# Potential Colon Cancer Biomarker Search using more than two Performance Measures in a Multiple Criteria Optimization Approach

Érika Watts-Oquendo, BS*; Matilde Sánchez-Peña, MS*; Clara E. Isaza, PhD*†; Mauricio Cabrera-Ríos, PhD*

**Objective: A new method using Multiple Criteria Optimization (MCO) proposed by our research group has shown evidence of being able to identify gene-based biomarkers for the detection of cancer using microarray data. Herein, we explore this method, considering more than two conflicting criteria for the MCO problem. Using this method would result in stronger outcomes when using different results from microarray analyses. It would also demonstrate that the method is suitable for carrying out meta-analysis.**

**Methods: Statistical comparisons between normal and cancer tissues were performed using a colon cancer microarray database. The different comparisons were carried out with a Mann–Whitney non-parametric test using partial permutations of the data. An MCO problem was built using the different p-values obtained. The associated solution was the set of genes reaching the best compromises between the p-values under consideration that were located in the so-called efficient frontier. Data Envelopment Analysis (DEA) was used to find the efficient frontier of the MCO problem. The capacity of DEA was explored using different numbers of p-values (criteria) in the model.**

**Results: The set of identified genes was consistent across the instances using different numbers of p-values in the DEA model, thereby providing evidence of the outcome stability of the proposed strategy. It was also observed that convergence to a larger number of potential biomarkers is faster with additional criteria, i.e., more p-values.**

**Conclusion: The MCO problem proposed for the cancer biomarker search using microarray data can be solved efficiently with DEA using more than two conflicting criteria. This approach can result in robust results when using different analyses of microarray data and, indeed, in a faster convergence to highly potential biomarkers. [*P R Health Sci J 2012;2:59-63*]**

*Key words: Microarray data, Potential cancer biomarkers, Performance measures, Multiple Criteria Optimization, Data Envelopment Analysis*

Cancer is the second most relevant cause of death worldwide (1). Microarray experiments aim to measure the change in the genetic expression of tens of thousands of genes simultaneously and have been used to generate many of the genetic pipelines in cancer research (2). When considering normal and cancer tissues, genes with the highest differential expression between these states are potential cancer biomarkers. Many and varied methodologies have been proposed for the identification of these genes (3). Our research group has proposed that the identification of potential cancer biomarkers using microarray data be carried out through Multiple Criteria Optimization (MCO) techniques (4).

An MCO problem aims to find the best compromises between two or more conflicting criteria considered. The best compromises are located in the so-called "efficient frontier" of the MCO problem. Results of two or more analyses for a set of genes can be used as conflicting criteria that can be accommodated in an MCO problem. Our hypothesis is that the genes located in the efficient frontier of the related MCO problem are potential cancer biomarkers.

*BioIE Lab, Department of Industrial Engineering, University of Puerto Rico Mayagüez Campus, Mayagüez, Puerto Rico; †Immunology Department, School of Biology, Universidad Autónoma de Nuevo León, México

*The authors have no conflicts of interest to disclose.*

Address correspondence to: Mauricio Cabrera-Ríos, PhD, Department of Industrial Engineering, University of Puerto Rico Mayagüez Campus, Call Box 9000, Mayagüez, PR 00681-9000. Email: mauricio.cabrera1@upr.edu

Data Envelopment Analysis (DEA) has been identified as being particularly well suited to the task of identifying the efficient frontiers of MCO problems (5). Here, the ability of the proposed method is explored using more than two performance measures and solved through DEA. Results show that the method identifies a consistent set of genes when increasing the number of performance measures in the MCO problem. It is also observed that convergence of a larger number of potential biomarkers is faster with additional criteria, i.e., more p-values.

## Methods

### Microarray Data

A colon cancer microarray database was selected for the described exploration. This database was first reported in Alon et al. (6) and is available at www.molbio.princeton.edu/colondata. It contains the measured expression of 6,500 genes in 22 normal tissues and 40 cancer tissues, all of which were characterized using Affymetrix Hum6000 arrays.

### Statistical Analysis

Statistical comparisons between the normal and cancer replicates were performed using the Mann–Whitney non-parametric test. The procedure is illustrated in Figure 1. P-values from different statistical analyses were obtained using partial permutations leaving one, two, or three tissues out of each state; the excluded tissues were selected randomly. From these comparisons, a total of ten different p-values were obtained for each gene. A p-value in the Mann–Whitney test is understood as representing the probability of finding a particular difference of medians between the two states by pure chance. Thus, to favor finding truly significant differences, low p-values are sought.

### Multiple Criteria Optimization and Data Envelopment Analysis

Considering that the aim is to find those genes that change their expressions to the greatest degree between the different states, a p-value can be seen as a criterion to be minimized: smaller p-values show stronger evidence for the rejection of the stated null hypothesis, which relates to not having a significant difference between the two states. Thus, one can build an MCO problem considering the different p-values available for each gene as criteria intended to be minimized simultaneously.

The case of an MCO problem using two different p-values is presented in Figure 2(a). Given the minimization objective for both p-values, the efficient frontier is located in the southwest corner. In order to use DEA to find the efficient frontier, it is necessary to maximize at least one of the conflicting criteria, so a transformation (shown in Figure 2b) should be performed on at least one of the considered p-values. For the instances presented here, half of the p-values were transformed.

Graphical representation becomes complicated when using more than two p-values; however, the use of DEA to find the efficient frontier can be extended to the desired number of dimensions easily without a loss of generality. Banker-Charnes-Cooper (BCC) input- and output-oriented DEA models were used for the frontier search. For this search, genes identified in the previous frontier were removed from the original list and the search process repeated until the tenth frontier was reached. The instances presented here correspond to the use of 2, 4 and 8 p-values for the MCO problem. Results obtained from the different combinations were compared.

In order to express the idea of the method in a simple manner, one can think of a small p-value for a particular gene as being an indicator of its importance. Taking tissues out of the dataset creates a series of somewhat different datasets that allow the computation of multiple p-values for all genes. If all of the p-values for a particular gene are small, then it is likely that that gene will be significantly differentially expressed. Genes with these characteristics tend to cluster along the particular edges of the set of genes under analysis. MCO´s objective is to find this specific edge (efficient frontier) and the genes lying on it.

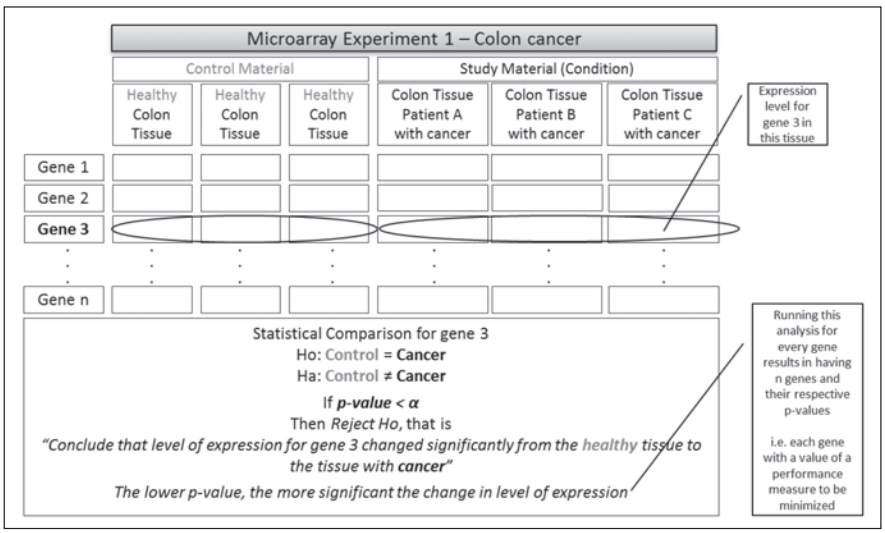


**Figure 1**. Illustration of the statistical comparisons executed with the gene expression on microarray experiments.

### MCO Gene Selection Validation

Since the purpose of this study was to see how useful it would be to model the microarray data analysis for potential biomarker identification as an MCO problem, the results needed to be validated. The validation was performed by undertaking a literature search for the genes identified by this method that changed their expression to the greatest degree between normal and cancer tissues.

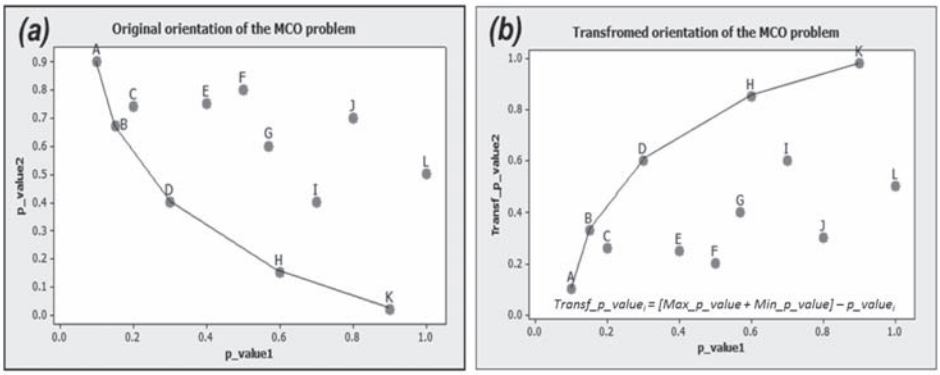


**Figure 2**. (a) Graphic representation of the Multiple Criteria Optimization (MCO) problem that was proposed for the finding of biomarkers using statistical p-values; each point represents a different gene. (b) The corresponding transformation of the MCO problem in order to find the efficient frontier using DEA.

## Results

One of the most notable results was that the number of genes identified in the efficient frontiers increased as the number of p-values that were considered in the model also increased. Table 1 shows the genes found in the different combinations of p-values; the information about each identified gene is followed by the frontier where each gene was localized in the corresponding run.

All of the genes selected by the analysis but one, GTF3A, have been previously reported to change their expressions in colorectal cancer and/or other cancer types (Table 2). These reports are based on in vitro and/or in vivo experiments. Even though the role of GTF3A in cancer is still not confirmed, it is quite possible that changes in its expression could be related to cancer development. The GTF3A gene product is a transcription factor that regulates expression of 5 S ribosomal RNA.

## Discussion

The MCO method of searching for biomarkers using available microarray data was demonstrated to be robust through the

**Table 1**. List of genes identified using multiple p-values in the Multiple Criteria Optimization (MCO) problem for the biomarker search. The number describes the frontier where each gene was found according to the different number of p-values executions used.

| Accession Number | Gene Symbol | Gene Name | p-values used | | | |
|---|---|---|---|---|---|---|
| | | | 2-pv | 4-pv | 6-pv | 8-pv |
| R87126 | yq31b10.s1 | Soares fetal liver spleen 1NFLS | 1 | 1 | 1 | 1 |
| H08393 | WDR77 | WD repeat domain 77 | 2 | 1 | 1 | 1 |
| R36977 | GTF3A | General Transcription factor IIIA | 3 | 2 | 2 | 1 |
| M22382 | HSPD1 | Heat Shock 60kDa protein 1 (chaperonin) | 4 | 3 | 2 | 1 |
| M26383 | IL8 | Interleukin 8 | 5 | 4 | 3 | 2 |
| X63629 | CDH3 | Cadherin 3, type 1, P-cadherin (placental) | 5 | 4 | 3 | 2 |
| H40095 | yn85b03.s1 | Soares adult brain N2b5HB55Y | 5 | 4 | 3 | 2 |
| X12671 | HNRNPA1 | Heterogeneous nuclear ribonucleoprotein A1 | 5 | 4 | 3 | 2 |
| J05032 | DARS | Aspartyl-tRNA synthetase | 6 | 5 | 4 | 2 |
| U09564 | SRPK1 | SRSF protein kinase 1 | 6 | 5 | 4 | 2 |
| Z50753 | GUCA2B | Guanylate cyclase activator 2B (uroguanylin) | 6 | 4 | 3 | 2 |
| J02854 | MYL9 | Myosin, light chain 9, regulatory | 7 | 6 | 3 | 2 |
| T47377 | S100P | Calcium binding protein P | 7 | 5 | 4 | 3 |
| T86473 | NME1 | Non-metastatic cells 1, protein (NM23A) | 7 | 6 | 5 | 3 |
| H43887 | CFD | Complement factor D (adipsin) | 7 | 5 | 4 | 3 |
| M36634 | VIP | Vasoactive intestinal peptide | 8 | 7 | 4 | 3 |
| R08183 | HSPE1 | Heat Shock 10kDa protein 1 (chaperonin 10) | 8 | 6 | 5 | 3 |
| T71025 | MT1G | Metallothionein 1G | 8 | 7 | 5 | 3 |
| U30825 | SRSF9 | Serine/arginine-rich splicing factor 9 | 9 | 7 | 5 | 3 |
| X14958 | HMGA1 | High mobility group AT-hook 1 | 9 | 7 | 5 | 3 |
| M26697 | NPM1 | Nucleophosmin (nucleolar phosphoprotein B23, numatrin) | 9 | 7 | 6 | 3 |
| R84411 | SNRPB | Small nuclear ribonucleoprotein polypeptides B and B1 | 10 | 8 | 6 | 4 |
| X12466 | SNRPE | Small nuclear ribonucleoprotein polypeptide E | 10 | 8 | 7 | 4 |
| M63391 | DES | Desmin | 10 | 8 | 3 | 2 |

**Table 2**. List of MCO-identified genes with examples of different types of cancer that have been shown to change their expression

| Gene | Cancer type involvement (not a comprehensive list) | References |
|---|---|---|
| WDRF77 | Ovarian, prostate | 7, 8 |
| HSPD1 | Colorectal | 9 |
| IL8 | Colorectal | 10, 11 |
| CDH3 | Biliary tract, esophageal | 1, 13 |
| HNRNPA1 | Involved in the switch to aerobic glycolysis, a process common to cancer cells. | 14 |
| DARS | Leukemia | 15 |
| SRPK1 | Colorectal | 16 |
| GUCA2B | Colorectal | 17 |
| MYL9 | Chicken sarcoma model for metastasis, breast cancer cell motility | 18, 19 |
| S100P | Pancreatic, lung adenocarcinomas, breast, colon | 20, 21, 22, 23 |
| NME1 | Colon | 24 |
| CFD | Gastric, tongue, colon | 25, 26, 27 |
| VIP | Colorectal | 28 |
| HSPE1 | Proposed to have a role in cancer etiology | 29 |
| MTG1 | Lung adenocarcinoma, colorectal | 30, 31 |
| SRSF9 | Regulation of procancerous proteins | 32, 33 |
| HMGA1 | Prostate, apoptosis inhibition | 34, 35 |
| NPM1 | Acute myeloid leukemia | 36, 37 |
| SNRPB | Proposed metastasis suppressor gene for prostate cancer | 38 |
| SNRPE | Hepatocellular carcinoma | 39 |
| DES | Colorectal | 40 |

MCO: Multiple Criteria Optimization

use of a different number of statistical p-values identifying a consistent set of genes. This approach may contribute to the rapid identification of genes by their biological validation as contributors to cancer. The method can also be explored using different DEA models and different types of available data, thus opening several opportunities for the meta-analysis of microarray experiments.

## Resumen

Objetivo: Nuestro grupo de investigación ha propuesto un método para identificar genes potencialmente biomarcadores para la detección de cáncer basado en el análisis de datos de microarreglos por medio de la Optimización de Criterios Múltiples (OCM). Este método ha demostrado ser muy efectivo. En este trabajo, se explora la capacidad de este método para involucrar más de dos criterios en conflicto dentro del problema de OCM. La capacidad de manejar esta situación sería un paso positivo hacia la utilización del método para meta-análisis. Métodos: Inicialmente, se llevaron a cabo comparaciones estadísticas a nivel de genes para contrastar su expresión relativa en tejidos de colon normales y tejidos de colon con cáncer. Estas comparaciones se hicieron a través de la prueba no-paramétrica Mann-Whitney usando permutaciones parciales de los datos disponibles. Los valores p obtenidos se utilizaron después para formular el problema de OCM. En la solución de este problema, la frontera eficiente, se identificaron los genes que correspondían a los mejores balances entres los valores p considerados. La técnica de Análisis Envolvente de Datos (AED) se utilizó para encontrar tal frontera eficiente. En este estudio, se trataron diferentes números de valores p en la formulación de AED para establecer su capacidad. Resultados: El conjunto de genes en la solución de las diferentes instancias correspondientes a casos con diferente número de valores p en la formulación de AED fue consistente. Ésto evidencia la estabilidad de las soluciones de la estrategia propuesta. Se observó también una convergencia más rápida a una cantidad mayor de biomarcadores potenciales cuando se incrementó el número de criterios a utilizar dentro del problema de OCM. Conclusión: El problema de OCM propuesto para la búsqueda de biomarcadores de cáncer a través del análisis de datos de microarreglos se puede resolver eficientemente a través de AED utilizando más de dos criterios en conflicto. Esto indica que existe robustez en los resultados arrojados por la estrategia y que es posible aumentar la rapidez de convergencia a biomarcadores altamente potenciales.

## Acknowledgments

M.S.P was supported by a research assistantship from the Industrial Engineering Department at UPRM. Authors acknowledge the support of BioSEI Grant 33 010 3080 301 (awarded to M.C.R.) and that of the PROMEP project (103.5/07/2523, granted to C.I.).

## References

1. American Cancer Society. Cancer Facts & Figures - 2010. Available at: Url: http://www.cancer.org/acs/groups/content/@nho/documents/document/acspc-024113.pdf.
2. Berns A. Gene expression in diagnosis. Nature 2000;403:491-2.
3. Tainsky MA. Genomic and proteomic biomarkers for cancer: A multitude of opportunities. Biochimica Biophys Acta 2009;1796:176-93.
4. Isaza C, Sanchez-Peña M, Rodriguez C, Cabrera M. Abstract B45: An optimization-based approach to potential biomarker identification with microarray data [Internet]. In: Poster Presentations - Other Biomarkers and Early Detection Topics. Philadelphia PA: Cancer Prevention Research - AACR; p. Supplement 2.
5. Castro C, Cabrera-Ríos M, Lilly B, Castro JM, Mount-Campbell CA. Identifying the best compromises between multiple performance measures in injection molding (IM) using Data Envelopment Analysis (DEA). J Integr Des Process Sci 2003;7:77-86.
6. Alon U, Barkai N, Notterman DA, Gish K, Ybarra S, Mack D, et al. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proc Natl Acad Sci U S A 1999;96:6745-50.

7. Ligr M, Patwa RR, Daniels G, Pan L, Wu X, Li Y, et al. Expression and function of androgen receptor coactivator p44/Mep50/WDR77 in ovarian cancer. PLoS One 2011;6:e26250.
8. Gu Z, Zhou L, Gao S, Wang Z. Nuclear transport signals control cellular localization and function of androgen receptor cofactor p44/WDR77. PLoS One 2011;6:e22395.
9. Ruan W, Wang Y, Ma Y, Xing X, Lin J, Cui J, Lai M. HSP60, a protein downregulated by IGFBP7 in colorectal carcinoma. J Exp Clin Cancer Res 2010;29:41.
10. McLean MH, Murray GI, Stewart KN, Norrie G, Mayer C, Hold GL, Thomson J, Fyfe N, Hope M, Mowat NA, Drew JE, El-Omar EM. The inflammatory microenvironment in colorectal neoplasia. PLoS One 2011;6:e15366.
11. Ning Y, Manegold PC, Hong YK, Zhang W, Pohl A, Lurje G, Winder T, Yang D, LaBonte MJ, Wilson PM, Ladner RD, Lenz HJ. Interleukin-8 is associated with proliferation, migration, angiogenesis and chemosensitivity in vitro and in vivo in colon cancer cell line models. Int J Cancer 2011;128:2038-49.
12. Riener MO, Vogetseder A, Pestalozzi BC, Clavien PA, Probst-Hensch N, Kristiansen G, Jochum W. Cell adhesion molecules P-cadherin and CD24 are markers for carcinoma and dysplasia in the biliary tract. Hum Pathol 2010;41:1558-65.
13. Boonstra JJ, van Marion R, Douben HJ, Lanchbury JS, Timms KM, Abkevich V, Tilanus HW, de Klein A, Dinjens WN. Mapping of homozygous deletions in verified esophageal adenocarcinoma cell lines and xenografts. Genes Chromosomes Cancer 2012;51:272-82.
14. Chen M, Zhang J, Manley JL. Turning on a fuel switch of cancer: hnRNP proteins regulate alternative splicing of pyruvate kinase mRNA. Cancer Res 2010;70:8977-80.
15. Chen SH, Yang W, Fan Y, Stocco G, Crews KR, Yang JJ, Paugh SW, Pui CH, Evans WE, Relling MV. A genome-wide approach identifies that the aspartate metabolism pathway contributes to asparaginase sensitivity. Leukemia 2011;25:66-74.
16. Thorsen K, Mansilla F, Schepeler T, Øster B, Rasmussen MH, Dyrskjøt L, Karni R, Akerman M, Krainer AR, Laurberg S, Andersen CL, Ørntoft TF. Alternative splicing of SLC39A14 in colorectal cancer is regulated by the Wnt pathway. Mol Cell Proteomics 2011;10:M110.002998.
17. Nagaraj SH, Reverter A. A Boolean-based systems biology approach to predict novel genes associated with cancer: Application to colorectal cancer. BMC Syst Biol 2011;5:35.
18. Cermák V, Kosla J, Plachý J, Trejbalová K, Hejnar J, Dvorák M. The transcription factor EGR1 regulates metastatic potential of v-src transformed sarcoma cells. Cell Mol Life Sci 2010;67:3557-68.
19. Götte M, Mohr C, Koo CY, Stock C, Vaske AK, Viola M, Ibrahim SA, Peddibhotla S, Teng YH, Low JY, Ebnet K, Kiesel L, Yip GW. miR-145-dependent targeting of junctional adhesion molecule A and modulation of fascin expression are associated with reduced breast cancer cell motility and invasiveness. Oncogene 2010;29:6569-80.
20. Dowen SE, Crnogorac-Jurcevic T, Gangeswaran R, Hansen M, Eloranta JJ, Bhakta V, Brentnall TA, Lüttges J, Klöppel G, Lemoine NR. Expression of S100P and its novel binding partner S100PBPR in early pancreatic cancer. Am J Pathol 2005;166:81-92.
21. Rehbein G, Simm A, Hofmann HS, Silber RE, Bartling B. Molecular regulation of S100P in human lung adenocarcinomas. Int J Mol Med 2008;22:69-77.
22. Guerreiro Da Silva ID, Hu YF, Russo IH, Ao X, Salicioni AM, Yang X, Russo J. S100P calcium-binding protein overexpression is associated with immortalization of human breast epithelial cells in vitro and early stages of breast cancer development in vivo. Int J Oncol 2000;16:231-40.
23. Jiang L, Lai YK, Zhang J, Wang H, Lin MC, He ML, Kung HF. Targeting S100P inhibits colon cancer growth and metastasis by Lentivirus-mediated RNA interference and proteomic analysis. Mol Med 2011;17:709-16. doi: 10.2119/molmed.2011.00008.
24. Bertucci F, Salas S, Eysteries S, Nasser V, Finetti P, Ginestier C, Charafe-Jauffret E, Loriod B, Bachelart L, Montfort J, Victorero G, Viret F, Ollendorff V, Fert V, Giovaninni M, Delpero JR, Nguyen C, Viens P, Monges G, Birnbaum D, Houlgatte R. Gene expression profiling of colon cancer by DNA microarrays and correlation with histoclinical parameters. Oncogene 2004;23:1377-91.
25. Claerhout S, Lim JY, Choi W, Park YY, Kim K, Kim SB, Lee JS, Mills GB, Cho JY. Gene expression signature analysis identifies vorinostat as a candidate therapy for gastric cancer. PLoS One 2011;6:e24662.
26. Ye H, Yu T, Temam S, Ziober BL, Wang J, Schwartz JL, Mao L, Wong DT, Zhou X. Transcriptomic dissection of tongue squamous cell carcinoma. BMC Genomics 2008;9:69.
27. Notterman DA, Alon U, Sierk AJ, Levine AJ. Transcriptional Gene Expression Profiles of Colorectal Adenoma, Adenocarcinoma, and Normal Tissue Examined by Oligonucleotide Arrays. Cancer Res 2001;61:3124-30.
28. Hong Y, Ho KS, Eu KW, Cheah PY. A susceptibility gene set for early onset colorectal cancer that integrates diverse signaling pathways: implication for tumorigenesis. Clin Cancer Res 2007;13:1107-14.
29. Czarnecka AM, Campanella C, Zummo G, Cappello F. Heat shock protein 10 and signal transduction: a "capsula eburnea" of carcinogenesis? Cell Stress Chaperones 2006;11:287-94.
30. Chung JH, Lee HJ, Kim BH, Cho NY, Kang GH. DNA methylation profile during multistage progression of pulmonary adenocarcinomas. Virchows Arch 2011;459:201-11.
31. Arriaga JM, Levy EM, Bravo AI, Bayo SM, Amat M, Aris M, Hannois A, Bruno L, Roberti MP, Loria FS, Pairola A, Huertas E, Mordoh J, Bianchini M. Metallothionein expression in colorectal cancer: relevance of different isoforms for tumor progression and patient survival. Hum Pathol 2012;43:197-208.
32. Somberg M, Li X, Johansson C, Orru B, Chang R, Rush M, Fay J, Ryan F, Schwartz S. Serine/arginine-rich protein 30c activates human papillomavirus type 16 L1 mRNA expression via a bimodal mechanism. J Gen Virol 2011;92(Pt 10):2411-21.
33. Cloutier P, Toutant J, Shkreta L, Goekjian S, Revil T, Chabot B. Antagonistic effects of the SRp30c protein and cryptic 5' splice sites on the alternative splicing of the apoptotic regulator Bcl-x. J Biol Chem 2008;283:21315-24.
34. Wei JJ, Wu X, Peng Y, Shi G, Basturk O, Yang X, Daniels G, Osman I, Ouyang J, Hernando E, Pellicer A, Rhim JS, Melamed J, Lee P. Regulation of HMGA1 expression by microRNA-296 affects prostate cancer growth and invasion. Clin Cancer Res 2011;17:1297-305.
35. Esposito F, Tornincasa M, Chieffi P, De Martino I, Pierantoni GM, Fusco A. High-mobility group A1 proteins regulate p53-mediated transcription of Bcl-2 gene. Cancer Res 2010;70:5379-88.
36. Bacher U, Haferlach T, Fehse B, Schnittger S, Kröger N. Minimal residual disease diagnostics and chimerism in the post-transplant period in acute myeloid leukemia. ScientificWorldJournal 2011;11:310-9.
37. Bacher U, Schnittger S, Haferlach T. Molecular genetics in acute myeloid leukemia. Curr Opin Oncol 2010;22:646-55.
38. Yi Y, Nandana S, Case T, Nelson C, Radmilovic T, Matusik RJ, Tsuchiya KD. Candidate metastasis suppressor genes uncovered by array comparative genomic hybridization in a mouse allograft model of prostate cancer. Mol Cytogenet 2009;2:18.
39. Jia D, Wei L, Guo W, Zha R, Bao M, Chen Z, Zhao Y, Ge C, Zhao F, Chen T, Yao M, Li J, Wang H, Gu J, He X. Genome-wide copy number analyses identified novel cancer genes in hepatocellular carcinoma. Hepatology 2011;54:1227-36.
40. Arentz G, Chataway T, Price TJ, Izwan Z, Hardi G, Cummins AG, Hardingham JE. Desmin expression in colorectal cancer stroma correlates with advanced stage disease and marks angiogenic microvessels. Clin Proteomics 2011;8:16.