levels.

Methods

Cell lines, induction of differentiation and de-differentiation

The Caco-2 cell line was obtained from the SAP Enstitusu (Ankara, Turkey). HCT116 (colorectal) and Mahlavu (hepatocellular) cancer cell lines were obtained from LGC Standards,

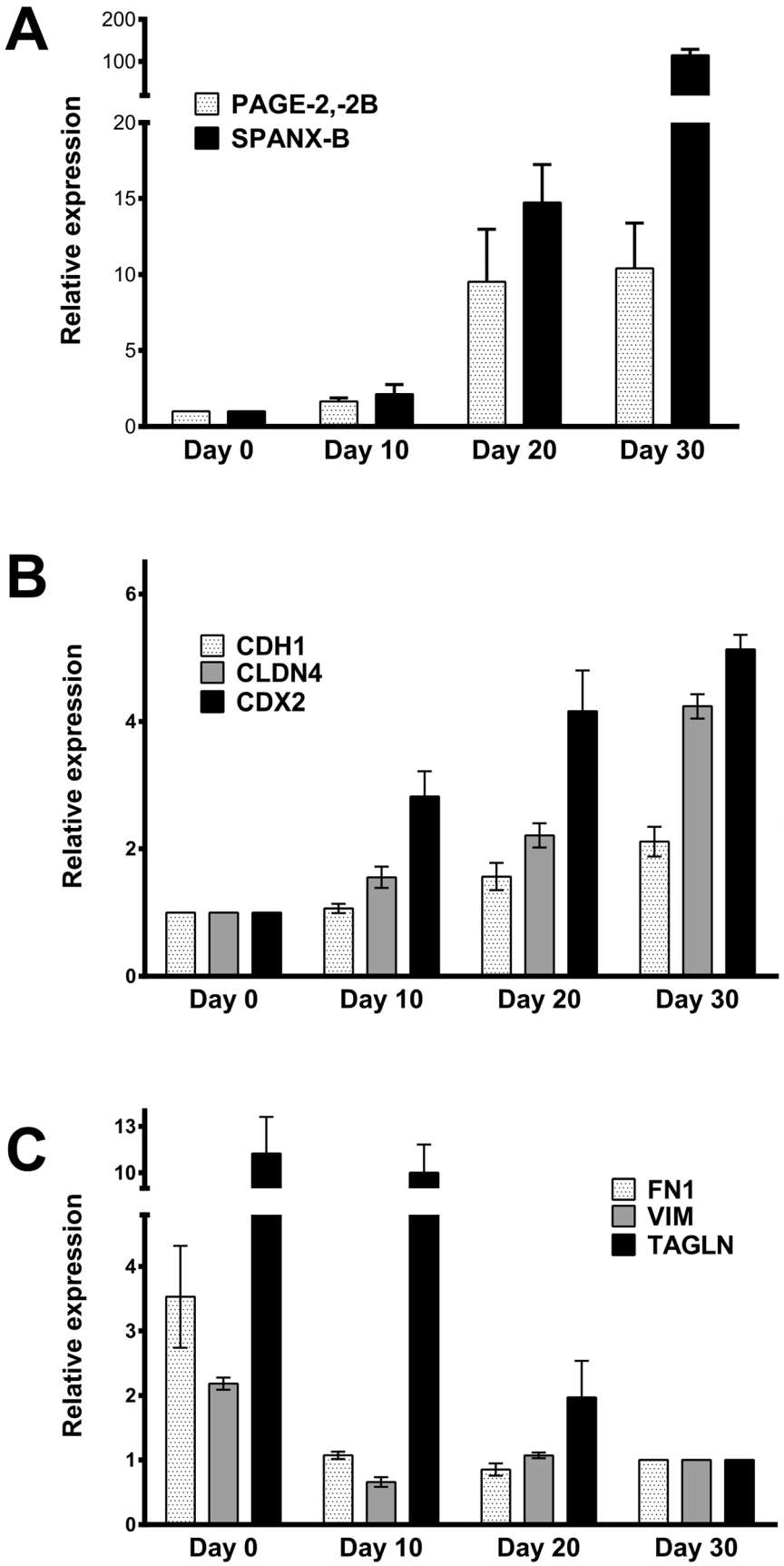


Figure 1. Up-regulation of CT genes in parallel to MET in the Caco-2 SD model. Relative mRNA expression of CT genes (*PAGE2*, *-2B* and *SPANXB*) (**A**), epithelial genes (*E-cadherin (CDH1)*, *claudin 4(CLDN4)*, *CDX2*) (**B**), and mesenchymal genes (*fibronectin 1 (FN1)*, *vimentin (VIM)*, *transgelin (TAGLN)*) (**C**) as determined by quantitative PCR at days 0, 10, 20 and 30 post-confluence. Data represent average of two experiments. Change in expression levels for all genes between days 0 and 30 is statistically significant ($P<0.0001$, by two way ANOVA with Tukey's post hoc test).
doi:10.1371/journal.pone.0107905.g001

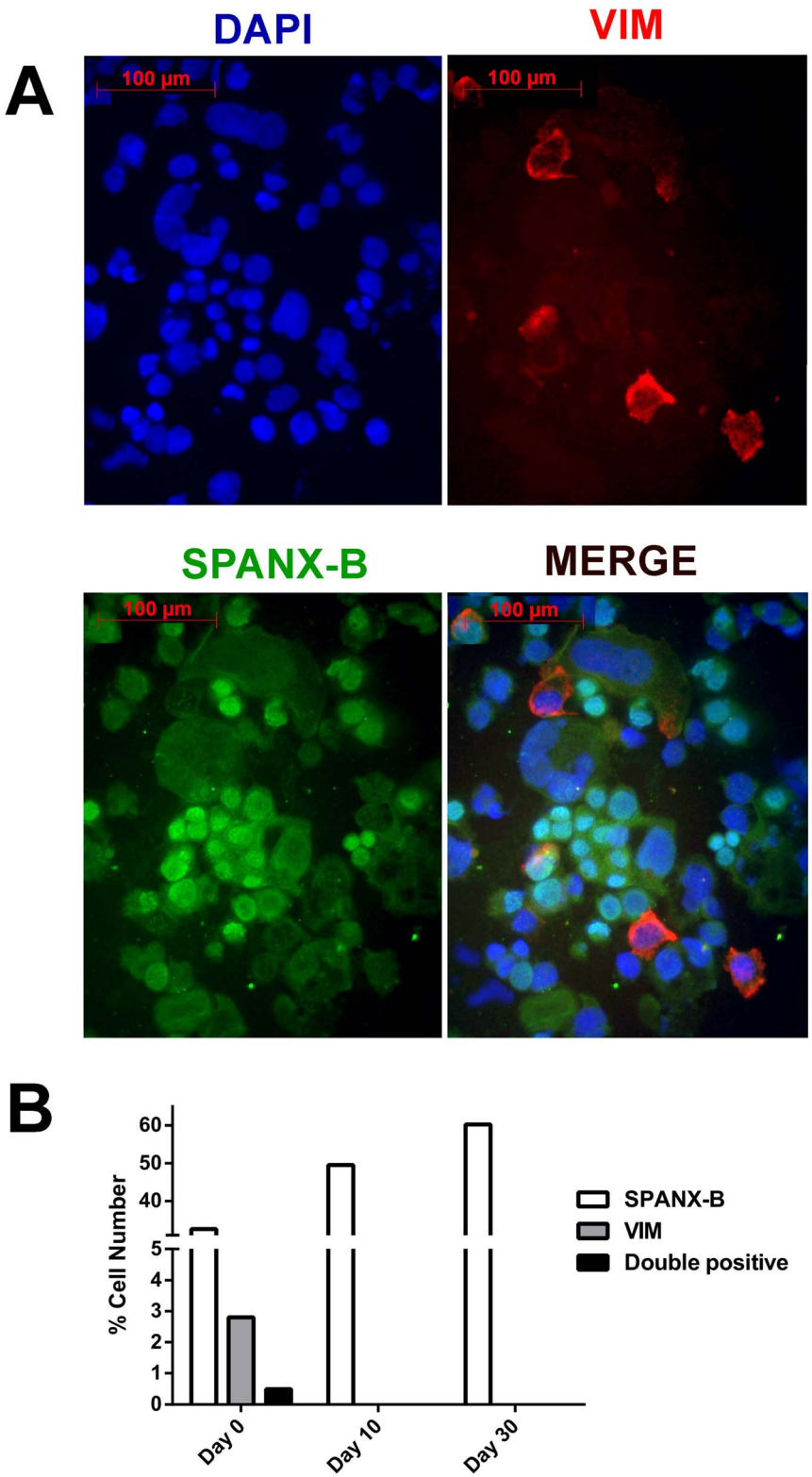


Figure 2. SPANX-B and vimentin expression are mutually exclusive in differentiating Caco-2 cells. Immunofluorescent staining of differentiating Caco-2 cells with DAPI counterstaining reveals a gradual increase in nuclear SPANX-B (green) with a concomitant decrease in cytoplasmic vimentin expression (red); (20× magnification) (**A**). Less than 1% of SPANX-B positive cells showed staining for vimentin on day 0 (**B**).
doi:10.1371/journal.pone.0107905.g002

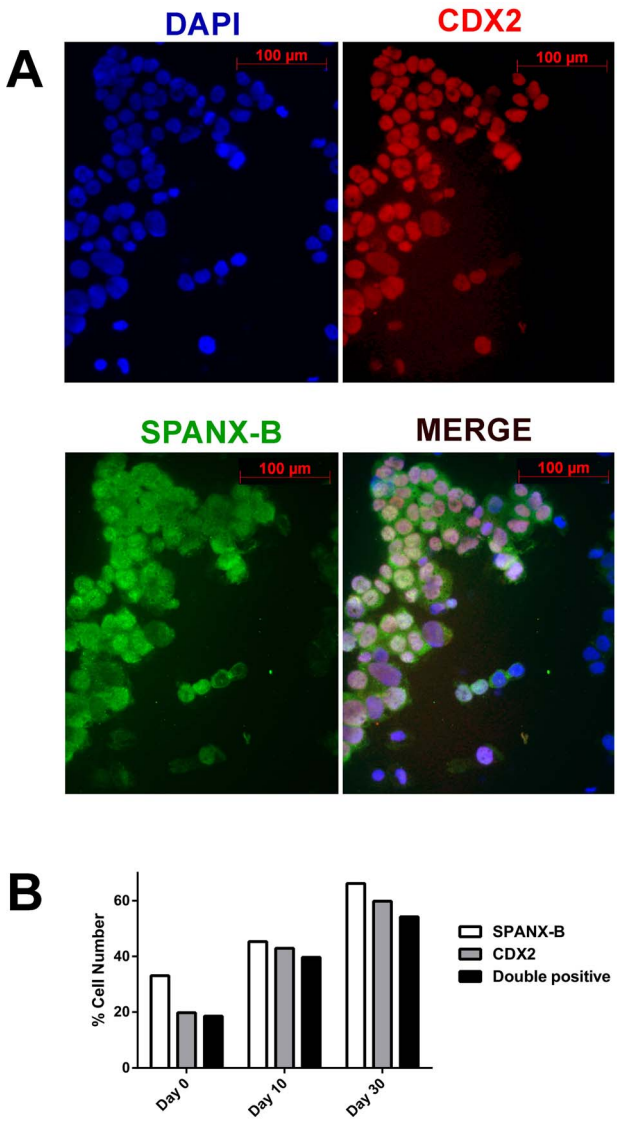


Figure 3. Nuclear co-localization of CDX2 and SPANX-B in differentiating Caco-2 cells. Immunofluorescent staining of differentiating Caco-2 cells with DAPI counterstaining reveals overlapping SPANX-B (Alexa Fluor 488: green) and CDX2 (Alexa Fluor 568: red) expression; (20× magnification) (**A**). More than 60 to 80% of the cells show double-labeling when analyzed quantitatively (**B**).
doi:10.1371/journal.pone.0107905.g003

Middlesex, UK. A lung cancer cell line (SK-LC-17) was from the Memorial Sloan Kettering Cancer Center, NY, USA. Caco-2 cells were grown in EMEM and others in RPMI, supplemented with 20% FBS, 2 mM L-glutamine, 0.1 mM non-essential amino acids, 1.5 g/L sodium bicarbonate and 1 mM sodium pyruvate. All cell culture media and supplements were purchased from Biochrom AG, Berlin, Germany. The first day cells reached confluence was designated day 0. Cells grown in parallel cultures were used to determine phenotypic changes at days 0, 5, 10, 20 and 30, post-confluence. Additional measures of differentiation for cells used in this study have been reported elsewhere [15]. To induce dedifferentiation, cells at the 20th day of differentiation were detached and replated at about 50% confluence and RNA and protein were harvested 5 days following replating.

In silico analysis of CT and EMT gene expression

Expression data contained within GSE1614 [16] was GC-RMA normalized using GeneSpring v. 11.0. CT gene expression was analyzed based on 31 probesets in corresponding to 23 CT genes from 7 families (**Figure S2**). An interpretation was generated with an entity list composed of EMT related genes as defined by Loboda et al. [17], at three different time points. Genes for validation were selected among those for which significant differences of expression ($p<0.05$) was observed by one way ANOVA test and Bonferroni FWER correction, when proliferating cells were compared to those at day 15.

Quantitative RT-PCR

Total RNA was isolated using the Trizol reagent (Ambion, Foster City, CA, USA) and treated with DNAse I (Ambion, Foster City, CA, USA). 200 ng of RNA was reverse transcribed using Revert-Aid first strand cDNA synthesis kit (Thermo Fisher Scientific, Boston, MA, USA). PCR reactions were performed using an ABI 7500 thermal cycler (Applied Biosystems, Carlsbad, CA, USA). All reactions were performed according to manufacturer's recommendations. TaqMan Gene Expression Assays (Applied Biosystems, Carlsbad, CA, USA) were used for the following: GAPDH (4352934E), SPANX-B (Hs02387419_gH), PAGE2 and -2B (Hs03805505_mH), GAGE (Hs00275620_m1), SSX4 (Hs023441531_m1), NY-ESO-1 (Hs00265824_m1), and MAGE-A3 (Hs00366532_m1). SYBR Green master mix with ROX reference dye (Applied Biosystems, Carlsbad, CA, USA) was used to determine *CDH1, CDX2, CLDN4, VIM, FN1* and *TAGLN* expression (**Table S1**). Cycling conditions for these assays were 50°C for 2 min., 95°C for 10 min., followed by 40 cycles of 94°C for 15 sec., 60–65°C for 1 min. Relative expression was calculated by the ΔΔCt method [18]. All samples were analyzed in triplicates and all experiments were repeated at least twice.

Promoter methylation analysis

Genomic DNA was isolated by Proteinase K treatment, following a phenol-chloroform extraction protocol. Bisulphite treatment of 200 ng genomic DNA was performed using Zymo DNA Methylation Gold Kit (Zymo Research, Irvine, CA, USA). Bisulphite modified DNA was stored at −20°C and used for PCR within 2 months. Two rounds of DNA amplification were performed using One Taq Hot Start DNA polymerase (New England Bioscience/NEB, Ipswich, MA, USA) using a Perkin Elmer 9700 thermal cycler (Applied Biosystems, Carlsbad, CA, USA). Primers used are given in **Table S1**. PCR products were gel extracted using the QIAGEN gel extraction kit (Qiagen, Hilden, Germany) and cloned into pCR2.1 (Invitrogen, Carlsbad, CA, USA). Plasmid DNA was purified using the QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany) from at least ten clones, and sequence analyzed by IONTEK (Istanbul, Turkey).

5-hydroxymethyl cytosine analysis

Caco-2 genomic DNA (gDNA) was sheared by probe sonication (30 sec. on, 30 sec. off, 5 cycles) to obtain 200–600 bp. fragments assessed by 1% agarose gel electrophoretic analysis. Immunoprecipitation was performed using the hMEDIP kit (Abcam, Cambridge, UK) according to manufacturer's instructions. 5 pg of control DNA was spiked into 500 ng of gDNA to use as an internal control. Positive and negative controls of the kit were included in all experiments. 2 µl from the eluted DNA was used as template for quantitative RT-PCR using 2 X SYBR Green master mix with ROX reference dye (Applied Biosystems, Carlsbad, CA,

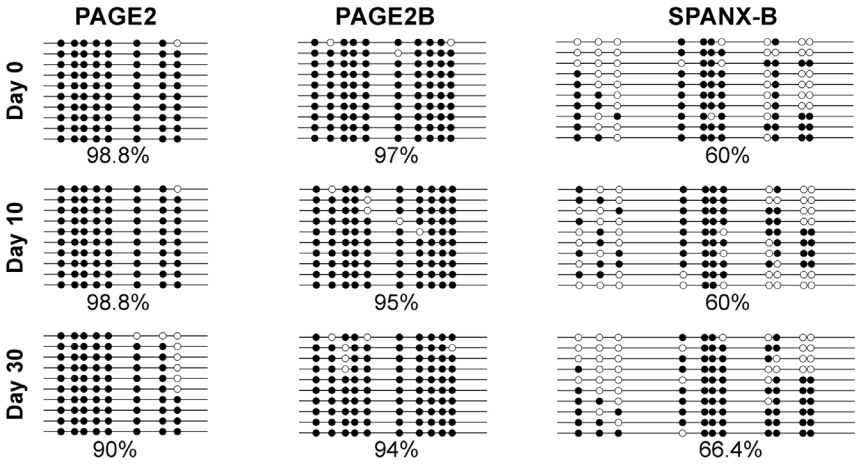


Figure 4. Bisulfide sequencing of *PAGE2, -2B* and *SPANX-B* promoter-proximal regions. Filled and empty circles represent methylated and unmethylated cytosines, respectively. % methylation within each analysed region, based on the 10 clones sequenced is indicated. CpG residues proximal to *PAGE2, -2B* and *SPANX-B* promoters during Caco2 differentiation at days 0, 10 and 30 reveals no statistically significant change (by one-way ANOVA).
doi:10.1371/journal.pone.0107905.g004

USA) with the primers given in **Tables S1 and S2**. Primer efficiencies were controlled. Cycling conditions were 50°C for 2 min., 95°C for 10 min. followed by 40 cycles of 94°C for 15 sec., 60°C for 1 min. Shared genomic DNA was included in quantitative RT-PCR to calculate % input.

Immunofluorescence microscopy

Cells attached to glass slides by centrifugation using the Shandon CytoSpin3 (Thermo Scientific, Waltham, MA, USA) were immediately fixed in 2% formaldehyde/PBS at room temperature for 15 min. Fixed cells were permeabilized in 0.2% Triton X-PBS for 10 min. followed by blocking with 1% BSA in 0.1% PBS-Tween for 1 hour. Incubations with the primary antibody, diluted at 1:50, were performed overnight at 4°C. Secondary antibody was added at 1:200, following washing in 0.1% PBS-Tween for 5 min. for 3 times, and incubated with cells for 45 min. at room temperature (see **Table S3** for the complete antibody list). Stained samples were mounted with mounting medium (Santa Cruz Biotechnology, Santa Cruz, CA, USA) containing DAPI solution. Cell lines used as positive controls were Mahlavu (PAGE-2,-2B and SPANXB), MDA-MB 231 (VIM), MCF-7 (TAGLN) and SW620 (CDX2, FN1). Negative controls were combinations of primary antibodies with un-related secondary antibodies. All images were obtained using an AxioCam MRc5 image capture device (Carl Zeiss, Oberkochen, Germany).

Western analysis

Cell lysates (extracted with RIPA buffer) separated on 4–12% Novex Bis-Tris SDS gels (Invitrogen, Carlsbad, CA, USA) were transferred to Immobilon-PSQ membranes (Millipore Corp. Bedford, MA, USA) with an Invitrogen western blotting system (Invitrogen. Carlsbad, CA, USA). Following blocking with 5% milk powder in 0.02% PBS-T, blots were incubated with primary antibody overnight at 4°C. Primary antibody dilutions were 1:1000 for CDX2, fibronectin, vimentin and transgelin, 1:2500 for β-actin and 1:100 for SPANX-B and PAGE-2,-2B antibodies. HRP conjugated secondary antibody (Abcam, Cambridge, UK) was used at a 1:5000 dilution and incubated at room temperature for 1 hour. Signals were detected using the ECL luminescence assay (BioRad, Hercules, CA, USA).

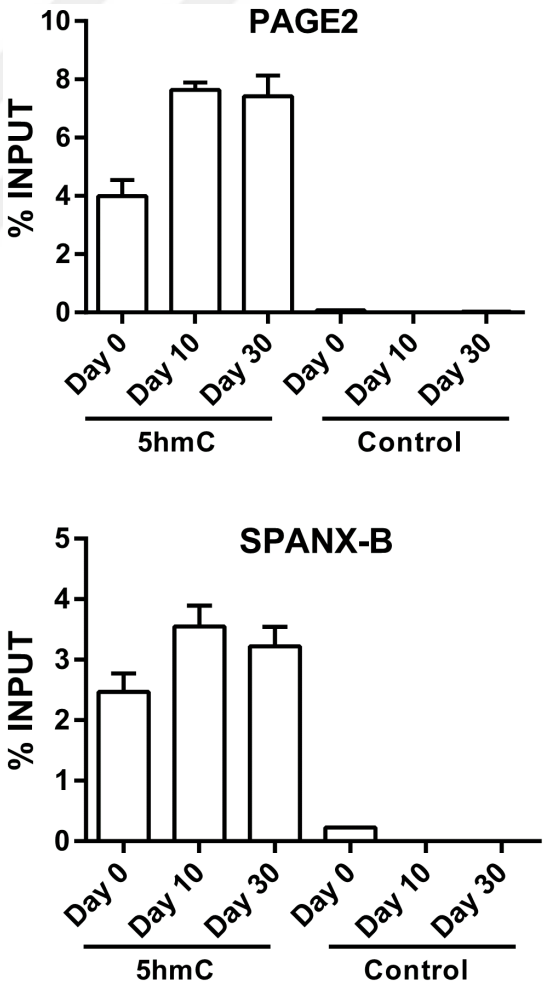


Figure 5. Increased hydroxymethylation of *PAGE2* and *SPANX-B* during Caco-2 spontaneous differentiation. CHIP experiments using an anti-hmC antibody and primers corresponding to +31 to +182 and +68 to +184 bp from the transcription start site of the *PAGE2* and *SPANX-B* genes, respectively. P values (one-way ANOVA) are 0.001 and 0.07 for PAGE2 and SPANX-B, respectively.
doi:10.1371/journal.pone.0107905.g005

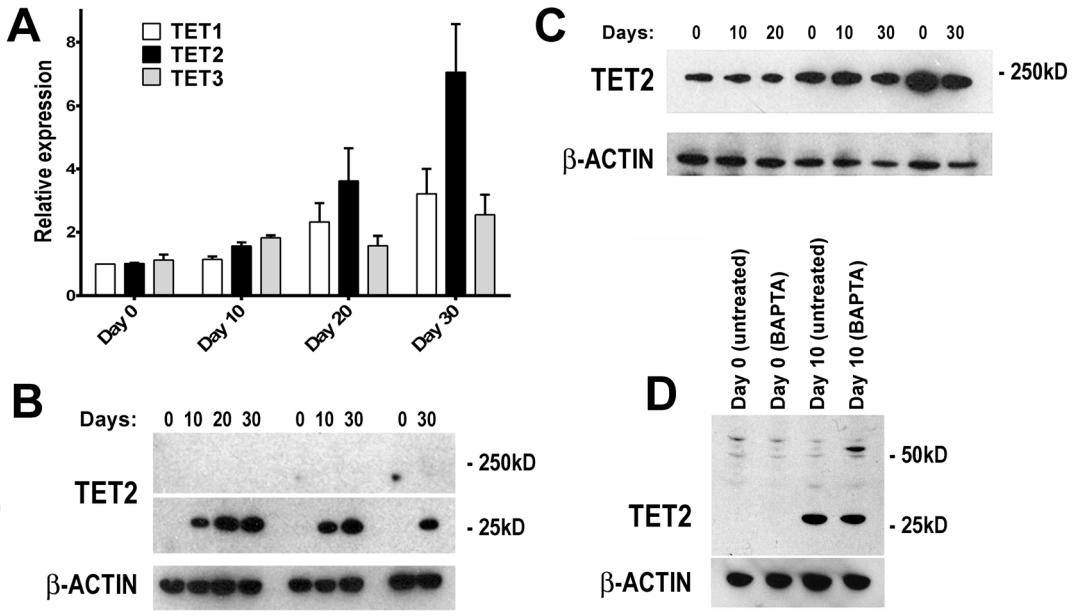


Figure 6. *TET* expression during Caco-2 SD. mRNA of all three *TET* genes increase gradually during Caco-2 SD (**A**). An increase in only a 25 kD version (**B**), but not the full-length TET2 protein (**C**) occurs simultaneously with the increase in mRNA. BAPTA-AM treatment results in a modest decrease in the 25 kD TET2 protein, with the generation of a larger mw version (**D**). P values, as determined by one-way ANOVA, are 0.03, 0.01, and 0.07, for Tet1, Tet2, and Tet3, respectively.
doi:10.1371/journal.pone.0107905.g006

Chromatin Immunoprecipitation

Chromatin Immunoprecipitation (ChIP) was performed as previously described [15]. Briefly, formaldehyde cross-linked cell constituents were precipitated by proten A sepharose beads coupled to antibodies against EZH2, HP-1 or H3K27m3 (Abcam, Cambridge, UK), as well as isotype-specific control. Precipitated DNA was amplified using primers specific for *PAGE2*, *-2* or *SPANX-B* promoter sequences (**Tables S1 and S2**), following de-crosslinking.

Results

CT gene expression during Caco-2 spontaneous differentiation (Caco-2 SD)

The undifferentiated colorectal cancer cell line Caco-2 undergoes enterocytic differentiation upon reaching confluence *in vitro* [19,20]. Gradual differentiation has been observed up to 30 days post-confluence as evidenced by the up-regulation of various differentiation-associated genes including sucrase-isomaltase, alkaline phosphatase and carcinoembryogenic antigen (CEA), (**Figure S1**) [15]. An *in silico* analysis of CT gene expression as defined by 31 probesets in the GSE1614 dataset, which contains gene expression data for the Caco-2 SD model obtained during differentiation (proliferating (2^{nd} day), post-proliferation-undifferentiated (8^{th} day), and post-proliferation-differentiated (15^{th} day)) [16], revealed modest up-regulation of almost all CT genes during differentiation (**Figure S2**). We chose to validate the change in expression of six CT gene/gene families by quantitative RT-PCR in differentiating Caco-2 cells *in vitro*. *GAGE, MAGE-A3, NY-ESO-1* and *SSX4* transcripts were undetectable on the first day of confluence, as well as at later time points (data not shown). However, significant up-regulation of *PAGE2* (*2 and 2B*), and *SPANX-B* genes was evident (**Figure 1**).

PAGE2 and *SPANX-B* expression follow MET in the Caco-2 SD model

Spontaneous differentiation of Caco-2 *in vitro* has been reported to result in MET [21,22]. To determine if this occurred in parallel to the up-regulation of *PAGE2* and *SPANX-B*, we analyzed the GSE1614 dataset for the expression of genes representing EMT in colorectal cancer [17], and selected 6 genes to be validated in our model. Analysis of mRNA and protein expression of these revealed a decrease in mesenchymal genes (*vimentin, fibronectin 1* and *transgelin*) with a concomitant increase in expression of epithelial genes (*CDX2, claudin-4* and *E-cadherin*) as the cells differentiated, demonstrating that the increase in CT gene expression occurs simultaneously with MET in this model (**Figure 1 & Figure S3**).

PAGE2, SPANX-B and EMT gene expression *in situ*

To study if the changes in protein expression of CT and EMT genes occurred simultaneously in the same cells, we performed double immunofluorescence staining during differentiation. A gradual loss of mesenchymal markers was observed as cells differentiated, with a concomitant increase in epithelial genes and CT genes. SPANX-B and PAGE-2 were frequently co-expressed with the epithelial marker CDX2 in the same cells but almost never with VIM or FN1 (**Figures 2, 3** and **Figures S4, S5, S6, S7**).

PAGE2 and SPANX-B expression correlates with increased hmC and ten-eleven translocation methylcytosine dioxygenase (TET) up-regulation

Expression of all CT genes studied thus far including *PAGE2* and *SPANX-B* have been associated with the demethylation of CpG residues within regions proximal to the transcription start site [1,2,23–26]. In this line, both *PAGE2* and *SPANX-B* can be up-regulated by 5-aza 2′-deoxycytidine treatment (**Figure S8**).

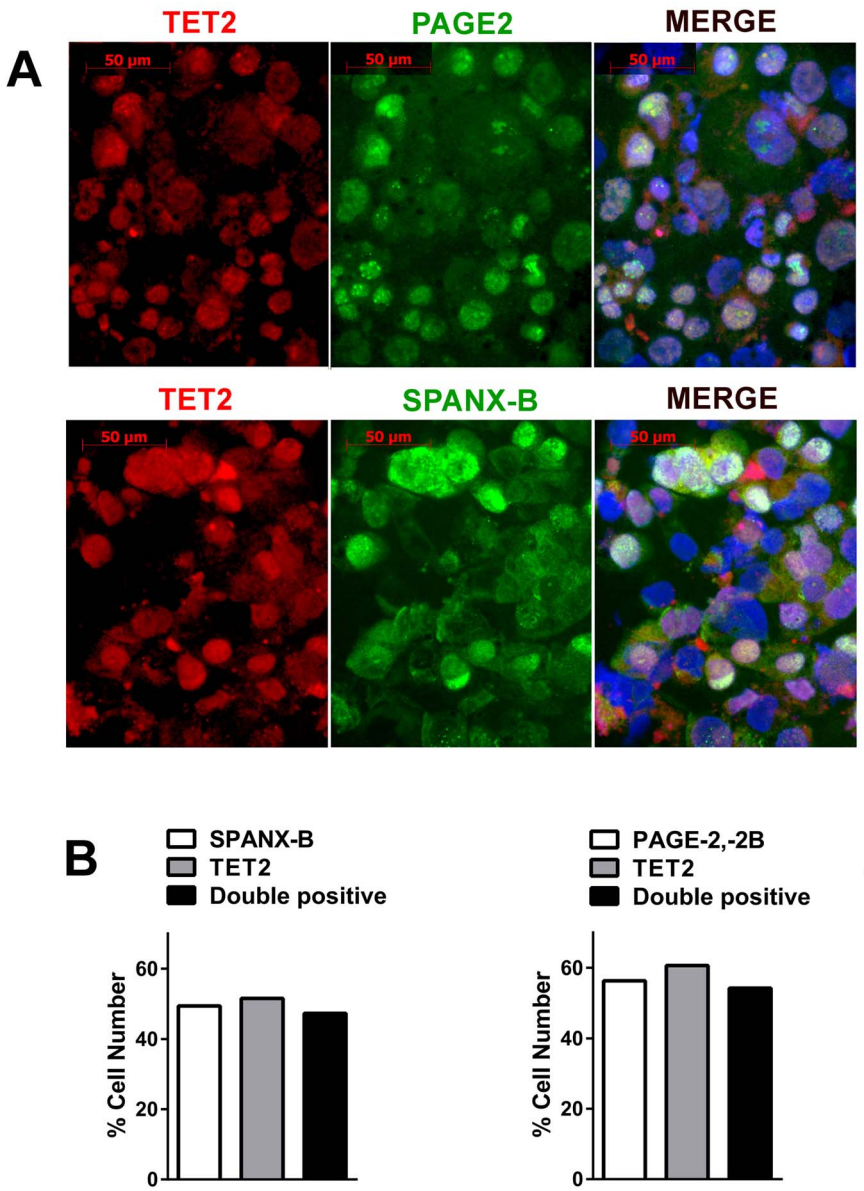


Figure 7. Overlapping TET2 and CT gene expression in differentiating Caco-2 cells. Double immunofluorescence staining for TET2 (Alexafluor 568: red) and SPANX-B or PAGE2 (Alexafluor 488: green) with DAPI counterstaining 20 days post-confluence show overlapping nuclear expression Magnification: 40× (**A**). More than 95% of cells expressing PAGE2 or SPANX-B were also positive for TET2 staining (**B**).
doi:10.1371/journal.pone.0107905.g007

However, bisulfite sequencing of promoter-proximal regions of both *PAGE2* and *SPANX-B* revealed no differences at different stages of Caco-2 SD (**Figure 4**). As bisulfite sequencing is unable to distinguish methyl cytosine (mC) from hmC, we asked whether the change in CT gene expression could be related to altered hmC/mC ratios within their promoters. In fact, chromatin immunoprecipitation (ChIP) with a hmC specific antibody revealed an increase in hmC during differentiation in both *PAGE2* and *SPANX-B2* promoters (**Figure 5**). We next asked if the increase in hmC was related to an increase in TET1, -2, and -3 expression as these proteins are responsible for converting mC to hmC [27,28]. Indeed, the increase in hmC of *PAGE2* and *SPANX-B2* promoters were correlated with an up-regulation of *TET2* mRNA expression, together with modest increases in *TET1* and *3* (**Figure 6A**). Double immunofluorescence staining revealed that the majority of cells expressing PAGE2 or SPANX-B were positive for TET2 staining; indicating these two events occurred in the same cells (**Figure 7**). It is therefore, likely that the increase in TET2 expression causes increased hmC in these genes. Interestingly, only a low molecular weight translation product (~25 kD) of TET2 was increased in the differentiating cells, when no clear difference in levels of the full-length TET2 protein was observed (**Figure 6B & C**). The peptide used for generating the commercial TET2 antibodies could specifically inhibit recognition of the 25 kD product confirming its identity with TET2 (data not shown). TET2 has recently been shown to undergo proteolytic cleavage by calpain 1 and 2, generating a 25 kD product *in vitro* [29]. To test if a cation-dependent protease is responsible for the generation of the 25 kD protein, we treated differentiating cells with an intracellular Ca^{2+} chelator (BAPTA-AM). This resulted in

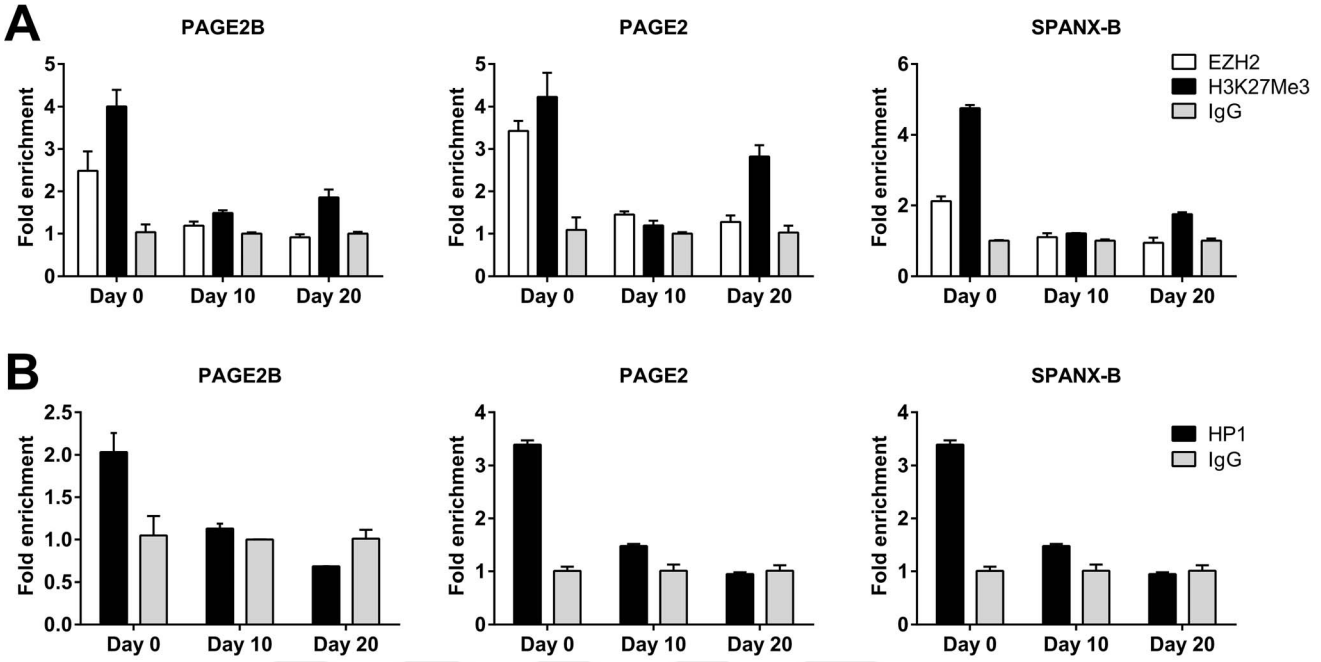


Figure 8. Chromatin modifications within *PAGE2* and *SPANX-B* during Caco-2 differentiation. CHIP analysis of *PAGE2, -2B* and *SPAN-X* transcription-start site-proximal regions reveals decreased EZH2 occupancy and H3K27m3 (**A**), as well as decreased HP-1 binding during differentiation (**B**). P values (one-way ANOVA) calculated for *PAGE2B, PAGE2*, and *SPANX-B*, are <0.001, 0.02, and 0.001 for EZH2; 0.003, <0.001, and <0.0001 for H3K27m3; and 0.001, <0.001, and <0.001, for HP1, respectively.
doi:10.1371/journal.pone.0107905.g008

a modest reduction in the 25 kD TET2 protein with the concomitant generation of a larger molecular weight product (~50 kD), suggesting that the 25 kD TET2 protein is a Ca^{+2} dependent protease cleavage product with a 50 kD intermediate (**Figure 6D**).

EZH2 and HP-1 occupancy of *PAGE2* and *SPANX-B* promoter proximal regions decrease during differentiation

Hydroxymethylation has been reported to prevail in promoters with dual H3K4 and H3K27 trimethylation that also bind PRC2 proteins [30]. The PRC2 complex protein EZH2 has been implied in the repression of GAGE, another CT gene [31]. We therefore, asked whether increased hmC within CT gene promoters resulted in altered EZH2 binding to the same sites. Indeed, ChIP experiments demonstrated a decrease in EZH2 occupancy, as well as a decrease in H3K27m3 in both *PAGE2* and *SPANX-B* promoters during Caco-2 SD (**Figure 8**). The PRC2 component SUZ12 has been reported to regulate H3K9 methylation and in turn, heterochromatin protein 1 (HP1α) binding. In fact, we observed a simultaneous decrease in HP1 binding to both *PAGE2* and *SPANX-B* promoters during differentiation, that correlated with *PAGE2* and *SPANX-B* upregulation (**Figure 8**). Thus, our data suggest that both PRC2 and HP-1 contribute to maintaining *PAGE2* and *SPANX-B* in a transcriptionally silent state when the

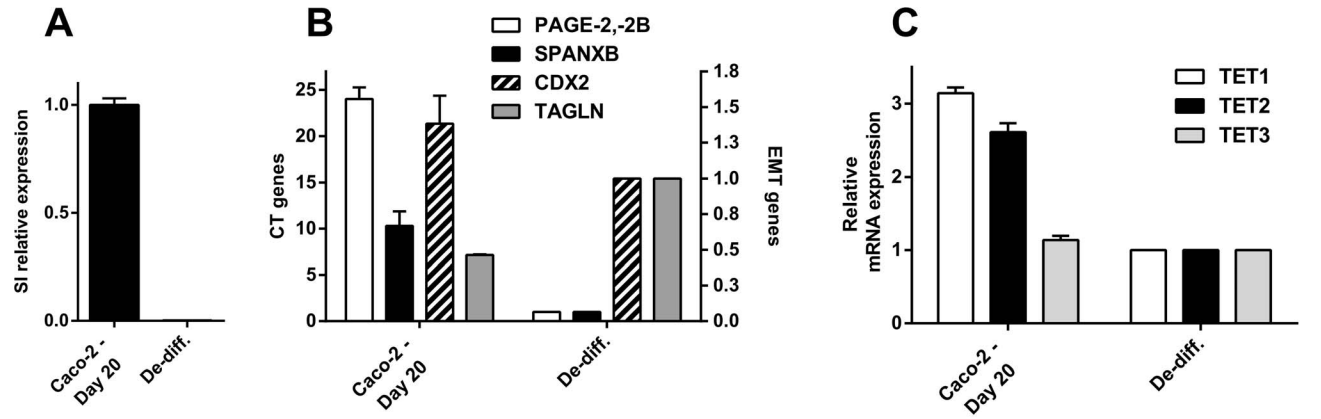


Figure 9. De-differentiation induced EMT and down-regulation of *CT* and *TET* genes. De-differentiation induced by growth under non-confluent conditions indicated by decreased *sucrose isomaltase* (SI) mRNA levels (**A**), leads to down-regulation of *CDX2*, *PAGE2, -2B* and *SPANX-B*, with concomitant up-regulation of *TGLN* (**B**). *TET1* and *-2* mRNAs are also down-regulated during de-differentiation (**C**).
doi:10.1371/journal.pone.0107905.g009

cells have a mesenchymal phenotype, and that increased TET2 expression and hmC mediated transcriptional activation are related to PRC2 and HP-1 dissociation from the promoters of these CT genes during differentiation.

PAGE2 and *SPANX-B* up-regulation is reversed during EMT

We hypothesized that if the epigenetic alterations underlying CT gene expression happened in parallel to MET, that this process could be reversed if cells entered EMT. To test this hypothesis, differentiated Caco-2 cells were detached and allowed to proliferate for 5 days. This resulted in their rapid de-differentiation as evidenced by down-regulation of SI. De-differentiated cells down-regulated *PAGE2* and *SPANX-B*, as well as *CDX*, as they up-regulated *TAGLN*, in line with ongoing EMT (**Figure 9**). Although transcription of all three *TET* genes decreased during de-differentiation (**Figure 9**), we did not observe a decrease in hmC during this period (data not shown). We, therefore, conclude that the up-regulation of CT gene expression is reversible in this model.

Discussion

Previous studies revealed that CT gene expression correlated with an epithelial rather than a mesenchymal phenotype, and showed the up-regulation of CT genes during MET [12,14]. To our knowledge, this is the first report describing alterations in several epigenetic mechanisms within promoters of two CT genes during MET-like differentiation concordant with a dynamic change in gene expression. As bisulfite sequencing of *PAGE2* and *SPANX-B* promoters revealed no change upon differentiation, the increased hmC must strictly involve methylated CpG residues. This is in line with the fact that TET enzymes are responsible for the conversion of 5-methyl cytosine to 5-hydroxymethyl cytosine [27,28]. Conversion of hmC to mC is a far more complex process and might not happen with similar kinetics [32,33]. This is likely the reason why we did not observe a change in hmC during the 5 day de-differentiation process of Caco-2 cells despite the decrease observed in global TET levels. Our finding that *PAGE2*, *SPANX-B* and *TET2* induction is reversible is similar to another study in embryonic stem cells where Vitamin C was shown to induce TET2 expression, which in turn, resulted in up-regulation of CT genes. Both events were reversible upon Vitamin C withdrawal [34].

Although hydroxymethylation within gene promoters has been reported to decrease during differentiation of normal cells, a recent study revealed that about 20% of all modified cytosines in most CT genes in human brain, where they are not expressed, consist of hmC [35]. Up-regulation of TET2 expression in cancer has been associated with MET [36,37]; and therefore, a more differentiated state [30]. Similar to the inverse correlation between EZH2 and CT/TET2 expression we report here, others have shown EZH2 and TET enzymes to repress and induce differentiation of neuronal precursors, respectively [38]. CT genes are up-regulated during the initial stages of development in the human embryo, but decrease as tissues differentiate further [39]. As adult colon tissue does not show PAGE2 or SPANX-B expression (data not shown), had Caco-2 cells the capability of differentiating further, both genes might have been down-regulated completely. On the other hand, the fact that we could not demonstrate up-regulation of *GAGE, MAGE-A3, NY-ESO-1* or *SSX4* expression in this model might be because these genes are expressed at earlier stages of differentiation. We believe this because SPANX-B expression is primarily in post-meiotic cells of the testis (i.e. spermatocytes, spermatids, or sperm), whereas GAGE, MAGE-A3, NY-ESO-1 or SSX expression is primarily in spermatogonia [11].

Our data and that of several others' indicate that cancer cells that express CT genes have more of an epithelial rather than a mesenchymal phenotype. We suggest that CT genes *PAGE2* and *SPANX-B* are induced during a window of differentiation that correlates with up-regulation of epithelial markers of differentiation. The Caco-2 SD model has made it possible to observe the actively changing epigenetic landscape within the promoters of these CT genes. However, as CT gene expression in tumors has closely been related to the methylation state of their promoter, the process that leads to CT gene induction *in vivo* might ultimately result in "fixing" of the epigenetic state which would in turn result in CpG methylation. Yet, via dynamic MET in tumors [40], it is conceivable that even this might change over the course of the disease.

From a clinical perspective, data from our lab as well as from others reveal that sub-grouping of tumors based on gene expression profiles can clearly identify cells with different chemosensitivity profiles [12,41,42]. In this line, we predict future studies will reveal distinct drug sensitivity profiles for colorectal cancer subtypes as possibly defined by *PAGE2* and *SPANX-B* expression, for which the Caco-2 SD model could be used.

Supporting Information

Figure S1 Post-confluence differentiation of Caco-2 *in vitro*. Up-regulation of *sucrase-isomaltase* (**A**), and carcinoembryonic antigen (CEA) (**B**) in cells collected at indicated days post confluence (DPC) as determined by quantitative RT-PCR, and Western analysis, respectively. Alkaline phosphatase expression is also upregulated as determined by immunohistochemistry revealing differentiation (**C**). Other measures of differentiation for the cells used in this study have been reported previously (ref. 17). *P<0.001 (ANOVA with Tukey's post hoc test).
(DOCX)

Figure S2 Up-regulation of CT gene expression during Caco-2 spontaneous differentiation *in vitro*. Heat map based on 31 probesets in GSE1614 corresponding to 23 CT genes from 7 families. As compared to proliferating cells, gene expression incrementally increases in at confluence (8^{th} day) and further during post-confluence differentiation (15^{th} day).
(DOCX)

Figure S3 Western analysis of differentially expressed genes during Caco-2 SD. A gradual increase in SPANX-B and CDX2 in parallel to a decrease in expression of FN, VIM and TGLN up to day 30 post-confluence. Results from 3 independent differentiation experiments are shown.
(DOCX)

Figure S4 SPANX-B and Fibronectin expression show limited overlap in differentiating Caco-2 cells. Immunofluorescent staining of differentiating Caco-2 cells with DAPI counterstaining reveals a gradual increase in nuclear SPANX-B (Alexa Fluor 488: green) with a concomitant decrease in cytoplasmic fibronectin expression (Alexa Fluor 568: red); (20× magnification) (**A**). Less than 10% of cells expressing SPANX-B stained for fibronectin at day 0 (**B**).
(DOCX)

Figure S5 PAGE2, -2B and Vimentin expression are mutually exclusive in differentiating Caco-2 cells. Immunofluorescent staining of differentiating Caco-2 cells with DAPI counterstaining reveals a gradual increase in nuclear PAGE2, -2B (Alexa Fluor 488: green) with a concomitant decrease in

cytoplasmic vimentin expression (Alexa Fluor 568: red); (20× magnification) (**A**). Less than 10% of cells showed double fluorescence when staining was analyzed quantitatively at day 0. At later time points, none of the cells showed double staining (**B**).
(DOCX)

Figure S6 PAGE2, -2B and fibronectin expression are mutually exclusive in differentiating Caco-2 cells. Immunofluorescent staining of differentiating Caco-2 cells with DAPI counterstaining reveals a gradual increase in nuclear PAGE2, -2B (Alexa Fluor 488: green) with a concomitant decrease in cytoplasmic fibronectin expression (Alexa Fluor 568: red); (20× magnification) (**A**). Less than 15% of cells showed double fluorescence when staining was analyzed quantitatively at day 0 (**B**).
(DOCX)

Figure S7 Nuclear co-localization of CDX2 and PAGE2, -2B in differentiating Caco-2 cells. Immunofluorescent staining of differentiating Caco-2 cells with DAPI counterstaining reveals overlapping PAGE2. -2B (Alexa Fluor 488: green) and CDX2 (Alexa Fluor 568: red) expression; (20× magnification) (**A**). More than 80% of the cells show double-labeling when analyzed quantitatively (**B**).
(DOCX)

Figure S8 Induction of PAGE2,-2B and SPANX-B gene expression by 5-aza-2′-deoxycytidine in HCT116 (top) and SK-LC-17 cell lines (bottom). Relative mRNA expression values at indicated time points compared to day 0, as determined by quantitative RT-PCR are shown.
(DOCX)

Table S1 PCR primers.
(DOCX)

Table S2 Primer locations.
(DOCX)

Table S3 Antibodies used for IF staining and western blot analysis.
(DOCX)

Acknowledgments

We thank Drs. Stefan Fuss and Isik Yulug for critical reviews and helpful suggestions.

Author Contributions

Conceived and designed the experiments: SYO AS BK YK KMS SB AOG. Performed the experiments: SYO AS BK YK KMS. Analyzed the data: SYO AS BK YK KMS SB AOG. Contributed reagents/materials/analysis tools: SB AOG. Contributed to the writing of the manuscript: SYO AS BK YK KMS SB AOG.

References

1. Scanlan MJ, Gure AO, Jungbluth AA, Old LJ, Chen YT (2002) Cancer/testis antigens: an expanding family of targets for cancer immunotherapy. Immunological Reviews 188: 22–32.
2. Caballero OL, Chen YT (2009) Cancer/testis (CT) antigens: potential targets for immunotherapy. Cancer Sci 100: 2014–2021.
3. Hofmann O, Caballero OL, Stevenson BJ, Chen YT, Cohen T, et al. (2008) Genome-wide analysis of cancer/testis gene expression. Proceedings of the National Academy of Sciences of the United States of America 105: 20422–20427.
4. Loriot A, Reister S, Parvizi GK, Lysy PA, De Smet C (2009) DNA methylation-associated repression of cancer-germline genes in human embryonic and adult stem cells. Stem Cells 27: 822–824.
5. Mashino K, Sadanaga N, Tanaka F, Yamaguchi H, Nagashima H, et al. (2001) Expression of multiple cancer-testis antigen genes in gastrointestinal and breast carcinomas. Br J Cancer 85: 713–720.
6. Gure AO, Chua R, Williamson B, Gonen M, Ferreira CA, et al. (2005) Cancer-testis genes are coordinately expressed and are markers of poor outcome in non-small cell lung cancer. Clinical Cancer Research 11: 8055–8062.
7. Woloszynska-Read A, Mhawech-Fauceglia P, Yu JH, Odunsi K, Karpf AR (2008) Intertumor and intratumor NY-ESO-1 expression heterogeneity is associated with promoter-specific and global DNA methylation status in ovarian cancer. Clinical Cancer Research 14: 3283–3290.
8. Jungbluth AA, Busam KJ, Kolb D, Iversen K, Coplan K, et al. (2000) Expression of MAGE-antigens in normal tissues and cancer. Int J Cancer 85: 460–465.
9. Jungbluth AA, Chen YT, Busam KJ, Coplan K, Kolb D, et al. (2002) CT7 (MAGE-C1) antigen expression in normal and neoplastic tissues. Int J Cancer 99: 839–845.
10. Jungbluth AA, Stockert E, Chen YT, Kolb D, Iversen K, et al. (2000) Monoclonal antibody MA454 reveals a heterogeneous expression pattern of MAGE-1 antigen in formalin-fixed paraffin embedded lung tumours. Br J Cancer 83: 493–497.
11. Chen YT, Chiu RT, Lee P, Beneck D, Jin BQ, et al. (2011) Chromosome X-encoded cancer/testis antigens show distinctive expression patterns in developing gonads and in testicular seminoma. Human Reproduction 26: 3232–3243.
12. Gupta PB, Onder TT, Jiang G, Tao K, Kuperwasser C, et al. (2009) Identification of selective inhibitors of cancer stem cells by high-throughput screening. Cell 138: 645–659.
13. Hoek KS (2007) DNA microarray analyses of melanoma gene expression: a decade in the mines. Pigment Cell Res 20: 466–484.
14. Argast GM, Krueger JS, Thomson S, Sujka-Kwok I, Carey K, et al. (2011) Inducible expression of TGF beta, Snail and Zeb1 recapitulates EMT in vitro and in vivo in a NSCLC model. Clinical & Experimental Metastasis 28: 593–614.
15. Astarci E, Sade A, Cimen I, Savas B, Banerjee S (2012) The NF-kappaB target genes ICAM-1 and VCAM-1 are differentially regulated during spontaneous differentiation of Caco-2 cells. FEBS J 279: 2966–2986.
16. Fleet JC, Wang L, Vitek O, Craig BA, Edenberg HJ (2003) Gene expression profiling of Caco-2 BBe cells suggests a role for specific signaling pathways during intestinal differentiation. Physiol Genomics 13: 57–68.
17. Loboda A, Nebozhyn MV, Watters JW, Buser CA, Shaw PM, et al. (2011) EMT is the dominant program in human colon cancer. BMC Medical Genomics 4.
18. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402–408.
19. Pinto M, Robineleon S, Appay MD, Kedinger M, Triadou N, et al. (1983) Enterocyte-Like Differentiation and Polarization of the Human-Colon Carcinoma Cell-Line Caco-2 in Culture. Biology of the Cell 47: 323–330.
20. Simon-Assmann P, Turck N, Sidhoum-Jenny M, Gradwohl G, Kedinger M (2007) In vitro models of intestinal epithelial cell differentiation. Cell Biology and Toxicology 23: 241–256.
21. Halbleib JM, Saaf AM, Brown PO, Nelson WJ (2007) Transcriptional modulation of genes encoding structural characteristics of differentiating Enterocytes during development of a polarized epithelium in vitro. Molecular Biology of the Cell 18: 4261–4278.
22. Saaf AM, Halbleib JM, Chen X, Yuen ST, Leung SY, et al. (2007) Parallels between global transcriptional programs of polarizing Caco-2 intestinal epithelial cells in vitro and gene expression programs in normal colon and colon cancer. Mol Biol Cell 18: 4245–4260.
23. Fratta E, Coral S, Covre A, Parisi G, Colizzi F, et al. (2011) The biology of cancer testis antigens: putative function, regulation and therapeutic potential. Mol Oncol 5: 164–182.
24. Cannuyer J, Loriot A, Parvizi GK, De Smet C (2013) Epigenetic Hierarchy within the MAGEA1 Cancer-Germline Gene: Promoter DNA Methylation Dictates Local Histone Modifications. Plos One 8.
25. Wang ZQ, Zhang J, Zhang YN, Lim SH (2006) SPAN-Xb expression in myeloma cells is dependent on promoter hypomethylation and can be upregulated pharmacologically. International Journal of Cancer 118: 1436–1444.
26. Wang ZQ, Zhang J, Zhang Y, Srivenugopal K, Lim SH (2006) SPAN-XB core promoter sequence is regulated in myeloma cells by specific CpG dinucleotides associated with MeCP2 protein. Blood 108: 628a–628a.
27. Iyer LM, Tahiliani M, Rao A, Aravind L (2009) Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. Cell Cycle 8: 1698–1710.
28. Loenarz C, Schofield CJ (2009) Oxygenase catalyzed 5-methylcytosine hydroxylation. Chem Biol 16: 580–583.
29. Wang Y, Zhang Y (2014) Regulation of TET Protein Stability by Calpains. Cell Rep 6: 278–284.

30. Jiang Y, Liu S, Chen X, Cao Y, Tao Y (2013) Genome-wide distribution of DNA methylation and DNA demethylation and related chromatin regulators in cancer. Biochim Biophys Acta 1835: 155–163.
31. Sun F, Chan E, Wu Z, Yang X, Marquez VE, et al. (2009) Combinatorial pharmacologic approaches target EZH2-mediated gene repression in breast cancer cells. Mol Cancer Ther 8: 3191–3202.
32. Tsukada Y (2012) Hydroxylation mediates chromatin demethylation. J Biochem 151: 229–246.
33. Solary E, Bernard OA, Tefferi A, Fuks F, Vainchenker W (2014) The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases. Leukemia 28: 485–496.
34. Blaschke K, Ebata KT, Karimi MM, Zepeda-Martinez JA, Goyal P, et al. (2013) Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature 500: 222–226.
35. Wen L, Li X, Yan L, Tan Y, Li R, et al. (2014) Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. Genome Biol 15: R49.
36. Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, et al. (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature 473: 398–402.
37. Song SJ, Poliseno L, Song MS, Ala U, Webster K, et al. (2013) MicroRNA-antagonism regulates breast cancer stemness and metastasis via TET-family-dependent chromatin remodeling. Cell 154: 311–324.
38. Hahn MA, Qiu R, Wu X, Li AX, Zhang H, et al. (2013) Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. Cell Rep 3: 291–300.
39. Gjerstorff MF, Harkness L, Kassem M, Frandsen U, Nielsen O, et al. (2008) Distinct GAGE and MAGE-A expression during early human development indicate specific roles in lineage differentiation. Hum Reprod 23: 2194–2201.
40. Yao D, Dai C, Peng S (2011) Mechanism of the mesenchymal-epithelial transition and its relationship with metastatic tumor formation. Mol Cancer Res 9: 1608–1620.
41. Isbilen M, Senses KM, Gure AO (2013) Predicting Chemotherapy Sensitivity Profiles for Breast Cancer Cell Lines with and Without Stem Cell-Like Features. Curr Signal Transduct Ther 8: 268–273.
42. Isbilen M, Gure AO (2014) Identifying effective molecularly targeted drugs for hematological cancers. Turkiye Klinikleri J Hematol-SpecialTopics 7: 1–7.

Hindawi Publishing Corporation
Journal of Immunology Research
Volume 2016, Article ID 8121985, 12 pages
http://dx.doi.org/10.1155/2016/8121985

Clinical Study

Adjuvant Autologous Melanoma Vaccine for Macroscopic Stage III Disease: Survival, Biomarkers, and Improved Response to CTLA-4 Blockade

Michal Lotem,[1] Sharon Merims,[1] Stephen Frank,[1] Tamar Hamburger,[1] Aviram Nissan,[2] Luna Kadouri,[1] Jonathan Cohen,[1] Ravid Straussman,[3] Galit Eisenberg,[1] Shoshana Frankenburg,[1] Einat Carmon,[2] Bilal Alaiyan,[2] Shlomo Shneibaum,[4] Zeynep Ozge Ayyildiz,[5] Murat Isbilen,[5] Kerem Mert Senses,[5] Ilan Ron,[6] Hanna Steinberg,[1] Yoav Smith,[7] Eitan Shiloni,[8] Ali Osmay Gure,[5] and Tamar Peretz[1]

[1]*Sharett Institute of Oncology, Hadassah Hebrew University Hospital, Ein Karem Campus, 91120 Jerusalem, Israel*
[2]*Departments of Surgery, Hadassah Hebrew University Hospital, Mount Scopus Campus, 91240 Jerusalem, Israel*
[3]*Department of Molecular Cell Biology, Weizmann Institute of Science, 7610001 Rehovot, Israel*
[4]*Department of Surgery, Tel Aviv Sourasky Medical Center, 64239 Tel Aviv, Israel*
[5]*Department of Molecular Biology and Genetics, Bilkent University, 06800 Ankara, Turkey*
[6]*Department of Oncology, Tel Aviv Sourasky Medical Center, 64239 Tel Aviv, Israel*
[7]*Genomic Data Analysis Unit, Hebrew University Medical School, 91120 Jerusalem, Israel*
[8]*Department of Surgery, Bnai Zion Medical Center, 31048 Haifa, Israel*

Correspondence should be addressed to Michal Lotem; mlotem@hadassah.org.il

Received 11 January 2016; Revised 30 March 2016; Accepted 17 April 2016

Academic Editor: Yi Zhang

Background. There is not yet an agreed adjuvant treatment for melanoma patients with American Joint Committee on Cancer stages III B and C. We report administration of an autologous melanoma vaccine to prevent disease recurrence. *Patients and Methods*. 126 patients received eight doses of irradiated autologous melanoma cells conjugated to dinitrophenyl and mixed with BCG. Delayed type hypersensitivity (DTH) response to unmodified melanoma cells was determined on the vaccine days 5 and 8. Gene expression analysis was performed on 35 tumors from patients with good or poor survival. *Results*. Median overall survival was 88 months with a 5-year survival of 54%. Patients attaining a strong DTH response had a significantly better ($p = 0.0001$) 5-year overall survival of 75% compared with 44% in patients without a strong response. Gene expression array linked a 50-gene signature to prognosis, including a cluster of four cancer testis antigens: CTAG2 (NY-ESO-2), MAGEA1, SSX1, and SSX4. Thirty-five patients, who received an autologous vaccine, followed by ipilimumab for progressive disease, had a significantly improved 3-year survival of 46% compared with 19% in nonvaccinated patients treated with ipilimumab alone ($p = 0.007$). *Conclusion*. Improved survival in patients attaining a strong DTH and increased response rate with subsequent ipilimumab suggests that the autologous vaccine confers protective immunity.

In loving memory of Olga Drize, Ph.D.

1. Introduction

The treatment of metastatic melanoma has been revolutionized in the last three years, with the FDA registration of Yervoy™, a monoclonal antibody blocking lymphocyte regulatory receptor cytotoxic T lymphocyte-associated antigen 4 (CTLA-4), and shortly afterward, the entry to the clinic of Zelboraf, a small molecule inhibitor of mutated B-RAF.

From being an incurable disease, stage IV melanoma has become an illness in which prolonged and even complete responses can be envisioned. However, while the prospects have improved for stage IV disease, for patients with American Joint Committee on Cancer (AJCC) stage III disease no new treatment options have been developed and validated since the approval of interferon α (IFNα) almost two decades ago [1–7]. A pegylated formulation did not offer improved tolerability and treatment was discontinued due to toxicity. Both the EORTC trials and the ECOG pooled analysis [8] showed that AJCC stage III patients with macroscopic lymph node involvement derived the smallest survival benefit, if any, from IFNα.

In this situation, melanoma vaccines could have been an alternative to IFNα, since they could induce a tumor-specific immune response to inhibit micrometastases at a stage when the suppressive effects of an advanced tumor are not yet an obstacle [9]. In the few controlled clinical trials reported to date, vaccinated patients did not experience a survival benefit [10]. This fact, together with the paucity of objective responses to active immunization in stage IV melanoma patients using different vaccination strategies, rendered a general impression of the futility of cancer vaccines [11, 12].

The resurgence of interest in cancer vaccination was the result of several clinical trials demonstrating that a component of active immunization could improve clinical outcome of immunotherapy protocols. Examples include the addition of a peptide vaccine to the administration of high dose interleukin-2 (IL-2) [13] and the use of a GMCSF-secreting tumor vaccine in combination with CTLA-4 blockade for metastatic prostate cancer [14]. These trials led to the increasing understanding that cancer immunotherapy is a multifaceted strategy and that a single treatment modality would not suffice. Noting that individuals exhibit heterogeneity of tumor antigens [15], the use of autologous tumor as a basis for vaccination can provide antigen authenticity. The unique expression profile of normal and mutated proteins in the patient's tumor cells is presented in conjunction with their own major histocompatibility complex (MHC) molecules, and this combination is necessary to induce antigen-specific reactive lymphocytes [16]. The coadministration of Bacillus Calmette-Guérin (BCG), a widely used immunological adjuvant, with the autologous vaccine, has previously been shown to enhance response to autologous vaccination protocols [17, 18].

In this paper we report our experience with an autologous melanoma vaccine as an adjuvant therapy for melanoma patients in the advanced categories of AJCC stage III disease: macroscopic lymph node involvement and resectable in-transit metastases. We demonstrate that clinical immune response (delayed type hypersensitivity, DTH) is linked to improved survival. Furthermore, using this vaccination protocol, molecular analysis of the melanoma showed that cancer testis antigens (CTAs), which are generally considered targets of immune response, are predictive of survival. Interestingly, patients who had previously received the melanoma cell vaccine had improved overall survival when treated with ipilimumab immunotherapy for metastatic disease.

2. Materials and Methods

2.1. Patients. This prospective phase II, single-institution, single-arm study included patients with operable AJCC stages IIIB and C melanoma. Clinical staging was based on palpable lymph nodes and/or satellites prior to surgery. Since we did not have pathological data on ulceration of the primary melanoma for all patients, AJCC stage IIIB was defined for any T (tumor) with pathological N1b, N2b, and N2c. Stage IIIC was defined as any T with pathological N3 [19].

2.2. Outcomes. The aim of the study was to document overall and disease-free survival (primary outcome) and to correlate skin reactivity to the autologous tumor of patients that were treated with the autologous melanoma cell vaccine with survival (secondary outcome).

2.3. Inclusion and Exclusion Criteria. To participate, patients had to undergo complete removal of their metastatic disease and have a normal CT scan within 30 days prior to vaccination. Additional inclusion criteria included an age of 18 years or older, normal liver and renal function tests, baseline lactate dehydrogenase (LDH) value below the laboratory upper limit of normal, and ability to provide informed consent. Exclusion criteria included primary ocular melanoma. Fertile patients were requested to use adequate anticonceptive measures throughout the study period.

Screening procedures included baseline blood analyses (hematological, chemical) and computed tomography (CT) and/or magnetic resonance imaging (MRI) of the entire body, including the brain, every 4 months in the first 2 years and then every 6 months for 10 years. The study was not sponsored and was conducted following approval by the institutional ethics committee; written informed consent was obtained from all patients. Seventy-five patients from this group were included in an immune monitoring study reported earlier [16].

2.4. Melanoma Cell Lines. For the preparation of the autologous vaccine, melanoma cell lines were established from resected metastases. All patients gave their informed consent to receive the vaccine. The melanoma cell lines were established and cultured as described [18]. Briefly, cells were extracted mechanically from fresh and sterile tumor specimens, frozen, and stored in liquid nitrogen in a medium containing 2.5% human albumin and 20% DMSO until needed. To assure melanocytic progeny, the expressions of S100, MART-1, and gp100 were determined by immunostaining using polyclonal rabbit anti-S100, monoclonal A-103, and HMB-45 Abs, respectively (Dako). Positive staining of more than 50% of cells with at least one of these antibodies was required. MHC class I-related chain A (MICA) expression was determined by flow cytometry of melanoma cell lines using anti-human MICA-APC (Allophycocyanin), R&D systems, Minneapolis, MN, USA.

Cell lines were routinely tested for mycoplasma contamination by EZ-PCR (Biological Laboratories, Beth Haemek, Israel), according to the manufacturer's instructions. Tumor

cultures that were found contaminated were incubated in the presence of 10 mg/mL Ciproxin 200 (Bayer) for two weeks, with change of medium every three days. The cells were retested after treatment and were used only after being found mycoplasma-free.

2.5. Vaccine Preparation and Vaccination Procedure. Melanoma cell lines were expanded to the required number necessary for preparation of at least 8 vaccine doses of $10–25 \times 10^6$ cells each and cryopreserved at −70°C. On the day of treatment, one dose of cells was thawed, washed, and irradiated to a dose of 230 Gy. At this stage, cells were still viable. Conjugation of melanoma cells with DNP was performed as described [20], leading to death of the cells (as determined by trypan blue exclusion). BCG (Statens Serum Institut, Denmark) was added to the vaccine prior to injection, diluted to 1 : 50 for the first three vaccine doses and up to 1 : 500 for the following doses, to avoid overreactivity at the injection site. BCG was reported to trigger dendritic cell maturation and to aid in diverting the CD4 T cell response towards a Th1 phenotype [21]. The vaccination procedure was carried out as described [18]. Briefly, patients were sensitized to DNP, to enhance the response to the vaccine, by applying 0.1 mL of 2% DNP dissolved in acetone-corn oil (Sigma) topically to the inner aspect of the arm ten days prior to injection of the first vaccine. Cyclophosphamide, 300 mg/m^2 per dose, was given 4 days preceding the first and second vaccines. This practice was based on the observation that cyclophosphamide prior to vaccination can reduce the proportion of T regulatory cells and enhance tumor-specific immune response [22–25]. Generally, vaccines were injected in 3 adjacent sites on the upper arm or thigh, avoiding limbs with dissected lymphatic basins. Each patient received eight vaccine doses, at three-week intervals.

Patients were evaluated periodically every 3 months and had a total body CT scan performed every 6 months, or as required according to their symptoms.

2.6. DTH Reaction. Evaluation of DTH to autologous melanoma cells was performed on the vaccine days 5 and 8, by intradermal injection of $1–3 \times 10^6$ irradiated (170 Gy) autologous melanoma cells. The DTH response was measured 48 h after injection. Erythema diameter of 5 mm and less was arbitrarily defined as negative; 5–10 mm weak; 10–15 mm positive; and ≥15 mm strong positive DTH.

2.7. Gene Expression and Connectivity Map Analysis (C-MAP). Thirty-five melanoma cell lines that were retrieved from 35 vaccinated patients were selected for gene expression analysis based on retrospective survival data of the donors. Gene expression profiling was performed using an assay based on a collection of cellular genomic signatures that produced a pattern-matching tool and formulation-based deduction of a wider expression profile. One thousand transcripts were identified, from which the remainder of the transcriptome could be computationally inferred. These 1,000 "landmark" transcripts were measured on Luminex beads, as part of the Connectivity Map (C-MAP) project (unpublished, R. Narayan, Broad Institute of Harvard University and MIT, Cambridge, MA). Cultures underwent a median of 4 passages (range 2–13) and mRNA was extracted from melanoma cell microcultures harvested at 90% to 100% confluence, produced in a synchronous way, under identical conditions of growth, in four replicates. Initial analysis was conducted with a bioconductor- (R-) based test (SSAT), which applies a Cox-regression analysis followed by a second test based on rank statistics. The analysis determines the best cut-off value that separates patients into those with favorable versus worse survival time.

2.8. Quantitative Real Time PCR (qRT-PCR). Applied Biosystems premade and custom primer probes designed with NCBI Primer–Blast Tool were used (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). RNA was normalized to GAPDH-RNA content using ABI 7500 SDS software, v1.2.2 (Applied Biosystems Inc., Foster City, CA). Positive and negative controls, as well as samples with no DNA, were included in every qRT-PCR experiment. PCR reactions were performed using ABI qRT-PCR thermocycler (7500 Real Time PCR System, Applied Biosystems Inc., Foster City, CA). The qRT-PCR program was run for 45 cycles, following an initial incubation at 95°C, 10 min. Each cycle consisted of 95°C × 15 sec and 60°C × 1 min.

B-RAF genotype was determined using Cobas® 4800 (Roche).

2.9. Patients Undergoing Ipilimumab Treatment. During the autologous vaccine study period, ipilimumab (Yervoy, BMS) was administered in our institution to patients with metastatic melanoma in the framework of BMS studies CA184-004, 024, and 025. Since its FDA approval as a standard second-line treatment, ipilimumab was given after a one- or two-dose course of DTIC. Patients from the autologous vaccine study which later developed nonresectable metastatic disease were among a larger group recruited for these protocols. Survival data is reported for all patients getting ipilimumab during 2007–2014.

2.10. Statistical Method. The comparison of survival curves between groups was carried out using the Kaplan-Meier Survival analysis with the log rank test. All tests applied were two-tailed, and a p value of 5% or less was considered statistically significant.

3. Results

3.1. Study Patients. Melanoma metastases were obtained from 159 eligible patients. From 33 patients (20%) we could not generate the number of cells required for the treatment. A total of 126 patients (55% male; median age, 59 years) with postoperative AJCC stages IIIB and C (45% stage IIIB; 55% stage IIIC) were enrolled. For patient characteristics see Table 1. Twenty-four patients (19%) presented with enlarged lymph nodes (LNs) at the time of diagnosis of the primary melanoma; 11 (9%) had unknown primary lesion; 22 (17%) had metastases *in transit*. Nineteen patients (15%)

Table 1: Patients characteristics.

Patients	126
Age, median (range)	59 (15–86)
Gender	
Male	69 (55%)
Female	57 (45%)
Primary melanoma	
Breslow	Median 3.05 mm (0.2–20)
Cutaneous	76 at least
Acral	10
Mucosal	0
Unknown	11
Ulcerated	29/58 documented
AJCC stage	
IIIB	57 (45%)
IIIC	69 (55%)
Satellites	22 (17%)
Number of involved LNs, mean (range)	2.4 (0–17)
Adjuvant radiotherapy	56 (44%)

had noninvolved sentinel LN but developed macroscopic disease later, in the same lymphatic basin. Forty-two patients (33%) had not undergone a sentinel LN biopsy and developed macroscopic LN involvement. The mean number of involved lymph nodes was 2.4, ranging from 0 (satellites) to 17. Fifty-six patients (44%) had undergone radiotherapy in addition to the surgical procedure, which was added when more than three nodes were involved or in cases of extracapsular invasion by melanoma cells.

3.2. Vaccine Safety. No grade 3-4 adverse events (CTCAE V4) were encountered among the 126 vaccinated patients. In all patients, an erythematous nodule developed at the vaccination site and resolved in the course of 3–6 months leaving a depressed scar.

3.3. Patient Survival Correlates with Intensity of Evolving DTH Response to Unmodified Melanoma Cells. The OS and disease-free survival (DFS) of participating patients were measured from the day of the first vaccine until the current analysis was performed. A total of 126 patients were included, with a median follow-up of 44.5 months (range 8–189 months). OS survival data was available for all 126 patients and DFS data was available for 107 patients. OS and DFS were analyzed for DTH < 15 mm (weak/positive DTH) versus DTH > 15 mm (strong positive DTH). Overall, for the whole cohort, the 5-year OS was 54% and DFS was 34%. There was no difference between OS of stage IIIB and IIIC patients ($p = 0.182$). Of 119 patients with recorded DTH response, 48 patients (40%) attained strong positive DTH (DTH > 15 mm), whereas 71 (60%) had a weak DTH response (<15 mm). The patients with strong DTH response had a 5-year OS of 75% and DFS of 47%. In contrast, patients with weak DTH had a significantly lower 5-year OS of 44% ($p < 0.0001$) and DFS of 26% ($p = 0.27$). Using the Kaplan-Meier analysis and the log rank test, the single parameter that most strongly correlated with OS and DFS in a univariate analysis was the DTH response (Figure 1). In Table 3, OS and DFS are compared between patients attaining DTH responses of 10 and 15 mm and patients who did not develop such response. For patients who attained strong positive DTH (>15 mm), the 5-year overall survival hazard ratio (HR) was 0.24 (95% CI 0.1–0.53; $p < 0.001$). The HR for 5-year disease recurrence was 0.4 (95% CI 0.1–0.83, $p = 0.015$ Pearson's chi square test), but in a longer follow-up, the protection from recurrence decreased to a HR of 0.63 (95% CI 0.3–1.32; $p = 0.2$). Age, gender, and depth of invasion of the primary melanoma had no impact on survival. In a survival analysis done for DTH cut-off of 10 mm, a similar trend was noted with a smaller p value (0.003) for improved 5-year OS in patients attaining a DTH response of >10 mm (64%) versus 32% in patients with DTH <10 mm. DFS was similar in the two groups ($p = 0.36$). Thus, the acquisition of powerful skin reactivity against nonmodified autologous melanoma cells, which reflects the development of specific cell mediated immunity, correlates favorably with survival, supporting previous results by us and by others, for example, [16, 18, 26].

For 56 patients in whom more than 3 involved lymph nodes were removed, radiation therapy was added to the resected lymphatic basin. Even though the patients requiring this added treatment belonged to a grave prognostic group, radiotherapy may enhance the immune response of the patients [27]. Indeed, the rate of strong DTH response in patients receiving radiotherapy was 42%. Five-year OS of these patients was 58% compared to 33% of those who received radiotherapy and had a weak DTH response ($p = 0.024$).

3.4. Cancer Testis Antigen mRNA Expression in Melanoma Cells Correlates with Improved OS. The C-MAP project was based on a collection of cellular genomic signatures to drugs, disease states, and cancer, in order to produce a pattern-matching tool and a formulation-based deduction of a wider expression profile. One thousand transcripts were identified from which the remainder of the transcriptome could be computationally inferred. These 1,000 "landmark" transcripts were measured on Luminex beads (unpublished, R. Narayan, Broad Institute of Harvard University and MIT, Cambridge, MA).

Thirty-five tumor samples, representing distinct subclasses of poor and good responders, were selected for C-MAP analysis: (1) eighteen poor responding patients with a median OS of 19 months (range 8–34), all of whom failed to develop strong skin reactivity to their autologous tumor, and (2) seventeen good responding patients with median OS of 105 months (range 46–194), 12 of whom also developed strong skin reactivity (DTH data missing for one). Figure 2 shows the hierarchical clustering of 50 genes expressed on patients' melanoma cells, which yield a significantly improved or worsened HR for survival. Several genes of interest are listed in Table 4. Cancer testis antigens CTAG2 (NY-ESO-2), MAGEA1, SSX1, and SSX4 clustered together in the hierarchical diagram (depicted in a circle in Figure 2). High

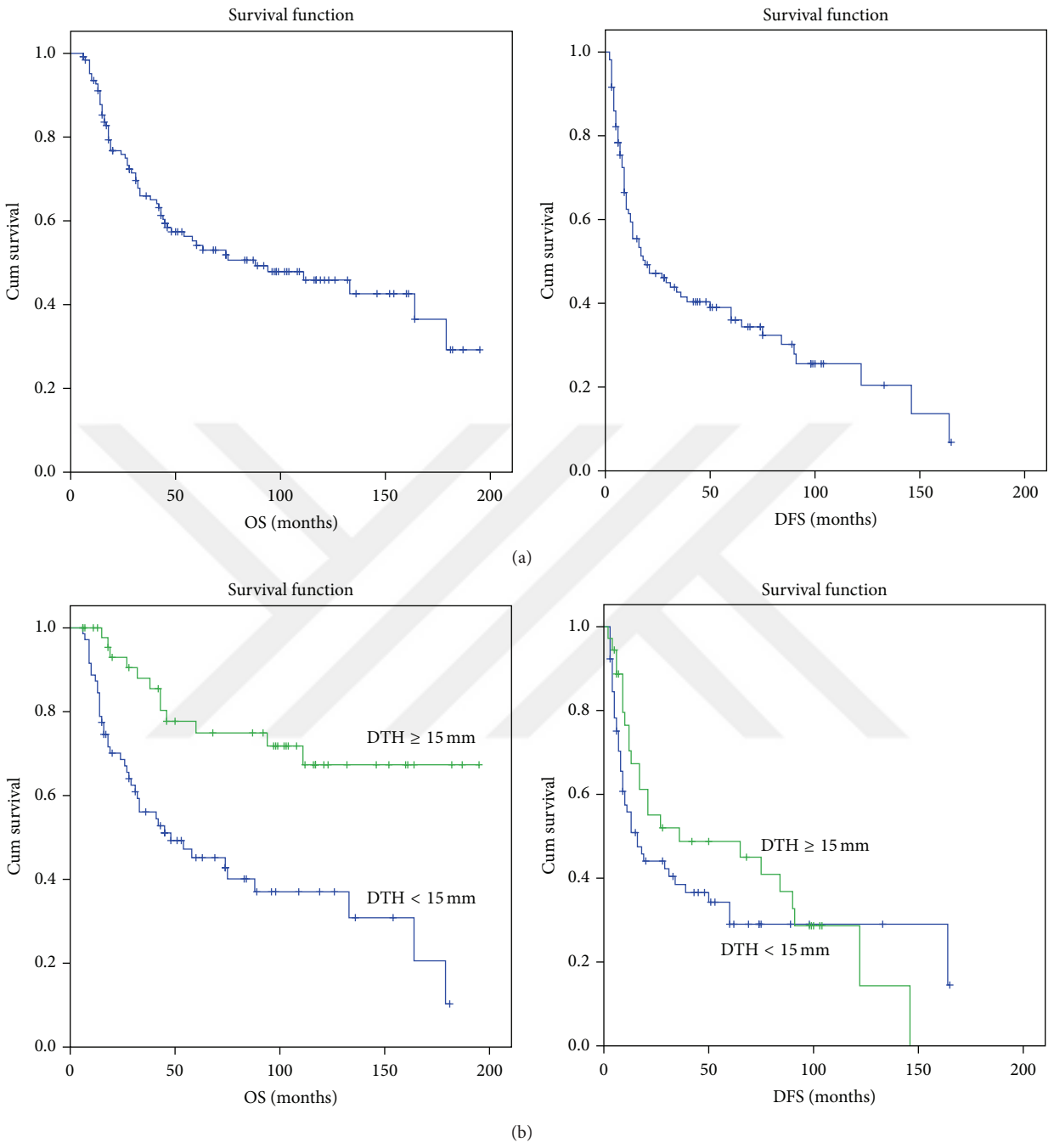


Figure 1: Kaplan-Meier survival curves of 126 melanoma patients with AJCC stages III B and C disease. (a) Survival data of all patients undergoing autologous vaccination. (b) Correlation of survival with delayed type hypersensitivity (DTH) response to unmodified melanoma attained following vaccination. OS: overall survival; DFS: disease free survival.

expression of each of these CTA genes was associated with a reduced risk of death. As CTA genes are coexpressed, we performed a principal component analysis (PCA) of the genes to stratify the patients, using the "princomp" function in R. We found the first principal component (PC1) to explain more than 80% of the variance when expression values from the 35 samples corresponding to all 51 probe sets from CTA genes in the C-MAP array were used (Figure 3(a)). To validate the predictive data extracted from C-MAP, we used qPCR results obtained from 21 melanoma line samples for MAGE-A1, SSX1, SSX4, and NY-ESO-1 (CTAG1B). Similar to the in silico data, a PCA analysis based on these four genes was able to explain 75% of the variance. In both analyses, stratifying the patients into low/intermediate and high expression based on PC1 values, we show worse outcome for low CTA expressors (below quartile 1) compared to high

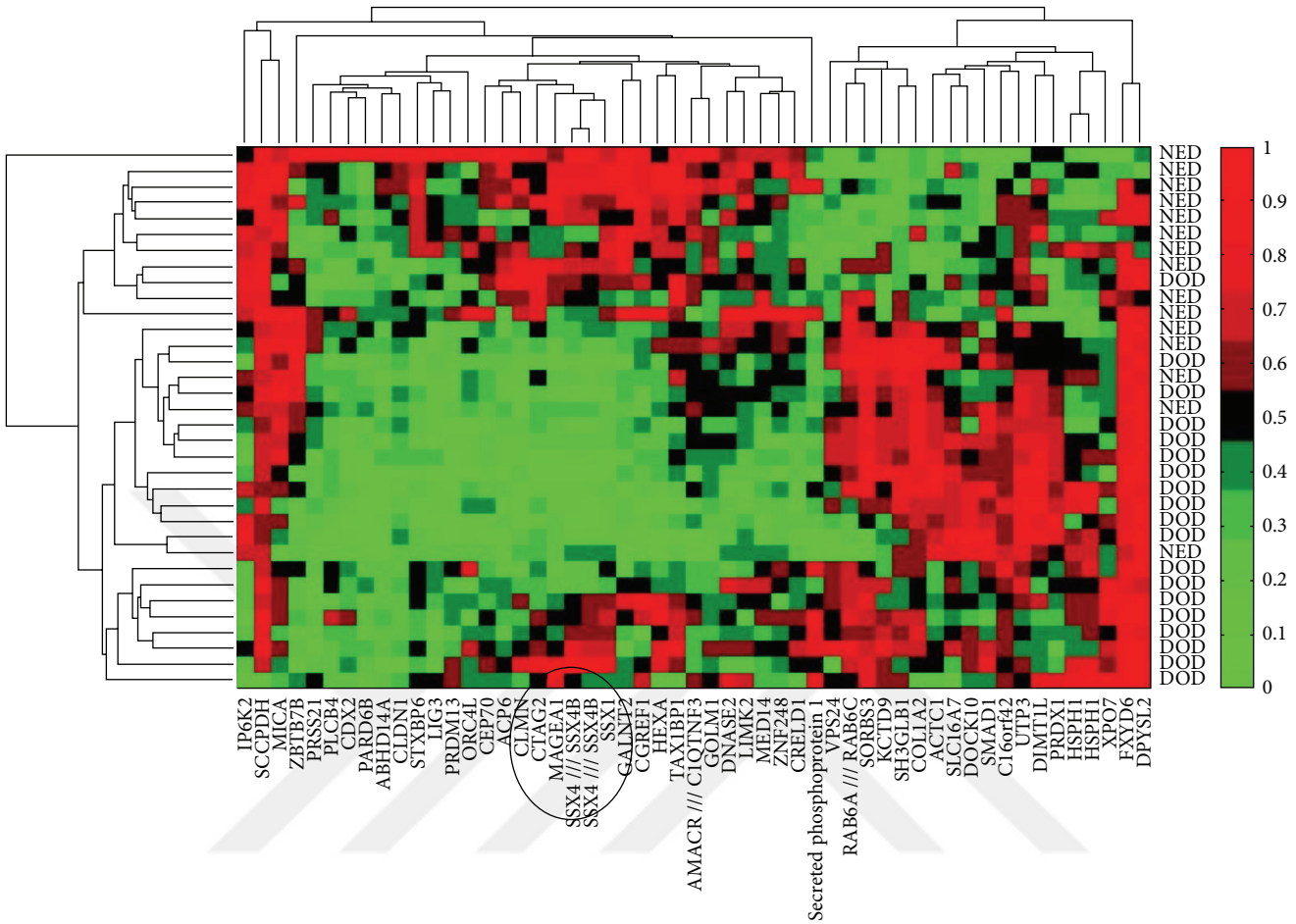


Figure 2: Hierarchical clustering gene expression of a 50-gene signature showing strongest association with prognosis in 35 stage III melanoma patients. Cancer testis antigens CTAG2, MAGEA1, SSX1, and SSX4 are circled in a cluster. NED: no evidence of disease at time of analysis; DOD: died of metastatic disease.

expressors (above quartile 4), $p = 0.02$ (Figure 3). A similar analysis performed using an unrelated melanoma cohort (GSE19234 [28]), which never had an autologous vaccine, did not reveal an association between CTA gene expression and survival (not shown), suggesting that CTA gene expression is not a prognostic factor by itself but becomes one in the context of autologous vaccination.

MICA (MHC class I related chain A), another gene which correlated with a reduced risk of death and for which there is an antibody for flow cytometry, was used to validate the expression data (Figure 4). Out of 11 samples analyzed, 10 yielded MICA protein expression (by flow cytometry) concordant with the gene expression value. In 8 samples the data was predictive of patient's current status, whether alive, no evidence of disease (NED), or died of disease (DOD).

3.5. B-RAF Status and Survival. B-RAF status was determined for 32 patients of the 35 that were analyzed for gene expression. Eleven patients (33%) harbored the V600E mutation. The median survival for patients with V600E mutation was 50 months compared with 45 months for the wild type (WT) group ($p = 0.9$). Since most patients in the series who had recurrent disease died before 2012, none of them were treated with B-RAF inhibitors when they developed metastases, and consequently differences in survival could not be attributed to better treatment options. Four out of 11 patients with V600E mutations had strong DTH response (36%), compared with seven out of 21 (33.3%) in the WT group ($p = 0.86$).

3.6. Patients Who Received Melanoma Vaccine Had Improved Survival following Ipilimumab Treatment for Advanced Disease. Thirty-five patients who received melanoma cell vaccine and developed nonresectable metastatic disease were referred to ipilimumab treatment in BMS studies CA 184-004, 024, and 025 and later as a standard second-line treatment. These patients were compared with a nonselected concurrent group of other 35 patients, who never received the vaccine (Table 2). The majority, 62 patients (89%), received ipilimumab at a dose of 3 mg/kg. Response Evaluation Criteria in Solid Tumors (RECIST) were used to define response to treatment. Complete response was achieved in six patients (8.6%), partial response in 14 (20%), and stable disease in 7 (10%). Median OS of the group was 14 months (95% CI

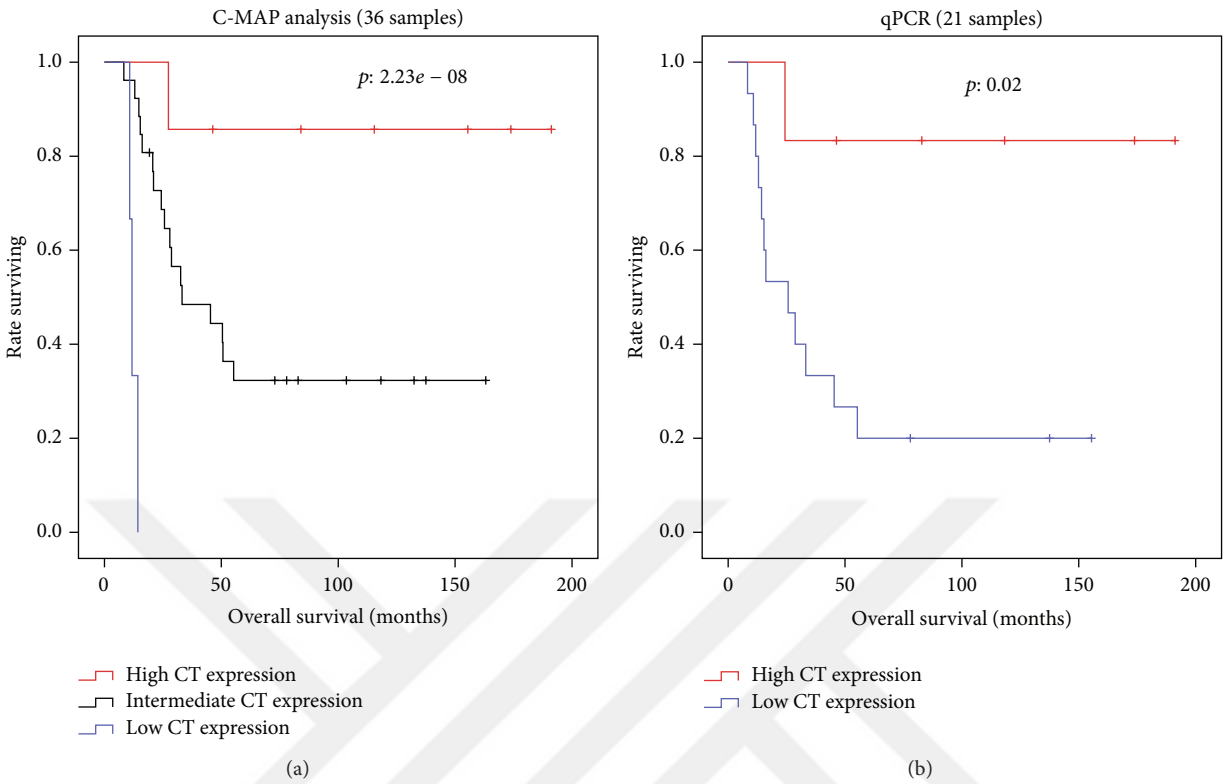


Figure 3: Overall survival curves stratifying the patients according to integrated cancer testis antigen genes expression. As CTA genes are coexpressed, we performed a principal component analysis (PCA) of the genes to stratify the patients into low/intermediate and high expression based on first principal component (PC1) values. (a) PC1 was determined in 35 melanoma lines based on all 51 probe sets from CTA genes in the C-MAP array. (b) PC1 was determined in 21 melanoma lines based on qPCR data generated for CTAs MAGE-A1, SSX1, SSX4, and NY-ESO-1 (CTAG1B).

5.7–22 months). Multivariate analysis revealed that the single parameter that correlated with an improved OS time was previous treatment with melanoma cell vaccine. In the group of 35 vaccinated patients, median OS was 31 months with a 3-year survival of 46%, compared with a median OS of 9 months and 3-year survival of only 19% in the nonvaccinated patients, ($p = 0.007$, Pearson's chi square test). The results from this retrospective cohort of patients suggest that response rate and survival are improved when ipilimumab treatment succeeds autologous melanoma vaccine.

4. Discussion

In this phase II study we administered a vaccine composed of the autologous tumor given to melanoma patients as a postoperative adjuvant for AJCC stage III disease following resection of macro metastases. The study was single-armed, as it was initiated prior to the registration of IFNα as standard of care for the adjuvant treatment of melanoma stages IIB and III. In an initial cohort, which included patients of worse prognoses (stages IIIB, IIIC, and IV) we observed an unexpected good survival rate with this vaccine [18]. In view of the reduced toxicity of the vaccination regimen, we did not offer patients a high dose interferon arm after its registration. On the other hand, the inclusion of an observation arm was ethically questionable. Thus, we opted to continue with the protocol in its single-arm design, to record survival rates in adjuvant stage IIIB and C patients. Eventually, these patients were the less likely ones to derive benefit from adjuvant IFNα therapy, since treatment with the standard of care IFNα-2b or PEG-IFN did not yield any survival advantage for them [29]. The projected 5-year OS for these patients was estimated at 40% and the DFS at 30%, as shown in the Kaplan-Meier curves generated in a meta-analysis from EORTC trials 18952 and 18992 of patients with AJCC stage III-N2. In our group of patients with the same stage disease, the 5-year OS reached 54%, with a median OS of 88 months and a 5-year DFS of 34%; these were achieved with no major adverse effects. In another series, the Nordic study, the best median survival, 72 months, was derived from intermediate-high doses of IFNα-2b for one year. These results, worse than those we observed with the autologous vaccine, were generated in a group of patients in which 35% had earlier stage disease (IIIA).

Notably, in our study population, patients with a strong positive skin reaction to their unmodified melanoma had much better prognosis, with a 5-year OS of 75% and DFS of 44%. The link between immune parameters that attest to the development of antitumor response and improved survival has been previously observed [17, 30–36]. DTH is a crude measurement, but an easy test to apply on all patients.

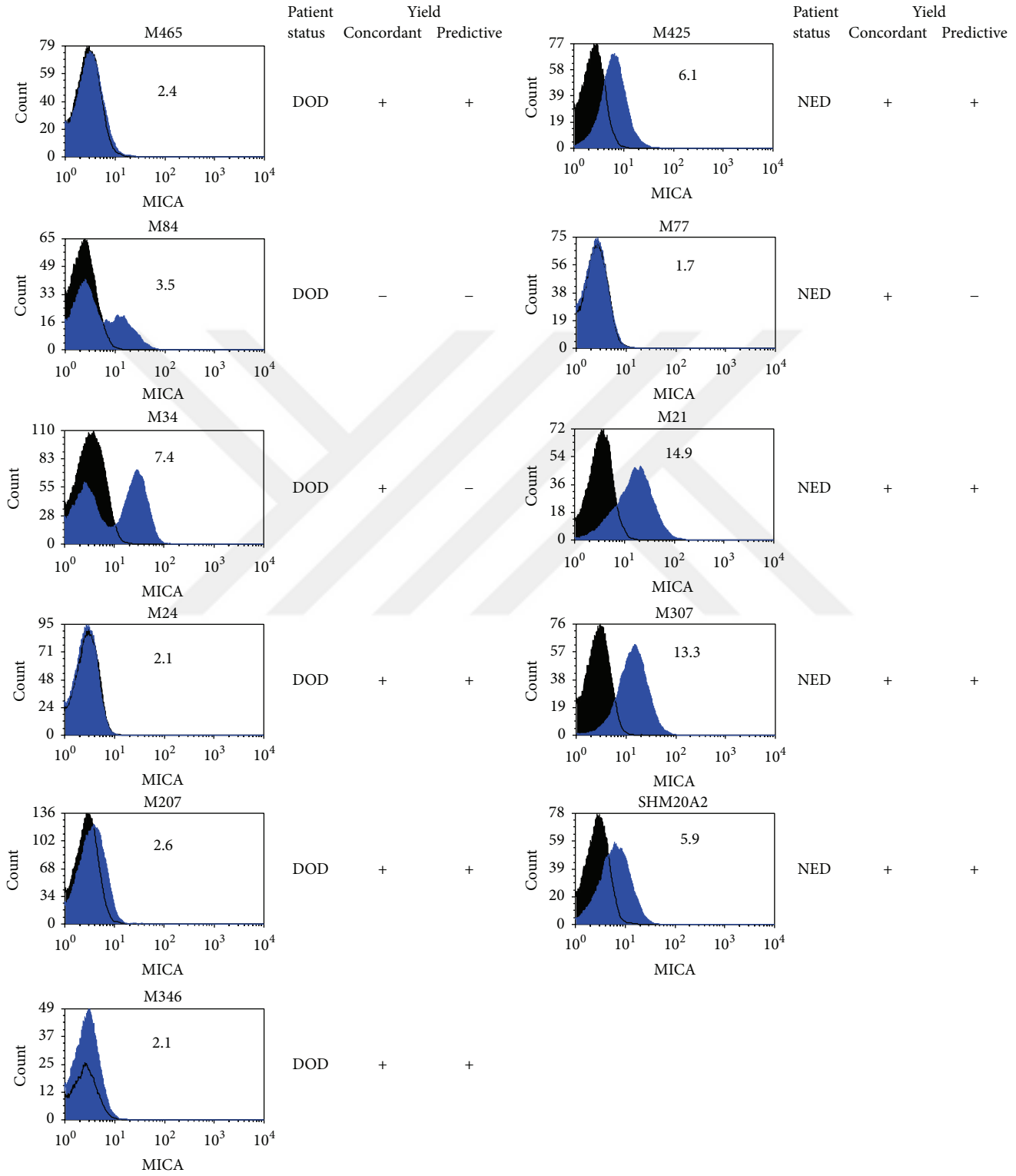


Figure 4: Flow cytometry analysis of MICA surface expression on melanoma cell lines. The number in the histogram is the Luminex 1000 expression value of MICA. Black histogram: background staining with isotype control Ab. Blue histogram: MICA staining with anti-MICA specific Ab.

Table 2: Clinical data and ipilimumab treatment results of 70 patients with AJCC stage IV melanoma who received or did not receive melanoma vaccine.

Prior vaccination	Yes		No		p^*
	n	%	*n*	%	
Patient number	35	50	35	50	
M stage					
M1A	8	23	3	9	
M1B	8	23	10	28	*0.3*
M1C	19	54	22	63	
Dose					
3 mg/kg	30	86	32	91	*0.45*
10 mg/kg	5	14	3	9	
Reinduction	4	12	3	9	
Treatment stopped for toxicity	5	14	2	6	
Objective response					
CR	3	9	3	9	
PR	10	29	4	11	*0.03***
S	6	17	1	3	
P	16	46	27	77	
Survival data					
Median OS (months)	31 (3–47)		9 (2–50)		*0.007*
3-year survival rate	46%		19%		

CR: complete response, PR: partial response, S: stable disease, and P: progression.
*Pearson chi square (two-sided) test.
** p value for progression (P) versus any benefit (CR, PR, and S).

Several vaccine clinical trials included a DTH test as part of their clinical evaluation and reported a positive correlation between DTH response and longer survival [37, 38]. Biopsies taken from skin injection sites revealed vaccine-induced antigen-specific T cells [39, 40]. We previously demonstrated effective antimelanoma CD4 T cell activity associated with improved survival in a cohort included in the present study [16]. Thus, a positive DTH reaction could be indicative of the emergence of antimelanoma immunity, and the improved survival of patients attaining strong DTH would attest to the vaccine's protective effects.

Since autologous tumor tissue is often not available for vaccine preparation, defined antigens, consisting of short or long peptides, are used as a substitution [41–43]. The enthusiasm for the use of defined antigens decreased when successful generation of antigen-specific T cells failed to protect against tumor progression [11, 44]. Loss of MHC and impaired peptide presentation by the melanoma were among the reasons for vaccine failure, but another limitation is that many tumor-progeny antigens are essentially self-antigens that evoke weak responses [45]. Unlike wildtype proteins, mutations have the potential to generate neoantigens which are better targets. The significance of mutation-derived neoantigens was illustrated when adoptively transferred tumor-infiltrating lymphocytes that destroyed melanoma in patients were surveyed and found to target mutated epitopes [46]. Furthermore, recent data has clearly demonstrated that these mutations are, in large part, the functional targets of immune checkpoint blockade [47]. It may not take long until "mutanome" libraries are generated for melanoma patients. But until the hurdles of preparing a totally individualized vaccine are overcome, the best source of mutated antigens, as we have previously shown, is still the autologous tumor and, ideally, those tumors that express both MHC class I and class II [16].

Another important component of melanoma cell vaccines, as reflected by our data, are a class of antigens known as cancer testis antigens. CTAs are products of several multigene families, many of which map to chromosome X, that have arisen through chromosomal duplications and were initially identified through immunologic assays [48]. When associated with disease outcome, CTAs sometimes confer worse prognosis [49], but when protective immune responses are recorded, CTAs are dominant targets. For example, rising antibody responses to CTA NY-ESO-1 (CTAG1B) were recorded following shrinkage of melanoma in a patient with abscopal tumor response [50], and among patients treated with ipilimumab, seropositivity to NY-ESO-1 with associated CD8 T cell response correlated with 77% clinical benefit [51].

Using a Luminex-based method to generate a gene expression array and qPCR validation, we showed increased expression of MAGEA1, CTAG1, CTAG2, SSX1, and SSX4 in patients with improved survival. We suggest that these patients' prolonged survival is attributed to the melanoma immunogenicity potentiated by the CTAs and augmented by the vaccination procedure.

Lastly, our data implies that patients who had been immunized against the autologous tumor prior to receiving ipilimumab survived longer than patients who did not receive an autologous melanoma vaccine. The precise mechanism of action of ipilimumab is not completely clear. CTLA-4 receptor blockade prevents inhibition of activated effector lymphocytes and selective increase in the ratio between T effectors and regulatory T cells within the tumor [52]. In patients with preexisting immune response, the removal of inhibitory signals boosts antitumor activity [51], leading to the hypothesis that prior vaccination might induce in a subset of patients a population of specific antitumor cytotoxic T cells and a potent immune response following immune checkpoint blockade. In this respect, it was encouraging to note that in a cohort of vaccinated patients that had received ipilimumab, the 3-year survival was 46%, compared with 19% for the nonvaccinated patients ($p = 0.007$).

5. Conclusions

In this noncontrolled phase II study we have shown that adjuvant treatment with autologous melanoma vaccine yields overall and disease-free survival rates that are not inferior to those obtained with interferon alpha. In a subgroup of patients who attained a strong positive skin response to unmodified autologous tumor, survival rates were exceptionally good, with a 5-year OS of 75%. Increased expression of CTAs by the tumor correlated with improved survival. Lastly, improved survival time following ipilimumab treatment was observed for patients who had previously been

Table 3: Patients survival data.

	Number (%)	1 year (%)	2 years (%)	5 years (%)	Median (mo, 95% CI)	p value
Overall survival						
All*	126	93	77	54	88 (40–137)	
DTH ≥ 15 mm	48 (40)	100	93	75	Not reached	
DTH < 15 mm	71 (60)	89	70	44	45 (13–20)	**<0001**
DTH ≥ 10 mm	75 (63)	97	83	64	181 (75–287)	
DTH < 10 mm	44 (37)	87	66	32	41 (23–59)	**0.003**
Disease-free survival						
All**	107	58	45	34	18 (5–31)	
DTH ≥ 15 mm	36 (36)	74	53	47	36 (0.00–87)	
DTH < 15 mm	65 (64)	55	42	26	15 (7–23)	**0.27**
Adjuvant radiotherapy Overall survival						
All	56	90	69	41	43 (27–59)	**0.055**
DTH ≥ 15 mm	22 (42)	100	87	58	111	
DTH < 15 mm	31 (58)	84	56	33	31 (20–42)	**0.024**

*DTH data was available for 119 of 126 patients.
**DTH data was available for 101 of 107 patients with recorded disease-free survival.

Table 4: Selected genes expressed on melanoma cells which correlate with overall survival. The genes were depicted by Maxstat-package utilizing a Cox-regression analysis followed by rank statistics to determine the best cut-off value which separates patients into favorable versus unfavorable survival groups. ↑ = higher expression correlates with prolonged survival time; ↓ = higher expression correlates with decreased survival time.

Gene	↑/↓	Full name	Maxstat cut point	Maxstat p value	CoxPH hazard ratio	CoxPH p value
MICA	↑	MHC class I-related chain A (NKG2D ligand)	7.8	0.018	0.5	0.015
CTAG2	↑	Cancer testis antigen 2 (LAGE-1, NY-ESO-2)	5.9	0.014	0.58	0.015
SSX4 /// SSX4B	↑	Synovial sarcoma, X breakpoint 4	3.9	0.008	0.58	0.017
TGFA	↑	Transforming growth factor α	4.08	0.011	0.66	0.018
KIR3DX1	↑	Killer cell Ig-like receptor	4	0.023	0.56	0.02
MAGEA1	↑	Melanoma antigen family A, 1	5.52	0.042	0.69	0.028
SSX1	↑	Synovial sarcoma, X breakpoint 1	3.28	0.035	0.63	0.043
SMAD1	↓	SMAD family member 1	5.5	0.046	2.58	0.0002
HSPH1	↓	Heat shock 105 kDa/110 kDa protein 1	10.7	0.008	1.9	0.007

vaccinated. Thus, we suggest that autologous melanoma vaccine induces protective immunity and may offer leverage for other immunotherapies.

Disclosure

The funding sources had no involvement in the study design or in its performance.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The authors wish to acknowledge the devoted technical work of Anna Kuznetz, Inna Ben David, and Yael Gelfand. This work was supported by research grants from Dr. Miriam and Sheldon G. Adelson Medical Research Foundation (AMRF), Israel Cancer Association, Deutsche Forschungsgemeinschaft (DFG), Rosetrees Trust, and the Perlstein Family Fund. Professor Ali O. Gure was supported by TUBITAK Grant no. 112S023.

References

[1] J. M. Kirkwood, M. H. Strawderman, M. S. Ernstoff, T. J. Smith, E. C. Borden, and R. H. Blum, "Interferon alfa-2b adjuvant therapy of high-risk resected cutaneous melanoma: the Eastern Cooperative Oncology Group trial EST 1684," *Journal of Clinical Oncology*, vol. 14, no. 1, pp. 7–17, 1996.

[2] J. M. Kirkwood, J. G. Ibrahim, V. K. Sondak et al., "High- and low-dose interferon alfa-2b in high-risk melanoma: first analysis of intergroup trial E1690/S9111/C9190," *Journal of Clinical Oncology*, vol. 18, no. 12, pp. 2444–2458, 2000.

[3] A. M. M. Eggermont, S. Suciu, R. MacKie et al., "Post-surgery adjuvant therapy with intermediate doses of interferon alfa 2b versus observation in patients with stage IIb/III melanoma (EORTC 18952): randomised controlled trial," *The Lancet*, vol. 366, no. 9492, pp. 1189–1196, 2005.

[4] N. Cascinelli, F. Belli, R. M. MacKie, M. Santinami, R. Bufalino, and A. Morabito, "Effect of long-term adjuvant therapy with interferon alpha-2a in patients with regional node metastases from cutaneous melanoma: a randomised trial," *The Lancet*, vol. 358, no. 9285, pp. 866–869, 2001.

[5] B. W. Hancock, K. Wheatley, S. Harris et al., "Adjuvant interferon in high-risk melanoma: the AIM HIGH study—United Kingdom Coordinating Committee on cancer research randomized study of adjuvant low-dose extended-duration interferon alfa-2a in high-risk resected malignant melanoma," *Journal of Clinical Oncology*, vol. 22, no. 1, pp. 53–61, 2004.

[6] C. Robert, L. Thomas, I. Bondarenko et al., "Ipilimumab plus dacarbazine for previously untreated metastatic melanoma," *The New England Journal of Medicine*, vol. 364, no. 26, pp. 2517–2526, 2011.

[7] S. Mocellin, M. B. Lens, S. Pasquali, P. Pilati, and V. Chiarion Sileni, "Interferon alpha for the adjuvant treatment of cutaneous melanoma," *The Cochrane Database of Systematic Reviews*, vol. 6, Article ID CD008955, 2013.

[8] J. M. Kirkwood, J. Manola, J. Ibrahim, V. Sondak, M. S. Ernstoff, and U. Rao, "A pooled analysis of eastern cooperative oncology group and intergroup trials of adjuvant high-dose interferon for melanoma," *Clinical Cancer Research*, vol. 10, no. 5, pp. 1670–1677, 2004.

[9] P. C. Rodríguez and A. C. Ochoa, "T cell dysfunction in cancer: role of myeloid cells and tumor cells regulating amino acid availability and oxidative stress," *Seminars in Cancer Biology*, vol. 16, no. 1, pp. 66–72, 2006.

[10] A. M. M. Eggermont, S. Suciu, P. Rutkowski et al., "Adjuvant ganglioside GM2-KLH/QS-21 vaccination versus observation after resection of primary tumor >1.5 mm in patients with stage II melanoma: results of the EORTC 18961 randomized phase III trial," *Journal of Clinical Oncolog*, vol. 31, no. 30, pp. 3831–3837, 2013.

[11] S. A. Rosenberg, R. M. Sherry, K. E. Morton et al., "Tumor progression can occur despite the induction of very high levels of self/tumor antigen-specific $CD8^+$ T cells in patients with melanoma," *The Journal of Immunology*, vol. 175, no. 9, pp. 6169–6176, 2005.

[12] S. A. Rosenberg, J. C. Yang, and N. P. Restifo, "Cancer immunotherapy: moving beyond current vaccines," *Nature Medicine*, vol. 10, no. 9, pp. 909–915, 2004.

[13] D. J. Schwartzentruber, D. H. Lawson, J. M. Richards et al., "gp100 peptide vaccine and interleukin-2 in patients with advanced melanoma," *The New England Journal of Medicine*, vol. 364, no. 22, pp. 2119–2127, 2011.

[14] S. Wada, C. M. Jackson, K. Yoshimura et al., "Sequencing CTLA-4 blockade with cell-based immunotherapy for prostate cancer," *Journal of Translational Medicine*, vol. 11, article 89, 2013.

[15] B. A. Fox, D. J. Schendel, L. H. Butterfield et al., "Defining the critical hurdles in cancer immunotherapy," *Journal of Translational Medicine*, vol. 9, no. 1, article 214, 2011.

[16] M. Lotem, A. Machlenkin, T. Hamburger et al., "Autologous melanoma vaccine induces antitumor and self-reactive immune responses that affect patient survival and depend on MHC class II expression on vaccine cells," *Clinical Cancer Research*, vol. 15, no. 15, pp. 4968–4977, 2009.

[17] A. Baars, A. M. E. Claessen, A. J. M. van den Eertwegh et al., "Skin tests predict survival after autologous tumor cell vaccination in metastatic melanoma: experience in 81 patients," *Annals of Oncology*, vol. 11, no. 8, pp. 965–970, 2000.

[18] M. Lotem, T. Peretz, O. Drize et al., "Autologous cell vaccine as a post operative adjuvant treatment for high-risk melanoma patients (AJCC stages III and IV)," *British Journal of Cancer*, vol. 86, no. 10, pp. 1534–1539, 2002.

[19] S. Edge, D. R. Byrd, C. C. Compton, A. G. Fritz, F. L. Greene, and A. Trotti, Eds., *AJCC Cancer Staging Manual*, Springer, New York, NY, USA, 7th edition, 2010.

[20] H. N. Claman and S. D. Miller, "Requirements for induction of T cell tolerance to DNFB: efficiency of membrane associated DNFB," *The Journal of Immunology*, vol. 117, no. 2, pp. 480–485, 1976.

[21] M. O. C. Ota, J. Vekemans, S. E. Schlegel-Haueter et al., "Influence of Mycobacterium bovis bacillus Calmette-Guérin on antibody and cytokine responses to human neonatal vaccination," *Journal of Immunology*, vol. 168, no. 2, pp. 919–925, 2002.

[22] J.-Y. Liu, Y. Wu, X.-S. Zhang et al., "Single administration of low dose cyclophosphamide augments the antitumor effect of dendritic cell vaccine," *Cancer Immunology, Immunotherapy*, vol. 56, no. 10, pp. 1597–1604, 2007.

[23] F. Ghiringhelli, N. Larmonier, E. Schmitt et al., "$CD4^+CD25^+$ regulatory T cells suppress tumor immunity but are sensitive to cyclophosphamide which allows immunotherapy of established tumors to be curative," *European Journal of Immunology*, vol. 34, no. 2, pp. 336–344, 2004.

[24] M. E. C. Lutsiak, R. T. Semnani, R. De Pascalis, S. V. S. Kashmiri, J. Schlom, and H. Sabzevari, "Inhibition of $CD4^+25^+$ T regulatory cell function implicated in enhanced immune response by low-dose cyclophosphamide," *Blood*, vol. 105, no. 7, pp. 2862–2868, 2005.

[25] S. Brode and A. Cooke, "Immune-potentiating effects of the chemotherapeutic drug cyclophosphamide," *Critical Reviews in Immunology*, vol. 28, no. 2, pp. 109–126, 2008.

[26] V. Russo, L. Pilla, F. Lunghi et al., "Clinical and immunologic responses in melanoma patients vaccinated with MAGE-A3-genetically modified lymphocytes," *International Journal of Cancer*, vol. 132, no. 11, pp. 2557–2566, 2013.

[27] B. Rini, "Future approaches in immunotherapy," *Seminars in Oncology*, vol. 41, supplement 5, no. 5, pp. S30–S40, 2014.

[28] R. Chen, G. Zhang, Y. Zhou, N. Li, and J. Lin, "A time course-dependent metastatic gene expression signature predicts outcome in human metastatic melanomas," *Diagnostic Pathology*, vol. 9, no. 1, article 155, 2014.

[29] A. M. M. Eggermont, S. Suciu, A. Testori et al., "Ulceration and stage are predictive of interferon efficacy in melanoma: results of the phase III adjuvant trials EORTC 18952 and EORTC 18991," *European Journal of Cancer*, vol. 48, no. 2, pp. 218–225, 2012.

[30] A. Escobar, M. López, A. Serrano et al., "Dendritic cell immunizations alone or combined with low doses of interleukin-2 induce specific immune responses in melanoma patients," *Clinical and Experimental Immunology*, vol. 142, no. 3, pp. 555–568, 2005.

[31] E. C. Hsueh, E. Famatiga, S. Shu, X. Ye, and D. L. Morton, "Peripheral blood CD4+ T-cell response before postoperative active immunotherapy correlates with clinical outcome in metastatic melanoma," *Annals of Surgical Oncology*, vol. 11, no. 10, pp. 892–899, 2004.

[32] D. Berd, T. Sato, H. C. Maguire Jr., J. Kairys, and M. J. Mastrangelo, "Immunopharmacologic analysis of an autologous, hapten-modified human melanoma vaccine," *Journal of Clinical Oncology*, vol. 22, no. 3, pp. 403–415, 2004.

[33] M. Carsana, G. Tragni, G. Nicolini et al., "Comparative assessment of TCRBV diversity in T lymphocytes present in blood, metastatic lesions, and DTH sites of two melanoma patients vaccinated with an IL-7 gene-modified autologous tumor cell vaccine," *Cancer Gene Therapy*, vol. 9, no. 3, pp. 243–253, 2002.

[34] W. H. J. Kruit, S. Suciu, B. Dreno et al., "Selection of immunostimulant AS15 for active immunization with MAGE-A3 protein: results of a randomized phase II study of the European Organisation for Research and Treatment of Cancer Melanoma Group in Metastatic Melanoma," *Journal of Clinical Oncology*, vol. 31, no. 19, pp. 2413–2420, 2013.

[35] U. E. Burkhardt, U. Hainz, K. Stevenson et al., "Autologous CLL cell vaccination early after transplant induces leukemia-specific T cells," *The Journal of Clinical Investigation*, vol. 123, no. 9, pp. 3756–3765, 2013.

[36] D. L. Morton, L. J. Foshag, D. S. B. Hoon et al., "Prolongation of survival in metastatic melanoma after active specific immunotherapy with a new polyvalent melanoma vaccine," *Annals of Surgery*, vol. 216, no. 4, pp. 463–482, 1992.

[37] T. Miyatake, Y. Ueda, A. Morimoto et al., "WT1 peptide immunotherapy for gynecologic malignancies resistant to conventional therapies: a phase II trial," *Journal of Cancer Research and Clinical Oncology*, vol. 139, no. 3, pp. 457–463, 2013.

[38] Y. Muragaki, T. Maruyama, H. Iseki et al., "Phase I/IIa trial of autologous formalin-fixed tumor vaccine concomitant with fractionated radiotherapy for newly diagnosed glioblastoma. Clinical article," *Journal of Neurosurgery*, vol. 115, no. 2, pp. 248–255, 2011.

[39] I. J. M. de Vries, M. R. Bernsen, W. J. Lesterhuis et al., "Immunomonitoring tumor-specific T cells in delayed-type hypersensitivity skin biopsies after dendritic cell vaccination correlates with clinical outcome," *Journal of Clinical Oncology*, vol. 23, no. 24, pp. 5779–5787, 2005.

[40] S. Wilgenhof, A. M. T. Van Nuffel, J. Corthals et al., "Therapeutic vaccination with an autologous mRNA electroporated dendritic cell vaccine in patients with advanced melanoma," *Journal of Immunotherapy*, vol. 34, no. 5, pp. 448–456, 2011.

[41] C. L. Slingluff Jr., G. R. Petroni, W. Olson et al., "Helper T-cell responses and clinical activity of a melanoma vaccine with multiple peptides from MAGE and melanocytic differentiation antigens," *Journal of Clinical Oncology*, vol. 26, no. 30, pp. 4973–4980, 2008.

[42] N. Leffers, A. J. A. Lambeck, M. J. M. Gooden et al., "Immunization with a P53 synthetic long peptide vaccine induces P53-specific immune responses in ovarian cancer patients, a phase II trial," *International Journal of Cancer*, vol. 125, no. 9, pp. 2104–2113, 2009.

[43] P. Sabbatini, T. Tsuji, L. Ferran et al., "Phase I trial of overlapping long peptides from a tumor self-antigen and poly-ICLC shows rapid induction of integrated immune response in ovarian cancer patients," *Clinical Cancer Research*, vol. 18, no. 23, pp. 6497–6508, 2012.

[44] C. Schaefer, L. H. Butterfield, S. Lee et al., "Function but not phenotype of melanoma peptide-specific $CD8^+$ T cells correlate with survival in a multiepitope peptide vaccine trial (ECOG 1696)," *International Journal of Cancer*, vol. 131, no. 4, pp. 874–884, 2012.

[45] D. B. Kazansky, "Intrathymic selection: new insight into tumor immunology," *Advances in Experimental Medicine and Biology*, vol. 601, pp. 133–144, 2007.

[46] P. F. Robbins, Y.-C. Lu, M. El-Gamil et al., "Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells," *Nature Medicine*, vol. 19, no. 6, pp. 747–752, 2013.

[47] N. Riaz, L. Morris, J. J. Havel, V. Makarov, A. Desrichard, and T. A. Chan, "The role of neoantigens in response to the immune checkpoint blockade," *International Immunology*, 2016.

[48] Y.-T. Chen, M. J. Scanlan, U. Sahin et al., "A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 5, pp. 1914–1918, 1997.

[49] A. O. Gure, R. Chua, B. Williamson et al., "Cancer-testis genes are coordinately expressed and are markers of poor outcome in non-small cell lung cancer," *Clinical Cancer Research*, vol. 11, no. 22, pp. 8055–8062, 2005.

[50] M. A. Postow, M. K. Callahan, C. A. Barker et al., "Immunologic correlates of the abscopal effect in a patient with melanoma," *The New England Journal of Medicine*, vol. 366, no. 10, pp. 925–931, 2012.

[51] J. Yuan, M. Adamow, B. A. Ginsberg et al., "Integrated NY-ESO-1 antibody and $CD8^+$ T-cell responses correlate with clinical benefit in advanced melanoma patients treated with ipilimumab," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 40, pp. 16723–16728, 2011.

[52] T. R. Simpson, F. Li, W. Montalvo-Ortiz et al., "Fc-dependent depletion of tumor-infiltrating regulatory t cells co-defines the efficacy of anti-CTLA-4 therapy against melanoma," *Journal of Experimental Medicine*, vol. 210, no. 9, pp. 1695–1710, 2013.

PERMISSIONS FOR COPYRIGHTED MATERIAL

AMERICAN SOCIETY FOR NUTRITION LICENSE TERMS AND CONDITIONS

Sep 03, 2016

This Agreement between Kerem Senses ("You") and American Society for Nutrition ("American Society for Nutrition") consists of your license details and the terms and conditions provided by American Society for Nutrition and Copyright Clearance Center.

License Number	3941230612995
License date	Sep 03, 2016
Licensed Content Publisher	American Society for Nutrition
Licensed Content Publication	The Journal of Nutrition
Licensed Content Title	Mechanisms of DNA Damage, DNA Hypomethylation, and Tumor Progression in the Folate/Methyl-Deficient Rat Model of Hepatocarcinogenesis
Licensed Content Author	S. Jill James, Igor P. Pogribny, Marta Pogribna, Barbara J. Miller, Stefanie Jernigan, Stepan Melnyk
Licensed Content Date	Nov 1, 2003
Licensed Content Volume Number	133
Licensed Content Issue Number	11
Type of Use	Thesis/Dissertation
Requestor type	Student
Portion	Figures/table/illustration
Number of Figures/table/illustration	1
List of figures/table/ illustration	FIGURE 1
Order reference number	
Title of your dissertation / thesis	CANCER TESTIS GENE EXPRESSION AS A BIOMARKER OF ONE-CARBON METABOLIC ACTIVITY, DRUG SENSITIVITY AND PHENOTYPIC HETEROGENEITY IN NON-SMALL CELL LUNG CANCER AND MALIGNANT MELANOMA
Expected completion date	Sep 2016
Estimated size(pages)	170
Requestor Location	Kerem Senses Bilkent University Dept. of Molecular Biology and Genetics SB Building 2nd floor, room 237 Ankara, 06800 Turkey Attn: Kerem Senses
Billing Type	Invoice
Billing Address	Kerem Senses Bilkent University Dept. of Molecular Biology and Genetics SB Building 2nd floor, room 237

Ankara, Turkey 06800
Attn: Kerem Senses

Total 0.00 USD

Terms and Conditions

Terms and Conditions for RightsLink Licenses

1) This license agreement ("Agreement") is between the Customer as identified by the user in **RightsLink** ("Customer") and The American Society for Nutrition, Inc, a nonprofit corporation with offices at 9211 Corporate Blvd., Suite 300, Rockville, MD 20850 USA ("ASN"), regarding content published in one of the following ASN journals: *The American Journal of Clinical Nutrition (AJCN)* or *The Journal of Nutrition (JN) or Advances in Nutrition*. The license administrator for this Agreement is the Copyright Clearance Center (CCC).

2) Introduction

The publisher of this copyrighted material is ASN. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by CCC, at the time that you opened your **RightsLink** account and that are available at any time at http://myaccount.copyright.com.

3) ASN occasionally publishes material for which it does not hold copyright; in such instances, ASN identifies and properly cites the appropriate copyright holder. It is the Customer's responsibility to verify that ASN holds copyright for which the Customer is requesting permission. If the material is credited to another source, then ASN does not have the authority to grant permission for its reuse and this license is to be considered invalid.

4) Limited License

ASN hereby grants to you a non-exclusive license to use this material. Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the licensing process. Except when material is licensed for reuse in a scholarly journal, magazine, newsletter, newspaper, book, or textbook, in which case permission extends to reprints of the product produced for publisher restocking, any form of republication must be completed within one year from the date hereof (although copies prepared before then may be distributed thereafter). Reprints must not be new products nor marketed as such: **reuse in new volumes or new editions must be licensed explicitly**. Any form of republication must be completed within one year from the date hereof (although copies prepared before then may be distributed thereafter); and any electronic posting is limited to the time specified in this request, if any. ASN asks that commercial organizations limit distribution of photocopies to within the organization requesting permission. Requests to distribute content copyrighted by ASN outside of the requesting organization should be processed as reprints or ePrints. Use of more than 50% of an article copyrighted by ASN in a single new work must be reviewed separately by ASN staff.

5) Geographic Rights

Licenses may be exercised solely in the country or countries specified by the Customer on the Quick Price or Additional Data page in **RightsLink**. However, in cases in which ASN has exclusive republication licensees in specific countries, licenses to reproduce full-text material in these countries may be revoked by ASN. In such cases, the Customer would be referred to the license agent in the respective country or countries. For questions about whether this would apply to your order, please contact ASN at publications@nutrition.org.

6) Alterations/Modifications

Except for adaptations of tables and figures and for translations of content as described below, material may not be modified. Translations of figures, tables, and excerpts are permitted as transactions within **RightsLink**. Translations of full-text materials must be approved by ASN before the Customer can proceed with the **RightsLink** transaction. Full-text translation requests processed through CCC will be considered invalid unless the

Customer enters the ASN-issued approval code for the translation. A copy of all full-text translations must be sent to ASN.

7) Reservation of Rights

ASN reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions, and (iii) CCC's Billing and Payment terms and conditions.

8) License Contingent on Payment

While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by ASN or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9) Copyright Notice: Disclaimer

You must give complete credit to the original source in connection with any reproduction or adaptation of the licensed material: 'Am J Clin Nutr (year;volume:page range), American Society for Nutrition,' or 'J Nutr (year;volume:page range), American Society for Nutrition;.'Adv Nutr (year;volume:page range), American Society for Nutrition.'For translations: Figures or tables cannot be used in advertisements.

Full-text translations of articles published in *The Journal of Nutrition* must include the following disclaimer: *Translated from the original into LANGUAGE by TRANSLATOR'S NAME. The translator assumes responsibility for the accuracy of the translation. The American Society for Nutrition is not responsible for translation errors. Readers are encouraged to access the original publication at http://jn.nutrition.org.*For photocopies and in-house reprints:

The article will be reprinted in its entirety without change.

No material will be attached to the reprints, and the reprint will be used for educational purposes only, not to promote, sell, or in any way endorse a product.

The citation appearing on the first page of the article must appear in the reprints.

As a courtesy, please notify the author of your intended reuse of his or her content.

All reprints must have reprinted on each:*The American Society for Nutrition, Inc., does not endorse any commercial enterprise.*

10) Warranties

ASN makes no representations or warranties with respect to the licensed material.

11) Indemnity

You hereby indemnify and agree to hold harmless ASN and CCC, and their respective officers, directors, employees, and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

12) No Transfer of License

This license is personal to you, but may be assigned or transferred by you to a business associate (or to your employer) if you give prompt written notice of the assignment or transfer to ASN. No such assignment or transfer shall relieve you of the obligation to pay the designated license fee on a timely basis (although payment by the identified assignee can fulfill your obligation).

13) No Amendment Except in Writing

This license may not be amended except in a writing signed by both parties (or, in the case of ASN, by CCC on ASN’s behalf).

14) Objection to Contrary Terms
ASN hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and ASN (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.
15) Jurisdiction
This license Agreement contains the entire understanding of the parties with respect to the licensed content and can be modified only by a signed, written agreement. This Agreement shall be construed in accordance with the laws of the State of Maryland and the US copyright laws. If any term of this Agreement shall be found invalid by any court of competent jurisdiction, such provision shall be enforced to the fullest extent that it is valid and enforceable under applicable law, and all other provisions of this Agreement shall remain in full force and effect. For state and local governments, the Terms and Conditions do not apply when in conflict with existing statues.
16) Other
There are additional terms and conditions, established by Copyright Clearance Center, Inc. ("CCC") as the administrator of this licensing service that relate to billing and payment for licenses provided through this service. Those terms and conditions apply to each transaction as if they were restated here. As a user of this service, you agreed to those terms and conditions at the time that you established your account, and you may see them again at any time at http://myaccount.copyright.com.

Terms and Conditions for Content Services

Subject to these terms of use, any terms set forth on the particular order, and payment of the applicable fee, you may make the following uses of the ordered materials:

Content Rental: You may access and view a single electronic copy of the materials ordered for the time period designated at the time the order is placed. Access to the materials will be provided through a dedicated content viewer or other portal, and access will be discontinued upon expiration of the designated time period. An order for Content Rental does not include any rights to print, download, save, create additional copies, to distribute or to reuse in any way the full text or parts of the materials.

Content Purchase: You may access and download a single electronic copy of the materials ordered. Copies will be provided by email or by such other means as publisher may make available from time to time. An order for Content Purchase does not include any rights to create additional copies or to distribute copies of the materials.

The materials may be accessed and used only by the person who placed the Order or the person on whose behalf the order was placed and only in accordance with the terms included in the particular order.
Special Terms: <%=specialTerms%>
v 2.7

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.