Research article

The influence of tumor size and environment on gene expression in commonly used human tumor lines

Michael A Gieseg*, Michael Z Man, Nicholas A Gorski, Steven J Madore, Eric P Kaldjian and Wilbur R Leopold

Address: Pfizer Global Research and Development, 2800 Plymouth Rd, Ann Arbor, Michigan, 48105, USA
Email: Michael A Gieseg* - michael.gieseg@pfizer.com; Michael Z Man - michael.man@pfizer.com; Nicholas A Gorski - ngorski@umich.edu; Steven J Madore - steven.madore@pfizer.com; Eric P Kaldjian - eric.kaldjian@umich.edu; Wilbur R Leopold - dick@molecularimaging.com
* Corresponding author

Abstract

Background: The expression profiles of solid tumor models in rodents have been only minimally studied despite their extensive use to develop anticancer agents. We have applied RNA expression profiling using Affymetrix U95A GeneChips to address fundamental biological questions about human tumor lines.

Methods: To determine whether gene expression changed significantly as a tumor increased in size, we analyzed samples from two human colon carcinoma lines (Colo205 and HCT-116) at three different sizes (200 mg, 500 mg and 1000 mg). To investigate whether gene expression was influenced by the strain of mouse, tumor samples isolated from C.B-17 SCID and Nu/Nu mice were also compared. Finally, the gene expression differences between tissue culture and in vivo samples were investigated by comparing profiles from lines grown in both environments.

Results: Multidimensional scaling and analysis of variance demonstrated that the tumor lines were dramatically different from each other and that gene expression remained constant as the tumors increased in size. Statistical analysis revealed that 63 genes were differentially expressed due to the strain of mouse the tumor was grown in but the function of the encoded proteins did not link to any distinct biological pathways. Hierarchical clustering of tissue culture and xenograft samples demonstrated that for each individual tumor line, the in vivo and in vitro profiles were more similar to each other than any other profile. We identified 36 genes with a pattern of high expression in xenograft samples that encoded proteins involved in extracellular matrix, cell surface receptors and transcription factors. An additional 17 genes were identified with a pattern of high expression in tissue culture samples and encoded proteins involved in cell division, cell cycle and RNA production.

Conclusions: The environment a tumor line is grown in can have a significant effect on gene expression but tumor size has little or no effect for subcutaneously grown solid tumors. Furthermore, an individual tumor line has an RNA expression pattern that clearly defines it from other lines even when grown in different environments. This could be used as a quality control tool for preclinical oncology studies.

Published: 15 July 2004

BMC Cancer 2004, 4:35 doi:10.1186/1471-2407-4-35

Received: 05 March 2004
Accepted: 15 July 2004

This article is available from: http://www.biomedcentral.com/1471-2407/4/35

© 2004 Gieseg et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.
Background

Preclinical animal models of human tumors represent a major tool for the selection and development of effective anticancer agents. There is a considerable number of well characterized human cancer cell lines, many of which can be grown as solid tumors (either subcutaneously or orthotopically) in immunodeficient mice. The ease of use and low cost of these models make them desirable for screening in vivo activity of anticancer compounds as compared to induced or transgenic rodent tumor models. Rodent models also allow pharmacodynamic and pharmacokinetic parameters to be directly measured and related to antitumor efficacy. Commonly used human cancer cell lines, such as the panel of 60 lines (NCI60) used by the Developmental Therapeutics Program at the National Cancer Institute, have been extensively studied at the molecular level in vitro and to a lesser extent in vivo. Recent advances in genomic technology allow molecular characterization of these models to an extent never before possible using RNA expression profiling, comparative genomic hybridization, proteomic and metabonomic profiling. Ross et al. [1] looked at the in vitro gene expression in the NCI60 panel. The gene expression pattern for many lines indicated a relationship with the tumor tissue of origin and also correlated with doubling time and drug metabolism. Virtanen et al. [2] have analyzed the expression patterns of 85 lung tumor samples from both clinical samples and established tumor lines. The fresh tumors clustered according to pathological subtype with many of the cell lines also clustering within the same groups. Pedersen et al. [3] performed a similar study with human small cell lung cancer cell lines and compared them to ressected tissue samples. Dan et al. [4] recently performed RNA expression profiling on 39 human cancer cell lines and related the gene expression patterns to chemosensitivity of 55 anticancer agents. Zembutsu et al. [5] performed a similar study but profiled 85 human cancer xenografts.

What was apparent in all these studies was that each tumor model was distinctly different and could be distinguished from related models using routine data analysis methods. Furthermore, it appears that it is possible to identify models that are no longer true to their origins. Ross et al. [1] identified MDA-MB-435 as a possible melanoma derived line, which is at odds with its supposed origins of a metastatic breast carcinoma [6]. However, such identification also raises the possibility that the sample that Ross et al. obtained was not the true "MDA-MB-435." Clearly there is the opportunity to use profiling technology as an advanced quality control method that could not only identify mislabeled tumor lines but possibly genetic drift.

While these reports have tremendous value they do not address some basic questions about human tumor models that could impact the design of in vivo drug development studies. For example, most, if not all, xenografts demonstrate Gompertzian growth kinetics with continuously increasing doubling times as they grow larger [7]. Despite this obvious change in biology, it is unknown whether xenografts alter their expression as they increase in size or whether expression differences exist for the same tumor line grown both as a solid tumor in a mouse or in tissue culture. To investigate these questions we grew human colon tumors (HCT-116 and Colo205) in Nu/Nu mice and harvested tumors at three sizes (200 mg, 500 mg, 1000 mg). Isolated RNA was analyzed on Affymetrix GeneChips. The effect of mouse strain on tumor gene expression was also investigated by comparing tumors at 500 mg grown in Nu/Nu and C.B-17 SCID mouse strains. Lastly we profiled additional tumor models to investigate the changes in gene expression that occur when a tumor line is grown in vivo or in vitro.

Methods

Xenograft and tissue culture methods

Cell lines were grown in DMEM/F12 supplemented with 10% fetal bovine serum (Invitrogen, Carlsbad, California). They were passaged in 75-cm² tissue culture flasks in an atmosphere of 5% CO₂ in air and were subcultured weekly, using 0.05% trypsin EDTA (Invitrogen, Carlsbad, California). RNA from tissue culture samples was harvested at mid-log phase. Human tumor xenografts were grown in either Nu/Nu mice or C.B-17 SCID female mice obtained from Charles River Laboratories (Wilmington, Massachusetts) between four and five weeks of age. Mice were housed five to a cage in animal rooms maintained at between 21–25°C with a 12 h alternating dark/light cycle. All animal studies were conducted under Veterinary Use Protocols approved by the Institutional Animal Care and Use Committee. Tumors were maintained by serial passage of 30 mg tumor fragments between animals, implanted subcutaneously into the right axillary region using a trocar needle aseptically. Tumors were passed when the primary had reached between 500 and 1000 mg and were never passed more than ten times. Tumor growth was followed by caliper measurements of perpendicular measures of the tumor. The weight in mg was estimated by the formula: tumor weight = a(b²)/2, where a and b are the tumor length and width respectively in mm. Tumor tissue was harvested immediately following animal sacrifice by excising the tumor and powdering it in a liquid nitrogen cooled crucible and pestle. Tumor powder was stored at -80°C until RNA isolation.

Affymetrix GeneChip

RNA was extracted using TRIZOL® reagent (Invitrogen, Carlsbad, California) according to the manufacturer's protocol. RNA integrity was monitored using denaturing agarose gel electrophoresis in 1X MOPS. Biotinylated target...
RNA was prepared from 15 μg of total RNA using the Affymetrix protocol. Briefly, double-stranded cDNA was prepared from the RNA template using a modified oligo-dT primer containing a 5’ T7 RNA polymerase promoter sequence and the Superscript Choice System for cDNA Synthesis (Invitrogen, Carlsbad, California). Following phenol-chloroform extraction and ethanol precipitation, one-half of the cDNA reaction (0.5 – 1.0 μg) was used as the template in an in vitro transcription reaction containing T7 RNA polymerase, a mixture of unlabeled ATP, CTP, GTP, and UTP, and biotin-11-CTP and biotin-16-UTP (BioArray High Yield Kit, ENZO, Farmingdale, New York). The resulting biotinylated-cRNA "target" was purified on an affinity resin (RNeasy, Qiagen, Valencia, California) and quantified using the convention that 1 O.D. 260 nm corresponds to 40 μg/mL of RNA. Typical yields ranged from 50 – 100 μg with transcript sizes between 3.0 to 0.25 kilonucleotides as determined by denaturing gel electrophoresis. Fifteen micrograms of biotinylated cRNA was randomly fragmented to an average size of 50 nucleotides by incubating at 94°C for 35 minutes in 40 mM TRIS-acetate, pH 8.1, 100 mM potassium acetate, and 30 mM magnesium acetate. The fragmented cRNA was hybridized in a solution containing 100 mM MES, 1 M [Na+], 20 mM EDTA, 0.01% TWEEN 20, 50 pM of Control Oligonucleotide B2, 0.1 mg/mL of sonicated herring sperm DNA, and 0.5 mg/mL BSA for 16 hours at 45°C on either the Human U95A or the Mouse U74A Affymetrix GeneChips® (Affymetrix, Santa Clara, California). Each hybridization included a mixture of four bacterial biotinylated-RNA transcripts (BioB, BioC, BioD, and cre) spiked at 1.5, 5, 25, and 100 pM, respectively. The hybridization reactions were processed and scanned according to the standard Affymetrix protocols.

Data analysis

All arrays were global scaled to a target intensity value of 600 using the standard Affymetrix protocol. Calculation of the scaling factor, background, noise and percent present, was performed according to Affymetrix protocols using the Data Mining Tool (Affymetrix Santa Clara, California). All resulting data sets were filtered using the absolute call metric (present or absent) using Microsoft Access (Microsoft Corporation, Redmond, Washington). Genes selected had expression levels classified as present at least once in the samples selected for the particular analysis.

To determine the relationship between tumor samples harvested at different sizes the filtered RNA profiling data was analyzed with classic multidimensional scaling (MDS), implemented in R [8,9]. MDS is an unsupervised learning technique that attempts to preserve the relationship between points from high dimensional space at lower dimensional spaces. The program R is a free integrated suite of software facilities for data manipulation, calculation and graphical display http://www.r-project.org.

Analysis of variance (ANOVA) was used to test the effects of tumor size and tumor line, using the following model: Expression of gene, ~ tumor.size + tumor.line + tumor.line*tumor.size. To test the effect due to mouse background, we used the following model: Expression of gene, ~ mouse background. RNA profiling data filtered on the absolute call metric was used for this analysis and ANOVA was implemented in R. ANOVA is a statistical linear modeling procedure that partitions the total variance into parts corresponding to various sources in the model [10,11]. It has been used previously in microarray data analysis [12-15].

To rigorously select genes with expression differences between samples, ANOVA p-values were adjusted using multiple comparison procedures. Multiple comparison procedures are tools to adjust p-values that might be inflated as a result of performing multiple hypothesis tests. The Benjamini and Hochberg procedure controls the false discovery rate, which is the expected fraction of false discoveries in all rejected hypothesis [16]. This procedure is less stringent than methods controlling the family wise error rate (e.g. the Bonferroni correction); hence it is more powerful.

Hierarchical clustering was performed in GeneSpring 5.0 (Silicon Genetics, Redwood City, California). The distinction calculation from Spotfire DecisionSite 6.2 (Spotfire Inc. Somerville, Massachusetts) was used to select genes differentially expressed in xenograft samples or tissue culture samples. All data from the tissue culture samples that had an in vivo pair (8 samples) were selected into one group and all data from the xenograft samples (8 samples) were selected into a second group. Genes were prefILTERED using the absolute call metric by selecting genes that were present at least once in the selected samples. A distinction value score and p-value was calculated for each gene. The score (≥1) and p-value (≤0.001) was then used to select genes that were differentially expressed between xenograft samples and tissue culture samples. To functionally classify gene lists, web resources such as NCBI (National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/) were searched and the data compiled. Further searching for gene associations in PubMed (NCBI) was also performed.

Results

Variation in tumor xenograft gene expression due to size

We focused on two human colon carcinoma xenografts (HCT-116 and Colo205) to investigate the effects of tumor size and mouse strain on gene expression. Samples were harvested in quadruplicate at three different tumor sizes (200 mg, 500 mg and 1000 mg) for both tumor
models grown in Nu/Nu mice (except for the 500 mg sample of Colo205 where five samples were harvested). These sizes were selected as they represent the range at which sensitivity to anticancer agents are traditionally tested and because most models approximate log-linear growth at these sizes. RNA expression profiling data was obtained from Affymetrix U95A GeneChips containing approximately 12600 genes. Genes present (above background) once or more across all samples were selected for further analysis (approximately 7600 genes).

Initial analysis of the expression data with multi-dimensional scaling (MDS) showed that samples from the same tumor line clustered together and that there was clear separation between samples from HCT-116 and Colo205 (Figure 1). Compared to the profound effect due to tumor

**Figure 1**

Multidimensional scaling plot of Colo205 and HCT-116 samples. Multidimensional scaling plot showing the relatedness of individual samples from Colo205 and HCT-116 to each other in 2D space. The color indicates the size of the tumor sample when harvested: red, 200 mg; blue, 500 mg; black, 1000 mg. The size of text indicates the mouse strain the tumor was grown in: small, Nu/Nu; large, C.B-17 SCID. Samples that are more related to each other are closer together. For Colo205 five samples were analyzed.
line, there was no clear separation among samples of different sizes in the MDS plot, suggesting that there was little alteration in gene expression due to differences in tumor size (tumor size effect).

The result from MDS was further confirmed by analysis of variance (ANOVA). Using ANOVA we modeled the effects of tumor line and tumor size on gene expression. Since in the ANOVA we conducted approximately 7000 statistical tests (on the selected genes), with a p-value cutoff of 0.01, we would expect approximately 70 genes (1% of 7000) scored as significantly changed due to chance alone. Indeed, the observed number of significantly changed genes (p \leq 0.01) due to tumor size effect was 154, approximately twice what was predicted by chance alone (Figure 2). In contrast, the number of significantly changed genes due to tumor line effect (4731 genes p-value < 0.01) was far greater than chance alone, indicating the two tumor lines were extremely different as suggested by MDS. The distribution of p-values confirmed the profound effect of tumor line on gene expression with genes affected at all expression levels (Figure 3).

To rigorously identify genes that may suggest functionally significant changes as a tumor increases in size from 200 mg to 1000 mg the ANOVA p-values were adjusted using multiple comparison procedures. Following this analysis none of the previously identified 154 genes had p-values < 0.05. This indicated that there was no change in gene expression as a human tumor xenograft increased in size from 200 mg to 1000 mg. This was true for both the HCT-116 and the Colo205 tumor lines.

Variation in tumor xenograft gene expression due to mouse strain

Additional RNA samples were prepared from 500 mg HCT-116 tumors grown in C.B-17 SCID mice and hybridized to Affymetrix U95A GeneChips. The expression data were compared to data from the same line grown in Nu/Nu mice harvested at 500 mg. The MDS analysis showed there was no distinct separation or grouping for samples from the two different mouse strains (Figure 1). This suggests that mouse strain played only a small role in altering the expression profiles of HCT-116 tumor samples. However, ANOVA did reveal that considerably more genes

Figure 2

Expected verses observed number of significantly changed genes. The graph shows the number of genes (y-axis) that fall into specific categories (x-axis) based on the ANOVA calculated p-values. The categories are as follows: Predicted, number of genes expected by chance alone; Tumor line effect, number of genes that fall within the specified ranges due to differences in the tumor line; Tumor size effect, number of genes that fall within the specified ranges due to changes in the tumor size; Mouse strain effect, number of genes that fall within the specified range due to changes in the mouse strain the tumors were grown in.
were altered in their expression due to the mouse background (493 genes) compared to chance alone (p < 0.01).

To identify genes that may suggest biologically significant differences due to the mouse background effect, the ANOVA p-values were adjusted using multiple comparison procedures. Using the Benjamini and Hochberg procedure [16] 63 genes were found to have a p-value < 0.05 with 32 genes increased in the C.B17-SCID strain and 31 increased in the Nu/Nu strain (Table 1 and 2). Functional classification of these genes using a gene ontology approach did not identify functions that could be linked to known biological pathways. Therefore a biological understanding of the changes in gene expression due to mouse strain remains elusive.

Figure 3
P-value distribution/volcano plot of line and size effect. Scatter plot with mean expression level (log 10) on the x-axis and – p-value (log 10) on y-axis. Each gene is plotted twice, once for the p-value resulting from the tumor-line effect ANOVA (black) and once for the p-value resulting from the tumor-size effect ANOVA (red). The blue line represents a p-value of 0.01. Genes with a lower p-value (more significant) have a higher – p-value log10. There are far more genes with significant p-values due to differences in the tumor lines than due to changes in tumor size.
Comparison of lines grown in both tissue culture and solid tumor xenografts

To explore the relatedness of different tumor lines to each other and to investigate the gene expression differences between growth in tissue culture and growth as a subcutaneous solid tumor, RNA samples were prepared from the 13 human tumor lines grown in tissue culture to mid-log phase (Table 3). An additional eight RNA samples were obtained of the same lines grown as xenografts in Nu/Nu mice. The gene expression profiling data resulting from hybridizing the 21 samples to Affymetrix U95A GeneChips was filtered for genes classified as present at least once across all samples.

Hierarchical clustering analysis using Spearman rank for samples and Pearson correlation for genes was used to determine the relationship between 13 different human tumor lines (Figure 4). The clustering analysis revealed that for each individual tumor line the xenograft and tissue culture profiles were more similar to each other than any other profile (with one exception). That is, tumor lines clustered together based on their genotypes rather than their growth conditions, suggesting “nature” (genotype) was more influential than “nurture” (growth conditions). This result was confirmed with MDS and principle component analysis (data not shown). The one exception from this pattern was the xenograft sample of ZR-75-1,

### Table 1: Genes with high expression in HCT-116 tumors grown in C.B-17 SCID mice

| Function                             | Affymetrix Probe Set ID | ANOVA p-value | BH p-value | Log2 FC | Gene Symbol | LocusLink ID | Sequence Probe Derived From | Gene Title                                                                 |
|--------------------------------------|-------------------------|----------------|------------|---------|-------------|--------------|---------------------------|---------------------------------------------------------------------------|
| Translation/RNA binding/RNA splicing | 34733_at                | 0.00029        | 0.046      | 0.5     | SF3A1       | 10291        | X85237                    | splicing factor 3a, subunit 1, 120 kDa                                   |
|                                      | 34829_at                | 0.00020        | 0.046      | 0.4     | DKCI        | 1736         | U59151                    | dyskeratosis congenita 1, dyskerin                                       |
|                                      | 35174.1_at              | 0.00020        | 0.046      | 0.4     | EEF1A2      | 1917         | X70940                    | eukaryotic translation elongation factor 1 alpha 2                       |
|                                      | 37462.1_at              | 0.00013        | 0.042      | 0.3     | SF3A2       | 8175         | L21990                    | splicing factor 3a, subunit 2, 66 kDa                                   |
|                                      | 39047_at                | 0.00007        | 0.030      | 0.5     | SART3       | 9733         | AB020880                  | squamous cell carcinoma antigen recognised by T cells 3                  |
|                                      | 40490_at                | 0.00041        | 0.049      | 0.7     | DDX21       | 9188         | U41387                    | DEAD/H box polypeptide 21                                               |
| Signal transduction                  | 33887_at                | 0.00029        | 0.046      | 0.5     | HGS         | 9146         | D84064                    | hepatocyte growth factor-regulated tyrosine kinase substrate             |
|                                      | 38019_at                | 0.00002        | 0.028      | 1.0     | CSNK1E      | 1454         | L37043                    | casein kinase 1, epsilon                                                 |
|                                      | 38779_r_at              | 0.00039        | 0.048      | 0.8     | HDGF        | 3068         | D16431                    | hepatoma-derived growth factor                                            |
|                                      | 40864_at                | 0.00029        | 0.046      | 0.4     | RAC1        | 5879         | D25274                    | ras-related C3 botulinum toxin substrate                                 |
| Small Molecule transport             | 32186_at                | 0.00023        | 0.046      | 0.7     | SLC7A5      | 8140         | M80244                    | solute carrier family 7, member 5                                       |
|                                      | 36557_at                | 0.00003        | 0.028      | 0.9     | CACNB1      | 782          | M99230                    | calcium channel, voltage-dependent, beta 1 subunit                       |
|                                      | 38029_at                | 0.00031        | 0.046      | 0.4     | SLC3A2      | 6520         | J02939                    | solute carrier family, member 2                                         |
| Protein Complex Assembly             | 31842_at                | 0.00044        | 0.050      | 0.5     | BCS1L       | 617          | AF038195                  | BCS1I-like (yeast)                                                       |
|                                      | 32575_at                | 0.00019        | 0.046      | 0.4     | NAP1L4      | 4676         | U77456                    | nucleosome assembly protein 1-like 4                                     |
| Protein transport and folding        | 36945_at                | 0.00036        | 0.047      | 0.5     | C12orf8     | 10961        | X94910                    | chromosome 12 open reading frame 8                                     |
|                                      | 40756_at                | 0.00017        | 0.046      | 0.4     | NPM3        | 10360        | AF081280                  | nucleophosmin/nucleoplasmin, 3                                           |
| Unknown or other functions           | 31670_s_at              | 0.00023        | 0.046      | 0.3     | SRFP72      | 6731         | U81554                    | signal recognition particle 72 kDa                                      |
|                                      | 34279_at                | 0.00005        | 0.030      | 2.0     | FLJ20719    | 55672        | AL050141                  | hypothetical protein FLJ20719                                           |
|                                      | 36192_at                | 0.00003        | 0.028      | 0.7     | KIAA019     | 9805         | DB3777                    | KIAAA0193 gene product                                                  |
|                                      | 36209_at                | 0.00003        | 0.028      | 0.8     | BRD2        | 6046         | S78771                    | bromodomain containing 2                                                 |
|                                      | 36210_g_at              | 0.00035        | 0.047      | 1.0     | BRD2        | 6046         | S78771                    | bromodomain containing 2                                                 |
|                                      | 36668_at                | 0.00035        | 0.047      | 0.4     | DIA1        | 1727         | M28713                    | diaphorase (NADH)                                                        |
|                                      | 37907_at                | 0.00006        | 0.030      | 0.4     | FBA         | 8263         | M34677                    | coagulation factor VIII-associated                                      |
|                                      | 38885_at                | 0.00000        | 0.007      | 1.0     | DNA2L       | 1763         | D42046                    | DNA2 DNA replication helicase 2-like                                    |
|                                      | 39436_at                | 0.00010        | 0.039      | 1.5     | BNIP3L      | 665          | AF079221                  | BCL2/adenovirus E1B 19 kDa interacting protein 3-like                    |
|                                      | 39758_f_at              | 0.00030        | 0.046      | 0.4     | LAMP1       | 3916         | J04182                    | lysosomal-associated membrane protein 1                                 |
|                                      | 39832_at                | 0.00007        | 0.030      | 0.5     | ARS11       | 51393        | AL096723                  | arsenate resistance protein ARS11                                       |
|                                      | 40589_at                | 0.00041        | 0.049      | 0.9     | SNTB2       | 6645         | U40572                    | syntrophin, beta 2                                                       |
|                                      | 40604_at                | 0.00012        | 0.040      | 1.1     | DYRK2       | 8445         | Y13493                    | dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 2         |
|                                      | 41062_at                | 0.00028        | 0.046      | 0.7     | NSP1C1      | 84759        | AA037278                  | likely ortholog of mouse nervous system polycyst 1                      |
|                                      | 982_at                  | 0.00011        | 0.039      | 0.4     | MCH5        | 4174         | X74795                    | MCM5 minichromosome maintenance deficient 5, cell division cycle 46     |

1p-value resulting from Benjamini and Hochberg procedure 2 log2 fold change calculated by dividing expression level from C.B-17 SCID sample by expression level of Nu/Nu sample
which did not cluster on the same node as the ZR-75-1 sample grown in tissue culture. However, subsequently we have found that our xenograft ZR-75-1 line has aberrant biological characteristics that are inconsistent with the known phenotypes of this line.

Although each tissue culture-xenograft sample did cluster together there were many gene changes between the paired samples (Table 4). When fold change values were calculated for each tissue culture xenograft pair and the number of genes with ≥2 fold-change tabulated there was a median of 425.5 genes increased in xenografts and 387 decreased. Again the ZR-75-1 sample was aberrant with many more gene changes than the other lines.

Another striking feature revealed by clustering analysis was that none of the tumor lines were particularly similar to each other. The tumor lines did not necessarily cluster based upon the tissue of origin from which each line was thought to be derived. However, a group of breast carcinoma lines that did cluster together (BT474, ZR-75-1, MDA-MB-453, MCF-7, SKBr-3 and MDA-MB-468), although MDA-MB-435 and MDA-MB-231 did not cluster with this group. The two lung carcinoma lines (H2009 and H125) did cluster closely, but MDA-MB-231 and HCT-116 (colon carcinoma) also clustered with this group. MDA-MB-435, MCF-7, SKBr-3 and MDA-MB-468),

The first part of our study was designed to address whether gene expression patterns in xenografts varied with tumor size. No significant changes (p-value < 0.05) were found in any genes after statistical analysis of the RNA profiling data. This was a somewhat surprising result as there is evidence to suggest that tumors show size-dependent biological variation in both a clinical setting and as implanted solid tumors in rodents. It is generally assumed clinically that the larger the tumor mass the lower the likelihood of curing the patient, regardless of the treatment [17,18]. Pathological and genetic heterogeneity increase with increasing tumor cell number and contribute to the emergence of clinical drug resistance [19]. As tumors grow larger their vascular surface decreases, intercapillary distance increases, interstitial pressure increases and necrotic foci develop [20]. Tumor doubling times and cell loss also increase with increasing tumor size [17]. Implanted animal tumor models most likely represent clinical end-stage disease and may not necessarily recapitulate clinical behavior seen during tumor development in humans. There are few reports on the molecular and biochemical changes that occur as implanted solid tumors increase in size. Massaad et al. [21] found that some drug metabolizing enzyme systems, including glutathione-S-transferases (GST) did change both activity and expression with increasing tumor size in the mouse colon adenocarcinoma Co38.

Preclinical tumor models show variability in response to treatment with a given anti-tumor agent despite being derived from a single tumor and being implanted into inbred mice [22]. This has been attributed to heterogeneity of the host or the tumor cell populations [23,24] along with tumor-to-tumor variations in perfusion and drug distribution [25]. We found there was no significant difference in gene expression between tumors of either the same size or different size as assessed in four replicate animals. If differences in tumor population, host metabolism or perfusion exist between animals implanted with the same tumors, our data suggests they are not detectable using gene expression analysis of the whole tumor.

Our results suggest that once tumor xenografts have grown to 200 mg, physiological influences on tumor transcription (such as the interaction with host stroma, nutrient supply and oxygenation) have reached steady state and do not change appreciably as the tumor grows to 1000 mg. It is also conceivable that tumors smaller and larger than the sizes investigated in this study may show gene expression changes. When a tumor is first implanted the hypoxic environment is likely to drive the expression of pro-angiogenic cytokines that will recruit new vessels to the tumor. As a tumor becomes very large (>2 g) angiogenic factors may again become highly expressed due to the inability for the existing vasculature to adequately support the large tumor mass. We have not excluded the possibility that there may be changes in protein activities.
Figure 4
Hierarchical clustering analysis of 13 tumor lines grown in different environments. Hierarchical clustering analysis showing the structure within the data of the 13 tissue culture samples (suffix – TC) and 8 xenograft samples (suffix – X). Samples are displayed vertically, genes are displayed horizontally. A dendrogram of relatedness of the samples is at the top in green. For any two samples, the vertical distance from the sample roots to the first node joining them is a measure of their similarity; the shorter the distance the more similar. The color in each cell of the table represents the median adjusted expression value of each gene. The color scale used to represent the expression ratios is shown on the right, with yellow indicating increased expression relative to the median and blue decreased.
Table 2: Genes with high expression in HCT-116 tumors grown in Nu/Nu mice

| Function                        | Affymetrix Probe Set ID | ANOVA p-value | BH p-value | Log2 FC | Gene Symbol | LocusLink ID | Sequence Probe Derived From | Gene Title                                                                 |
|---------------------------------|-------------------------|---------------|------------|---------|-------------|--------------|-----------------------------|-----------------------------------------------------------------------------|
| Translation/RNA binding/RNA splicing | 32276_at               | 0.00036       | 0.047      | -0.5    | RPL32       | 6161         | X03342                      | ribosomal protein L32                                                       |
|                                 | 32432_f_at             | 0.00032       | 0.046      | -0.6    | RPL15       | 6138         | L25899                      | ribosomal protein L15                                                      |
|                                 | 33614_at               | 0.00028       | 0.046      | -0.5    | RPL18A      | 6142         | X80822                      | ribosomal protein L18A                                                     |
|                                 | 33660_at               | 0.00039       | 0.048      | -0.5    | RPL5        | 6125         | U14966                      | ribosomal protein L5                                                       |
|                                 | 34117_g_at             | 0.00029       | 0.046      | -0.5    | RPS15A      | 620          | W52024                      | ribosomal protein S15a                                                    |
|                                 | 34345_at               | 0.00043       | 0.049      | -0.1    | C20orf14    | 24148        | AF026031                    | chromosome 20 open reading frame 14                                       |
|                                 | 398_at                 | 0.00029       | 0.046      | -0.6    | DDX18       | 8886         | X98743                      | DEAD/H box polypeptide 18 (Myc-regulated)                                  |
|                                 | 40887_g_at             | 0.00017       | 0.046      | -0.5    | EEF1A1      | 1915         | L41498                      | eukaryotic translation elongation factor 1 alpha 1                        |
| Signal Transduction             | 35772_at               | 0.00029       | 0.046      | -0.9    | ARHGEF12    | 23365        | AB002380                    | Rho guanine nucleotide exchange factor (GEF) 12                            |
|                                 | 36091_at               | 0.00038       | 0.048      | -0.6    | SCAP2       | 8935         | AF051323                    | src family associated phosphoprotein 2                                     |
|                                 | 40784_at               | 0.00003       | 0.028      | -0.7    | PPP2RSC     | 5527         | Z69030                      | protein phosphatase 2, regulatory subunit B (B56), gamma isoform           |
|                                 | 769_s_at               | 0.00032       | 0.046      | -0.3    | ANXA2       | 302          | D00017                      | annexin A2                                                                 |
|                                 | 777_at                 | 0.00001       | 0.039      | -0.6    | GDI2        | 2665         | D13988                      | GDP dissociation inhibitor 2                                               |
| Unknown or Other Functions      | 1160_at                | 0.00024       | 0.046      | -0.3    | CYC1        | 1537         | J04444                      | cytochrome c-1                                                            |
|                                 | 1226_at                | 0.00010       | 0.039      | -0.5    | ADAM17      | 6868         | U69611                      | a disintegrin and metalloproteinas domain 17                               |
|                                 | 31531_g_at             | 0.00007       | 0.030      | -0.5    | ACACB       | 32           | U89344                      | acetyl-Coenzyme A carboxylase beta                                         |
|                                 | 31684_at               | 0.00004       | 0.030      | -0.8    | ANXA2P1     | 303          | M62896                      | annexin A2 pseudogene 1                                                    |
|                                 | 32518_at               | 0.00014       | 0.042      | -0.6    | ZNF259      | 8882         | AF019767                    | zinc finger protein 259                                                    |
|                                 | 33363_at               | 0.00041       | 0.049      | -0.1    | JTVI        | 7965         | W25934                     | JTV1 gene                                                                  |
|                                 | 33389_at               | 0.00017       | 0.046      | -1.0    | CYPS1       | 1595         | U23942                      | cytochrome P450, 51                                                        |
|                                 | 33558_at               | 0.00012       | 0.039      | -0.5    | TBX5        | 6910         | Y09445                      | T-box 5                                                                    |
|                                 | 33820_g_at             | 0.00000       | 0.007      | -0.7    | HSPA8       | 3312         | X13794                      | heat shock 70 kDa protein 8                                                |
|                                 | 33937_at               | 0.00021       | 0.046      | -0.6    | E0F1       | 4288         | X65550                      | antigen identified by monoclonal antibody Ki-67                           |
|                                 | 34808_at               | 0.00022       | 0.046      | -0.4    | AB023216    | KIAA0999 protein         |                                                                           |
|                                 | 37218_at               | 0.00029       | 0.046      | -0.4    | BTG3        | 10950        | D64110                      | BTG family, member 3                                                      |
|                                 | 38553_g_at             | 0.00025       | 0.046      | -1.3    | BAP29       | 55973        | A998478                     | B-cell receptor-associated protein BAP29                                   |
|                                 | 38589_s_at             | 0.00042       | 0.049      | -0.6    | PTMA        | 5757         | M14630                      | prothymosin, alpha (gene sequence 2B)                                      |
|                                 | 39167_r_at             | 0.00045       | 0.050      | -1.2    | SERPINH1    | 872          | D83174                      | serine (or cysteine) proteinase inhibitor, clade H, member 2              |
|                                 | 418_at                 | 0.00004       | 0.030      | -0.5    | MK167       | 4288         | X65550                      |                                                                           |
|                                 | 893_at                 | 0.00003       | 0.028      | -0.4    | E2-EFF      | 27338        | M91670                      | ubiquitin carrier protein                                                  |
|                                 | 932_s_at               | 0.00031       | 0.046      | -0.8    | ZNF91       | 7644         | L11672                      | zinc finger protein 91                                                     |

1p-value resulting from Benjamini and Hochberg procedure 2 log2 fold change calculated by dividing expression level from C.B-17 SCID sample by expression level of Nu/Nu sample

Table 3: Human tumor lines used in this study

| Tumor        | Tissue of Origin | Xenograft sample | Tissue culture sample |
|--------------|------------------|------------------|-----------------------|
| BT-474       | Breast           | No               | Yes                   |
| MCF-7        | Breast           | No               | Yes                   |
| MDA-MB-231   | Breast           | Yes              | Yes                   |
| MDA-MB-435   | Breast           | Yes              | Yes                   |
| MDA-MB-453   | Breast           | No               | Yes                   |
| MDA-MB-468   | Breast           | Yes              | Yes                   |
| SKBr-3       | Breast           | No               | Yes                   |
| ZR-75-1      | Breast           | Yes              | Yes                   |
| Colo205      | Colon            | Yes              | Yes                   |
| HCT-116      | Colon            | Yes              | Yes                   |
| H125         | Lung             | Yes              | Yes                   |
| H2009        | Lung             | No               | Yes                   |
| A431         | Squamous cell    | Yes              | Yes                   |
Figure 5
Heat-map plot of differential expressed genes in xenograft or tissue culture samples. Heat-map plot showing the genes selected as differentially expressed between the tissue culture and xenograft samples. The samples are displayed vertically, genes are displayed horizontally. Yellow indicates high expression while blue indicates low expression, relative to the median expression for each gene across all samples.
or levels, or that other tumor lines may show changes as the size increases. But since the phenomenon was independently observed in two tumor lines it suggests it may be common to many other lines. The results also imply that tumors harvested within this size range without drug treatment should be very comparable. It remains possible that selective isolation and analysis of tumor regions or subpopulations by microdissection may identify expression differences.

By comparing the same tumor line grown in different mouse strains we identified 63 genes with a significant change in gene expression with biological functions ranging from cellular signaling, RNA processing and translation. Although different genes were differentially expressed in different mouse strains, particular gene functions did not associate with one strain. Therefore, the biological significance of these changes remains unclear. It should be noted that although the changes were statistically significant, in general the magnitude of the change was small. Only 12% of the significant genes displayed a change greater or equal to 2-fold. This raises the question of whether statistical significance equates to biological significance and how relevant thresholds can be determined for analyzing expression profiling data.

In the second part of the study we compared the gene expression patterns of tumor lines grown as xenografts with the same lines grown in tissue culture. We found that each xenograft – tissue culture pair was more similar to one another than any other line. This has been previously observed with RNA expression profiling of small cell lung cancer lines grown as both xenografts and tissue culture [3]. In our data, the one exception to this clustering pattern, ZR-75-1, exhibited aberrant growth characteristics; it had a rapid doubling time (2–3 days) and was not estrogen dependent as has been reported [26,27].

There was a gene expression pattern consistent with growth either in a xenograft or tissue culture environment, but remarkably it consisted of only a small number of genes. We identified 36 genes with high relative expression in xenograft samples and 17 genes with high relative expression in tissue culture samples. Their biological functions were consistent with the differences between a tissue culture sample and a xenograft sample. Genes expressed at increased levels in tissue culture samples were primarily involved in cell division, cell cycle, transcription and translation and were consistent with a greater percentage of cycling cells in the tissue culture samples, which we have previously observed with flow cytometry (data not shown). Many of the genes expressed in xenograft samples encoded for proteins involved in extracellular matrix, cell adhesion, cell surface receptors transcription and translation and suggest increased interactions with an in vivo stromal environment, including the development of a 3-dimensional matrix.

There are several reports in the literature of tumor line expression profiling studies demonstrating that lines derived from a common tissue of origin group together in most cases [1,4,5]. We observed that most of the breast tumor lines clustered together but the other lines generally did not show clustering patterns based upon tissue of origin. However, our study evaluated fewer samples and consisted predominantly of breast cancer tumor lines with only two representatives of non-small cell lung cancer and colon cancer.

Since each tumor line was substantially different to every other line tested it suggests that it would be possible to identify a group of genes that could be used to distinguish

| Tumor line | Number of genes with ≥2 fold increase | Number of genes with ≥2 fold decrease | Total with ≥2 fold change |
|------------|--------------------------------------|--------------------------------------|--------------------------|
| A431       | 542                                  | 650                                  | 1192                     |
| Colo205    | 447                                  | 328                                  | 775                      |
| H125       | 335                                  | 304                                  | 639                      |
| HCT-116    | 351                                  | 393                                  | 744                      |
| MDA-MB-231 | 489                                  | 381                                  | 870                      |
| MDA-MB-435 | 404                                  | 519                                  | 923                      |
| MDA-MB-468 | 381                                  | 204                                  | 585                      |
| ZR-75-1    | 889                                  | 1218                                 | 2107                     |
| Median     | 425.5                                | 387                                  | 822.5                    |
individual lines. Also, it has been our experience that it is possible to identify aberrant samples resulting from labeling error by expression profiling that were missed by other methods. What is less clear is whether genetic drift causes changes in RNA expression that can be identified using RNA profiling technology. Recently a report has been published showing that MCF-7 sublines were substantially different to each other both at the genetic and RNA expression levels [28] suggesting that variation within individual lines can be identified. Given this, it seems reasonable that RNA expression profiling could be used as a comprehensive methodology for identifying aberrant or incorrectly labeled samples. This would provide an additional quality control tool to standardize tumor models and the in vivo testing of therapeutic agents.

Conclusions
Our data suggest that the environment a tumor line is grown in can have a significant effect on gene expression even though tumor size has little or no effect for a subcutaneously grown solid tumor. Furthermore, an individual tumor line has an RNA expression pattern that clearly defines it from other lines even when grown in different environments such as tissue culture or in vivo. Routine RNA expression profiling of a selected set of genes could be used as a quality control tool for preclinical oncology

Table 5: Genes with high expression in xenograft samples

| Function | Affymetrix Probe Set ID | Gene Symbol | LocusLink ID | Sequence Probe Set Derived From | Gene Title |
|----------|-------------------------|-------------|--------------|---------------------------------|------------|
| Extracellular Matrix/Cell Adhesion | 31574_s_at | M14087 | Beta-galactoside-binding lectin, mRNA sequence |
| | 31810_g_at | CNTNI1 | 1272 | Z21488 | contactin I |
| | 32305_at | COL1A2 | 1278 | J03464 | collagen, type I, alpha 2 |
| | 32701_at | ARVC | 421 | U51669 | armadillo repeat gene deletes in velocardiofacial syndrome |
| | 36298_at | PRPH | 5630 | L14565 | peripherin |
| | 38181_at | MMP11 | 4320 | X57766 | matrix metalloproteinase 11 (stromelysin 3) |
| | 38570_at | HLA-DOB | 3112 | X03066 | major histocompatibility complex, class II, DO beta |
| | 38614_s_at | OTG | 8473 | U77413 | O-linked N-acetylglycosamine (GlcNAc) transferase |
| Cell surface receptor | 33950_g_at | CRHR2 | 1395 | AF011406 | corticotropin releasing hormone receptor 2 |
| | 35485_at | GRM4 | 2914 | X80818 | glutamate receptor, metabotropic 4 |
| | 35503_at | HTR1B | 3351 | M81590 | 5-hydroxytryptamine (serotonin) receptor 1B |
| | 36277_at | CD3E | 916 | M23323 | CD3E antigen, epsilon polypeptide (Tf3 complex) |
| | 41190_at | TRIP8 | 3932 | L40411 | tumor necrosis factor receptor superfamily, member 25 |
| Transcription/transcription factor | 34786_at | JMJDI | 55818 | AB018285 | zinc finger protein |
| | 36944_f_at | PLAGL1 | 5325 | U72621 | pleiomorphic adenoma gene-like 1 |
| | 37491_at | TAF1 | 6872 | D90359 | TAF1 RNA polymerase II, TATA box binding protein-associated factor |
| Intracellular transport | 38216_at | TRIPB | 9323 | L40411 | thyroid hormone receptor interactor 8 |
| Translation/RNA binding/RNA splicing | 32228_at | AP2A2 | 161 | AB020706 | adaptor-related protein complex 2, alpha 2 subunit |
| | 33779_at | VAPM1 | 6843 | AF060358 | vesicle-associated membrane protein 1 (synaptobrevin 1) |
| Unknown or other functions | 1155_at | J03069 | hypothetical protein CG003 |
| | 1529_at | 13CDNA73 | 10129 | U50534 | hypothetical protein CG003 |
| | 31525_s_at | KIAA1000 | 22989 | J00153 | KIAA1000 protein |
| | 32355_at | DKFZP564_D166 | 26115 | AL050270 | putative ankyrin-repeat containing protein |
| | 36891_at | OTG | 8473 | U24389 | neuroblastoma-amplified protein |
| | 37979_at | YAPI | 10413 | X80507 | Yes-associated protein 1, 65 kDa |
| | 39499_s_at | PARD3 | 56280 | W25794 | par-3 partitioning defective 3 homolog |
| | 40928_at | WSB1 | 26118 | W26496 | SOCS box-containing WD protein SWIP-1 |
| | 41113_at | KIAA0557 | 26048 | A1971396 | KIAA0557 protein |
| | 41328_s_at | EML2 | 24319 | AL096717 | echinoderm microtubule associated protein like 2 |
| | 40856_at | SERPING1 | 5176 | U29953 | serine (or cysteine) protease inhibitor, clade F, member 1 |
| | 32833_at | CLK1 | 1195 | M59827 | CDC-like kinase 1 |
| | 39451_s_at | IDS | 3423 | AF050145 | iduronate 2-sulfatase (Hunter syndrome) |
| | 36386_at | PDK1 | 5163 | L42450 | pyruvate dehydrogenase kinase, isoenzyme 1 |
Table 6: Genes with high expression in tissue culture samples

| Function                  | Affymetrix ProbeSet ID | Gene Symbol | LocusLink ID | Sequence ProbeSet Derived From | Gene Title                                                                 |
|---------------------------|------------------------|-------------|--------------|-------------------------------|-----------------------------------------------------------------------------|
| Cell division             | 34851_at               | STK6        | 6790         | AF011468                       | serine/threonine kinase 6                                                  |
|                           | 33346_r_at             | TUBG1       | 7283         | M61764                         | tubulin, gamma 1                                                           |
|                           | 35995_at               | ZWINT       | 11130        | AF067656                       | ZW10 interactor                                                            |
|                           | 37171_at               | KNSL5       | 9493         | X67155                         | kinesin-like 5 (mitotic kinesin-like protein 1)                             |
| Cell cycle                | 34736_at               | CCNB1       | 891          | M25753                         | cyclin B                                                                   |
|                           | 37458_at               | CDC45L      | 8318         | AJ223728                       | CDC45 cell division cycle 45-like (S. cerevisiae)                          |
|                           | 38804_at               | CSE1L       | 1434         | AF053641                       | CSE1 chromosome segregation 1-like (yeast)                                  |
| Translation/RNA binding/RNA splicing | 35323_at               | EIF3F9      | 8662         | U78525                         | eukaryotic translation initiation factor 3, subunit 9 eta, 116 kDa          |
|                           | 38679_g_at             | SNRPE       | 6635         | AA733050                       | small nuclear ribonucleoprotein polypeptide E                              |
|                           | 41460_at               | RBM14       | 10432        | AF080561                       | RNA binding motif protein 12                                                |
| Transcription/transcription factor | 39664_at               | CKN1        | 1161         | U28413                         | Cockayne syndrome 1 (classical)                                            |
|                           | 41332_at               | POLR2E      | 5434         | D38251                         | polymerase (RNA) II (DNA directed) polypeptide E, 25 kDa                   |
| Unknown or other functions | 35715_at               | DKFZP564M082| 25906        | AL080071                       | DKFZP564M082 protein                                                       |
|                           | 40758_at               | ICT1        | 3396         | X81788                         | immature colon carcinoma transcript 1                                     |
|                           | 1749_at                | GCDH        | 2639         | AD000092                       | glutaryl-Coenzyme A dehydrogenase                                          |
|                           | 40134_at               | ATP5J2      | 9551         | AF047436                       | ATP synthase, subunit f, isof orm 2                                         |
|                           | 41400_at               | TK1         | 7083         | KG0581                         | thymidine kinase 1, soluble                                                |

studies and would allow the standardization of tumor models used worldwide.

Competing interests
None declared.

Authors' contributions
MAG carried out sample preparation, RNA isolation, data analysis and drafted the manuscript. MZM performed the MDS analysis, ANOVA and statistical analysis. NAG carried out sample preparation, RNA isolation, data processing. EPK conceived part of the study and participated in its design. WRL conceived part of the study and participated in its design.

Acknowledgements
We kindly acknowledge Dr Sarah Fowler for assistance with proofing this manuscript and Dr William Elliott for helpful discussion about the growth characteristics of solid tumor xenografts.

References
1. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, Iyer V, Jeffrey SS, Van de RM, Waltham M, Parmigenshikov A, Lee JC, Lashkari D, Shalon D, Myers TG, Weinstein JN, Botstein D, Brown PO: Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet 2000, 24:227-235.
2. Virtanen C, Ishikawa Y, Honjoh D, Kimura M, Shimane M, Miyoshi T, Nomura H, Jones MH: Integrated classification of lung tumors and cell lines by expression profiling. Proc Natl Acad Sci USA 2002, 99:12357-12362.
3. Pedersen N, Mortensen S, Sorensen SB, Pedersen MW, Rieneck K, Bovin LF, Poulsen HS: Transcriptional gene expression profiling of small cell lung cancer cells. Cancer Res 2003, 63:1943-1953.
4. Dan S, Tsunoda T, Kitahara O, Yamagawa R, Zembutsu H, Katagiri T, Yamazaki K, Nakamura Y, Yamori T: An integrated database of chemosensitivity to 55 anticancer drugs and gene expression profiles of 39 human cancer cell lines. Cancer Res 2002, 62:1139-1147.
5. Zembutsu H, Ohnishi Y, Tsunoda T, Furukawa T, Katagiri T, Ueyama Y, Tamaoki N, Nomura T, Kitahara O, Yamagawa R, Hirata K, Nakamura Y: Genome-wide cDNA microarray screening to correlate gene expression profiles with sensitivity of 85 human cancer xenografts to anticancer drugs. Cancer Res 2002, 62:518-527.
6. Cailleau R, Olive M, Cruciger QV: Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. In vitro 1978, 14:911-915.
7. Jackson RC: From The Growth Kinetics of Large Tumors. In The Theoretical Foundations of Cancer Chemotherapy Introduced by Computer Models Edited by: Jackson RC. San Diego, California: Academic Press Inc; 1992:259-298.
8. Ihaka R, Gentleman RR: A Language for Data Analysis and Graphics. Journal of Computational and Graphical Statistics 1996, 5:299-314.
9. Venables WN, Ripley BD: Modern Applied Statistics with S-PLUS Second edition. Springer; 1997.
10. Fisher RD: Statistical methods for research workers London: Oliver and Boyd; 1950.
11. Fisher LD, van Belle G: Biostatistics: a methodology for the health sciences New York: John Wiley & Sons Inc; 1993.
12. Kerr MK, Martin M, Churchill GA: Analysis of variance for gene expression microarray data. J Comput Biol 2000, 7:819-837.
13. Goomes KR, Highsmith WE, Krogmann TA, Baggery KA, Stivers DN, Abruzzo LV: Identifying and quantifying sources of variation in microarray data using high-density cDNA membrane arrays. J Comput Biol 2002, 9:655-669.
14. Pritchard CC, Hsu L, Delrow J, Nelson PS: Project normal: defining normal variance in mouse gene expression. Proc Natl Acad Sci USA 2001, 98:13266-13271.
15. Pavlidis P, Noble WS: Analysis of strain and regional variation in gene expression in mouse brain. Genome Biol 2001, 2:1-15.
16. Benjamini Y, Hochberg Y: Controlling the false discovery rate; a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 1995, 57:289-300.
17. DeVre VT Jr: The James Ewing lecture. The relationship between tumor mass and resistance to chemotherapy. Implications for surgical adjuvant treatment of cancer. Cancer 1983, 51:1209-1220.
18. Michaelson JS, Silverstein M, Wyatt J, Weber G, Moore R, Halpern E, Kopans DB, Hughes K: Predicting the survival of patients with breast carcinoma using tumor size. Cancer 2002, 95:713-723.
19. Baylin SB: From Clonal selection and heterogeneity of human solid neoplasms. In Design of Models for Testing Cancer Therapeutic Agents Edited by: Fidler IJ, White RJ. New York: Van Nostrand Reinhold; 1981:50-63.
20. Jain RK: Delivery of novel therapeutic agents in tumors: physiological barriers and strategies. J Natl Cancer Inst 1989, 81:570-576.
21. Massaad L, Chabot GG, Toussaint C, Koscielny S, Morizet J, Bissery MC, Gouyette A: Influence of tumor size on the main drug-metabolizing enzyme systems in mouse colon adenocarcinoma Co38. Cancer Chemother Pharmacol 1994, 34:497-502.
22. Schabel FM, Griswold DP, Corbett TH, Laster VR, Lloyd HH, Rose WC: From Variable response of advanced solid tumors of mice to treatment with anticancer drugs. In Design of Models for Testing Cancer Therapeutic Agents Edited by: Fidler IJ, White RJ. New York: Van Nostrand Reinhold; 1981:95-113.
23. Nelson JA, Hokanson JA, Jenkins VK: Role of the host in the variable chemotherapeutic response of advanced Ridgway osteogenic sarcoma. Cancer Chemother Pharmacol 1982, 9:148-155.
24. Trope C: From Different susceptibilities of tumor cell sub-populations to cytotoxic agents. In Design of Models for Testing Cancer Therapeutic Agents Edited by: Fidler IJ, White RJ. New York: Van Nostrand Reinhold; 1981:64-79.
25. Simpson-Herren L, Noker PE, Wagoner SD: Variability of tumor response to chemotherapy. II. Contribution of tumo heterogeneity. Cancer Chemother Pharmacol 1988, 22:131-136.
26. Gutman M, Couillard S, Roy J, Labrie F, Candas B, Labrie C: Comparison of the effects of EM-652 (SCH57068), tamoxifen, toremifene, droloxifene, idoxifene, GW-5638 and raloxifene on the growth of human ZR-75-1 breast tumors in nude mice. Int J Cancer 2002, 99:273-278.
27. Cameron DA, Ritchie AA, Langdon S, Anderson TJ, Miller VR: Tamoxifen induced apoptosis in ZR-75-1 breast cancer xenografts antedates tumor regression. Breast Cancer Res Treat 1997, 45:99-107.
28. Nugoli M, Chuchana P, Vendrell J, Orsetti B, Ursule L, Nguyen C, Birnbaum D, Douzery EJ, Cohen P, Theillet C: Genetic variability in MCF-7 sublines: evidence of rapid genomic and RNA expression profile modifications. BMC Cancer 2003, 3:13.

Pre-publication history
The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2407/4/35/prepub