Published in final edited form as:
In Silico Biol. 2011 ; 11(1-2): 1–10. doi:10.3233/ISB-2012-0440.

Identification of common tumor signatures based on gene set enrichment analysis

Xiaosheng Wang
Biometric Research Branch, National Cancer Institute, National Institutes of Health, Rockville, MD 20852, U.S.A, Tel: +1 3014515907

Xiaosheng Wang: xiaosheng.wang@nih.gov

Abstract

The identification of common tumor signatures can discover the shared molecular mechanisms underlying tumorgenesis whereby we can prevent and treat tumors by a system intervention. We identified tumor-associated signatures including pathways, transcription factors, microRNAs and gene ontology categories by analyzing gene sets for differential expression between normal vs. tumor phenotypes classes in various tumor gene expression datasets. We obtained the common tumor signatures based on their identified frequencies for different tumor types. Some shared signatures important for various tumor types were uncovered and discussed. We proposed that the interventions aiming at both the shared tumor signatures and the tissue-specific tumor signatures might be a potential approach to overcoming cancer.

Keywords

tumor; gene expression profiling; gene set enrichment analysis; bioinformatics

1 Introduction

A large amount of studies have revealed that cancer has been associated with the genetic and genomic changes [1–2]. As the microarray technology have enabled the simultaneous measurement of the expression levels of tens of thousands of genes in a single experiment [3], the use of microarray technology to analyze gene expression changes in tumor tissues is a powerful tool for uncovering the molecular mechanisms underlying cancer [4]. At the same time, the biology of cancer is extremely intricate so that a simple genetic or genomic perspective is insufficient to understand it. Only by attaining more complete cancer-associated molecular profiles such as pathways and transcriptional regulatory circuits, could we comprehend the disease more clearly.

Gene expression profiling has been widely used for identification of cancerous biomarkers whereby we can improve cancerous diagnosis, treatment and prognosis [5–19]. Moreover, since it has been recognized that a gene set could be more biologically significant than individual genes considering gene interactions, the microarray-based gene set enrichment analysis has been investigated on the assumption that it could provide additional insights into the cancer biology [20–22]. Generally speaking, cancer is a systems biology disease [23–24]. To understand the disease at a system level, identification of common tumor signatures among multiple tumor tissues is a critical avenue, although a substantial number of tumor signatures might be tissue-specific.

In the present study, we identified the common tumor signatures closely associated with various tumor types. The signatures include four types: pathways, transcriptional factors (TFs), microRNAs (miRNAs) and gene ontology (GO) categories, which were identified

through the gene set enrichment analysis based on gene expression profiling. The signatures suggested some basic molecular mechanisms underlying tumor, and might imply potential routes of interventions for cancerous diagnosis and treatment.

2 Methods and Materials

2.1 Methods

We identified important pathways, TFs, miRNAs and GO categories by analyzing gene sets for differential expression between normal vs. tumor phenotypes classes. The LS or KS permuation test and Efron-Tibshirani's GSA maxmean test were used to determine the significant gene sets at 0.05 significance level for identification of pathways, TFs and miRNAs, and 0.0001 significance level for GO categories. The pathways (BioCarta) related to the significant gene sets were identified. The TFs were identified by the gene sets, in each of which all genes were experimentally verified to be targets of the same transcription factor (TF). Each miRNA potentially targeting all the genes in one of the gene sets was identified. The identification of important pathways, TFs and miRNAs was performed with the gene set expression class comparison tool in BRB-ArrayTools, which is an integrated software package for the visualization and statistical analysis of DNA microarray gene expression data [25].

2.2 Materials

We analyzed 23 human gene expression datasets involving 15 tumor types (Table 1) [26]. For each dataset, we carried out class comparison algorithm to identify informative pathways, TFs, miRNAs and GO categories relevant to the tumor(s).

3 Results and Analysis

3.1 Identification of tumor-associated pathways

In the total of 26 class comparisons, we identified 25 pathway sets significant at 0.05 threshold level. The 25 sets encompassed 304 different pathways, 17 of which appeared at least in 10 different sets, suggesting that they were associated with at least 10 different types of tumors. Table 2 lists the 17 most frequent identified pathways. The complete 304 pathways identified are presented in the supplementary Table S1. From Table 2, we can see that the most common tumor-associated pathways are often involved in cell cycle regulation, mitogen-activated protein kinase (MAPK) signaling, epidermal growth factor receptor (EGFR), metabolism, oxidative stress, cell motility etc. Many studies have come to the similar conclusions [27–43].

3.2 Identification of tumor-associated TFs

We identified 26 sets of TF targets significant at 0.05 threshold level. There were 99 different TFs identified relevant to the 26 sets, 22 of which were associated with more than 1/3 of the 26 target sets (Table 3). The most frequently identified TF was c-Myc with 62% occurrence rate, and the next ones were E2F-4, MYB and TP53 all with 58% occurrence rate. All the 99 TFs and their occurrence rates were provided in the supplementary Table S2.

Evidently, c-Myc is one of the most important TFs relevant to cancer [44]. Since c-Myc target genes are often involved in the critical mechanisms underlying cancer like cell cycle regulation, apoptosis, metabolism etc., the dysregulation of c-Myc greatly contributes to cancer [45–47]. Table 3 shows that two members of the MYB family of TFs: MYB and MYBL2, have important relevance to cancer. Indeed, many studies have strongly suggested that they played a role in tumorgenesis [48–58]. The two members of the E2F TF family: E2F-1 and E2F-4, have been revealed to be associated with cancer [59–63]. An extremely

important tumor-associated TF p53 is also presented in Table 3. The role played by p53 in tumorgenesis has been well-recognized [64–66].

3.3 Identification of tumor-associated miRNAs

We identified 24 sets of miRNA targets significant at 0.05 threshold level. The 24 sets were involved in 587 different miRNAs, 34 of which were associated with at least one half of the 24 sets. The 34 miRNAs are listed in Table 4 and the 587 miRNAs are provided in the supplementary Table S3. The most frequently identified three miRNAs were miR-29b, miR-29c and miR-29a, members of miR-29 miRNA gene family. The miR-29 family has been proven to be strongly involved in cancer [67–71]. Table 4 shows that another miRNA gene family miR-30 seems to be closely associated with various tumors. There has been some evidence to support this conclusion [72–74]. In addition, miR-19 and miR-526 miRNA gene families appear to be involved in various tumors (see Table 4). Some literatures have suggested their roles in tumorgenesis [75–77]. The other miRNAs with high frequencies like miR-181c, miR-590, miR-212, miR-338 and miR-202 have also been reported to be associated with tumorgenesis [78–82]. It should be noted that most of the cited support literatures were from recent publications, while all the gene expression datasets studied were from earlier publications, indicating that our inference and prediction were reliable to a certain degree.

3.4 Identification of tumor-associated GO categories

We identified 25 sets of GO categories significant at 0.0001 threshold level. The 25 sets were involved in 2273 different GO terms, 39 of which were concerned with at least five different tumor types. The 39 GO terms are listed in Table 5 and the complete 2273 GO terms are provided in the supplementary Table S4.

Table 5 shows that the genes involved in immune, metabolism, development, cell proliferation and differentiation, damage response etc., are most relevant to tumorgenesis.

4 Discussion and Conclusions

The microarray analysis of gene expression profiling of tumor tissues can not only discover the marker genes relevant to tumor malignancies, but also identify the informative gene sets to reveal the molecular mechanisms underlying tumor. The gene set enrichment analysis is a strong supplement to the individual gene analysis as it can potentially make use of the gene interaction information, which is often missed by the individual gene analysis. In this study, we used the gene set enrichment analysis to identify the shared tumor signatures whereby we could reveal the common mechanisms underlying different types of tumors, and therefore might provide a basic reference to tumor prevention and treatment.

Our tumor signatures included tumor-associated pathways, TFs, miRNAs and GO catagories. Each of the tumor signatures was related to multiple genes and identified based on gene set comparison. Therefore, these kinds of signatures may imply the mechanisms underlying tumor at a close system level. Since it has been recognized that cancer was a systems biology disease, the systems interventions aiming at cancer prevention and treatment could contribute to conquering cancer. Of course, some tumors might involve tissue-specific signatures, and therefore the tissue-specific interventions are necessary for treating the tumors in addition to the systems interventions.

The reliability of the results obtained by the present study is mainly affected by two factors: the quality of microarrays and the statistical power. Microarrays, especially poor-qualified microarrays, often contain a large amount of noises, which are prone to result to identification of false signatures. In addition, the discovery of common tumor signatures by

their occurrence frequencies in different tumor types is not based on sufficiently strong statistical power so that some signatures might be identified by chance. This is a work needed to be improved in the future.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med. 2004; 10(8):789–99. [PubMed: 15286780]
2. Balmain A, Gray J, Ponder B. The genetics and genomics of cancer. Nature genetics. 2003; 33(Suppl):238–44. [PubMed: 12610533]
3. Schena M, et al. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science. 1995; 270(5235):467–70. [PubMed: 7569999]
4. Segal E, et al. From signatures to models: understanding cancer using microarrays. Nature genetics. 2005; 37(Suppl):S38–45. [PubMed: 15920529]
5. Amundson SA, Smilenov LB. Integration of biological knowledge and gene expression data for biomarker selection: FN1 as a potential predictor of radiation resistance in head and neck cancer. Cancer biology & therapy. 2011; 10(12):1252–5. [PubMed: 20948301]
6. Abba MC, et al. Breast cancer biomarker discovery in the functional genomic age: a systematic review of 42 gene expression signatures. Biomarker insights. 2010; 5:103–18. [PubMed: 21082037]
7. Yang W, Ying D, Lau YL. In-depth cDNA library sequencing provides quantitative gene expression profiling in cancer biomarker discovery. Genomics, proteomics & bioinformatics / Beijing Genomics Institute. 2009; 7(1–2):1–12.
8. Alvarez H, et al. Serial analysis of gene expression identifies connective tissue growth factor expression as a prognostic biomarker in gallbladder cancer. Clinical cancer research: an official journal of the American Association for Cancer Research. 2008; 14(9):2631–8. [PubMed: 18451226]
9. Peng Y, Li W, Liu Y. A hybrid approach for biomarker discovery from microarray gene expression data for cancer classification. Cancer informatics. 2007; 2:301–11. [PubMed: 19458773]
10. Koh MS, et al. CDC4 gene expression as potential biomarker for targeted therapy in prostate cancer. Cancer biology & therapy. 2006; 5(1):78–83. [PubMed: 16357515]
11. Yousef GM, et al. Identification of new splice variants and differential expression of the human kallikrein 10 gene, a candidate cancer biomarker. Tumour biology: the journal of the International Society for Oncodevelopmental Biology and Medicine. 2005; 26(5):227–35. [PubMed: 16103744]
12. Statnikov A, et al. GEMS: a system for automated cancer diagnosis and biomarker discovery from microarray gene expression data. International journal of medical informatics. 2005; 74(7–8):491–503. [PubMed: 15967710]
13. Turashvili G, et al. P-cadherin expression as a prognostic biomarker in a 3992 case tissue microarray series of breast cancer. Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc. 2011; 24(1):64–81.
14. Thomson TA, et al. Tissue microarray for routine clinical breast biomarker analysis. The British Columbia Cancer Agency 2008 experience. American journal of clinical pathology. 2010; 133(6): 909–14. [PubMed: 20472849]
15. Luo J, Ellis MJ. Microarray data analysis in neoadjuvant biomarker studies in estrogen receptor-positive breast cancer. Breast cancer research: BCR. 2010; 12(4):112. [PubMed: 20804563]
16. Dahinden C, et al. Mining tissue microarray data to uncover combinations of biomarker expression patterns that improve intermediate staging and grading of clear cell renal cell cancer. Clinical cancer research: an official journal of the American Association for Cancer Research. 2010; 16(1): 88–98. [PubMed: 20028743]

17. Yang N, et al. MicroRNA microarray identifies Let-7i as a novel biomarker and therapeutic target in human epithelial ovarian cancer. Cancer research. 2008; 68(24):10307–14. [PubMed: 19074899]
18. Shimoyama T, Sasaki T. Genomic and post-genomic approaches to cancer biomarker discovery - microarray standards at last. Gan to kagaku ryoho Cancer & chemotherapy. 2007; 34(10):1527–34.
19. Rubin MA, et al. Tissue microarray sampling strategy for prostate cancer biomarker analysis. The American journal of surgical pathology. 2002; 26(3):312–9. [PubMed: 11859202]
20. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102(43):15545–50. [PubMed: 16199517]
21. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research. 2009; 37(1):1–13. [PubMed: 19033363]
22. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols. 2009; 4(1):44–57.
23. Wang E, Lenferink A, O'Connor-McCourt M. Cancer systems biology: exploring cancer-associated genes on cellular networks. Cell Mol Life Sci. 2007; 64(14):1752–62. [PubMed: 17415519]
24. Hornberg JJ, et al. Cancer: a Systems Biology disease. Biosystems. 2006; 83(2–3):81–90. [PubMed: 16426740]
25. Simon R, et al. Analysis of Gene Expression Data Using BRB-Array Tools. Cancer Informatics. 2007; 3:11–17. [PubMed: 19455231]
26. Zhao Y, Simon R. BRB-ArrayToolsData Archive for human cancer gene expression: a unique and efficient data sharing resource. Cancer Inform. 2008; 6:9–15. [PubMed: 19259398]
27. Hartwell LH, Kastan MB. Cell cycle control and cancer. Science (New York, N Y). 1994; 266(5192):1821–8.
28. Massague J. G1 cell-cycle control and cancer. Nature. 2004; 432(7015):298–306. [PubMed: 15549091]
29. Malumbres M, Barbacid M. To cycle or not to cycle: a critical decision in cancer. Nature reviews Cancer. 2001; 1(3):222–31.
30. Dhillon AS, et al. MAP kinase signalling pathways in cancer. Oncogene. 2007; 26(22):3279–90. [PubMed: 17496922]
31. Mimeault M, et al. Recent advances in cancer stem/progenitor cell research: therapeutic implications for overcoming resistance to the most aggressive cancers. Journal of cellular and molecular medicine. 2007; 11(5):981–1011. [PubMed: 17979879]
32. Martinez-Outschoorn UE, et al. Stromal-epithelial metabolic coupling in cancer: Integrating autophagy and metabolism in the tumor microenvironment. The international journal of biochemistry & cell biology. 2011; 43(7):1045–51.
33. Erez A, et al. Insights into the pathogenesis and treatment of cancer from inborn errors of metabolism. American journal of human genetics. 2011; 88(4):402–21. [PubMed: 21473982]
34. Kurhanewicz J, et al. Analysis of cancer metabolism by imaging hyperpolarized nuclei: prospects for translation to clinical research. Neoplasia (New York, N Y). 2011; 13(2):81–97.
35. Whitaker-Menezes D, et al. Evidence for a stromal-epithelial "lactate shuttle" in human tumors: MCT4 is a marker of oxidative stress in cancer-associated fibroblasts. Cell cycle (Georgetown, Tex). 2011; 10(11):1772–83.
36. Saw CLL, Wu Q, Kong ANT. Anti-cancer and potential chemopreventive actions of ginseng by activating Nrf2 (NFE2L2) anti-oxidative stress/anti-inflammatory pathways. Chinese medicine. 2010; 5:37. [PubMed: 20979613]
37. Reuter S, et al. Oxidative stress, inflammation, and cancer: how are they linked? Free radical biology & medicine. 2010; 49(11):1603–16. [PubMed: 20840865]
38. Martinez-Outschoorn UE, et al. Oxidative stress in cancer associated fibroblasts drives tumor-stroma co-evolution: A new paradigm for understanding tumor metabolism, the field effect and genomic instability in cancer cells. Cell cycle (Georgetown, Tex). 2010; 9(16):3256–76.

39. Essick EE, Sam F. Oxidative stress and autophagy in cardiac disease, neurological disorders, aging and cancer. Oxidative medicine and cellular longevity. 2010; 3(3):168–77. [PubMed: 20716941]
40. Donkena KV, Young CYF, Tindall DJ. Oxidative stress andDNA methylation in prostate cancer. Obstetrics and gynecology international. 2010; 2010:302051. [PubMed: 20671914]
41. Pavlides S, et al. The autophagic tumor stroma model of cancer: Role of oxidative stress and ketone production in fueling tumor cell metabolism. Cell cycle (Georgetown, Tex). 2010; 9(17): 3485–505.
42. Carter SB. Principles of cell motility: the direction of cell movement and cancer invasion. Nature. 1965; 208(5016):1183–7. [PubMed: 5331254]
43. Brabek J, et al. The role of the tissue microenvironment in the regulation of cancer cell motility and invasion. Cell communication and signaling: CCS. 2010; 8:22. [PubMed: 20822526]
44. Little CD, et al. Amplification and expression of the c-myc oncogene in human lung cancer cell lines. Nature. 1983; 306(5939):194–6. [PubMed: 6646201]
45. Dang CV. c-Myc target genes involved in cell growth, apoptosis, and metabolism. Molecular and cellular biology. 1999; 19(1):1–11. [PubMed: 9858526]
46. Prendergast GC. Mechanisms of apoptosis by c-Myc. Oncogene. 1999; 18(19):2967–87. [PubMed: 10378693]
47. Evan GI, Vousden KH. Proliferation, cell cycle and apoptosis in cancer. Nature. 2001; 411(6835): 342–8. [PubMed: 11357141]
48. Astbury K, et al. MYBL2 (B-MYB) in cervical cancer: putative biomarker. International journal of gynecological cancer: official journal of the International Gynecological Cancer Society. 2011; 21(2):206–12. [PubMed: 21270603]
49. Thorner AR, et al. Potential tumor suppressor role for the c-Myb oncogene in luminal breast cancer. PloS one. 2010; 5(10):e13073. [PubMed: 20949095]
50. Stenman G, Andersson MK, Andren Y. New tricks from an old oncogene: gene fusion and copy number alterations of MYB in human cancer. Cell cycle (Georgetown, Tex). 2010; 9(15):2986–95.
51. Drabsch Y, Robert RG, Gonda TJ. MYB suppresses differentiation and apoptosis of human breast cancer cells. Breast cancer research: BCR. 2010; 12(4):R55. [PubMed: 20659323]
52. Chen L, et al. c-myb activates CXCL12 transcription in T47D and MCF7 breast cancer cells. Acta biochimica et biophysica Sinica. 2010; 42(1):1–7. [PubMed: 20043041]
53. Ramsay RG, Gonda TJ. MYB function in normal and cancer cells. Nature reviews Cancer. 2008; 8(7):523–34.
54. Kim SY, et al. Adenovirus-mediated expression of dominant negative c-myb induces apoptosis in head and neck cancer cells and inhibits tumor growth in animal model. Oral oncology. 2008; 44(4):383–92. [PubMed: 17690006]
55. Sala A. B-MYB, a transcription factor implicated in regulating cell cycle, apoptosis and cancer. European journal of cancer (Oxford, England: 1990). 2005; 41(16):2479–84.
56. Funato T, et al. Use of c-myb antisense oligonucleotides to increase the sensitivity of human colon cancer cells to cisplatin. Oncology reports. 2001; 8(4):807–10. [PubMed: 11410788]
57. Kauraniemi P, et al. MYB oncogene amplification in hereditary BRCA1 breast cancer. Cancer research. 2000; 60(19):5323–8. [PubMed: 11034064]
58. Introna M, Golay J. How can oncogenic transcription factors cause cancer: a critical review of the myb story. Leukemia: official journal of the Leukemia Society of America, Leukemia Research Fund, U K. 1999; 13(9):1301–6.
59. Engelmann D, Putzer BM. Translating DNA damage into cancer cell death-A roadmap for E2F1 apoptotic signalling and opportunities for new drug combinations to overcome chemoresistance. Drug resistance updates: reviews and commentaries in antimicrobial and anticancer chemotherapy. 2010; 13(4–5):119–31. [PubMed: 20675184]
60. Poplawski P, Nauman A. Thyroid hormone -triiodothyronine -has contrary effect on proliferation of human proximal tubules cell line (HK2) and renal cancer cell lines (Caki-2, Caki-1) -role of E2F4, E2F5 and p107, p130. Thyroid research. 2008; 1(1):5. [PubMed: 19014670]

61. Macaluso M, et al. pRb2/p130-E2F4/5-HDAC1-SUV39H1-p300 and pRb2/p130-E2F4/5-HDAC1-SUV39H1-DNMT1 multimolecular complexes mediate the transcription of estrogen receptor-alpha in breastcancer. Oncogene. 2003; 22(23):3511–7. [PubMed: 12789259]
62. Garneau H, et al. Nuclear expression of E2F4 induces cell death via multiple pathways in normal human intestinal epithelial crypt cells but not in colon cancer cells. American journal of physiology Gastrointestinal and liver physiology. 2007; 293(4):G758–72. [PubMed: 17656449]
63. Garneau H, et al. E2F4 expression is required for cell cycle progression of normal intestinal crypt cells and colorectal cancer cells. Journal of cellular physiology. 2009; 221(2):350–8. [PubMed: 19562678]
64. Junttila MR, et al. Selective activation of p53-mediated tumour suppression in high-grade tumours. Nature. 2010; 468(7323):567–71. [PubMed: 21107427]
65. Berns A. Cancer: The blind spot of p53. Nature. 2010; 468(7323):519–20. [PubMed: 21107421]
66. Melino G. Journal club. Acancer biologist weighs up p53, metabolism and cancer. Nature. 2010; 466(7309):905. [PubMed: 20725003]
67. Mott JL, et al. mir-29 regulates Mcl-1 protein expression and apoptosis. Oncogene. 2007; 26(42): 6133–40. [PubMed: 17404574]
68. Pekarsky Y, et al. Tcl1 expression in chronic lymphocytic leukemia is regulated by miR-29 and miR-181. Cancer research. 2006; 66(24):11590–3. [PubMed: 17178851]
69. Fabbri M, et al. MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104(40):15805–10. [PubMed: 17890317]
70. Park SY, et al. miR-29 miRNAs activate p53 by targeting p85 alpha and CDC42. Nature structural & molecular biology. 2009; 16(1):23–9.
71. Zhao JJ, et al. microRNA expression profile and identification of miR-29 as a prognostic marker and pathogenetic factor by targeting CDK6 in mantle cell lymphoma. Blood. 2010; 115(13):2630–9. [PubMed: 20086245]
72. Li J, et al. miR-30 regulates mitochondrial fission through targeting p53 and the dynamin-related protein-1 pathway. PLoS genetics. 2010; 6(1):e1000795. [PubMed: 20062521]
73. Yu F, et al. Mir-30 reduction maintains self-renewal and inhibits apoptosis in breast tumor-initiating cells. Oncogene. 2010; 29(29):4194–204. [PubMed: 20498642]
74. Wu F, et al. MicroRNA-mediated regulation of Ubc9 expression in cancer cells. Clinical cancer research: an official journal of the American Association for Cancer Research. 2009; 15(5):1550–7. [PubMed: 19223510]
75. Zhang X, et al. MicroRNA-19 (miR-19) regulates tissue factor expression in breast cancer cells. The Journal of biological chemistry. 2011; 286(2):1429–35. [PubMed: 21059650]
76. Cao Y, et al. MicroRNA-dependent regulation of PTEN after arsenic trioxide treatment in bladder cancer cell line T24. Tumour biology: the journal of the International Society for Oncodevelopmental Biology and Medicine. 2011; 32(1):179–88. [PubMed: 20857258]
77. Brase JC, et al. Serum microRNAs as non-invasive biomarkers for cancer. Molecular cancer. 2010; 9:306. [PubMed: 21110877]
78. Hashimoto Y, et al. Involvement of epigenetically silenced microRNA-181c in gastric carcinogenesis. Carcinogenesis. 2010; 31(5):777–84. [PubMed: 20080834]
79. Huang XH, et al. Bead-based microarray analysis of microRNA expression in hepatocellular carcinoma: miR-338 is downregulated. Hepatology research: the official journal of the Japan Society of Hepatology. 2009; 39(8):786–94. [PubMed: 19473441]
80. Incoronato M, et al. miR-212 increases tumor necrosis factor-related apoptosis-inducing ligand sensitivity in non-small cell lung cancer by targeting the antiapoptotic protein PED. Cancer research. 2010; 70(9):3638–46. [PubMed: 20388802]
81. Shohet JM, et al. A Genome-Wide Search for Promoters That Respond to Increased MYCN Reveals Both New Oncogenic and Tumor Suppressor MicroRNAs Associated with Aggressive Neuroblastoma. Cancer research. 2011; 71(11):3841–51. [PubMed: 21498633]
82. Lee YN, et al. KIT signaling regulates MITF expression through miRNAs in normal and malignant mast cell proliferation. Blood. 2011; 117(13):3629–40. [PubMed: 21273305]

83. Sun L, et al. Neuronal and glioma-derived stem cell factor induces angiogenesis within the brain. Cancer cell. 2006; 9(4):287–300. [PubMed: 16616334]
84. Wong YF, et al. Expression genomics of cervical cancer: molecular classification and prediction of radiotherapy response by DNA microarray. Clinical cancer research: an official journal of the American Association for Cancer Research. 2003; 9(15):5486–92. [PubMed: 14654527]
85. Skotheim RI, et al. Differentiation of human embryonal carcinomas in vitro and in vivo reveals expression profiles relevant to normal development. Cancer research. 2005; 65(13):5588–98. [PubMed: 15994931]
86. Kimchi ET, et al. Progression of Barrett's metaplasia to adenocarcinoma is associated with the suppression of the transcriptional programs of epidermal differentiation. Cancer research. 2005; 65(8):3146–54. [PubMed: 15833844]
87. Chen X, et al. Variation in gene expression patterns in human gastric cancers. Molecular biology of the cell. 2003; 14(8):3208–15. [PubMed: 12925757]
88. Hippo Y, et al. Global gene expression analysis of gastric cancer by oligonucleotide microarrays. Cancer research. 2002; 62(1):233–40. [PubMed: 11782383]
89. Cromer A, et al. Identification of genes associated with tumorigenesis and metastatic potential of hypopharyngeal cancer by microarray analysis. Oncogene. 2004; 23(14):2484–98. [PubMed: 14676830]
90. Beer DG, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nature medicine. 2002; 8(8):816–24.
91. Jones MH, et al. Two prognostically significant subtypes of high-grade lung neuroendocrine tumours independent of small-cell and large-cell neuroendocrine carcinomas identified by gene expression profiles. Lancet. 2004; 363(9411):775–81. [PubMed: 15016488]
92. Talantov D, et al. Novel genes associated with malignant melanoma but not benign melanocytic lesions. Clinical cancer research: an official journal of the American Association for Cancer Research. 2005; 11(20):7234–42. [PubMed: 16243793]
93. Gordon GJ, et al. Identification of novel candidate oncogenes and tumor suppressors in malignant pleural mesothelioma using large-scale transcriptional profiling. The American journal of pathology. 2005; 166(6):1827–40. [PubMed: 15920167]
94. Ishikawa M, et al. Experimental trial for diagnosis of pancreatic ductal carcinoma based on gene expression profiles of pancreatic ductal cells. Cancer science. 2005; 96(7):387–93. [PubMed: 16053509]
95. Dhanasekaran SM, et al. Delineation of prognostic biomarkers in prostate cancer. Nature. 2001; 412(6849):822–6. [PubMed: 11518967]
96. Lapointe J, et al. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101(3):811–6. [PubMed: 14711987]
97. Nanni S, et al. Epithelial-restricted gene profile of primary cultures from human prostate tumors: a molecular approach to predict clinical behavior of prostate cancer. Molecular cancer research: MCR. 2006; 4(2):79–92. [PubMed: 16513839]
98. Singh D, et al. Gene expression correlates of clinical prostate cancer behavior. Cancer cell. 2002; 1(2):203–9. [PubMed: 12086878]
99. Varambally S, et al. Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression. Cancer cell. 2005; 8(5):393–406. [PubMed: 16286247]
100. Boer JM, et al. Identification and classification of differentially expressed genes in renal cell carcinoma by expression profiling on a global human31,500-element cDNA array. Genome research. 2001; 11(11):1861–70. [PubMed: 11691851]
101. Lenburg ME, et al. Previously unidentified changes in renal cell carcinoma gene expression identified by parametric analysis of microarray data. BMC cancer. 2003; 3:31. [PubMed: 14641932]
102. Detwiller KY, et al. Analysis of hypoxia-related gene expression in sarcomas and effect of hypoxia on RNA interference of vascular endothelial cell growth factor A. Cancer research. 2005; 65(13):5881–9. [PubMed: 15994966]

103. Reyes, I., et al. Identification of kallikrein 7, kallikrein 10 and secreted frizzled-related protein 2 as candidate molecular markers for papillary thyroid carcinoma using microarray analysis [abstract]. the 96th Annual Meeting of the American Association for Cancer Research; 2005; Anaheim, CA.

104. Hoffman PJ, et al. Molecular characterization of uterine fibroids and its implication for underlying mechanisms of pathogenesis. Fertility and sterility. 2004; 82(3):639–49. [PubMed: 15374708]

105. Quade BJ, et al. Molecular pathogenesis of uterine smooth muscle tumors from transcriptional profiling. Genes, chromosomes & cancer. 2004; 40(2):97–108. [PubMed: 15101043]

Table 1

Summary of human tumor gene expression datasets

Tumor Type	# Datasets	Reference
Brain Cancer	1	[83]
Cervical Cancer	1	[84]
Embryonal Cancer	1	[85]
Esophageal Cancer	1	[86]
Gastric Cancer	2	[87–88]
Head and Neck Cancer	1	[89]
Lung Cancer	2	[90–91]
Melanoma	1	[92]
Mesothelioma	1	[93]
Pancreatic Cancer	1	[94]
Prostate Cancer	5	[95–99]
Renal Cancer	2	[100–101]
Soft Tissue Sarcoma	1	[102]
Thyroid Cancer	1	[103]
Uterine Leiomyoma	2	[104–105]

Table 2

Seventeen pathways frequently identified in tumors

Pathway Symbol[a]	Pathway Name[a]	Frequency[b]
Vitcb	Vitamin C in the Brain	14
Mhc	Antigen Processing and Presentation	13
No1	Actions of Nitric Oxide in the Heart	13
Keratinocyte	Keratinocyte Differentiation	12
Cbl	Mediated ligand-induced downregulation of EGF receptors	12
Biopeptides	Bioactive Peptide Induced Signaling Pathway	12
Cell Cycle	Cyclins and Cell Cycle Regulation	11
Cftr	Cystic fibrosis transmembrane conductance regulator (CFTR) and beta 2 adrenergic receptor (b2AR) pathway	11
Erad	ER-associated degradation Pathway	11
Arenrf2	Oxidative Stress Induced Gene Expression Via Nrf2	11
Vobesity	Visceral Fat Deposits and the Metabolic Syndrome	11
PlateletApp	Platelet Amyloid Precursor Protein Pathway	11
Mcalpain	mCalpain and friends in Cell motility	10
Ranms	Role of Ran in mitotic spindle regulation	10
P38 Mapk	p38 MAPK Signaling Pathway	10
Tcra	Lck and Fyn tyrosine kinases in initiation of TCR Activation	10
Dsp	Regulation of MAP Kinase Pathways Through Dual Specificity Phosphatases	10

[a] BioCarta pathway symbol and name.

[b] The occurrence times of the pathway in the 25 pathway sets identified.

Table 3

Twenty-two TFs frequently identified in tumors

TF	Frequency
c-Myc	16
E2F-4	15
MYB	15
P53	15
SMAD1	14
TAL1	13
TFAP2A	11
JUN	11
PPARA	11
MYBL2	10
SP2	10
EPAS1	10
FLI1	10
STAT3	10
E2F-1	9
PPARD	9
STAT5B	9
NFIC	9
POU2F2	9
PGR	9
ETS2	9
HIF1A	9

Table 4

Thirty-four miRNAs frequently identified in tumors

miRNA	Frequency
miR-29b	18
miR-29c	17
miR-29a	16
miR-30e-3p	14
miR-547	14
miR-181c	13
miR-30e-5p	13
miR-590	13
miR-212	13
miR-603	13
miR-669b	13
miR-338	13
miR-202	13
miR-1	12
miR-103	12
miR-181a	12
miR-19a	12
miR-19b	12
miR-200b	12
miR-21	12
miR-30a-5p	12
miR-30c	12
miR-30d	12
miR-330	12
miR-518a	12
miR-526c	12
miR-562	12
miR-31	12
miR-128a	12
miR-30a-3p	12
miR-182	12
miR-669c	12
miR-526b	12
miR-95	12

Table 5

Thirty-nine GO terms frequently identified in tumors

GO Terms	Frequency
antigen processing and presentation of peptide antigen	7
MHC protein complex	7
antigen processing and presentation of peptide antigen via MHC class I	6
MHC class I receptor activity	6
cytosolic ribosome	6
extracellular matrix part	6
collagen	6
fibrillar collagen	6
cell proliferation	6
regulation of cell proliferation	6
regulation of biological quality	6
basement membrane	5
collagen metabolic process	5
ribosomal subunit	5
extracellular matrix structural constituent	5
peptide cross-linking	5
glycosaminoglycan binding	5
small ribosomal subunit	5
blood vessel development	5
vasculature development	5
response to external stimulus	5
response to wounding	5
anatomical structure morphogenesis	5
tissue development	5
response to organic substance	5
cell differentiation	5
response to chemical stimulus	5
cell development	5
organ development	5
cellular developmental process	5
proteinaceous extracellular matrix	5
extracellular space	5
extracellular matrix	5
extracellular region part	5
structural molecule activity	5
cytoskeletal protein binding	5
cytoskeleton	5

GO Terms	Frequency
identical protein binding	5
response to stress	5