Genomics Data 4 (2015) 158–161

Contents lists available at ScienceDirect

Genomics Data

journal homepage: http://www.journals.elsevier.com/genomics-data/

Data in Brief

Global gene expression profiling analysis reveals reduction of stemness after B-RAF inhibition in colorectal cancer cell lines

Melanie Boerries [a,b,c,*], Ricarda Herr [d,e], Tilman Brummer [d,f], Hauke Busch [a,b,c]

[a] Systems Biology of the Cellular Microenvironment Group, Institute of Molecular Medicine and Cell Research (IMMZ), Albert-Ludwigs-University (ALU), Freiburg, Germany
[b] German Cancer Consortium (DKTK), Freiburg, Germany
[c] German Cancer Research Center (DKFZ), Heidelberg, Germany
[d] Signal Transduction in Tumour Development and Drug Resistance Group, IMMZ, ALU, Freiburg, Germany
[e] Faculty of Biology, ALU, Freiburg, Germany
[f] Centre for Biological Signalling Studies BIOSS, ALU, Freiburg, Germany

ARTICLE INFO

Article history:
Received 5 April 2015
Accepted 9 April 2015
Available online 28 April 2015

Keywords:
Colon cancer
Gene expression
Differentiation
B-RAF
Bioinformatics

ABSTRACT

Cancer cell differentiation is an important field of discussion in the light of cancer stem cells. In a recent study by Herr et al. (2015) "B-RAF inhibitors induce epithelial differentiation in BRAF-mutant colorectal cancer cells" we described how inhibition of mutant BRAF in colorectal cancer cell lines induces cell re-differentiation that is correlated with the loss of tumor growth *in vitro* and *in vivo*. We used Illumina HumanHT-12 v4 Expression BeadChip to characterize the gain of differentiation of PLX4720-treated 3D cultures of HT29 and Colo-205 cells. Here, we describe the experimental design and statistical analysis that were performed on the data set leading to the above hypothesis. The data are publicly available at the Gene Expression Omnibus (GEO) database under the accession number GSE50791.

Specifications	
Organism/cell line/tissue	*Homo sapiens*/colorectal cancer (CRC) cell lines HT29 and Colo-205
Sex	HT29 cells are derived from the primary tumor of a 44-year-old female CRC patient. Colo-205 is a metastatic cell line derived from the malignant ascites of a 70-year-old male CRC patient.
Sequencer or array type	Illumina HumanHT-12 v4 Expression BeadChip
Data format	Raw and quantile normalized
Experimental factors	I.e. tumor vs. normal, any pretreatment of samples HT29 and Colo-205 cells were grown in three-dimensional (3D) Matrigel culture *in vitro* and treated with DMSO or the B-RAF inhibitor PLX4720.
Experimental features	We performed a time-resolved microarray analysis of PLX4720- *versus* DMSO-treated HT29 and Colo-205 3D tissue cultures to determine differentially expressed transcripts. Cells were seeded into Matrigel, cultivated for 4 days and subsequently treated with DMSO or 3 μM PLX4720. Cells were harvested at different time points, total RNA was isolated and Biotin-labeled cRNA samples were hybridized on Illumina HumanHT-12 v4 Expression BeadChip.
Consent	Not applicable
Sample source location	Not applicable

* Corresponding author at: Systems Biology of the Cellular Microenvironment Group, IMMZ, ALU, Freiburg, Germany.
E-mail address: m.boerrries@dkfz.de (M. Boerries).

Direct link to deposited data

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE50791.

Experimental design, materials and methods

Tissue culture

Three-dimensional (3D) cultures of the colon cancer cell lines Colo-205 [2] (CLS Cell Lines Service GmbH, Heidelberg, Germany) and HT29 [3] (kind gift of Prof. Dr. T. Brabletz, Erlangen, Germany) were set-up as described previously [1,4]. In brief, 4-well chamber slides (BD Biosciences) were coated with a thin layer of ice-cold Matrigel (75 μl) and incubated at 37 °C for at least 30 min to allow solidification of the Matrigel. Subsequently, cells ($3–5 \times 10^3$ cells/well) were resuspended in culture medium containing 2% Matrigel and seeded on top of the solidified Matrigel. The Matrigel-supplemented medium was replaced every 2–3 days.

Inhibitor treatment and cell extraction

HT29 and Colo-205 3D cultures were treated with DMSO or with 3 μM of the B-RAF inhibitor PLX4720 dissolved in DMSO 4 days after seeding. Cells were harvested at three different time points after starting the treatment: at 1, 3 and 8 days for HT29 and at 1,3 and

http://dx.doi.org/10.1016/j.gdata.2015.04.015

10 days for Colo-205 cells. The different growth behaviors of the cell lines required different experimental end time points. The HT29 samples were measured in duplicates. To extract the cells the medium was aspirated and 500 µl recovery solution (BD Biosciences) were added to each well. The cell/gel mixture was scraped into a 15 ml tube using a pipette tip. Following 1 h of incubation on ice the cells were washed twice with 15 ml ice-cold DPBS (centrifugation: 10 min, 1200 rpm, 4 °C) and the cell pellet used for total RNA isolation.

RNA isolation and microarray analysis

RNA was extracted using the Universal RNA Purification Kit (GeneMatrix) from Roboklon. RNA quality and integrity were verified using the Agilent 2100 Bioanalyzer system (Agilent Technologies). Biotin-labeled cRNA samples for hybridization on Illumina HumanHT-12 v4 Expression BeadChip (Illumina, Inc.) were prepared according to the Illumina's recommended sample labeling procedure. In brief, 250 ng total RNA was used for complementary DNA (cDNA) synthesis, followed by an amplification/labeling step (*in vitro* transcription) to synthesize biotin-labeled cRNA according to the MessageAmp II aRNA Amplification kit (Ambion, Inc., Austin, TX). Quality of cRNA was controlled using the RNA Nano Chip Assay on an Agilent 2100. Microarray scanning was performed using a Beadstation array scanner, setting adjusted to a scaling factor of 1 and PMT settings at 430. Data extraction was done for all beads individually, and outliers are removed when >2.5 MAD (median absolute deviation). All remaining data points are used for the calculation of the mean average signal for a given probe, and standard deviation for each probe was calculated. All arrays were quantile normalized without background subtraction using the Illumina BeadStudio software.

The Illumina probes were annotated with the custom mappings from the Bioconductor R package illuminaHumanv4.db [5]. Probes that were flagged as 'bad', *i.e.* matching repeat sequences, intergenic or intronic regions, or having 'No match', *i.e.* without match for any genomic region or transcript, according to the illuminaHumanv4PROBEQUALITY mapping were discarded. The remaining probes were annotated to the EntrezIDs provided by the illuminaHumanv4ENTREZREANNOTATED map. The 'findLargest' function from the R Bioconductor package genefilter [6] resolved the problem, if multiple probes matched the same Entrez gene. The function finds all replicates and selects the one with the largest interquartile rank (IQR) in expression across all samples, finally retaining 19,178 Entrez genes in the Colo-205 and the HT29 array samples.

RNA of both cell lines was hybridized to different Illumina bead chips. After normalization there was a marked difference in the normalized expression values (Fig. 1A). Due to the unknown contributions of either technical or biological variability we restricted our analysis to comparisons of within cell line differences with respect to treatment and time. A principal component analysis (PCA) [7] on both cell line samples (Fig. 1B and C) confirmed 3D culture growth as the largest influence on gene expression followed by the effect of PLX4720 treatment, which showed in the separation along PC1 and PC2, and is marked in the plots by 'Culture' and 'Treatment', respectively. Furthermore, the duplicates for the HT29 lie close to each other, which demonstrates the reproducibility of the 3D culture.

Due to the similar PCA response patterns between Colo-205 and HT29 we hypothesized that the culture conditions and treatment might elicit similar genes. As we did not have replicate samples for the Colo-205, we performed for both cell lines a moderated F-test using the R 'limma' package [8] on the treatment *versus* control samples to identify genes that responded differently over time. p-Values from the F-statistic were corrected for false discovery rate estimation according to Benjamini and Hochberg. There was a significant Pearson correlation of $r = 0.71$ for all differentially genes between PLX4720 and control in both cell lines (adjusted p-value <0.001), supporting our initial hypothesis. A hypergeometric test on the Biological Processes from the human Gene Ontology using the R GOstats package [9] showed a strong involvement of these genes in developmental and differentiation processes (Fig. 2).

B-RAF inhibition leads to induction of differentiation

The above findings led to the assumption that the effect of PLX4720 reintroduced a differentiation phenotype from the otherwise more stem cell like cancer phenotype. To corroborate in detail the loss of stemness/pluripotency we used the PluriNet gene set [10] as marker for the stemness or inversely the differentiation status of the cells. The marker set consists of 299 genes that were derived from classifying human pluripotent, multipotent and differentiated cells. To assess the overall stemness/differentiation status of the Colo-205 and HT29 cells, we compared the PluriNet gene expression among all NCI60 cell lines [11]. Gene expression had been measured using the Affymetrix Human Genome U133 Plus 2.0 arrays, which we normalized together using robust multichip averaging [12] in conjunction with the custom definition file from Brainarray in Version 17.0 [13]. A principal component analysis separated the samples along the PC1 according to their marker gene expression, predicting the acute lymphoblastic leukemia cell line MOLT-4 as least and the ovary adenocarcinoma derived SK-OV-3 as the most differentiated cells (Fig. 3A). The significant higher expression of the PluriNet marker genes in Colo-205 cells relative to HT29 (p-value $< 10^{-10}$, one-sided t-test) likewise indicated a higher stemness of the former (Fig. 3B). To assess the loss of stemness we performed a gene set enrichment using the Generally Applicable Gene-set Enrichment (GAGE) [14]. For analysis we performed an

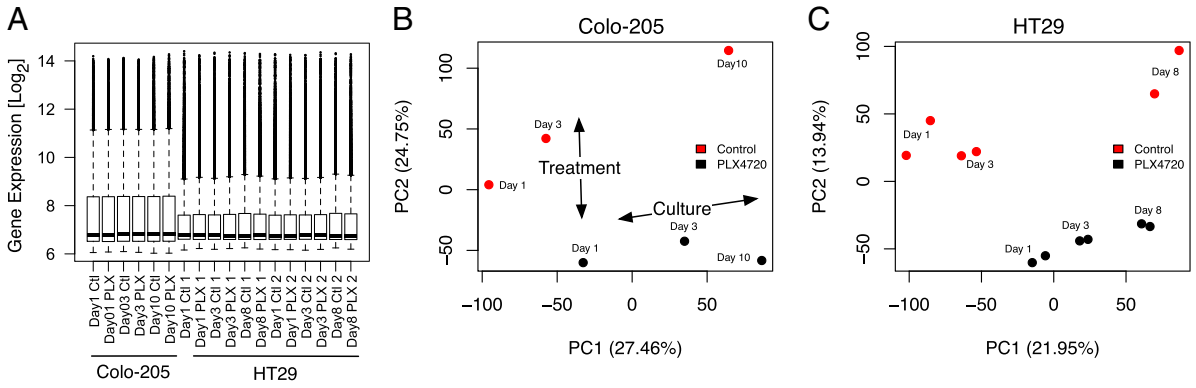

Fig. 1. Bead chip normalization and principal component analysis. (A) Boxplot of the quantile normalized gene expression data. Colo-205 and HT29 have been hybridized to different Illumina Human-HT12 bead chips resulting in different gene expression distributions. (B, C) Principal component analysis of the Colo-205 and HT29 cell line samples. Sample separation along the first principal component (PC1) result from gene expression changes due to 3D culture over time, while PC2 separates samples due to different treatments. Reproducibility of the results is confirmed by the proximity of the HT29 sample duplicates.

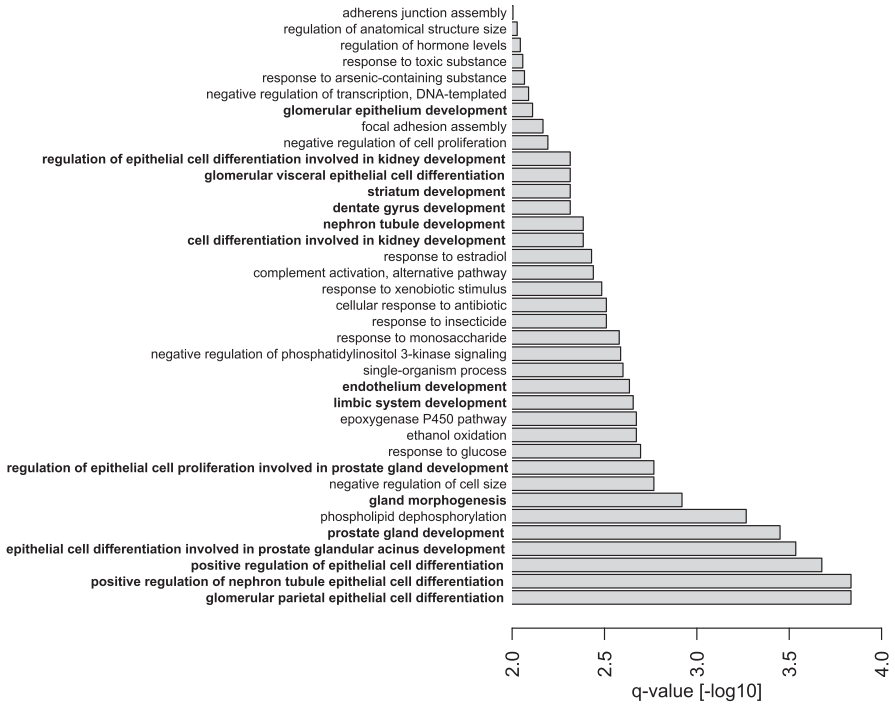

Fig. 2. Hypergeometric test for biological processes enrichment. Hypergeometric test for biological processes enrichment of genes differentially regulated between all PLX4720 treatment and control samples in both Colo-205 and HT29 cells (FDR corrected p-value < 0.001, moderated F-test). GO terms related to differentiation and development are marked in bold. Terms related to metabolism are not considered for clarity. p-Value cutoff for the enrichment p < 0.01.

unpaired sample comparison of PLX4720 treated cells on days 3 and 8/10 relative to day 1. We tested for changes in a gene set in the same direction using the per gene fold change and summarizing individual p-values using Stouffer's method. Not surprisingly, GAGE confirmed the reduction of stemness with time (Fig. 3C) in both cell lines with a more significant effect in the Colo-205 cells.

Discussion

We described analysis on the two colon cancer cell lines that harbor *BRAF*^{V600E} mutations. To elucidate the effect of this mutation on carcinogenesis we compared the changes in gene expression of Colo-205 and HT29 3D cultures under B-RAF inhibition over time. While the cell

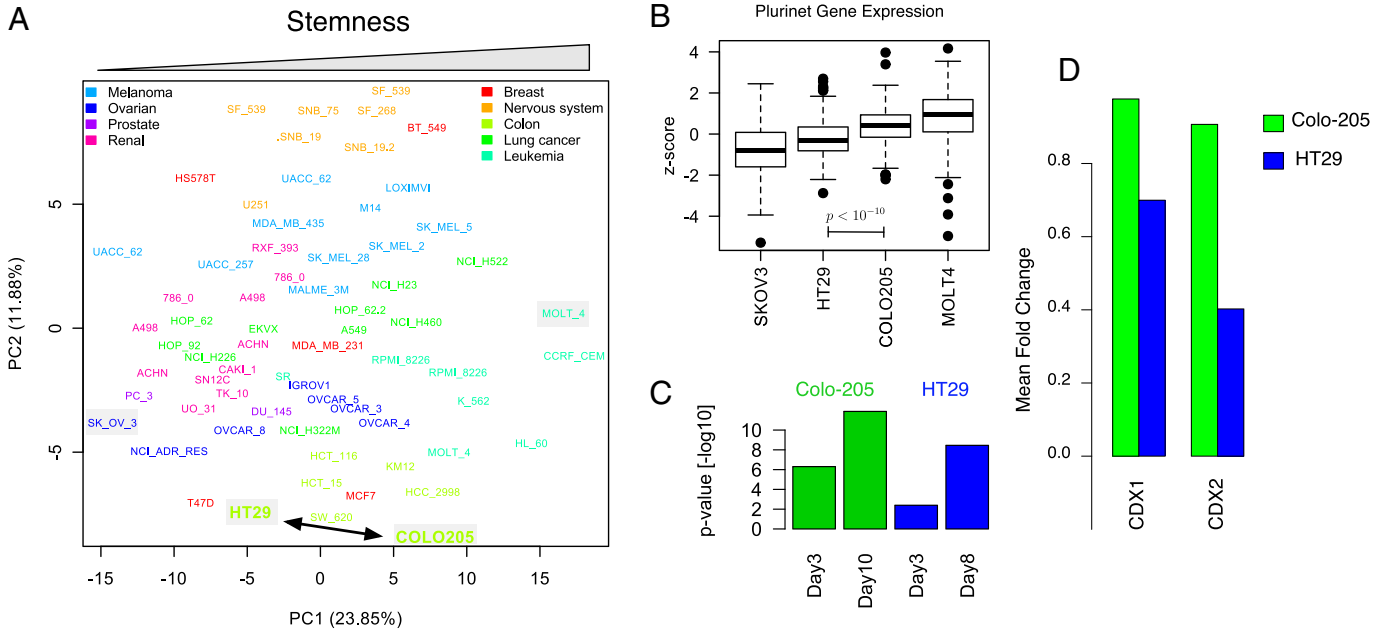

Fig. 3. PLX4720 induces loss of stemness in Colo-205 and HT29. (A) Principal component analysis of all NCI-60 cell lines based on the PluriNet marker genes for stemness. The first principal component separates the samples according to their gene set expression. The cell line names are color coded according to their tissue origin. The cell lines used in this paper are highlighted at the bottom. (B) Boxplot of the scaled expression of the PluriNet marker genes. Median gene expression and therefore stemness increase from left to right. (C) Significance of loss of stemness of the treated cell lines estimated by gene set enrichment of the PluriNet gene set taking the PLX4720 treatment at day 1 as reference. (D) Fold change in the expression of differentiation marker genes for Colo-205 and HT29 cell lines.

culture data indicated a high reproducibility, technical and/or biological batch effects hindered us from direct comparison of the cell line data. Therefore, we had to analyze each cell type individually. Yet, functional enrichment of the gene expression clearly correlated with the treatment and cell culture phenotype and suggested the restoration of differentiated epithelia. The analysis predicted novel players involved in the differentiation process and inversely suggested stemness properties inherent in the cancer cell lines. Including a broader data base from the NCI60 cell line panel allowed quantifying the relative differentiation status of the cells. Gene set enrichment then demonstrated the loss of stemness under B-RAF inhibition in line with the observed differentiation phenotype. Based on Fig. 3C, it is even tempting to speculate that the Colo-205 cells lose their stemness faster than HT29, as the latter are more differentiated than the former according to Fig. 3B. This is further supported by the stronger up-regulation over time of differentiation markers in Colo-205 that are not part of the PluriNet gene set, e.g., the transcription factors Caudal Type Homeobox 1 and 2 (Fig. 3D) or the various gene products associated with epithelial differentiation and effector functions [1].

In conclusion, we demonstrated the use of functional analysis in combination with public database data to elucidate the oncogenic mechanisms underlying *BRAF* mutations in Colon cancer that led to the discovery of novel players and novel therapeutic rationales for using pathway inhibitors in the treatment of this disease.

Acknowledgment

All authors acknowledge financial support by the DFG funded SFB 850 (projects B4 and Z1). We thank the microarray unit of the DKFZ Genomics and Proteomics Core Facility for providing the Illumina Whole-Genome Expression BeadChips and related services.

References

[1] R. Herr, M. Köhler, H. Andrlová, F. Weinberg, Y. Möller, S. Halbach, et al., B-Raf inhibitors induce epithelial differentiation in BRAF-mutant colorectal cancer cells. Cancer Res. 75 (2015) 216–229, http://dx.doi.org/10.1158/0008-5472.CAN-13-3686.
[2] T.U. Semple, L.A. Quinn, L.K. Woods, G.E. Moore, Tumor and lymphoid cell lines from a patient with carcinoma of the colon for a cytotoxicity model. Cancer Res. 38 (1978) 1345–1355.
[3] J. Fogh, G. Trempe, New Human Tumor Cell Lines. in: J. Fogh (Ed.), Hum. Tumor Cells Vitro, Springer, US 1975, pp. 115–159 (http://link.springer.com/chapter/10.1007/978-1-4757-1647-4_5 (accessed April 3, 2015)).
[4] M. Röring, R. Herr, G.J. Fiala, K. Heilmann, S. Braun, A.E. Eisenhardt, et al., Distinct requirement for an intact dimer interface in wild-type, V600E and kinase-dead B-Raf signalling. EMBO J. 31 (2012) 2629–2647, http://dx.doi.org/10.1038/emboj.2012.100.
[5] M. Dunning, A. Lynch, M. Eldridge, illuminaHumanv4.db: Illumina HumanHT12v4 annotation data (chip illuminaHumanv4), R package version 1.26.0.
[6] R. Gentleman, V. Carey, W. Huber, F. Hahne, genefilter: genefilter: methods for filtering genes from high-throughput experiments, R package version 1.50.0.
[7] D.W. Roberts, labdsv: ordination and multivariate analysis for ecology. http://CRAN.R-project.org/package=labdsv2013.
[8] G.K. Smyth, Limma: linear models for microarray data. in: R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, W. Huber (Eds.),Bioinforma. Comput. Biol. Solut. Using R BioconductorSpringer, New York 2005, pp. 397–420.
[9] S. Falcon, R. Gentleman, Using GOstats to test gene lists for GO term association. Bioinform. Oxf. Engl. 23 (2007) 257–258, http://dx.doi.org/10.1093/bioinformatics/btl567.
[10] F.-J. Müller, L.C. Laurent, D. Kostka, I. Ulitsky, R. Williams, C. Lu, et al., Regulatory networks define phenotypic classes of human stem cell lines. Nature 455 (2008) 401–405, http://dx.doi.org/10.1038/nature07213.
[11] R.H. Shoemaker, The NCI60 human tumour cell line anticancer drug screen. Nat. Rev. Cancer 6 (2006) 813–823, http://dx.doi.org/10.1038/nrc1951.
[12] L. Gautier, L. Cope, B.M. Bolstad, R.A. Irizarry, affy—analysis of Affymetrix GeneChip data at the probe level. Bioinform. Oxf. Engl. 20 (2004) 307–315, http://dx.doi.org/10.1093/bioinformatics/btg405.
[13] M. Dai, P. Wang, A.D. Boyd, G. Kostov, B. Athey, E.G. Jones, et al., Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res. 33 (2005)http://dx.doi.org/10.1093/nar/gni179 e175.
[14] W. Luo, M.S. Friedman, K. Shedden, K.D. Hankenson, P.J. Woolf, GAGE: generally applicable gene set enrichment for pathway analysis. BMC Bioinform. 10 (2009) 161, http://dx.doi.org/10.1186/1471-2105-10-161.