



Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN FARMACOLOGIA E ONCOLOGIA MOLECOLARE

CICLO XXV

COORDINATORE Prof. ANTONIO CUNEO

A systems biology approach to non-coding RNAs: the networks of cancer

Settore Scientifico Disciplinare BIO/17

Dottorando

Dott. Galasso Marco

Tutore

Prof. Volinia Stefano

Anni 2010/2012

LIST OF PUBLICATIONS

The thesis is based on the following papers and reviews, plus on the projected AIRC 2010 (IG8588):

- I Identification of microRNA activity by Targets' Reverse EXpression.
Volinia S, Visone R, **Galasso M**, Rossi E, Croce CM. *Bioinformatics*. 2010 Jan 1;26(1):91-7. Epub 2009 Nov 6.
- II Non-coding RNAs: a key to future personalized molecular therapy?
Galasso M, Elena Sana M, Volinia S. *Genome Med*. 2010 Feb 18;2(2):12
- III Reprogramming of miRNA networks in cancer and leukemia. Volinia S, **Galasso M**, Costinean S, Tagliavini L, Gamberoni G, Drusco A, Marchesini J, Mascellani N, Sana ME, Abu Jarour R, Despons C, Teitell M, Baffa R, Aqeilan R, Iorio MV, Taccioli C, Garzon R, Di Leva G, Fabbri M, Catozzi M, Pleviati M, Ambs S, Palumbo T, Garofalo M, Veronese A, Bottoni A, Gasparini P, Harris CC, Visone R, Pekarsky Y, de la Chapelle A, Bloomston M, Dillhoff M, Rassenti LZ, Kipps TJ, Huebner K, Pichiorri F, Lenze D, Cairo S, Buendia MA, Pineau P, Dejean A, Zaneni N, Rossi S, Calin GA, Liu CG, Palatini J, Negrini M, Vecchione A, Rosenberg A, Croce CM. *Genome Res*. 2010 May;20(5):589-99.
- IV GAMES identifies and annotates mutations in next-generation sequencing projects.
Sana ME, Iascone M, Marchetti D, Palatini J, **Galasso M**, Volinia S. *Bioinformatics*. 2011 Jan 1;27(1):9-13. Epub 2010 Oct 22. PMID: 20971986
- V Prion proteins (PRNP and PRND) are over-expressed in osteosarcoma.
Sollazzo V, **Galasso M**, Volinia S, Carinci F. *J Orthop Res*. 2011 Dec 6. doi: 10.1002/jor.22034. [Epub ahead of print]
- VI Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA.
Volinia S, **Galasso M**, Sana ME, Wise TF, Palatini J, Huebner K, Croce CM. *Proc Natl Acad Sci U S A*. 2012 Feb 21;109(8):3024-9. Epub 2012 Feb 6.
- VII MicroRNA signatures associate with pathogenesis and progression of osteosarcoma.
Jones KB, Salah Z, Del Mare S, **Galasso M**, Gaudio E, Nuovo GJ, Lovat F, Leblanc K, Palatini J, Randall RL, Volinia S, Stein GS, Croce CM, Lian JB, Aqeilan RI. *Cancer Res*. 2012 Feb 20.

VIII MicroRNA expression signatures in solid malignancies. **Galasso M**, Sandhu SK, Volinia S. *Cancer J*. 2012 May-Jun;18(3):238-43. PMID:22647360

IX miR-155 targets histone deacetylase 4 (HDAC4) and impairs transcriptional activity of B-cell lymphoma 6 (BCL6) in the E μ -miR-155 transgenic mouse model. Sandhu SK, Volinia S, Costinean S, **Galasso M**, Neinast R, Santhanam R, Parthun MR, Perrotti D, Marcucci G, Garzon R, Croce CM. *Proc Natl Acad Sci U S A*. 2012 Dec 4;109(49):20047-52. doi: 10.1073/pnas.1213764109. Epub 2012 Nov 19. PMID: 23169640 [PubMed - in process]

X Highly-Specific Transcribed-Ultra Conserved Regions in Human Embryonic Stem Cells. **Galasso M**, Dama P, Previati M, Sandhu S, Palatini J, Wise T, Warner S, Sana ME, Zanella R, Abujarour R, Despons C, Teitell MA, Garzon R, Calin G, Croce CM & Volinia S. *Submitted*

XI Estrogen mediated-activation of miR-191/425 cluster modulates tumorigenicity of breast cancer cells depending on Estrogen Receptor status. Di Leva G, Piovan C, Gasparini P, Ngankeu A, Taccioli C, Briskin D Douglas GC, Bolon B, Anderlucci L, Hansjuerg A, Nuovo G, Li M, Iorio MV, **Galasso M**, Ramasamy S, Marcucci G, Perrotti D, Powell AK, Brasatz A, Garofalo M, Nephew PK, and Croce MC. *In Press PLOS Genetics*

CONTENTS:

SUMMARY	vii
ACKNOWLEDGMENTS	ix

1 INTRODUCTION

1.1 Gene regulation and post-transcriptional regulation	1
1.2 Non-coding RNAs	2
1.3 Transcribed ultraconserved region	5
1.4 microRNA biogenesis and function	6
1.5 microRNAs : oncogene and tumor suppressor roles	7
1.6 microRNA signatures in solid cancer and leukemia	9
1.7 microRNA target genes	15
1.8 High-throughput technology	18
1.9 Regulatory networks	24

2 THESIS OBJECTIVES	26
3 THE TISSUE SPECIFICITY OF T-UCRs AND miRNAs	
3.1 Material and Methods	27
3.1.1 Data analysis	28
3.1.2 Calculation of Information Content	29
3.1.3 Decision Tree	30
3.2 miRNAs tissue specificity in normal and embryonic samples	32
3.3 T-UCRs tissue specificity in normal and embryonic samples	35
4 CLASSIFICATION OF SOLID CANCER BY miRNA EXPRESSION	
4.1 Material and Methods	41
4.1.1 Statistical Tests and Algorithms	41
4.2 miRNAs signatures in cancer samples	42
4.3 T-UCRS signatures in cancer samples	44

4.4 Identification of putative “classifier” miRNA	45
4.5 Feature selection and classification	49
5 REPROGRAMMING OF miRNA NETWORK	
5.1 Material and Methods	53
5.1.1 Data analysis	53
5.1.2 Network generation	53
5.2 miRNA Bayesian network of normal samples	56
5.3 miRNA Bayesian network of solid cancers	58
5.4 miRNA Bayesian network of leukemia	63
6 IN SILICO, IN-VITRO AND IN-VIVO VALIDATION	
6.1 Material and Methods	66
6.1.1 Cell line culture	67
6.1.2 Isolation and Reverse transcription of a total RNA	67

6.1.3 Real-time PCR	68
6.2 Real-Time PCR confirms the uc.283 plus specificity for ESC	68
6.3 Functional genome analysis: in silico validation	72
6.4 In vivo validation	75
7 FINAL DISCUSSION	77
REFERENCES	79
APPENDIX A	88
APPENDIX B	105

SUMMARY

A non-coding RNA is a functional RNA molecule that is not translated into a protein. This class of molecules is involved in many cellular processes and includes highly abundant and functionally important RNAs such as transfer RNA (tRNA), ribosomal RNA (rRNA), as well as small interfering RNAs (siRNAs), microRNAs (miRNAs) and transcribed ultraconserved regions (T-UCRs). First of all, we investigate the specificity of each of the classes studied of two selected non-coding RNA: Transcribed UltraConserved Region and microRNAs. second, we want to find whether these non-coding RNAs can be candidates as features for the selection of specific cancers, using statistical algorithms and bioinformatics tools. Third, we generate miRNA gene networks in normal and different cancer and leukemia. The term “ultraconserved” refer to genomic regions longer than 200 bp that are absolutely conserved (100% homology with no insertions or deletions) in human, mouse, and rat genomes. There are 481 T-UCRs. The reason for this extreme conservation remains a mystery; T-UCRs may play a functional role in the ontogeny and phylogeny of mammals and other vertebrates. Genome-wide profiling revealed that UCRs are frequently located on overlapping exons in genes involved in RNA processing and can be found in introns or at fragile sites and in cancer-associated genomic regions. We investigate the expression of T-UCRs in 374 normal samples from 46 different tissues, grouped by 16 systems. Moreover, we analyzed the tissue specificity of T-UCRs in cancers. Tissue specific T-UCRs can differentiate cell types. We then examine the expression of T-UCRs in human embryonic stem cells, induced pluripotent stem cells, and a series of differentiated cell types (trophoblast, embryoid bodies at 7 and 14 days of differentiation, definitive endoderm, and spontaneous differentiated monolayers). One T-UCR in particular, uc.283 plus, is highly specific for embryonic and induced pluripotent stem cells, as confirmed by real time PCR (RT-PCR). miRNAs are global regulators of protein output. Each miRNA has been studied for its single contribution to differential expression or to a compact

predictive signature. Thus, we propose a study of miRNAs in cancer by applying a systems biology approach. We study miRNA profiles in 4419 human samples (3312 neoplastic, 1107 non-malignant), corresponding to 50 normal tissues (grouped by 17 systems) and 51 cancer types. We calculate tissue specificity and cancer type specificity, a small set of miRNAs were tissue-specific while many others were broadly expressed. Then we find whether non-coding RNAs can be candidates as features for the selection of specific cancers, using statistical algorithms and bioinformatics tools, as decision trees. Afterwards, we build miRNA gene networks by using our very large expression miRNA database. The complexity of our expression database enable us to perform a detailed analysis of coordinated miRNA activities. We also build specialized miRNA networks for different solid tumors and leukemias. Combining differential expression, genetic networks, DNA copy number alterations and other systems biology approaches we confirm or discovered, miRNAs with comprehensive roles in cancer. Normal tissues are represented by single complete miRNA networks. Cancers instead show separate and unlinked miRNA sub-networks. miRNAs independent from the general transcriptional program were often known as cancer-related. We validate our finding by *in silico*, *in vitro* and *in vivo* analysis. The target genes of these uncoordinated miRNA involve in specific cancer-related pathways.

ACKNOWLEDGMENTS

*To Stefano and Sara,
to my family and my friends,
to all those believe in me...*

“ The future belongs to those who believe in the beauty of their dreams”

Anna Eleanor Roosevelt

1 INTRODUCTION

1.1 Gene regulation and Post-transcriptional regulation

Different cell types cells in a multicellular organism are capable of altering their patterns of gene expression in response to extracellular signals [1]. There are many steps in the pathway leading DNA to protein, and all of them can in principle be regulated. Mechanisms that underlie the control of gene expression are becoming increasingly well understood. There were principally six steps at which eukaryote gene expression can be controlled: 1) transcriptional control; 2) RNA processing control; 3) RNA transport control; 4) post-transcriptional control; 5) translational control; 6) post-translational modification of the protein into its mature, functional form control. We will focus on the mechanism of post-transcriptional regulation mediated by non-coding RNAs.

1.2 Non-Coding RNAs

A non-coding RNA (ncRNA) is a functional RNA molecule that is not translated into a protein. This class of molecules is involved in many cellular processes and includes highly abundant and functionally important RNAs such as transfer RNA (tRNA) and ribosomal RNA (rRNA), as well as small interfering RNAs (siRNAs), microRNAs (miRNAs), PIWI-associated RNAs (piRNAs), small nucleolar RNAs (snoRNAs), promoter-associated RNAs (PARs), the recently identified telomere specific small RNAs (tel-sRNAs), long non-coding RNAs (lncRNAs) and transcribed ultraconserved region (T-UCRs). Moreover, this broad family of ncRNAs contributes to molecular alterations in a number of pathological conditions [2].

SiRNAs are small RNAs, 21–22 nucleotides (nt) long, produced by Dicer cleavage of complementary dsRNA duplexes. siRNAs form complexes with Argonaute proteins and are involved in gene regulation, transposon control and viral defense.

PiRNAs, 24- 30 nt in length, are produced by a Dicer-independent mechanism, associate with Piwi-class Argonaute proteins and are principally restricted to the germline and bordering somatic cells. piRNAs are important for transposon control [3, 4] and regulate chromatin state [5]. A recent study suggests that an antisense RNA may trigger transcriptional silencing of a partner sense tumor suppressor gene; this effect occurs both in cis and in trans, and is Dicer-independent. The biochemical mediators of this silencing involve a Piwi-like protein, as their role in mammals is just beginning to be understood [6, 7]. In zebrafish, piRNAs have been implicated in germ cell maintenance and many of them were mapped to transposons suggesting that they play a role in silencing mechanisms of repetitive elements in vertebrates [8]. At the moment, no relations between piRNAs and diseases have yet been disclosed.

SnoRNAs are small RNA molecules, approximately 60-300 nt long, which generally serve as guides for the catalytic modification of selected ribosomal RNAs [9, 10]. In vertebrates, most snoRNAs have been shown to reside in introns of protein coding host genes and are processed out of the excised introns. snoRNAs can be divided

into two large groups, called box C/D snoRNAs and box H/ACA snoRNAs. Most of them are responsible for guiding the 2'-O-ribose methylations (box C/D snoRNAs) and pseudouridylations (box H/ACA snoRNAs) of ribosomal RNAs, transfer RNAs and small nuclear RNAs [11]. Many snoRNAs have been described as retrogenes [12]. Some snoRNA is processed to a small RNA that can function like a miRNA [13]. PARs encompass a suite of long and short RNAs, including promoter-associated small RNAs (PASRs) and transcriptional initiation RNAs, that overlap promoter regions. Their function is so far unknown but they possibly regulate transcription, as exogenous PASRs have been observed to reduce expression of genes with homologous promoter sequences [14].

Tel-sRNAs in mouse embryonic stem cells are approximately 24 nucleotides in length, Dicer-independent, and 2'-O-methylated at the 3' terminus. The tel-sRNAs are asymmetric with specificity toward telomere G-rich strand, are evolutionarily conserved from protozoan to mammals and they may have a role in telomere maintenance [15].

LncRNAs are often longer than 2 kb in length with a coding potential of less than 100 amino acids and the number of lncRNAs exceeds that of protein coding genes [16, 17]. Recently, several studies identify that many lncRNAs are known to be retained in various sub-nuclear compartments [18] suggesting that such RNAs may have a potential function in the compartment where they are localized [19].

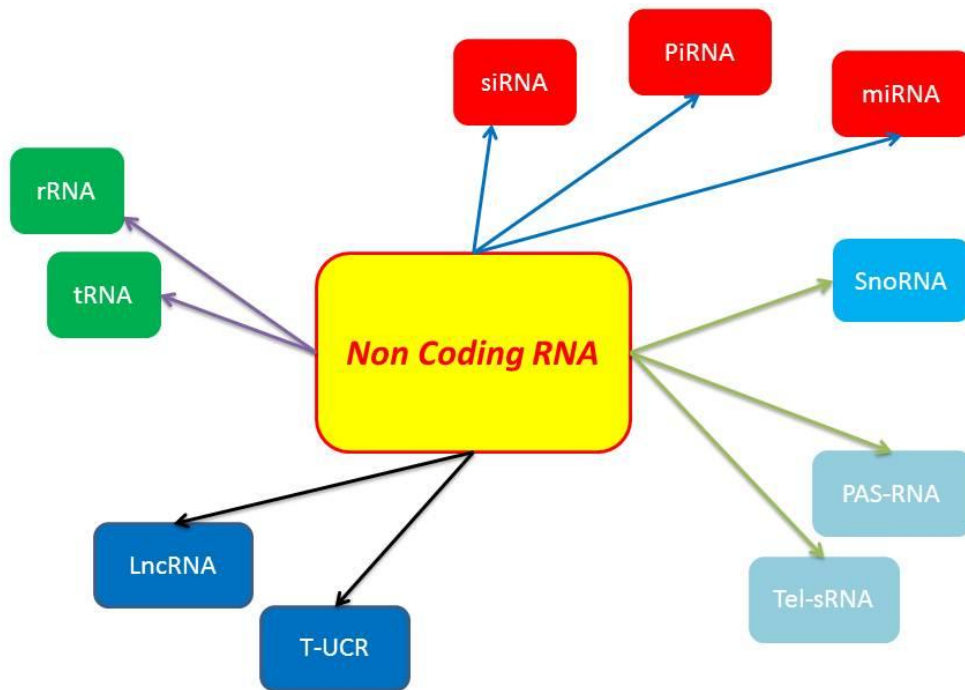


Figure 1 Class of Non Coding RNAs.

1.3 Transcribed Ultraconserved Regions

The term “ultraconserved” was originally proposed by Bejerano and colleagues to refer to genomic regions longer than 200 bp that are absolutely conserved (100% homology with no insertions or deletions) in human, mouse, and rat genomes [20]. Many of these elements possess tissue-specific enhancer activity [21-23] and others have been shown to associate with splicing regulators. Evolutionary conservation has become a powerful tool to identify functionally important regions in the human genome[24] . A greater percentage of the shorter elements show extreme conservation within mammals only and nearly 47% of ultraconserved elements in human have been localized to exons of genes involved in RNA processing or in the regulation of transcription and development [25].

The reason for this extreme conservation remains a mystery; but it suggests that UCRs may play a functional role in the ontogeny and phylogeny of mammals and other vertebrates. This idea is supported by findings that have identified a distal enhancer and an ultraconserved exon derived from a novel retroposon active more than 400 million years ago in lobe-finned fishes and terrestrial vertebrates, and maintained as active in the “living fossil” coelacanth. A recent study showed the concurrent presence of enhancer and transcript functions in non-exonic UCRs, and suggested they may be long ncRNAs [26].

Scientific interest in T-UCRs has significantly increased when they were shown to be involved in human carcinogenesis. Studies suggested that UCRs could contribute to the development of several types of malignancies [27]. Genome-wide profiling revealed that UCRs have distinct signatures in human leukemias and carcinomas [28] and are frequently located at fragile sites and in cancer associated genomic regions [29]. Recent clinical findings also suggest that T-UCR signatures can have prognostic value in high-risk neuroblastoma patients [30]. It has been shown that

some T-UCRs are deregulated in neuroblastomas, which could provide additional prognostic value in conjunction with MYC activity/amplification, currently used as a clinical prognostic indicator [31].

1.4 MicroRNA biogenesis and functions

miRNAs are a conserved class of non-coding RNAs that regulate the translation of mRNAs (messenger RNAs) by inhibiting ribosome function, decapping the 5'Cap structure, de-adenylating the poly(A) tail and degrading the target mRNAs [32]. MiRNAs are involved in mechanisms of gene regulation in both normal and diseased conditions and play a role during development, regulation of cell proliferation and apoptosis. The first miRNAs were identified in the nematode *Caenorhabditis elegans*, as small RNAs that interacted with the 3' untranslated region (UTR) of the *lin-14* mRNA to inhibit its expression [33]. miRNAs are single-stranded RNAs of 19-24 nucleotides (nt) in length generated through a complex maturation process. MiRNAs are transcribed by RNA polymerase II as molecules of variable length, the so-called primary-miRNA (pri-miRNA). The pri-miRNA is cleaved by Drosha - a RNase III endonuclease that recognizes internal hairpin structures - and is then actively exported by Exportin-5 to the cytoplasm as a miRNA precursor (pre-miRNA) of approximately 70-120 nt. Once in the cytoplasm, the pre-miRNA is further digested by the enzyme Dicer (another Ribonuclease III), that yields a 21-22 nt long double-stranded RNA (dsRNA) with 2- nt long overhangs [34]. The active single-stranded molecule(s) from this dsRNA associate(s) with several proteins to form the RNA-induced silencing complex (RISC), which includes proteins of the Argonaute family [6]. miRNAs bind mostly to mRNA segments originating from the 3' UTRs of genes [35] but the mechanism of translational repression is only partially understood.

1.5 microRNAs : oncogene and tumor suppressor roles

The number of miRNAs in the human genome has been steadily increasing, and is now of 1600 (miRBase Release 19). Since their discovery, miRNAs quickly became the protagonists in the limelight of cancer research. By targeting and controlling the expression of mRNA, miRNAs can control highly complex signal-transduction pathways and other biological pathways. The biologic roles of miRNAs in cancer suggest a correlation with prognosis and therapeutic outcome. Dysregulation of physiologic microRNA (miR) activity has been shown to play an important role in tumor initiation and progression, the identification of differentially regulated miRNAs in cancer suggested a functional involvement in carcinogenesis, assigning them at times roles as tumor suppressors or as oncogenes.

Those miRNAs whose expression is increased in tumors may be considered as oncogenes - also called "oncomiRs" - which promote tumor development by inhibiting tumor suppressor genes and/or genes controlling cell differentiation or apoptosis.

The miR-17-92 cluster is a miRNA polycistron located at chromosome 13q31, a genomic locus that is amplified in lung cancer and several kinds of lymphoma, including diffuse large B-cell lymphoma. This cluster has been found to be regulated by c-Myc, an important transcription factor that is over-expressed in many human cancers .

Microarray analysis revealed that these miRNAs may act as anti-apoptotic factors in human malignant disease. In several types of lymphomas, including Burkitt's lymphoma, the expression of miR-155 is increased. Furthermore, miR-155 is located in the only phylogenetically-conserved region of the BIC (B-cell receptor inducible) gene, suggesting that miR-155 may be responsible for BIC's oncogenic activity [36]. Conversely, under-expressed miRNAs in cancers, such as some members of the let-7 family, may function as tumor suppressor genes by regulating oncogenes and/or genes that control cell differentiation or apoptosis. Studies indicate that the RAS oncogene is a direct target of let-7, which negatively regulates RAS by pairing to its

3' UTR for translational repression. Some recent studies have focused on microRNA-binding site polymorphisms. Let-7 microRNA-binding site polymorphisms in the KRAS 3' UTR have been associated with reduced survival in oral cancers [37]. Other miRNAs that play a potential role as tumor suppressors include miR-15 and miR-16, which induce apoptosis by targeting anti-apoptotic gene BCL-2 (B-cell lymphoma 2) mRNA [38], MCL1 (myeloid cell leukemia sequence 1), CCND1 (cyclin D1), and WNT3A (wingless-type MMTV integration site family, member 3A) [39]. MiRNAs are also involved in advanced stages of tumor progression, stressing their roles as metastasis activators or suppressors [40].

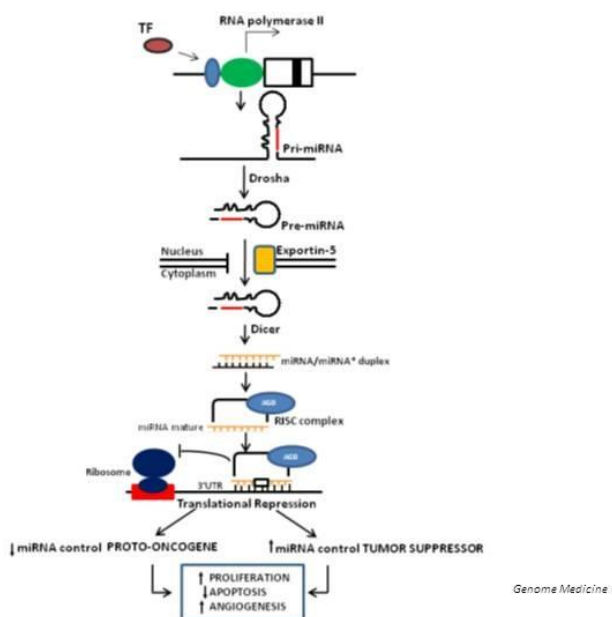


Figure 2 Schematic overview of microRNA processing and functional roles in cancer.

1.6 microRNAs signatures in solid cancers and leukemia

The identification of differentially regulated miRNAs in cancer suggested a functional involvement in carcinogenesis, assigning them at times roles as tumor suppressors or as oncogenes. Such involvement exerts different effects based on the identity of the regulated target genes within the affected tissues. To date, several studies have addressed the different roles of miRNA in cancer establishment and progression. There are many relevant discoveries for each one of the major human solid cancer types. Table A1 is an overview of prominent miRNAs associated to solid cancers. Lung cancer is one of the first solid tumors in which microRNA expression patterns have been studied extensively. Lebanony et al. developed a diagnostic assay based on miR-205 expression that can distinguish squamous from adenocarcinoma non-small-cell lung cancer (NSCLCs), with high sensitivity and specificity [41]. Knockdown of miR-31 represses lung adenocarcinomas cell clonal growth and in vivo tumorigenicity, via LATS2 and PPP2R2A, two protein regulating tumor growth [42]. miR-21 was found to drive lung tumorigenesis through inhibition of negative regulators of the Ras/MEK/ERK pathway and inhibition of apoptosis. Let-7 was found to play the opposite role, acting as a tumor suppressor, partially through functional targeting of RAS, HMAG2 [43] and MYC [44]. Furthermore, reduced *let-7* gene expression in NSCLC patients was correlated with poor prognosis [45, 46]; and a SNP in a *let-7* complementary site in the KRAS 3' UTR region increased NSCLC risk [47]. Pancreatic cancer (PC) is a highly aggressive malignancy with worst prognosis. The first approach used for identification of miRNA was the expression profiling of human specimens pancreatic adenocarcinoma paired normal pancreas, chronic pancreatitis and pancreatic cancer cell lines. Several studies have recently shown the deregulation of miRNA expression in PC tumor tissues using miRNA array technology and real-time PCR analysis [48-50]. These works identified, through independent different studies, a group of miRNAs up-regulated that increase

proliferation, tumorigenesis and invasion: miR-21 [51], miR-155 [52], miR-27a [53], miR-210 [54], miR-221/222 [50] and miR-196a-2 and miR-10b, that also may be negative survival predictor [49, 55]. On the other hand, there is a group of microRNA down-regulated having the opposite role: miR-146a [56], miR-20a [57], miR-96 [58] and miR-375 [49]. Aberrant expression of this set of miRNAs were independently associated with reduced survival. In plasma and in serum of patient with PC microRNAs has been suggested as a novel noninvasive approach of prognostic biomarkers [59, 60]. Iorio et al [61] reported the first 13 miRNA expression signature in breast cancer that was able to precisely discriminate between normal and tumor tissue. MicroRNA expression has also been shown to relate to some histopathologic features of breast carcinoma, such as estrogen receptor (ER) and progesterone receptor expression, grade and stage, and presence of invasion. Studies have been undertaken to investigate the correlation between microRNA expression and the classification in different subtypes of breast cancer. Among the differentially expressed and consistently deregulated miRNAs in breast cancer, miR-10b, miR-125b, and miR-145 are down-regulated, while miR-21 and miR-155 are up-regulated which indicates their tumor suppressor or oncogenic roles, respectively. miR-21 is associated with advanced clinical stage, lymph node metastasis, and poor patient prognosis in breast cancer [62, 63]. miR-21 is known for regulation of cell survival and proliferation through direct targeting of tumor suppressor genes like PTEN, PDCD4 and TPM1. Iorio et al [61] reported the first 13 miRNA expression signature in breast cancer that was able to precisely discriminate between normal and tumor tissue. MicroRNA expression has also been shown to relate to some histopathologic features of breast carcinoma, such as estrogen receptor (ER) and progesterone receptor expression, grade and stage, and presence of invasion. Studies have been undertaken to investigate the correlation between microRNA expression and the classification in different subtypes of breast cancer. Among the differentially expressed and consistently deregulated miRNAs in breast cancer, miR-10b, miR-125b, and miR-145 are down-regulated, while miR-21 and miR-155 are up-regulated

which indicates their tumor suppressor or oncogenic roles, respectively. miR-21 is associated with advanced clinical stage, lymph node metastasis, and poor patient prognosis in breast cancer [62, 63]. miR-21 is known for regulation of cell survival and proliferation through direct targeting of tumor suppressor genes like PTEN, PDCD4 and TPM1. Next generation sequencing experiment proposed by Farazi et al. presented a comparative analysis of miRNA read frequencies from different histological subtypes of breast cancer (normal breast samples, most non-invasive ductal carcinoma in situ and invasive carcinomas) by deep sequencing. They showed and confirmed the increment of level of miR-21 and multiple decreased miRNA families, including miR-98/let-7 [64]. Afterward, Volinia et al. identified nine-microRNA that differentiated invasive from in situ ductal carcinoma: let-7d, miR-210, and -221 were down-regulated in the in situ and up-regulated in the invasive transition. miR-210 had also a prognostic meaning for overall survival and time to metastasis [65]. MiRNA signatures obtained from different histotypes of ovarian carcinomas (serous, endometrioid, clear cell, and mixed) uncovered several miRNAs which appear to be specific to the histotypes compared to the normal tissues. In endometrioid tumors, the four most significantly up-regulated miRs include miR-200a, miR-200b, miR-200c and miR-141 which were also present in other types of ovarian cancer [66]. In addition, miR-21, overexpressed in several solid tumors, where it exerts anti-apoptotic effect, was also found to be highly up-regulated. Alternatively, down modulation of microRNAs such as miR-222, which is known to target c-Kit, has been observed in endometrioid tumors. C-Kit is involved in cancer and is down-regulated in folate deficient conditions. Aberrant expression of several miRs: miR-214, miR-199a*, miR-200a, and miR-100 have been detected in a near or over half of ovarian cancers, especially in late-stage and high-grade tumors. Hepatocellular carcinomas (HCCs) may have a distinct miRNA expression fingerprint according to malignancy, risk factors, and oncogene/tumor suppressor gene alterations. Dissecting these relationships provides a new hypothesis to understand the functional impact of miRNA deregulation in liver tumorigenesis and the promising use of miRNAs as

diagnostic markers. Ladeiro et al. [67] analyzed the expression levels of 250 miRNAs in 46 benign and malignant hepatocellular tumors and compared to those of 4 normal liver samples with qRT-PCR. miRNAs associated with genetic and clinical characteristics were validated in a second series of 43 liver tumor samples and 16 non-tumor samples. Their study identified and validated miR-224 overexpression in all tumors and miR-200c, miR-200, miR-21, miR-224, miR-10b, and miR-222 specific deregulated in benign or malignant tumors. Lately, Pineau et al. identified a set of 12 miRNAs linked to disease progression and demonstrated that the over-expression of miR-221 was protumorigenic in vivo, stimulated growth of tumorigenic murine hepatic progenitor cells [68]. Altered expression of microRNAs have a functional role in the initiation and progression of colon cancer. Schetter et al., in one of the first studies with a large cohort of colon adenocarcinoma patients samples, identified thirty-seven microRNAs differentially expressed in tumors from the test cohort. Again, high levels of miR-21 was associated with poor survival and poor therapeutic outcome [69]. Lately, in a subsequent study they found systematic changes in eight inflammatory gene (*IL-8*, *IL-23a*, *IL-1a*, *IL-1b*, *IL-17a*, *INF γ* , *IL-6* and *FOXP3*) expression in colon tumors; that was associated with *miR-21* expression. They suggested that the combination of inflammatory genes and *miR-21* expression can be a better predictor of prognosis [70].

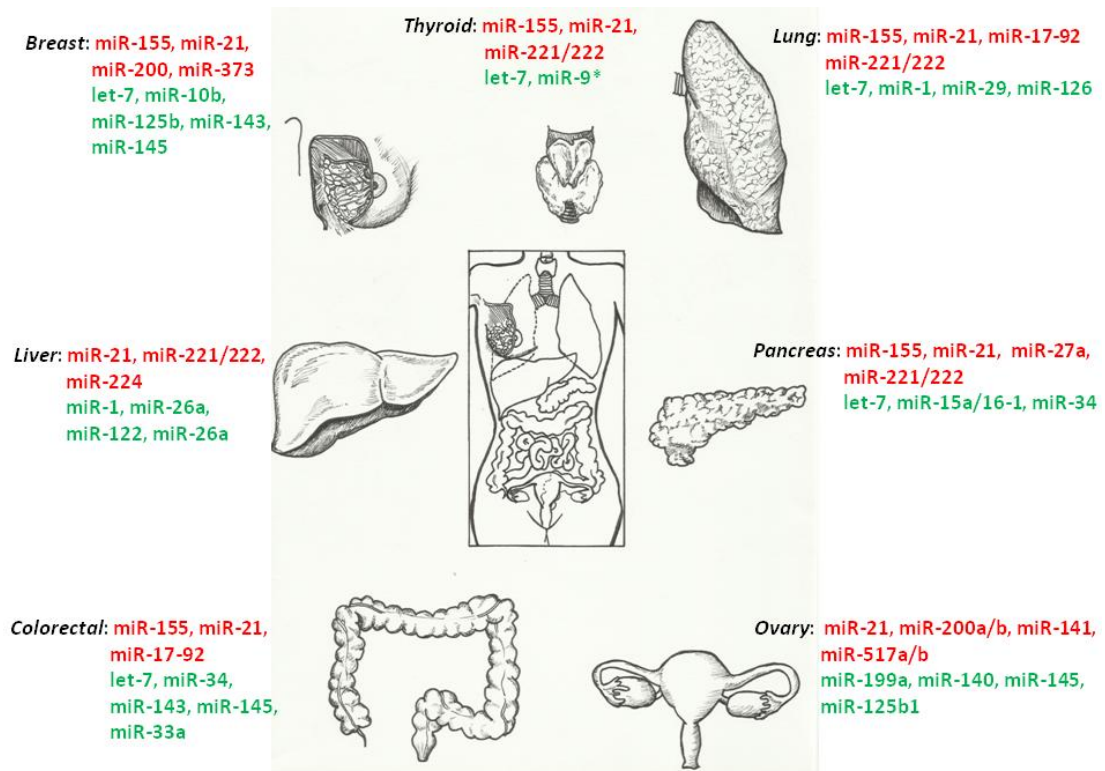


Figure 3 The most important miRNAs differentially regulated in solid cancers. miRNAs are either red (over-expressed) or green (down-regulated).

The role of miRNA in cancer was first discovered in leukemia. Calin et al. [71] reported evidence for the role of miRNAs in the pathogenesis of CLL deletions and/or down-regulation of miR-15a and miR-16-1 at chromosome 13q14 were associated with CLL. Afterward, a study demonstrated that this cluster can regulate the expression of the BCL-2 (B-cell lymphoma 2) oncogene [72]. Other relevant alterations of miRNAs in CLL include down-regulation of miR-181a, let-7a, miR-30d, miR-150 and miR-92, and over-expression of miR-155. In Table A2 are resumed the principal miRNAs associated to leukemia and hematologic diseases. miR-155 and miR17-92 cluster are commonly amplified in B-cell lymphoma patients and together

were among the earliest ncRNAs to be linked cancer. miRNA expression signatures have revealed differences between acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) [73]. Specific miRNA were correlated to karyotype alterations in AML , that have reported microRNA signatures associated with cytogenetics and prognosis of AML, with molecular abnormalities like translocations t(11q23), trisomy 8 and FLT3 (receptor tyrosine kinase) internal tandem duplication (ITD) mutations. Down-regulation of miR-221 and miR-222 was observed in AML [74].

1.7 microRNA target genes

The explication of the miRNA activity is strictly related to mRNA targets. In order to predict miRNA target interactions, many computational approaches have been developed. The topic of these software predicts biological targets of miRNAs by searching for the presence of conserved 8mer and 7mer sites that match the seed region of each miRNA [75]. Experimental evidence were needed to validate the target prediction. Below there are list some of the published and most used algorithms.

TargetScan [76] was one of the first RNA fold program and a total free binding energy is calculated with the RNAeval algorithm. The free energy is converted to a zscore and predictions in each organism are ranked. The algorithm takes three parameters: one that defines the relation between the binding energy and the z-score, a z-score cut-off, and a ranking cutoff value. As an option, nonconserved sites are also predicted . TargetScanHuman considers matches to annotated human UTRs and their orthologs, as defined by UCSC whole-genome alignments. Conserved targeting has also been detected within open reading frames (ORFs). “TargetScan” is one of the earliest published algorithms, it has maintained its role as a gold standard in many experiments.

RNA22 [77] is a pattern-based algorithm for the discovery of microRNA target sites and the corresponding heteroduplexes Statistical significance of each individual motif is assessed by training a second-order Markov chain. The key idea of rna22 is that the reverse complement of any salient sequence features that one can identify in mature microRNA sequences (using pattern discovery techniques) should allow one to identify candidate microRNA target sites in a sequence of interest: once a candidate microRNA target site has been located, the targeting microRNA can be identified with the help of any of several algorithms able to compute RNA:RNA heteroduplexes.

miRanda miR-SVR [78] are incorporated in microRNA.org , as a target searcher tool. mirSVR is a machine learning method for ranking microRNA target sites by a down-regulation score. The algorithm trains a regression model on sequence and contextual features extracted from miRanda-predicted target sites. It uses the same basic principles of TargetScan, but there are some difference in the calculation of the score.

PITA [79] is an algorithm, published in 2007 by Kertesz and coworkers, mainly based on secondary RNA structure. This application can be run as a standalone version or as a webserver. The main assumption is based on the fact that the mRNA structure plays a role in target recognition by thermodynamically promoting or disfavoring the interaction, using complementarity analysis within seed regions (single mismatch or G:U wobble pairing can be allowed) and then compares the free energy gained from the formation of the miRNA-target duplex and the energetic cost of unpairing the target to make it accessible to the miRNA.

PicTar [80] searches for nearly but not fully complementary regions of conservative 3'UTRs and then calculates the free energy of created duplexes. Each result is scored using Hidden-Markov Model . miRNAs with multiple alignments are favored., It was the first method that considered a parallel expression on the cellular level or an action in a common biological pathways of a miRNA and transcript, co-expressed miRNAs are more likely to target the same mRNAs. According to the authors, adding probabilistic knowledge about co-expression significantly increases specificity. It uses sequence alignment to eight vertebrate species to eliminate false positive results and it scores the candidate genes of each species separately to create a combined score for a gene.

Diana microT [81, 82] is an algorithm based on the use of uses a 38nt-long frame that is moved along 3'UTR , searching for sites with canonical central bulge and it requires 7, 8 or 9 nt-long complementarity in 5' region of miRNA. 6 nt-long matches

within seed region or with one wobble pairing are also considered while enhanced by additional base pairing in 3' region of miRNA. DIANA-microT use conservative alignment for scoring but also considers non-conservative sites. It gives unique signal-to-noise ratio (SNR) which is a ratio between a total of predicted targets by single miRNA in searched 3'UTR and a total of predicted targets by artificial miRNA with randomized sequence in searched 3'UTR. Another tool of the same platform is DIANA-mirPath [83], that is able to integrate human and mouse microRNAs in known cellular pathways.

1.8 High-throughput technology

One of the first high-throughput technology was microarray. Microarray technology evolved from Southern blotting, where fragmented DNA is attached to a substrate and then probed with a known DNA sequence. A DNA microarray is used in molecular biology and medicine, it consists of a collection of microscopic DNA spots attached to a solid surface. Scientists use DNA microarrays to measure the expression levels of large numbers of genes simultaneously or to genotype multiple regions of a genome. Each DNA spot contains picomoles (10⁻¹² moles) of a specific DNA sequence, known as probes (or reporters or oligos). These can be a short section of a gene or other DNA element that are used to hybridize a cDNA or cRNA (also called anti-sense RNA) sample (called target) under high-stringency conditions. Probe-target hybridization is usually detected and quantified by detection of fluorophore-, silver-, or chemiluminescence-labeled targets to determine relative abundance of nucleic acid sequences in the target.

The core principle behind microarrays is hybridization between two DNA strands, the property of complementary nucleic acid sequences to specifically pair with each other by forming hydrogen bonds between complementary nucleotide base pairs. A high number of complementary base pairs in a nucleotide sequence means tighter non-covalent bonding between the two strands. After washing off of non-specific bonding sequences, only strongly paired strands will remain hybridized. Fluorescently labeled target sequences that bind to a probe sequence generate a signal that depends on the hybridization conditions (such as temperature), and washing after hybridization. Total strength of the signal, from a spot (feature), depends upon the amount of target sample binding to the probes present on that spot. Microarrays use relative quantization in which the intensity of a feature is compared to the intensity of the same feature under a different condition, and the identity of the feature is known by its position. Many types of arrays exist and the broadest distinction is whether they are spatially arranged on a surface or on coded beads: traditional solid-phase array

is a collection of orderly microscopic "spots", called features, each with a thousands of identical and specific probes attached to a solid surface, such as glass, plastic or silicon biochip (commonly known as a genome chip, DNA chip or gene array). Thousands of these features can be placed in known locations on a single DNA microarray. The alternative bead array is a collection of microscopic polystyrene beads, each with a specific probe and a ratio of two or more dyes, which do not interfere with the fluorescent dyes used on the target sequence.

DNA microarrays can be used to detect DNA (as in comparative genomic hybridization), or detect RNA (most commonly as cDNA after reverse transcription) that may or may not be translated into proteins. There are many different types of microarray-based assays:

- DNA microarrays for transcriptomics research
- DNA tiling arrays for ChIP on chip experiments
- DNA microarrays for genotyping
- DNA microarrays for resequencing
- Protein microarrays to detect protein-protein interactions
- Antibody-Based Protein Array for protein quantification
- Chromatin Immunoprecipitation Assay (ChIP)
- Single Nucleotide Polymorphism (SNP) array
- Tissue Microarray.

Microarray Expression Analysis: In this experimental setup, the cDNA derived from the mRNA of known genes is immobilized. The sample has genes from the normal as well as the diseased tissues, or from cells in differentiation. Spots with more intensity are obtained for diseased tissue gene if the gene is over expressed in the diseased condition. This expression pattern is then compared to the expression pattern of a gene differentially expressed, in this context microarray analysis could be used as a "phenotyping tool". Furthermore, It can be useful for the determination of

molecular pathways: analysis of gene-expression changes in biological systems that model disease might lead to the identification of pathways that play a crucial role in disease pathogenesis. Another analysis is represented by the global assessment of drug effects in appropriate cell, or animal model systems or patient samples. Besides steady-state effects, it is also possible to determine the kinetics of the expression response as different genes respond in a time-dependent fashion at different intervals following a perturbation, or it is possible determinate comprehensively how cells or tissues respond to a particular drug might make it possible to predict either efficacy or adverse effects to a drug, eventually providing the presence of biomarkers that can be used to readily monitor drug efficacies as well as side effects and disease progression. In order to identify genomic copy number variations (CNV) across the whole genome at a different resolution levels, it is possible use an array, named Comparative genomic hybridization array (aCGH).

In order to analyze miRNAs profiling, many microarray based on miRNA were developed too. There are many platforms that can perform miRNA array profiling, for example Agilent, Affymetrix, Exiqon, Illumina, etc.

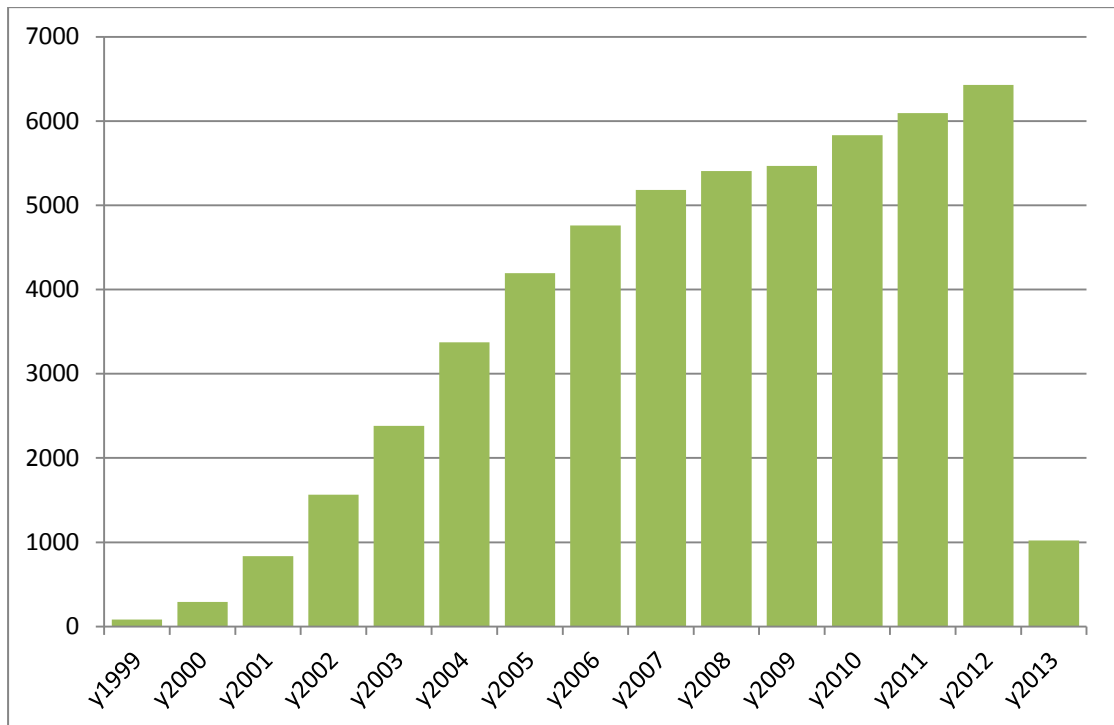


Figure 4 Graphic result of the research of "microarray" term per year in citation (Figure is obtained using information from <http://www.ncbi.nlm.nih.gov/pubmed/?term=microarray>)

The number of scientific articles that have used microarray technology had exponentially grew up in the last fifteen years (Figure 4). The huge amount of information obtained by this technology are available online for all user, there are many public database for microarray data. The database meant to be central repositories and the most known are:

- GEO- Gene expression omnibus (<http://www.ncbi.nlm.nih.gov/geo/>):

It uses the annotation standard Minimum Information About a Microarray Experiment (MIAME) and serves as a public repository for a wide range of high-throughput experimental data. These data include single and dual channel

microarray-based experiments, measuring the abundance of mRNA, genomic DNA and protein molecules. Data from non-array-based high-throughput functional genomics and proteomics technologies are also archived. Several tools are provided to assist with the visualization and exploration of GEO data, including hierarchical cluster heat maps and searchable individual gene expression profiles.

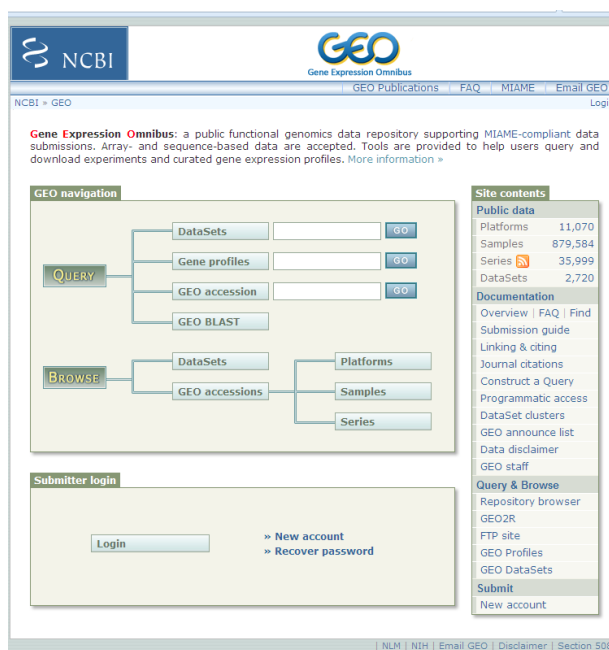






Figure 5 Screenshot of GEO interface (<http://www.ncbi.nlm.nih.gov/geo/>).

- ArrayExpress Archive (<http://www.ebi.ac.uk/arrayexpress/>):

It uses the annotation standard MIAME and the associated XML data exchange format Microarray Gene Expression Markup Language (MAGE-ML) and it is designed to store well annotated data in a structured way. Consists of the database itself, data submissions in MAGE-ML format or via an online submission tool MIAMExpress, online database query interface, and the Expression Profiler online analysis tool.

EMBL-EBI [Terms of Use](#) | [Privacy](#) | [Cookies](#)

[Databases](#) | [Tools](#) | [Research](#) | [Training](#) | [Industry](#) | [About Us](#) | [Help](#) | [Site Index](#)  


The **ArrayExpress Archive** is a database of functional genomics experiments including gene expression where you can query and download data collected to MIAME and MINSEQE standards. **Gene Expression Atlas** contains a subset of curated and re-annotated Archive data which can be queried for individual gene expression under different biological conditions across experiments.

Experiments Archive

33885 experiments, 1036569 assays

Experiment, citation, sample and factor annotations

[Browse experiments](#) | [platform designs](#) | [protocols](#) | [files](#)

[Submitter/reviewer login](#)  [ArrayExpress Query Help](#)

Gene Expression Atlas

3596 experiments, 99866 assays, 20862 conditions

Genes up/down in Conditions

Any species

[Gene Expression Atlas Home](#)

News

31 Jan 2013 - **Data Submission and Quality in Microarray-Based MicroRNA Profiling**

A new paper, *Data Submission and Quality in Microarray-Based MicroRNA Profiling*, has been published online. It discusses how reluctance to share data is associated with low study quality.

Links

- [ArrayExpress User Survey](#)
- [Help](#) | [Training](#) | [FAQ](#) | [Citing](#)
- [Submit Data](#) (array based and re-sequencing)
- [Programmatic Access](#) | [FTP Access](#)
- [Software Downloads](#)
- [EFO](#) | [Bioconductor Package](#) | [Quality Metrics](#)
- [ArrayExpress Scientific Advisory Board](#)
- [Funding](#) | [Acknowledgements](#)
- [Functional Genomics Group](#)

Figure 6 ArrayExpress interface from EMBL-EBI (<http://www.ebi.ac.uk/arrayexpress/>)

1.9 Regulatory networks

A biological networks use a directed graph representation in an effort to model the way that proteins and other biological molecules are involved in gene expression. A gene regulatory network (GRN) is a collection of DNA segments in a cell which interact with each other (indirectly through their RNA and protein expression products) and with other substances in the cell, thereby governing the rates at which genes in the network are transcribed into RNA.

A typical gene regulatory network might be defined like this: the nodes of the network are proteins, their corresponding mRNAs, or protein/protein complexes. Edges between nodes represent individual molecular reactions, the protein/protein or protein/mRNA interactions through which the products of one gene affect those of another. These interactions can be inductive, with an increase in the concentration of one leading to an increase in the other, or inhibitory, with an increase in one leading to a decrease in the other. A series of edges indicates a chain of such dependences, with cycles corresponding to feedback loops. The network structure is an abstraction of the system's chemical dynamics, describing the manifold ways in which one substance affects all the others to which it is connected. In one instance, such GRNs can be inferred from the biological literature on a given system and represent a distillation of the collective knowledge about a set of related biochemical reactions. Genes can be viewed as nodes in the network, with input being proteins such as transcription factors, and outputs being the level of gene expression. The node itself can also be viewed as a function which can be obtained by combining basic functions upon the inputs (in the Boolean network described below these are Boolean functions, typically AND, OR, and NOT). These functions have been interpreted as performing a kind of information processing within the cell, which determines cellular behavior. The basic drivers within cells are concentrations of a macromolecule, which determine both spatial (location within the cell or tissue) and temporal (cell cycle or developmental stage) coordinates of the cell. The gene networks are only beginning

to be understood, and it is a next step for biology to attempt to deduce the functions for each gene "node", to help understand the behavior of the system in increasing levels of complexity, from gene to signaling pathway, cell or tissue/organ level ("systems biology"). Mathematical models of GRNs have been developed to capture the behavior of a system, and in some cases generate predictions corresponding to experimental observations. We used a Bayesian approach.

Bayesian networks (BNs) or Bayes nets for short, are considered a kind of probabilistic graphical models. These graphical structures are used to represent knowledge about an uncertain domain. In particular, each node in the graph represents a random variable, while the edges between the nodes represent probabilistic dependencies among the corresponding random variables. Using known statistical and computational methods the conditional dependencies in the graph are estimated; so it is possible combine principles from graph theory, probability theory, computer science, and statistics.

Graphical models with undirected edges are generally called Markov random fields or Markov networks. These networks provide a simple definition of independence between any two distinct nodes. BNs correspond to another graphical models structure known as a directed acyclic graph. BNs are both mathematically rigorous and intuitively understandable. They enable an effective representation and computation of the joint probability distribution over a set of random variables.

The structure of a directed acyclic graph is defined by two sets: the set of nodes (vertices) and the set of directed edges.

The nodes represent random variables and are drawn as circles labeled by the variable names. The edges represent direct dependence among the variables and are drawn by arrows between nodes. In particular, an edge from node 1 to node 2 represents a statistical dependence between the corresponding variables [84, 85].

Thus, the arrow indicates that a value taken by variable depends (influences) on the value taken by another variable.

2 THESIS OBJECTIVES

The aim of the thesis was to answer the question whether the non-coding RNAs expression data can be used for a classification of the normal tissues and of the several cancers using system biology approaches. Therefore we want to investigate the specificity of each of the classes studied of non-coding RNA, as Transcribed UltraConserved Region and microRNAs. Furthermore, we want to find whether non-coding RNAs can be candidates as features for the selection of specific cancers, using statistical algorithms and bioinformatics tools. Then, we decided to infer Bayesian networks for miRNAs in normal and cancer samples to identify the relations between their expression and pathological role. In order to prove our expression data analysis and the networks of cancers, we want to apply: *in silico* validation using several bioinformatic approaches on data available in our and public database, experimental validations using biologic and molecular technics of laboratory. We expect our results to show evidences which sustain the use of non-coding RNAs expression data for the characterization of cancers and eventually the identification of new markers of cellular subpopulations of cancer.

Consequently, the aims of the thesis were :

- i. Calculate tissue specificity of the non-coding RNAs in normal samples, using our large expression database;
- ii. Classification of solid cancer by non-coding RNAs;
- iii. Elaboration of the miRNA bayesian networks of cancers;
- iv. *In silico*, *in vitro* and *in vivo* validation.

3 THE TISSUE SPECIFICITY OF T-UCRs AND miRNAs

3.1 Material and Methods

Microarray analysis was performed miRNA microarray assay on a CodeLink platform. The miRNA array consists of 4,104 probes printed in duplicate. This array can simultaneously profile more than 1,500 mature miRNAs and their corresponding precursor [86]. The Ohio State University Comprehensive Cancer Center custom microarray was used for T-UCR expression profiling following previously published protocols. For each UCR two 40-mer probes were designed, one corresponding to the sense genomic sequence (named “+” or “plus”) and the other to the complementary sequence (named “+A” or “minus”). Each oligo was printed in duplicate in two different slide locations, and therefore quadruplicate numerical values were available for analysis. Briefly, 5 micrograms of total RNA were used for hybridization of miRNA microarray chips. These chips contain gene-specific oligonucleotide probes, spotted by contacting technologies and covalently attached to a polymeric matrix. The microarrays were hybridized in 6X SSPE (0.9 M NaCl/60 mM NaH₂PO₄ ·H₂O/8 mM EDTA, pH 7.4)/30% formamide at 25°C for 18 hrs, washed in 0.75X TNT (Tris·HCl/ NaCl/ Tween 20) at 37°C for 40 minutes, and processed by using a method of detection of the biotin-containing transcripts by streptavidin-Alexa647 conjugate. Processed slides were scanned using a microarray scanner (Axon), with the laser set to 635 nm, at fixed PMT setting, and a scan resolution of 10 mm. Microarray images were analyzed by using GenePix Pro and post-processing was performed [87] . Briefly, average values of the replicate spots of each miRNA were background-subtracted and subject to further analysis. miRNAs were retained when present in at least 20% of samples and when at least 20% of the miRNA had fold change of more than 1.5 from the gene median. Absent calls were thresholded prior to normalization and statistical analysis. Normalization was performed by using the quantiles method. MicroRNA nomenclature was according to

the microRNA database at Sanger Center [88] . The expression data analyzed in this study can be accessed at the NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) under accession nos. GSE8126 and GSE7055, GSE17155, GSE3467, GSE7828, GSE6857, GSE16654, and GSE14936, and at ArrayExpress (www.ebi.ac.uk/microarray-as/ae) under accession nos. E-TABM-866, E-TABM-664 E-TABM-762 and E-TABM-763, E-TABM-508, E-TABM-429, E-TABM-434, E-TABM-405, E-TABM-343, E-TABM-41 and E-TABM-42, E-TABM-48, E-TABM-22 , E-TABM-23, ETABM-46, E-TABM-47, E-TABM-49, and E-TABM-50, and E-MEXP-1796, E-TABM-37, E-TABM-341, E-TABM-969 and E-TABM-970 for normal tissues, E-TABM-971 for breast cancer, E-TABM-972 and E-TABM-974 for acute myeloid leukemia, E-TABM-973 for chronic lymphocytic leukemia and E-TABM-975 for ovarian cancer.

3.1.1 Data analysis

We have developed an SQL miRNA internal database that contained the data retrieved from a large number of different experiments performed with miRNA microarray. Bad spots were removed. Non-Expressed spots were averaged for each gpr files (chip). For each mature miRNA, we computed the geometric mean of its multiple reporters in the chip. A NaN value was assigned to miRNAs with more than 50% of corrupted spots, as reported by the GenePix image analysis software. All the results were log₂-transformed. The Normalization was performed by using the quantiles normalization, as implemented in Bioconductor "affy" package [89]. Quantile normalization is a non-parametric procedure and the algorithm mapped every value on any one chip to the corresponding quantile of the standard distribution. BRB Arraytools was used to perform *t-test* over 2-classes experiments of F-tests over multiple classes.

3.1.2 Calculation of Information Content

First, all samples were classified according to their organ-, tissue- and celltype; then the normal samples were grouped in specific systems and the disease samples in specific pathological states. To assess the specificity of miRNA expression across groups, we needed to estimate what fraction of the total, for a given miRNA belonged to each single group. We used the procedure described in the first miRNA expression atlas [90], with the only exception that we called $E_{m,t}$ the value of miRNA m in the group t referred to as "mean expression value" (subtracted of the background value, 100). From here onwards, we essentially proceeded as the reference, therefore: To allow for comparisons between tissue types, we first normalized the counts in each tissue type t :

$$F_{m,t} = E_{m,t} / \sum_{m'} E_{m',t}$$

If a miRNA m is expressed with high specificity in tissue type t , then the value $F_{m,t}$ is large (close to 1) not only relative to other miRNAs m' in the same tissue type t (or column of F), but also relative to the same miRNA m in other tissue types $t' \neq t$. It is then natural to normalize the lines of F as:

$$G_{m,t} = F_{m,t} / \sum_{t'} F_{m,t'}$$

Each line m of G contains a distribution of normalized frequencies of miRNA m across samples. When G_{m,t^*} is close to 1 in tissue t^* , we may infer that miRNA m is specifically expressed in tissue type t^* , meaning that a large fraction of clones in t^* and much smaller fractions of clones in other $t \neq t^*$ correspond to miRNA m . We use the information-theoretic concept of "information content" to quantify how strongly biased the distribution of $G_{m,t}$ is for a given m :

$$sm = \log_2(\text{number of tissue types}) + \sum_t G_{m,t} \log_2(G_{m,t}).$$

The specificity score varies between 0, when the expression level of the miRNA m is the same across all tissues, and \log_2 of the number of tissue types, when only one tissue expresses the miRNA. To minimize artifacts from miRNAs or tissues with very small expression levels, we considered only miRNAs with a total expression value above 10 times the number of normal tissues and above 100 times the number of cancer types; with a minimal expression value (after background subtraction) of 100. For the calculation of overall specificity, we thus included 130 and 133 different mature miRNAs for normal tissues and cancer, respectively. Therefore, it is possible that we missed some specifically expressed miRNAs that in our data had either very low expression or were specific to tissues/diseases that we did not sample sufficiently. The same approach is used for the T-UCR information content.

3.1.3 Decision Tree

The expression matrix of T-UCRs values was used to infer an ultra-selection system using weka software [91]. First, for all the samples we have applied a supervised attribute filter to select attributes (in this case the T-UCRs expressions value). This process has selected 16 attributes; after we used the weka classifiers trees J48 graft -C 0.25 -M 2, in order to calculate how these attributes produce an accurate classification of the tissue selected. We used also JRip (RIPPER) algorithms. Classes were examined in increasing size and an initial set of rules for the class was generated using incremental reduced error JRip proceeds by treating all the examples of a particular judgment in the training data as a class, and finding a set of rules that cover all the members of that class. So, it proceeded to the next class and repeat the same, repeating this until all classes have been covered. J48 is an open source Java implementation of the C4.5 algorithm in the weka data mining tool. C4.5

builds decision trees from a set of training data, using the concept of information entropy. The training data is a set $S = (s_1, s_2, \dots)$ of already classified samples. Each sample s_i consists of a p -dimensional vector and the x_j represent attributes or features of the sample, as well as the class in which s_i falls. At each node of the tree, C4.5 chooses the attribute of the data that most effectively splits its set of samples into subsets enriched in one class or the other. The splitting criterion is the normalized information gain (difference in entropy). The attribute with the highest normalized information gain is chosen to make the decision. The C4.5 algorithm then recurses on the smaller sub-lists. All the samples in the list belong to the same class. When this happens, it simply creates a leaf node for the decision tree saying to choose that class. None of the features provide any information gain. In this case, C4.5 creates a decision node higher up the tree using the expected value of the class. Work flow of the information that we have used for performing analysis using weka software are in APPENDIX B.

3.2 miRNAs tissue specificity in normal and embryonic samples

The first microRNA that we found specific for the embryonic cell types was the hsa-miR-302: embryonic cell types, such as embryonic bodies, trophoblasts, endoderm, or other stem cells including induced pluripotent stem cells (iPS) also harbored high levels of this miRNA. In the figure 7, after hsa-miR-302, the most specific miRNAs were hsa-miR-338-5p, hsa-miR-323-3p, and hsa-miR-335, with presented highest expression in epidermis /nervous system, nervous system and breast respectively. All the values of the IC were listed in the Table A3.

Hsa-miR-371-5p was here well represented in embryonic cell types, in particular there was a high hsa-miR-371-5p expression in embryoid bodies. Hsa-miR-129-3p is specific for nervous system in both datasets. We detected hsa-miR-142-5p in hematopoietic system and connective tissues. Hsa-miR-9, hsa-miR-128 and hsa-miR-138 were expressed in nervous system.

Nevertheless, the IC distributions were very similar: 25 miRNAs with IC higher than 1.35 for the microarrays. The most notable IC was without doubt the hsa-miR-302a/b/c.

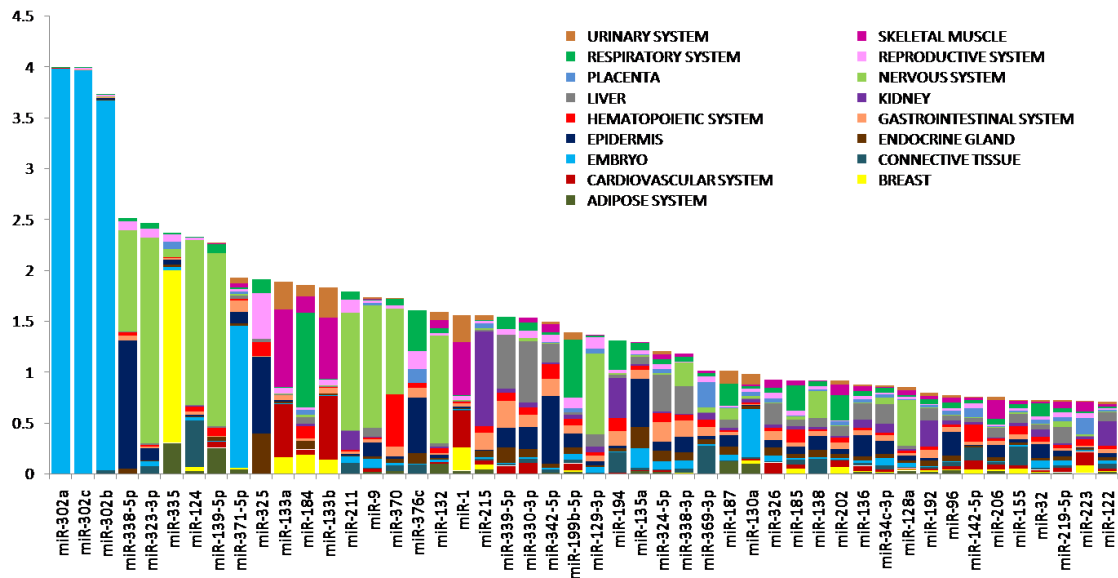


Figure 7. miRNA specificity in 50 normal tissues grouped by system. The tissue specificity was calculated by using the information content (IC), value expressed on y-axis; each color represents a system. miR-302 cluster is the most representative for embryo tissues.

Since hsa-miR- 302 expression only partially decreased throughout these stages of ES differentiation, its IC in embryonic tissues was low (0.5). This variation might be due to the fact that we assayed different cell types of embryonic origin, in comparison to only ES cells. The miRNAs which had highest tissue specificity in embryos were: hsa-miR-211 in 14 day embryoid bodies (EBs), hsa-miR-10b, hsa-miR-218, hsa-miR-122, and hsa-miR-148a in spontaneous differentiated monolayers, hsa-miR-138 and hsa-miR-338-3p in 7 day EBs, and hsa-miR-99a in trophoblast (Figure 8 and Table A4).

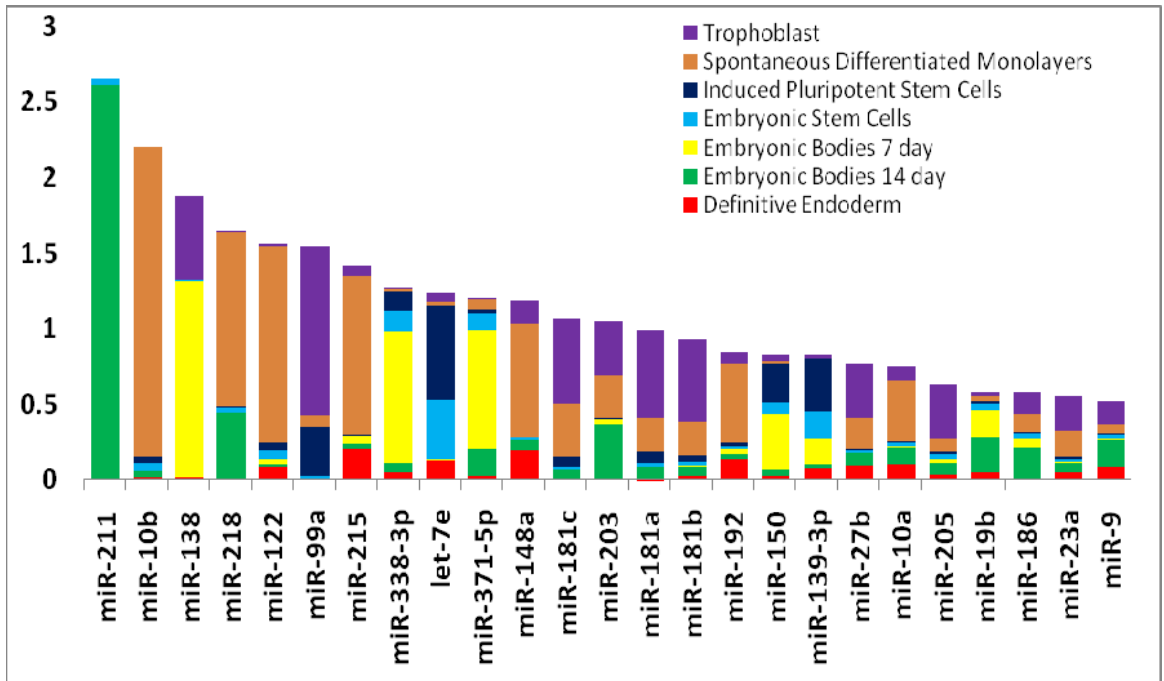


Figure 8 The specificity was calculated by using the information content (IC). The miRNA specificity during ES cell differentiation.

3.3 T-UCRs tissue specificity in normal and embryonic samples

We used the same approach for check the expression of the T-UCRs. We tested the expression in 374 samples from 46 types of normal tissue derived from 16 histological groups. Tissue specificity was calculated using the information content (IC) as described previously in 3.1.2 Material and Methods section. For this profile we used data obtained with the OSUMC microarray platform, where there were 974 different probes for putative T-UCRs in both genomic strands (sense, designated 'plus' and antisense, designated 'minus').

First, a wide analysis of expression showed that only a portion among the tested UCRs (224 out of 974) was expressed in human tissues ($p\text{-value} < 0.05$). Of these: 48% were non-exonic, 26% possibly exonic and 26% exonic, thus reflecting previously published studies. Fifty-seven T-UCRs were bi-directionally transcribed (57 on 224) and their frequencies were: 37% non-exonic, 23% possibly exonic, and 40% exonic (Table A5).

The most tissue specific T-UCRs were represented by a group of 13 T-UCRs, exclusively expressed in epidermis, with an IC value ranging from 3.23 to 1.70, composed of 67% non-exonic, 19% possibly exonic, and 14% exonic sequences (Figure 9 and Table A6).

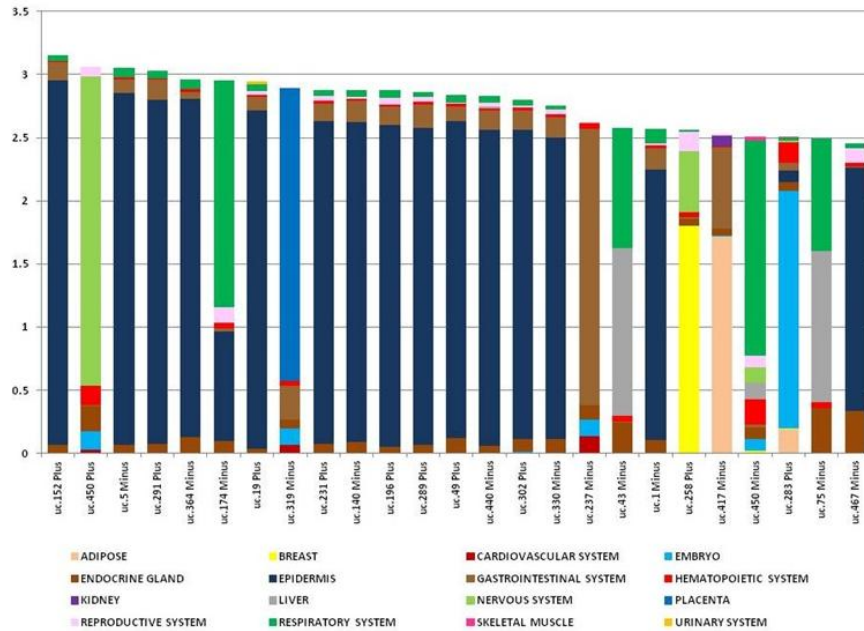


Figure 9. T-UCR specificity in 46 normal tissues grouped by 16 systems. The tissue specificity was calculated by using the information content (IC), value expressed on y-axis; different colors represents different systems.

These T-UCRs were not located in the same genomic cluster. Another tissue-specific UCR was uc.450 plus, highly expressed in the central nervous system. This confirms previous data showing that uc.450 was expressed in the dorsal root ganglion and the neural tube [92]. In particular, researcher hypothesized an enhancer function for hs385, which fully includes uc.450. Interestingly, the opposite probe, uc.450 minus, was specific for the respiratory system. Uc.174 minus, located on the opposite strand of the *MATR3* exon, was expressed only in the respiratory system and epidermis.

The placenta was characterized by the expression of uc.319 minus (intergenic), while uc.237 minus (intronic) was restricted to the gastrointestinal system.

Two T-UCRs expressed both in liver and in the respiratory system were uc.43 minus (intronic, our probe is homologous to *Mus Musculus* CN668140 EST) uc.75 minus, located on the opposite strand of the last exon of *ZEB2*, and uc.42 plus (intronic). In addition, uc.417 minus was expressed in adipose tissue and in the gastrointestinal system. The only T-UCR specific for the embryo was uc.283 plus. This unique uc.283 plus, a 277 nucleotide-long sequence located at chr10:50,604,757-50,605,033, was sufficient to discriminate between adult differentiated tissues and pluripotent stem cells (Figure 10).

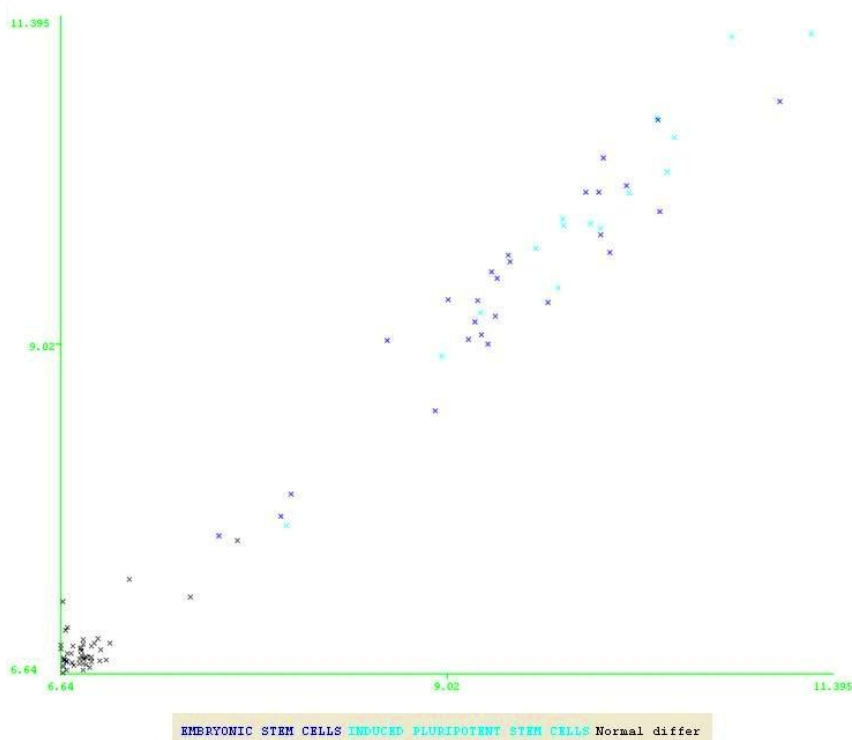


Figure 10 The spot represented the value of the probe uc.283plus , blue and cyan were plotted the ESC and IPS; black were the pool of normal adult differentiated cells.

Second, we found which were the T-UCR specific for the stages of ES differentiation (Figure 11).

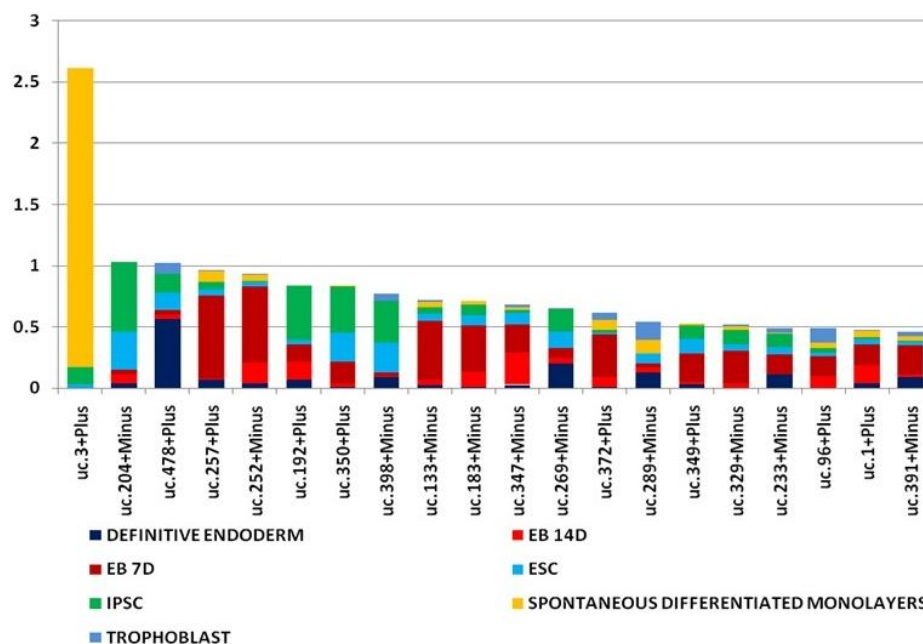


Figure 11 The T-UCRs specificity during ES cell differentiation, seven different types of embryonic condition. The specificity was calculated by using the information content (IC).

There were four T-UCRs expressed only in ESC and IPSC. We checked the enhancer activity using VISTA tool from Genome Browser track. Uc.204 minus, was located in a intergenic genomic area, identified by human element hs293, showed enhancer function (for branchial arch, forebrain and midbrain). Uc.192 plus, located in intergenic position, hs217, showed an enhancer function (for forebrain and midbrain). Moreover, in our data uc.192 plus highly expressed in IPS and in the second node of the decision tree, we noted that this over-expression can discriminate IPS from ESC (Figure 12).

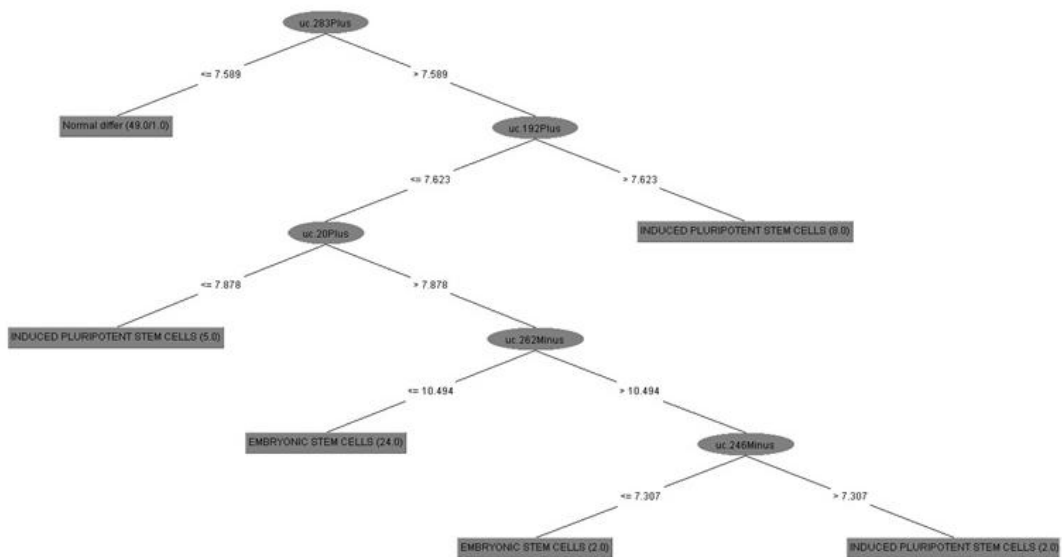


Figure 12 Decision tree is obtained by analysis of the expression data.

Uc.350 plus didn't show enhancer function, it was located in intronic region of DACH1, a gene that encodes a chromatin-associated protein associated with other DNA-binding transcription factors regulating gene expression and cell fate determination during development. Uc.398 minus was intergenic and located in a CpG island (374 nucleotides), there were to EST homologous in Xenopus and Danio Rerio .

For definitive endoderm the most representative T-UCR was uc.478 plus, but it was the unique probe that cover two genome sequences on two different chromosome: 4 and X, intronic/exon containing of Gria 2 (chr.4) and Gria 3 (chr. X). There was a little cluster of UCRs for embryonic bodies : uc.257 plus (intronic) , uc.252 minus was

intronic and homologous to an EST of *Petromyzon marinus* , this EST was extracted by fertilized egg to completion of digestive tract retained in a library of Lamprey EST Embryo. Uc.133 minus was intronic and located in the RSRC1 gene. Uc.183 minus: intronic/exon containing FBXW11. Uc.347 minus was intergenic, while uc.372 plus is located to an intron of RALGAPA1. Finally, uc.269 minus was present in definitive endoderm, ESC and IPSC; it was located in an intronic region of KIAA1608 and the respective hs314 has shown enhancer function for midbrain (mesencephalon) in VISTA Enhancer Browser.

4 CLASSIFICATION OF SOLID CANCER BY miRNAs EXPRESSION

4.1 Material and Methods

Microarray analysis was performed miRNA microarray assay on a CodeLink platform. The miRNA array consists of 4,104 probes printed in duplicate. This array can simultaneously profile more than 1,500 mature miRNAs and their corresponding precursor, all the other information were described in 3.1 material nad Methods section. We used our SQL miRNA internal database that contained the data retrieved from a large number of different experiments performed with miRNA microarray.

4.1.1 Statistical Tests and Algorithms

The Calculation of the Information Content was discussed in the section 3.1.2. Normalized, by using the quantiles method, microarray data were managed and analyzed by BRB-ArrayTools, version 3.8.1. Genes whose expression differed by at least 1.5-fold from the median, in at least 20% of the arrays were used. A stringent significance threshold was used to limit the number of false positive findings. The result of this approach was determined by 2 sample t test with nominal significance level at 0.001. The false discovery rate (FDR) is the expected proportion of positive results that are false positives at the various levels of significance and was controlled using the step-up method of Benjamini and Hochberg. In this analysis, at any selected FDR level, the expected proportion of false positives was determined. Class prediction algorithms determined whether miRNA expression patterns could accurately differentiate between the classes selected. We developed models based on the compound covariate predictor, nearest neighbor classification, and support vector machine. The models incorporated genes that were differentially expressed

among genes at the significance level (0.001) as assessed by the random variance t test. We used the prediction test to identify the classifier signature with the lowest misclassification error. We performed an unsupervised hierarchical clustering of the samples. The explanation of the decision tree and algorithm were described in section 3.1.3. Work flow of the experimental information that we have used for performing analysis using weka software are in APPENDIX B.

4.2 miRNAs signatures in cancer samples

In order to identify miRNA differentially expressed in several cancer, we investigated 51 types of cancer in 3312 samples (2532 solid cancers and 780 leukemia samples). First, we compared all the solid cancer versus the normal tissues, this class comparison identified a list of miRNA differentially expressed in solid cancers (Table A6). MicroRNA described as oncogenes, for example: hsa-miR-21 and hsa-miR-17/20/25/92/103/106/146a were over-expressed, whereas microRNA characterized for a tumor-suppressor role, as miR-203/205 and miR-143/145, were down-regulated. Second, we defined miRNA particularly expressed in some specific cancer calculating the information content. Hsa-miR-369-3p was the most specific for the ovarian cancer, the second (in order of IC value) was the hsa-miR-325, that it was expressed in pancreas non-functioning endocrine cancer (Figure 13 and Table A7).

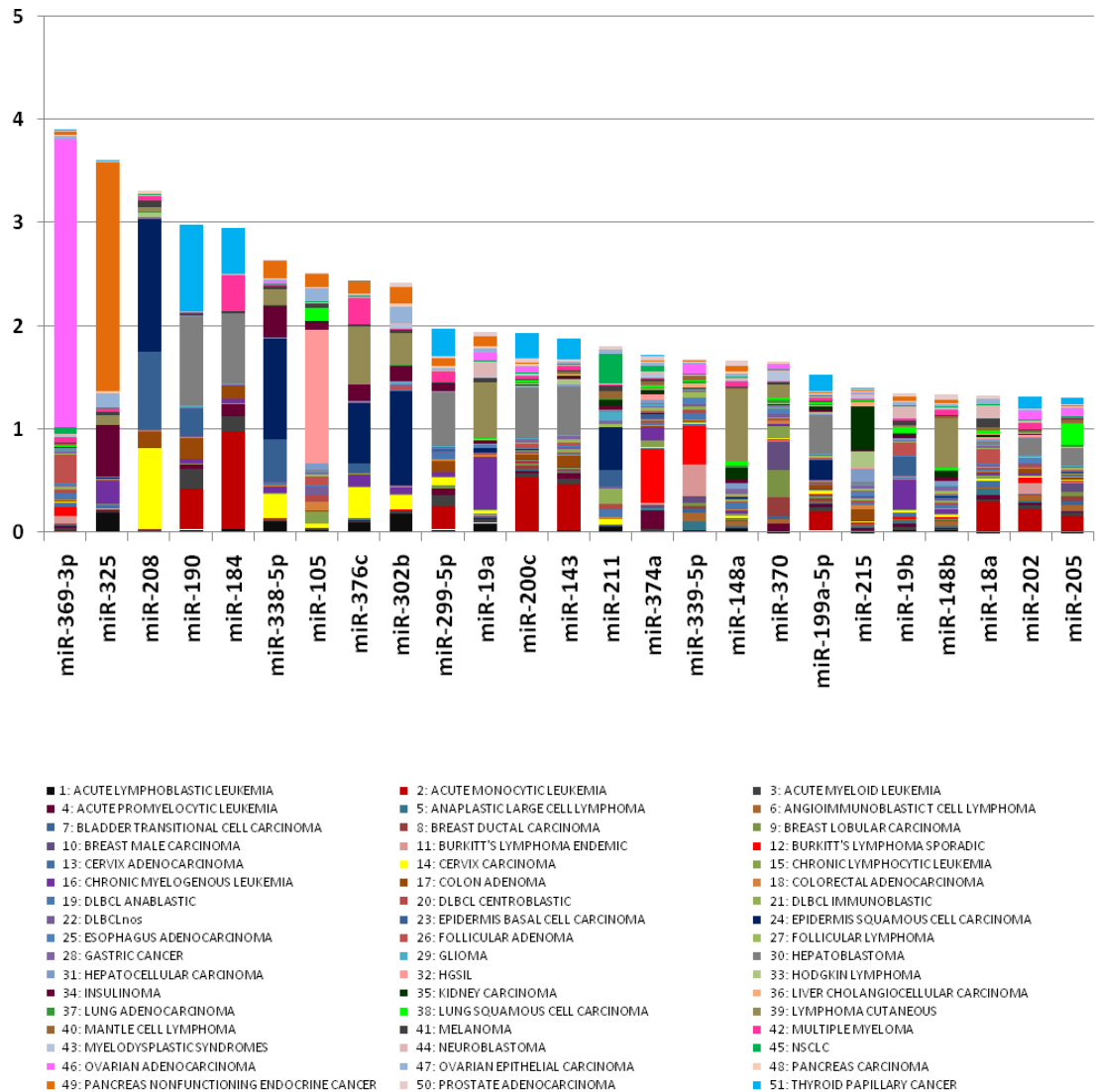


Figure 13 miRNA specificity in tumors and leukemia (31 solid tumors and 20 leukemia types) sorted by Information Content (IC).

4.3 T-UCRs signatures in cancer samples

We investigated the expression of T-UCR in several cancer samples in our database. After, a preliminary class comparison analysis between solid cancer versus normal counterparts we identified a list of T-UCR differentially expressed. Then, we calculated information content of every probes for the 27 cancer types (Figure 14).

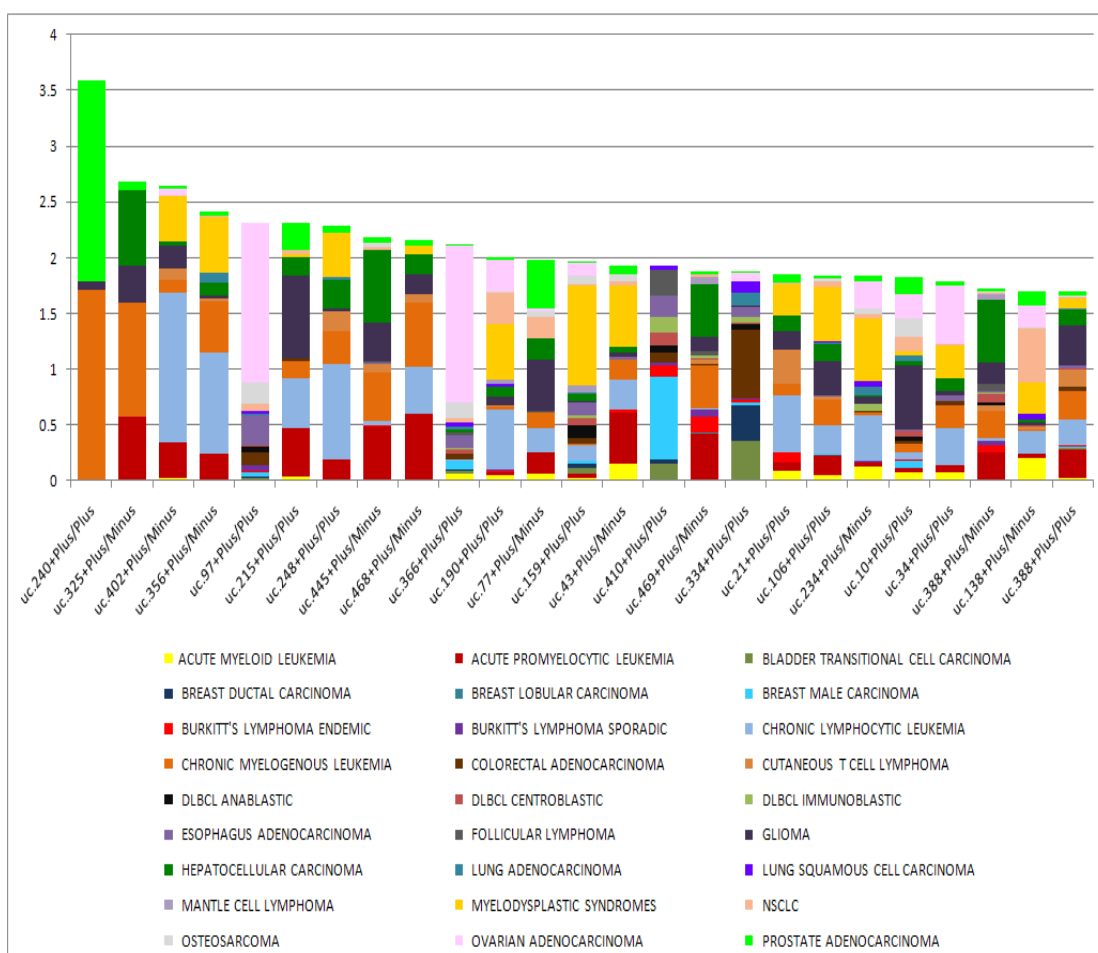


Figure 14 T-UCRs specificity in tumors and leukemia sorted by Information Content (IC).

Uc.240 plus was the T-UCR presented the highest value of IC, and it was expressed in chronic myelogenous leukemia and prostate adenocarcinoma. The second, in order of IC value, was the uc.325 minus, and it was principally expressed in chronic myelogenous leukemia and hepatocellular carcinoma. This T-UCR is classified as intronic for the gene *ELP4*. Uc.402 minus and uc.356 minus were both specific for chronic lymphocytic leukemia, uc.402 minus was located in an intronic region of the gene *RPGRIP1L* and uc.356 minus was located in an intronic region of the gene *MBNL2*.

4.4 Identification of putative “miRNA classifier”

The first phase of this study consisted in the data collection process which involved the data cleaning and data preprocessing. Using our complex and large expression database, we tested the expression profile of microRNAs in a first cohort of 140 training samples.

First, we random selected these 140 samples (training), that were represented by ten samples for fourteen several tumors. We inferred class prediction test between a specific cancer versus the other type of cancer, to increase the specificity of our approach we compared single solid tumor with other types of solid tumors (recursive feature elimination method was used to select 10 genes and Leave-one-out cross-validation method was used to compute mis-classification rate, $p\text{-value} < 0.001$).

These selection system was repeated for every class of cancer, identifying a list of miRNAs that were up- and down-regulated in these first cohort. These data from first array platforms identified a subset of 10 candidate microRNAs for any class of tumors: BLADDER CANCER, BREAST CANCER, COLON CARCINOMA, ESOPHAGUS CARCINOMA, GASTRIC CANCER, GLIOMA, HEPATOCELLULAR CARCINOMA, LUNG SQUAMOUS CELL CARCINOMA, MELANOMA, NSCLC

(NON-SMALL CELL LUNG CANCER), OVARIAN CARCINOMA, PANCREAS CARCINOMA, PROSTATE CARCINOMA.

We repeated the same analysis, but this time we used other 140 randomized samples (test) selected in our database. Therefore, we obtained two different lists of significant miRNAs. After, we intersected the two lists: we finally selected 100 microRNAs, listed in Table A8.

The matrix of values of these 100 miRNAs that we have previously selected, was imported in weka software to infer an ultraselction system. First, for the all samples we have applied a supervised attribute filter. We used as evaluator: CfsSubsetEval and as searcher: BestFirst -D1 -N5. Then, these processes have selected 31 miRNAs, listed in Table A9 and after a final unsupervised attribute selection we had a list of 15 miRNAs, see table A10.

To have a vision of the impact of expression of this 15 miRNAs in tumours, we performed two hierarchical clustered heatmaps. The expression profile of this group of miRNAs showed clearly pattern of differential up or down-regulated miRNAs. There were two heatmaps for the two cohorts of samples analyzed in the twenty solid cancers (Figure 15 and Figure 16).



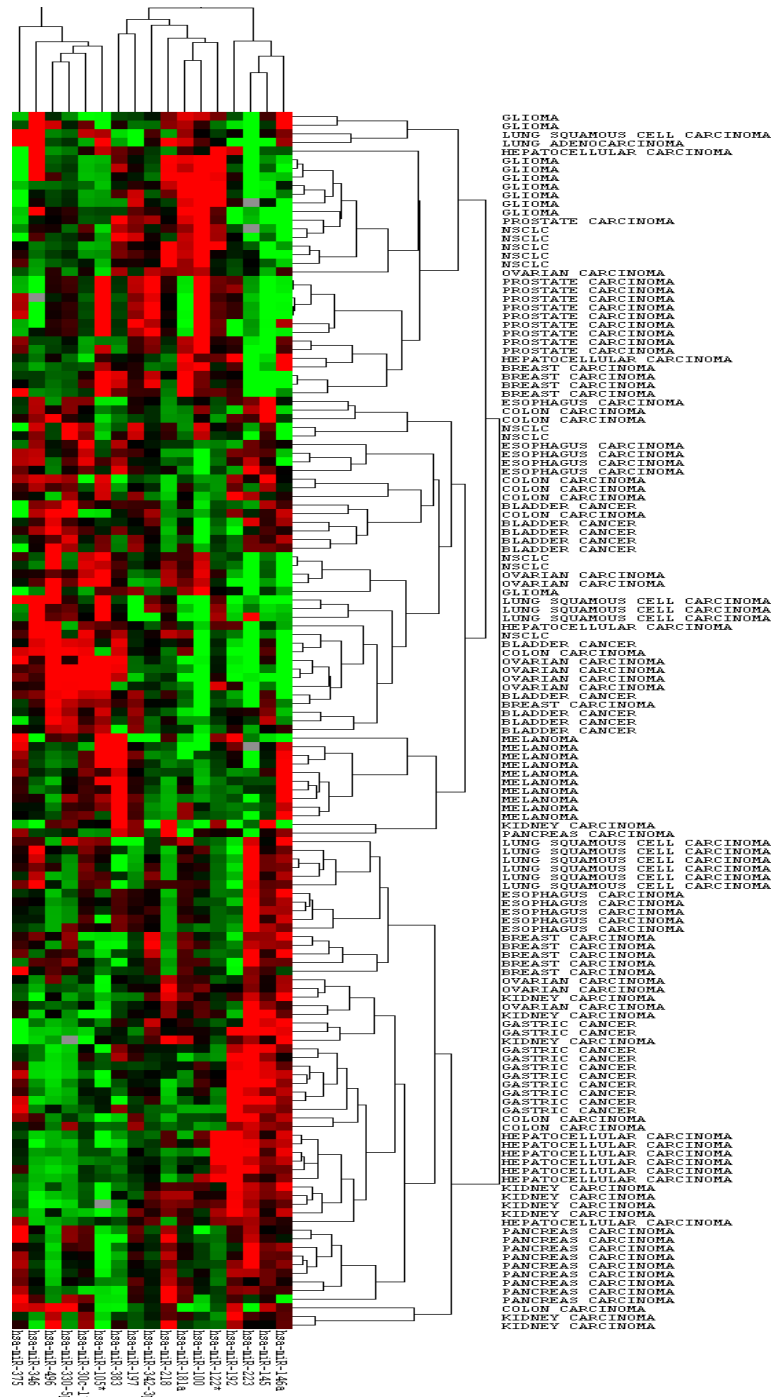


Figure 15 Training test cohort of 15 miRNA

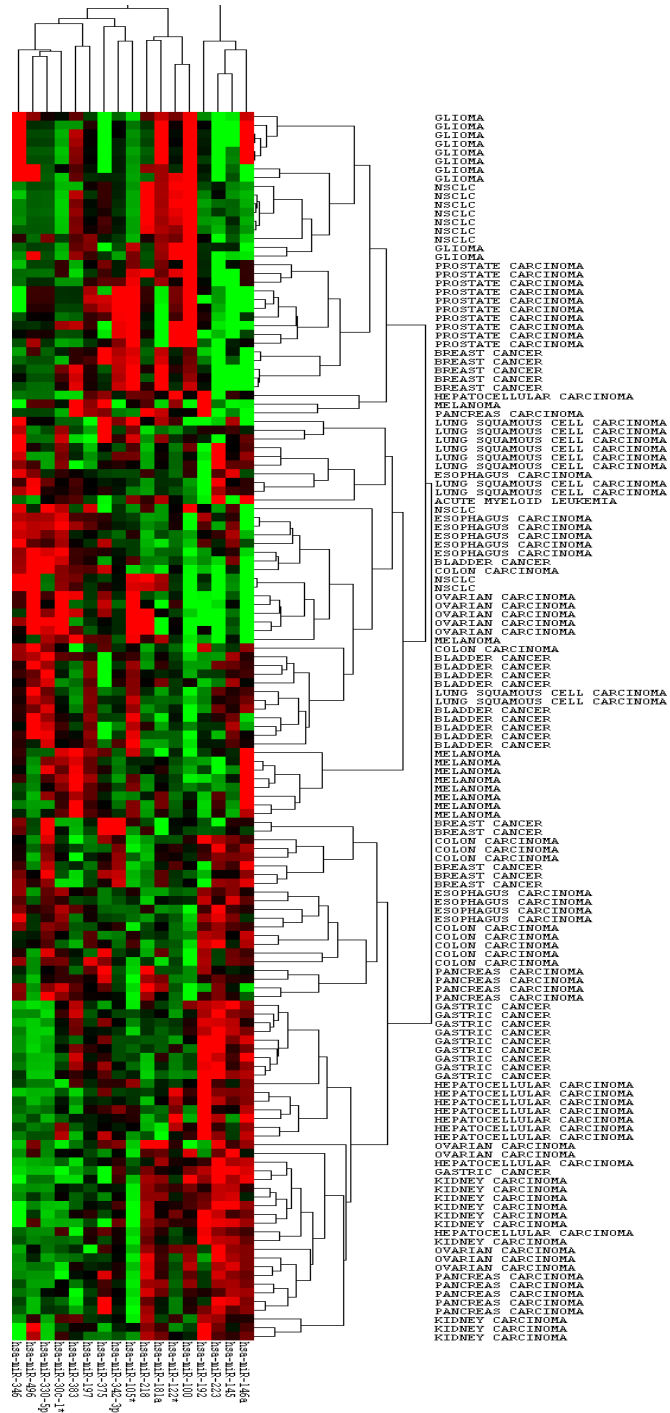


Figure 16 Test set cohort of 15 miRNA

4.5 Feature selection and classification

The second part of this analysis was focused on the choice of the classification for the solid cancer samples. We used two different algorithms: J48graft and JRip, two classifier able to investigate our dataset of solid cancer, that are respectively a decision-tree and a rule induction algorithm. First, we used the decision tree for our samples the J48graft algorithm (an implementation of C4.5 classification algorithm), implemented in the weka software. The C4.5 classification algorithm is easy to understand as the derived rules have a very straightforward interpretation. In our study, we used all the selected attributes defined previously by supervised attribute filter. We used both the list of 31 miRNAs that the list of 15 miRNAs. We classified using the first 140 samples random selected as training set. All the output information are reported in Appendix B. After, we used the other cohort of 140 samples random selected for the test set. The decision tree developed describes 26 number of leaves. It was able to define the classes correctly with a weighted average of accuracy of 79.3 %. We reported two considerations that appear in the testing set: the list of 15 miRNAs' list selected was more precise of the 2% than the 31 miRNAs' list. When we analyzed all the leaves, we noted that the more difficulty was represented by the classification of esophagus cancer, that it was often confused with colon cancer. As reported in the confusion matrix below.

```

a b c d e f g h i j k l m n <-- classified as
18 0 0 0 0 0 0 0 1 1 0 0 0 0 | a = BLADDER CANCER
2 15 2 0 0 0 0 0 0 0 1 0 0 0 | b = BREAST CARCINOMA
0 1 17 1 0 0 0 0 0 0 1 0 0 0 | c = COLON CARCINOMA
0 0 5 12 0 0 0 0 1 0 2 0 0 0 | d = ESOPHAGUS CARCINOMA
0 0 0 1 17 0 1 1 0 0 0 0 0 0 | e = GASTRIC CANCER
0 1 0 0 0 9 0 0 0 0 6 0 0 4 | f = GLIOMA
0 0 0 0 0 1 18 0 0 0 0 0 1 0 | g = HEPATOCELLULAR CARCINOMA
0 0 0 0 0 1 0 18 0 0 0 0 1 0 | h = KIDNEY CARCINOMA
4 0 1 0 0 1 0 0 13 0 1 0 0 0 | i = LUNG SQUAMOUS CELL CANCER
1 0 0 0 0 0 0 0 0 18 1 0 0 0 | j = MELANOMA
0 0 0 0 0 0 0 0 0 0 20 0 0 0 | k = NSCLC
0 0 0 0 0 0 0 3 0 0 4 11 2 0 | l = OVARIAN CARCINOMA
0 0 0 0 0 0 0 1 0 0 0 0 19 0 | m = PANCREAS CARCINOMA
1 0 0 0 0 1 0 0 0 0 0 0 1 17 | n = PROSTATE CARCINOMA

```

We used JRip following the same data and passage: it generated the rules for the classification, that we reported below.

(hsa-miR-383 <= 4.417) and (hsa-miR-30c-1* >= 4.024) and (hsa-miR-215 <= 5.322) => Class=LUNG SQUAMOUS CELL CARCINOMA (11.0/3.0)
(hsa-miR-9 >= 8.487) => Class=GLIOMA (10.0/2.0)
(hsa-miR-122 >= 11.89) => Class=HEPATOCELLULAR CARCINOMA (9.0/0.0)
(hsa-miR-145 <= 3.852) and (hsa-let-7e >= 9.894) => Class=PROSTATE CARCINOMA (9.0/0.0)
(hsa-miR-383 >= 9.519) => Class=MELANOMA (9.0/1.0)
(hsa-miR-15b >= 8.367) and (hsa-miR-218 <= 4.893) => Class=BREAST CARCINOMA (8.0/1.0)
(hsa-miR-145 >= 11.285) and (hsa-miR-383 >= 6.18) => Class=GASTRIC CANCER (8.0/0.0)
(hsa-miR-145 <= 7.731) and (hsa-miR-193b <= 8.717) => Class=NSCLC (9.0/1.0)
(hsa-miR-330-5p >= 8.863) => Class=NSCLC (2.0/0.0)
(hsa-miR-330-5p >= 7.761) and (hsa-miR-122 <= 7.731) => Class=BLADDER CANCER (4.0/0.0)
(hsa-miR-218 <= 3.513) and (hsa-miR-10a <= 9.198) => Class=BLADDER CANCER (4.0/0.0)
(hsa-miR-145 <= 9.173) and (hsa-miR-375 <= 8.103) => Class=OVARIAN CARCINOMA (8.0/1.0)
(hsa-miR-330-5p <= 3.533) => Class=KIDNEY CARCINOMA (5.0/0.0)
(hsa-miR-218 >= 7.41) => Class=PANCREAS CARCINOMA (12.0/2.0)
(hsa-miR-383 >= 7.298) => Class=ESOPHAGUS CARCINOMA (11.0/3.0)
=> Class=COLON CARCINOMA (21.0/12.0)

Observing the rules we noted that for a classification of cancers sample, these system utilized only a single miRNA value. For example, hsa-miR-9 is sufficient to discriminate the Glioma class; this finding was well characterized in literature. In fact, recently miR-9, basing on data from the Tissue Cancer Genome Atlas (TCGA

<https://tcga-data.nci.nih.gov/tcga/>), was specific to a particular subtype of glioma, characterized by oligoneural precursor [93]. Similarly, the hsa-miR-122 was specific for the Hepatocellular carcinoma with a high value of expression [68, 94] .

We noted that the accuracy for a corrected classification of the classes was 68.2%.

Therefore the J48 system have a better classification accuracy. These type of classification, used an interpretation based on combinations of one-to three microRNAs in each node to make a decision, it will be useful for the identification of the tissue origin of cancers of unknown primary origin (CUP), a significant fraction of cancers that presented one or more metastases but where it was not possible identify the tissue of origin. Other previous study have been shown models based on miRNAs expression for the identification of tumor tissue origin of CUP [95, 96].

5 REPROGRAMMING OF miRNA NETWORK

5.1 Material and Methods

We used several bioinformatic tools that we described in the sections above. An SQL miRNA internal database was built with the data retrieved from a large number of different experiments performed in our laboratory.

5.1.1 Data analysis

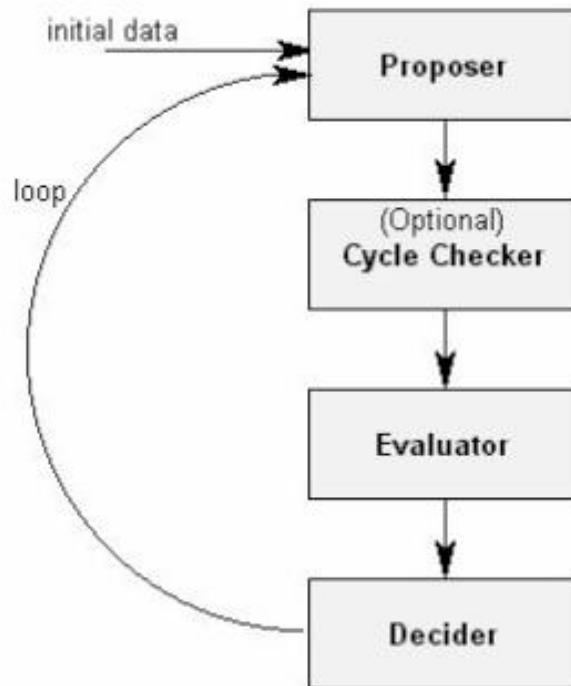
The union of the target mRNAs was used as an input to DAVID EASE, using the David Bioinformatics Resources system (<http://david.abcc.ncifcrf.gov>). We compared the list of terms related to the predicted targeted mRNAs.

The terms were evaluated by p value ($P < 0.05$) and Benjamini– Hochberg correction for multiple testing controlled the P values. Target genes selection was carried out by Diana microT.

5.1.2 Network generation

We used Banjo to infer the Bayesian network for the different tissues and diseases. Banjo is a software application and framework for structure learning of static and dynamic Bayesian networks and it was designed from the ground up to provide efficient structure inference when analyzing large, research-oriented data sets. Banjo focuses on score-based structure inference. Banjo have as core of the algorithm a “searcher”.

Banjo Searcher



The first step, Proposer, consists in selecting a graph structure, Grough, to be evaluated according to the data. The strategies currently implemented in Banjo are the simulated annealing and the hill climbing. The graph Grough is then scanned for cycles, Cycle Checker, and the acyclic graph G generated. The two possible strategies to visit the graph and remove the cycles are the Breadth First Search e Depth First Search. The acyclic graph G is then evaluated, Evaluator, according to a scoring function. Banjo implements a Bayesian score, dened as the logarithm of the probability of the proposed graph model given the data. The Decider decides,

possibly stochastically, whether to accept the proposed network (as the new current network) and best scored networks are then reported. For each tissue or disease all the mature expressed and varying miRNAs were used as input to Banjo.

The expression values were preprocessed with Gene Pattern to only filter out non varying miRNAs, according to the following parameters: #filter.flag = filter (Variation filter and thresholding flag); #preprocessing.flag = no disc or norm (Discretization and normalization flag); #minchange = 10 (Minimum fold change for filter); #mindelta = 512 (Minimum delta for filter); #threshold = 64 (Value for threshold); #ceiling = 20000 (Value for ceiling); #max.sigma.binning = 1 (Maximum sigma for binning); #prob.thres = 1 (Value for uniform probability threshold filter); #num.excl = 2% of total chips (Number of experiments to exclude (max & min) before applying variation filter); #log.base.two = no (Whether to take the log base two after thresholding); #number.of.columns.above.threshold = 17 1% of total chips (Remove row if n columns not >= than given threshold above.threshold); #column.threshold = 512 (Threshold for removing rows). We then performed a quality control step to remove chips with abnormal expression distribution across miRNAs: chips were retained only if less than 25% of miRNAs were absent (expression value <64). Similarly, miRNAs were retained only when less than 25% of samples had absent expression (value <64). The static Bayesian network inference algorithm was run on the miRNA expression matrix by using standard parameters, with a discretization policy of q6. Consensus graphs, based on top 100 networks, were obtained from at least 8 x10⁹ searched networks. We applied the MCL graph-based clustering algorithm to extraction of clusters (i.e. groups of densely connected nodes) from miRNA networks. We used MCL (Neat), an unsupervised cluster algorithm for graphs based on simulation of (stochastic) flow in graphs, that has been shown to enable good performances in extracting co-regulated genes from transcriptome networks. We used yEd graph editor (yFiles software, Tübingen, Germany) for graphs visualization.

5.2 miRNA Bayesian network of normal samples

We applied Banjo to infer the Bayesian network for normal tissues. miRNA relations were modeled as graphs where nodes represent the miRNAs and colored edges the relationships between them. We extracted coherent groups of nodes by adopting clustering algorithms, in order to discover miRNA groups with highly related expression patterns. We used the MCL graph-based algorithm (Enright et al. 2002). Figure 17 displays the miRNA network of normal tissues, obtained from over 1000 samples and 50 cell types/tissues. We used all of the expressed miRNAs to build the Bayesian networks (rather than only the differentially expressed ones). The MCL clusters with high co-expression patterns throughout normal tissues are linked by specific colored edges.

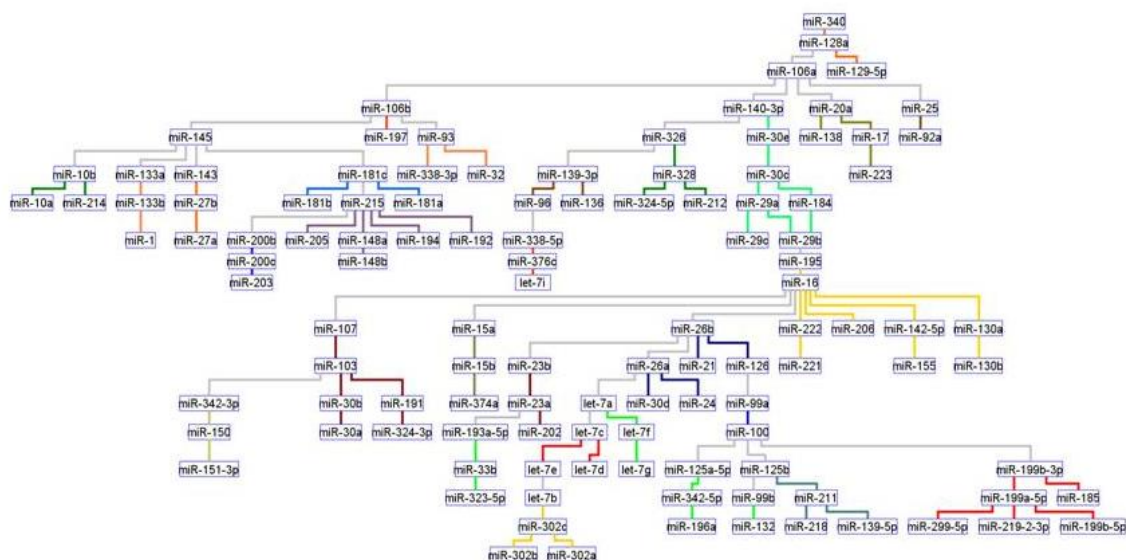


Figure 17 miRNA network in normal tissues (1107 samples, 50 tissues, 115 miRNAs)

We find in literature some confirmations about our network developed. The proliferation cluster hsa-miR-106a/b/93 linked to hsa-miR-20a/17 and to hsa-miR-25/92a (MYC associated) includes hsa-miR-223 is involved in cell cycle progression [97]. In the opposite side, on the left of the image, we find hsa-miR-133a/b were in a cluster with hsa-miR-1 (light orange) and all were involved in skeletal muscle proliferation and differentiation. A close cluster is hsa-miR-10a/b and hsa-miR-214 (green). hsa-miR-214 can modulate the expression of genes regulated by Hedgehog. Inhibition of hsa-miR-214 resulted in a reduction or loss of slow-muscle cell types, so these muscle/differentiation clusters are linked to hsa-miR-143/145 [98]. The hsa-miR-145 node links the proliferation clusters described above to the muscle differentiation clusters.

The hsa-miR-29 family, targets genes are induced in quiescence [99], is linked to hsa-miR-30 and to hsa-miR-15/16, miRNAs that target the anti-apoptotic protein BCL2. Hsa-miR-221-222, are considered regulators of the cell cycle, together with hsa-miR-206, hsa-miR-155 (pre-B cell proliferation), and hsa-miR-130a/b are in a yellow cluster [74].

In one of the center branch the hsa-miR-200 family, hsa-miR-203 and hsa-miR-205 are known to be directly involved in TGF- β mediated EMT and differentiation [100]. hsa-miR-181 family members are involved in hematopoietic differentiation [101].

Has-miR-145 act as an hub for hsa-miR-181, hsa-miR-200, hsa-miR-205, hsa-miR-215, so the loss of the hsa-miR-145 might also impact the TP53/EMT/differentiation branch. Another important hub appears to be hsa-miR-16, which is located in the other half of the net and coordinates hsa-miR-29 and hsa-miR-221/206/155/130 clusters. Together, miR-15 and miR-16 function by targeting multiple oncogenes, including BCL2, MCL1, CCND1, and WNT3A [102]. The hsa-miR-16 hub also feeds the hsa-miR-26/let-7/hsa-miR-302 branch. hsa-miR-26 and their host genes cooperate to inhibit the G1/S transition by activating the pRb protein [103]. Let-7 regulates Ras and hsa-miR-302 are expressed in ES cells and other early embryonic tissues.

5.3 miRNA Bayesian network of solid cancers

Initially, we started inferring a global miRNA expression network for solid cancers (Figure 18). The most connected hub was represented by hsa-miR-30c (degree 10), followed by hsa-miR-16 (degree 6). In the normal condition, hsa-miR-16 was the most connected node (degree 8) and hsa-miR-30c had only a low degree of 3. In the opposite way we identified: hsa-miR-215 (degree 6 in normal tissues and 3 in cancer) and hsa-miR-103/106a (degree 5 in normal tissues and only 1 in cancer). The exchanges of hubs between non-malignant and cancer tissues was the first notable sign of divergences in their respective miRNA programs. The MCL clustering algorithm was employed to map the sub-networks with high co-expression patterns (these MCL clusters, or cliques, are linked by specific colored edges). For simplify the visualization of the miRNA according to their differential expression in tumors, we color-coded the miRNA nodes (red = over-expressed; green = down-regulated). Neighbors preferentially appeared with the same trend, reflecting clustered miRNAs were either over-expressed or down-regulated. For example, hsa-miR-17/20a, hsa-miR-106a/b, and hsa-miR-93 were all up-regulated in cancers. Conversely, hsa-miR-143/145, hsa-miR-133a/b, hsa-miR-214 and hsa-miR-138, all in the same co-expression clique, were down-regulated.

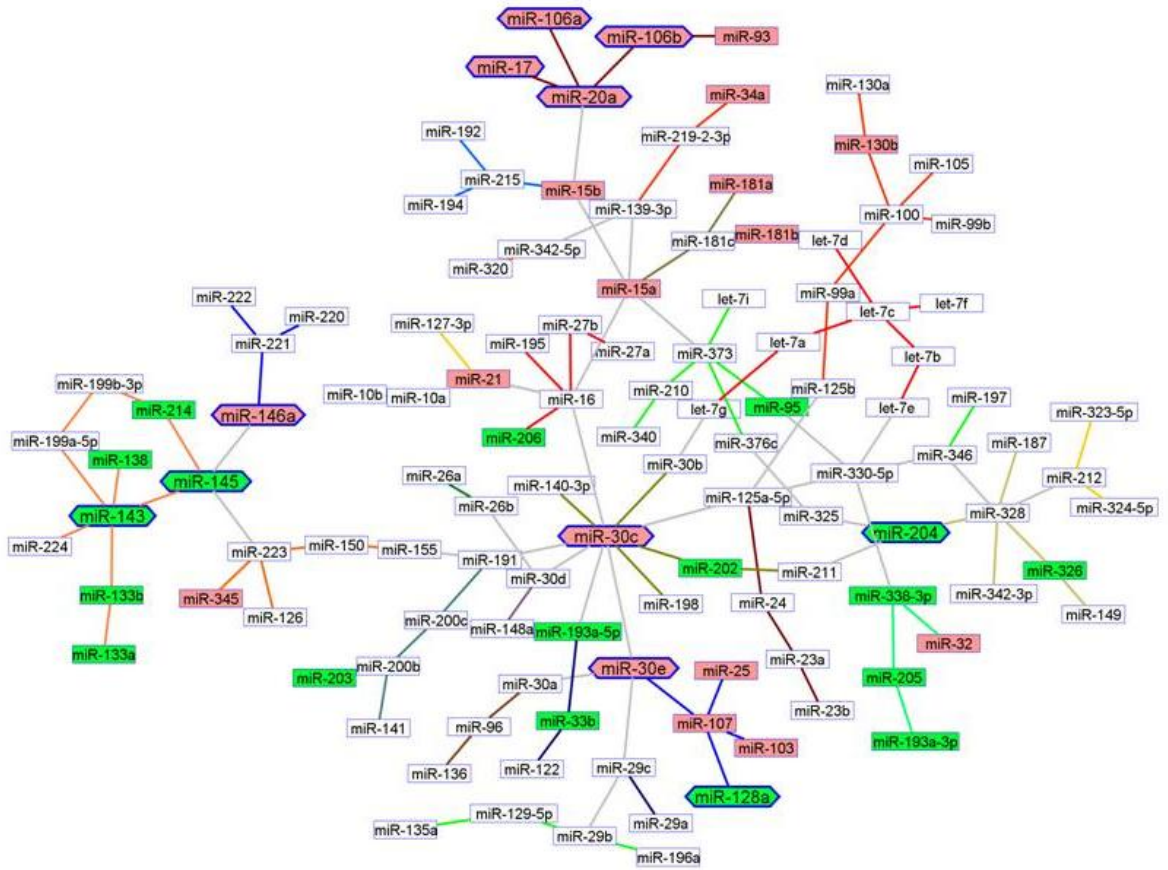


Figure 18 The miRNA network in solid cancers (2532 samples, 31 cancer types, 120 miRNAs). The network was inferred for all expressed and varying miRNAs, without preselecting for differential expression. miRNAs are either red (over-expressed) or green (down-regulated).

The single graph in the overall solid cancer net can be explained only by the same miRNAs having variable roles in a range of cancers, as in normal condition it can regulated different expressed targets in a specific tissues. Afterwards, we focused on the investigation of selected cancer types (lung, colon, breast, prostate). For each cancer tissues we built graphs for normal and cancer tissues and applied MCL to

extract miRNA co-expression clusters. Lung cancer the first that we analyzed; the adenocarcinomas' miRNome was represented by 1 major and 8 unconnected sub-networks, while the normal counterpart was a single complete miRNA network Figure 19).

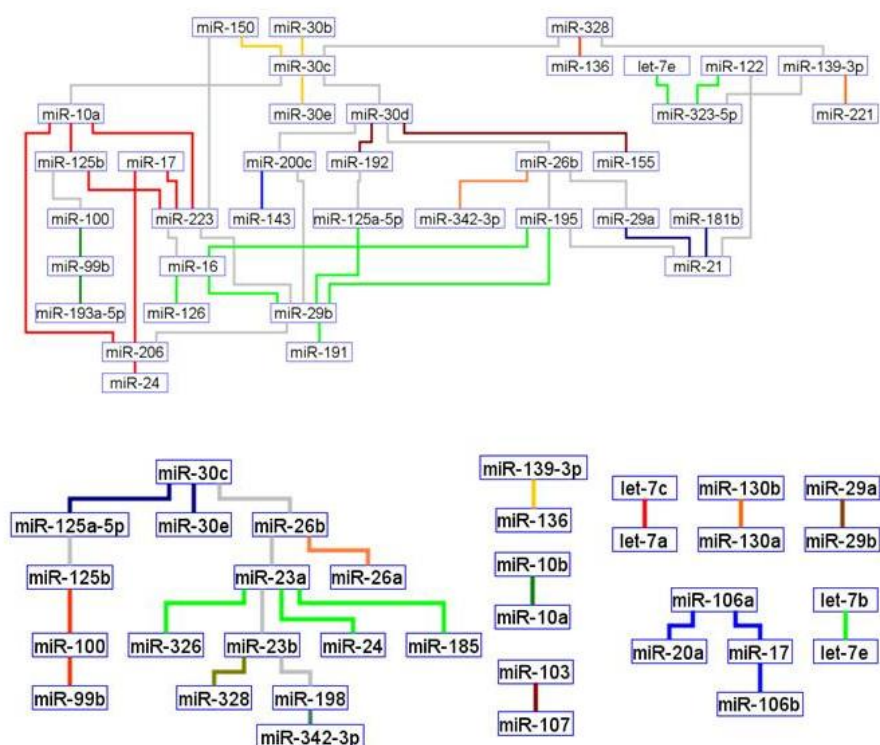


Figure 19 Comparison of miRNA networks in normal lung and adenocarcinoma. A) Normal lung (71 samples). The figure shows a single complete miRNA network. B) Lung adenocarcinoma (125 samples).

Notably, for the other cancer investigated: colon, breast, and prostate cancers (Figure 20,21 and 22 respectively) the disjoint trends appear similar to lung cancer,

consequently we have many disjoint miRNA network. In order to understand if this miRNA nets were involved in a similar pathway and which role that can exhibit in cancer, basically we compare the disjoint nets' tumors and we identified a number of notable, and often common, miRNAs in the unconnected clusters.

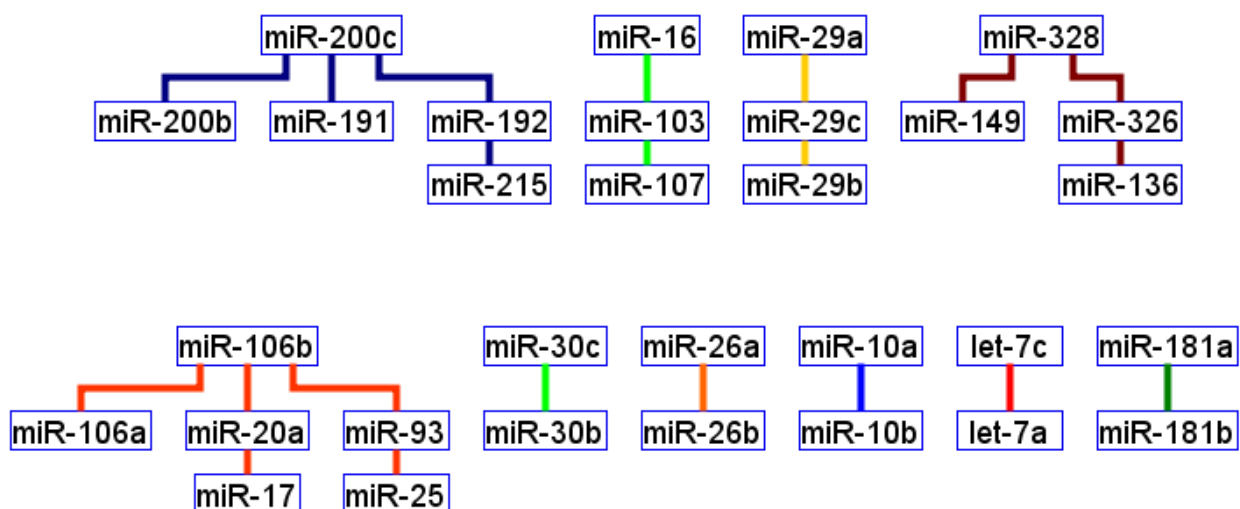


Figure 20 MicroRNA genetic network in colon adenocarcinoma

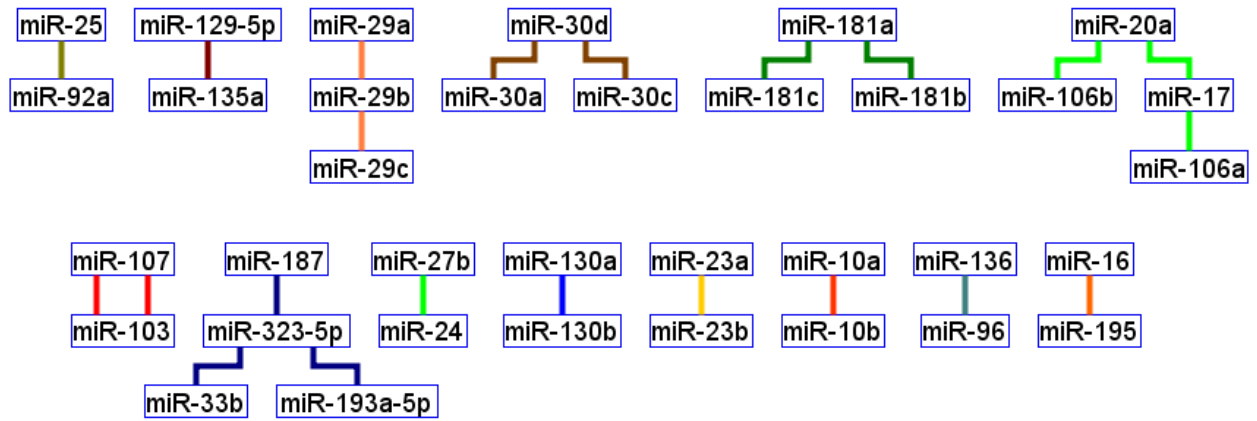


Figure 21 MicroRNA genetic network in breast cancer

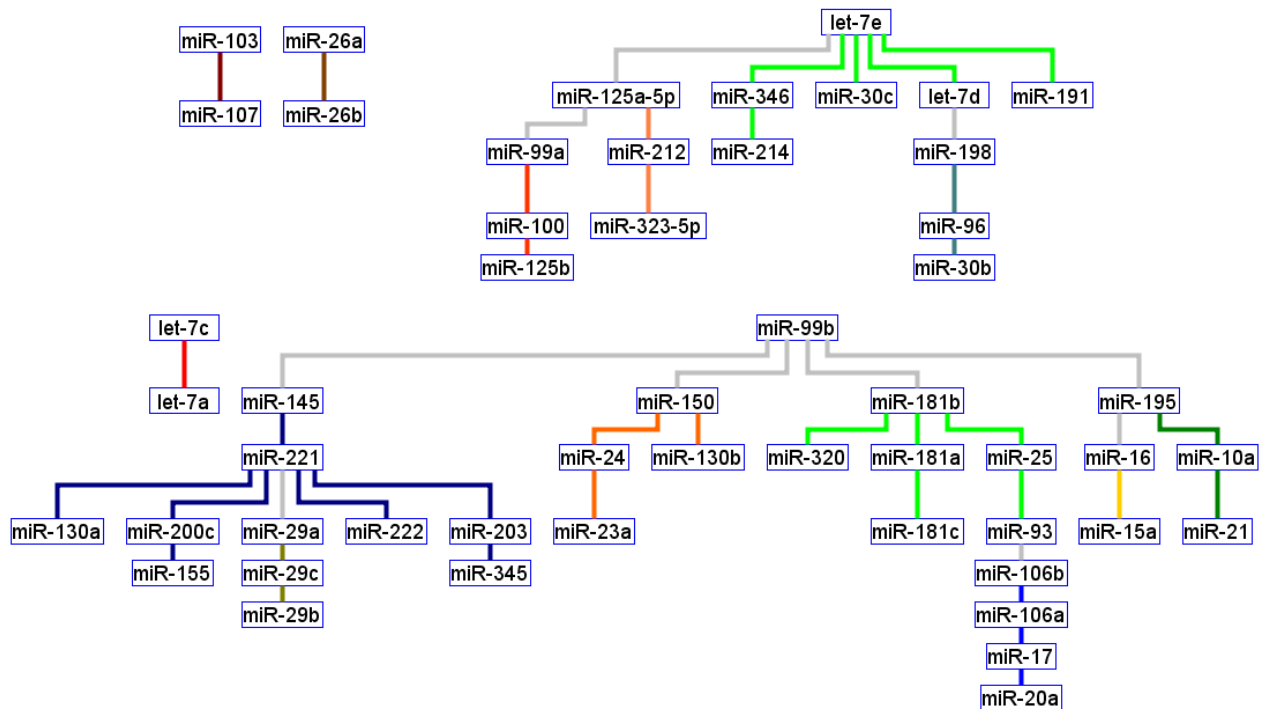


Figure 22 MicroRNA genetic network in prostate cancer

The microRNA found in common with the cancer were:

- hsa-miR-10a/b : identified in lung, colon and breast cancers;
- hsa-miR-26a/b: identified in colon and prostate cancers;
- hsa-miR-29a/b: identified in breast, colon, and lung cancers;
- hsa-miR-181 family members: identified in breast, colon and prostate cancers.

The hsa-miR-10a/b and hsa-miR-26a/b were well known key miRNAs implicated in several solid cancer [104]. Hsa-miR-29a/b and hsa-miR-181 were two most studied over-expressed miRNAs in multiple cancers and their role was associated with tumorigenesis and cancer progression [105]. At last, these two miRNAs' family were implicated in leukemia too, thus we can suppose a coordinated roles depending on the context and that can miRNAs can act as either oncogenes or tumours suppressors. Other miRNA cliques included hsa-miR-200c [106], linked to TP53-associated hsa-miR-192/215 in a colon sub-network, and high levels of miR-215 conferred chemoresistance due to cell cycle arrest and reduced cell proliferation in colon cancer patients [107]. Finally, we concluded that miRNA networks were reprogrammed in solid cancer and the expression of few notable miRNAs was independent from the major network.

5.4 miRNA Bayesian network of leukemia

We generated the networks for two hematological cancers, acute myeloid leukemia (AML) and chronic lymphocytic leukemia (CLL). MiRNA profiling of these two hematological malignancies have been well investigated and many data have generated a miRNA signatures about diagnosis, cytogenetics and prognosis [108]. The miRNA network in AML also had disjointed cliques (Figure 23). The most prominent finding here was that hsa-miR-155 and hsa-miR-181, two miRNAs with

clinical relevance were positioned in two separated sub-networks. In fact hsa-miR-155 was miR-155 has been shown to promote myeloid development and function and play a positive role in the promotion of cancer and autoimmunity [109], while the increment of the expression levels of hsa-miR-181 family members have been shown to be associated with favorable outcome in patients with cytogenetically normal acute myeloid leukemia , in opposite the up-regulation of a HOXA-PBX3 homeobox-gene signature following down-regulation of hsa-miR-181 is associated with adverse prognosis in patients with cytogenetically abnormal AML_ [110]. In fact, hsa-miR-181 was associated to hsa-miR-146a in a detached yellow mini-clique [111]. Has-miR-155 was in the same green clique of hsa-miR-223, hsa-miR-92a, hsa-miR-25 and hsa-miR-32. Finally, hsa-miR-29b has a key role in AML and, in accordance, it acts as a hub in the AML net [112].

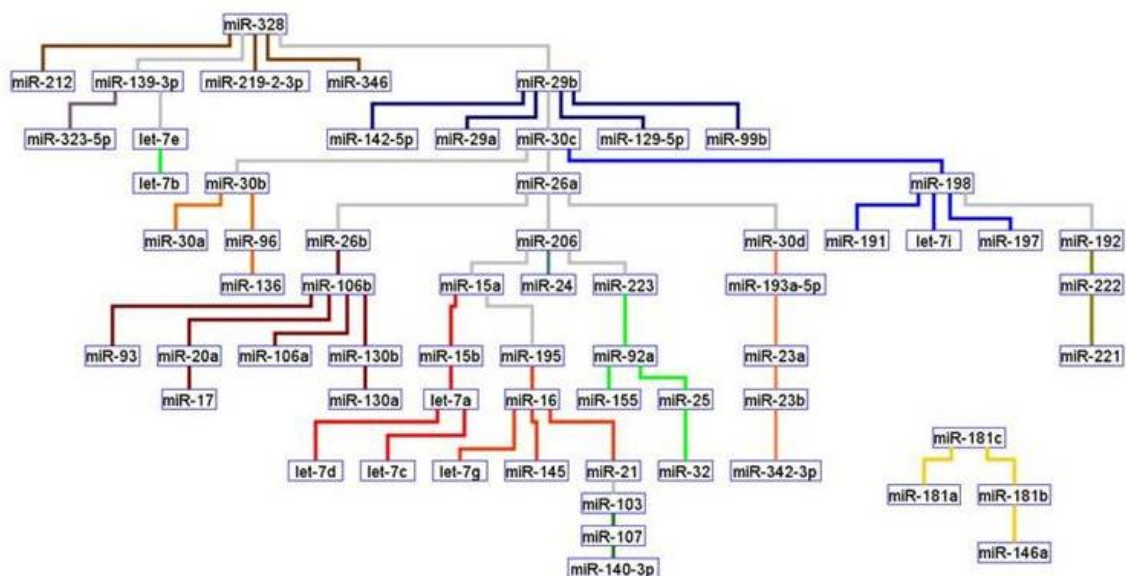


Figure 23 The miRNA network in acute myeloid leukemia (589 samples)

In chronic lymphocytic leukemia (CLL) two small cliques were separated from the main net (Figure 24): hsa-miR-23a/b [113] and the second one embracing the hsa-miR-15a and has-miR-195 and has-miR-16. miR-15/16 were miRNAs frequently deleted in CLL, functioned by targeting multiple oncogenes, including BCL2, MCL1, CCND1, and WNT3A [102]. Thus, the network topologies for these two leukemias could recapitulate their respective molecular pathology, with the key hsa-miR-29b acting as a hub in AML, but it was presented only in a branch in CLL. AML prognostic hsa-miR-181 was disjointed in AML but not in CLL, with the reverse being true for the CLL prognostic hsa-miR-15/16 pair.

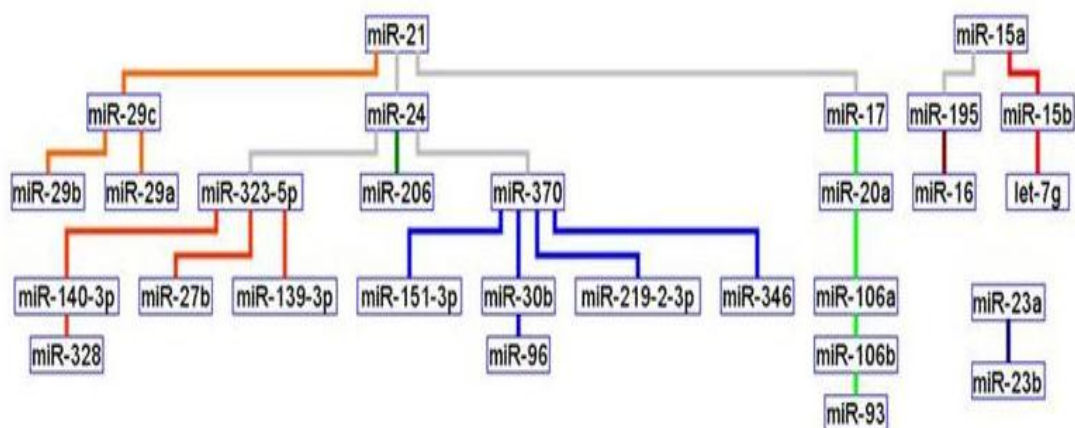


Figure 24 The miRNA network in chronic lymphocytic leukemia (254 samples)

6 IN SILICO, IN-VITRO AND IN-VIVO VALIDATION

6.1 Material and Methods

The *in vitro* analyses were performed in cell line. For *in silico* validation we used several bioinformatic tools. We used array comparative genomic hybridization (aCGH) using samples from selected GEO datasets and Stanford Microarray Database (SMD) resumed in Table A11. 744 arrays were studied (537 samples from GEO and 207 from SMD). All platforms were 2-channel based, data were downloaded as normalized values, and genes were annotated according to the gene symbol. All normalized log ratios were converted to log₂ ratios. Bootstrap analysis was used (10,000 random swaps of cancer and control channels) to obtain p-values and confidence limits for deletion and amplifications. We investigated 306 miRNA loci; 168 miRNA loci were associated to a host gene, and 138 miRNA loci to the flanking genes. miRNA families were defined according to TargetScan. The threshold p-value for a miRNA family was set at 0.05 to the number of family members. To control for multiple testing, we performed 100 bootstrapping cycles and used the results to calculate the FDR.

We used for target genes selection a software (described in introduction) DIANA-miRpath. After, the union of the target mRNAs with a score above 3 was used as an input to ClueGO [114]. ClueGO visualizes the selected terms in a functionally grouped annotation network that reflects the relationships between the terms based on the similarity of their associated genes. ClueGO was used to relate differential expression in cancer to functional pathways (KEGG). The size of the nodes reflects the statistical significance of the terms. The degree of connectivity between terms (edges) is calculated using kappa statistics. The calculated kappa score is also used for defining functional groups. A term can be included in several groups. The re-occurrence of the term is shown by adding " n". The not grouped terms are shown in white color. The group leading term is the most significant term of the group. The

network integrates only the positive kappa score term associations and is automatically laid out using Organic layout algorithm supported by Cytoscape. Right-sided hyper-geometric test yielded the enrichment for GO-terms. Benjamini-Hochberg correction for multiple testing controlled the p-values.

6.1.1 Cell line culture

Embryonic stem (ES) cell lines were cultured at 37°C and 5% CO₂ in specific medium. Total RNA was extracted by using TRIZOL method. To remove of contaminating genomic DNA from RNA samples, 1 µg each of them are treated with 1U DNase I Ampli Grade (Invitrogen) in final reaction volume of 10 µl. The incubation time was 10 minutes at room temperature. The enzyme was inactivated by addition of 25mM of EDTA solution followed by heating at 65°C for 10 min.

6.1.2 Isolation and reverse transcription of a total RNA

We collected tissue samples from mice. Sections from each tissue sample were lysed in TRIZOL reagent (Invitrogen Life Technologies, Carlsbad, CA USA). Total RNA was extracted using a standard protocol followed by purification with the Qiagen RNeasy Mini Kit (QIAGEN Inc, Valencia, CA USA). RNA purity was assessed by the ratio of spectrophotometric absorbance at 260 and 280 nm (A_{260}/A_{280}) using NanoDrop ND-1000 (NanoDrop Inc, Wilmington, DE USA).

Total cellular RNA (800 ng) was reverse transcribed using random hexamers (20pmoles Invitrogen Life Technologies, Carlsbad, CA) and 1mM dNTP mix in a total reaction volume of 10 µl. The mixture was incubated at 65°C for 5 minutes for denaturization, and snaped cooled. SuperScript III enzyme (200 U ,Invitrogen) was added to the mixture, with 100 mM MgCl₂, 0.2 M DDT and 40 U of RNase OUT

(Invitrogen Life Technologies) in final reaction volume of 20 μ l. The reaction was incubated at 25°C for 10 minutes followed by 50 minutes at 50°C, followed by 85°C for 5 minutes and chilled in ice. To increase the sensitivity of the PCR reaction we added 2 U of RNase H to remove the RNA template from cDNA:RNA hybrid molecule. Finally, the resulting cDNA was diluted to a final concentration of 10 ng/ μ l.

6.1.2 Real-time PCR

In order to quantify T-UCRs we used real-time PCR, with SYBR green chemistry. The qPCR reaction was carried out in a total volume of 10 μ l containing 10 ng of cDNA, 2X SYBR Green Mastermix (BioRAD) and 100nM of the specific UCR primer mix. The 18S rRNA was used as reference gene. Reactions were set up in 384-well plates. All reactions were performed in triplicate. Cycling conditions were as follows: 95°C for 3 minutes, 40 cycles at 95°C for 15 seconds and at 60°C for 1 minutes, followed by a melting curve. For each cDNA template, the triplicate T-UCR Ct values were averaged, and the normalized Ct (Δ Ct) was calculated by subtracting the mean Ct value for 18S rRNA from each T-UCR mean Ct value.

6.2 Real-Time PCR confirms the uc.283 plus specificity for ESC

To validate our microarray data that have showed uc.283 plus specific for embryo stem (ES) cells, we used quantitative real time polymerase chain reaction (RT-qPCR) on RNAs extracted from ESCs and five mouse adult tissues (spleen, seminal vesicles, thymus, lung, and bone marrow). We used miR-302 as a positive control, a miRNA specifically expressed in ESCs, as we have shown previously. The RT-qPCR (Figure 25) confirmed strong and ESC-specific expression for both miR-302 and uc.283 plus (Wilcoxon test p-value <0.01).

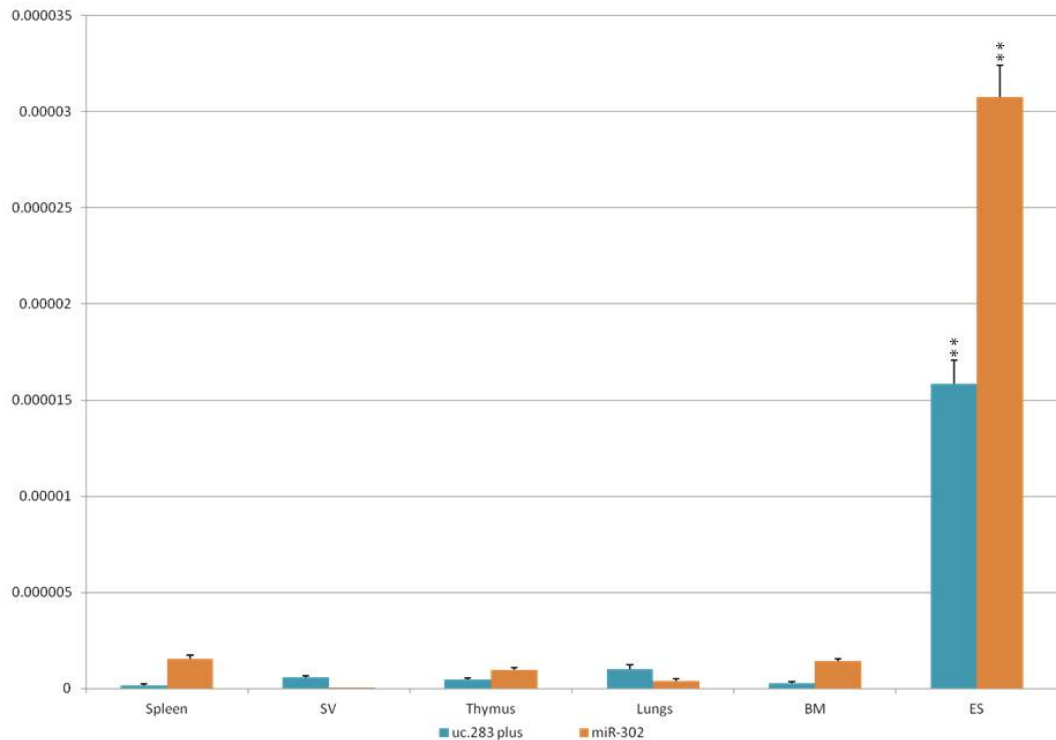


Figure 25 Real time PCR confirms the results of the microarray analysis. The uc.283 plus is expressed at higher levels in the embryonic stem cells (ES) than in the adult tissues, such as spleen, seminal vesicles (SV) , thymus, lungs and bone marrow (BM). The miR-302 was used as a positive control. Graph include labeled error bars and ** represents the p-value < 0.01

Expression data values were listed in Table A12. No reported enhancer activity was associated using the Vista Enhancer Browser with uc.283 at E11.5. In a further quest to investigate uc.283 function, we studied a 2,000 bp region surrounding this UCR on the UCSC Genome Browser (Figure 26). In the H1-hESC Chromatin State Segmentation (HMM track from ENCODE/Broad), this sequence might be an inactive or poised promoter. Moreover the uc.283 locus presented a low signal for H3K4me1 in the histone modification ChIP-Seq data of the H1 ESC line, suggesting possible monomethylation. Monomethylation at H3K4me1 is a marker for the

presence of an active enhancer, in contrast to promoters, marked by trimethylation of histone H3 at lysine residue 4 (H3K4me3).

In addition, the edges of uc.283 overlap the initial tracts of two open chromatin regions (OpenChrom_15681 and OpenChrom_15682). Data from the methyl 450K bead array track (ENCODE/HAIB) showed that, in close proximity to the start of uc.283 plus, the CpG sequence was unmethylated in H1-hESC and Human Umbilical Vein Endothelial Cells (HUVECs), and partially or totally methylated in other cell types. These findings, together with the proximity to open chromatin regions, strengthen the possibility that during embryonic development and in ESCs uc.283 plus could be located in a euchromatic region and be transcribed.

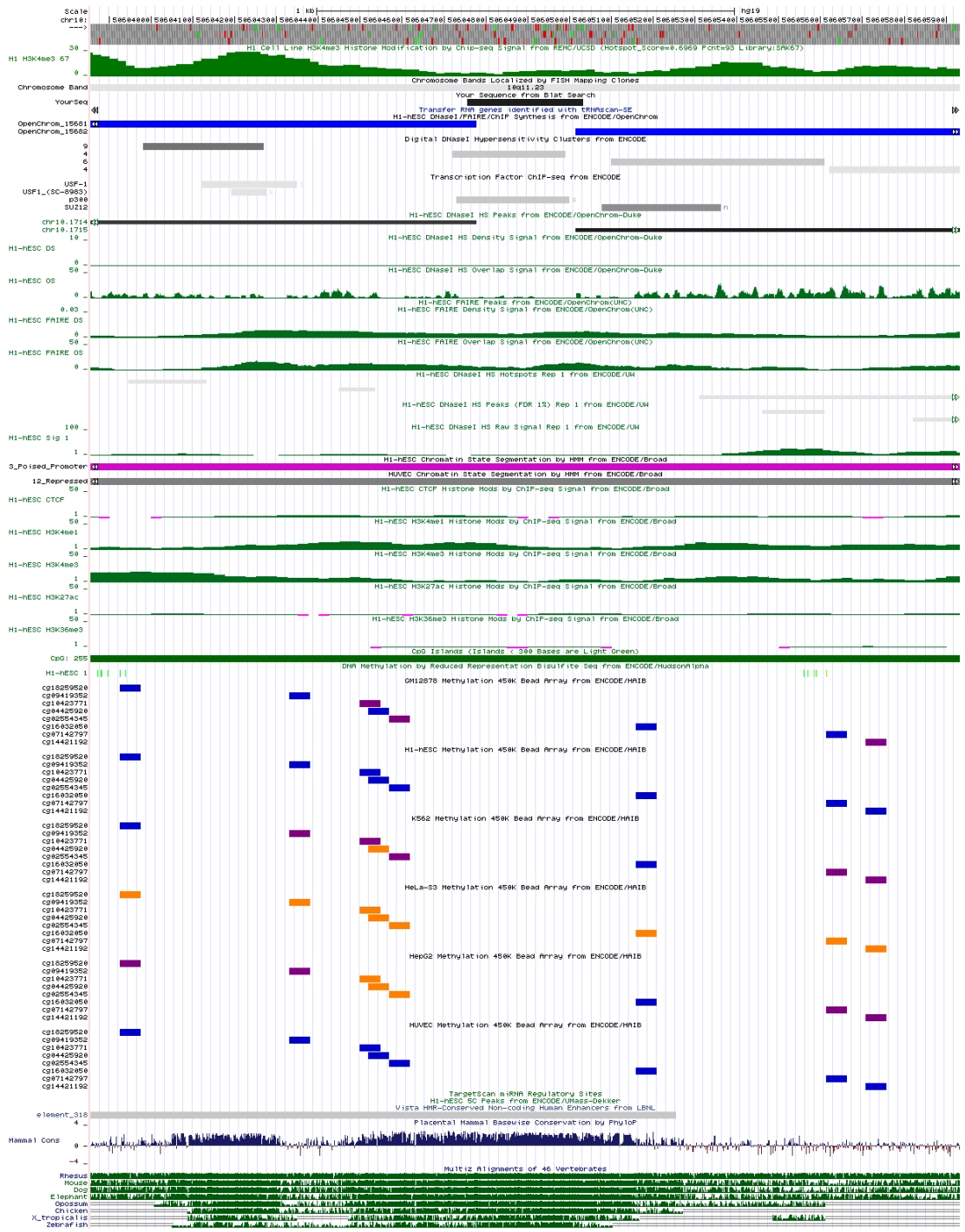


Figure 26 Analysis of uc.283 plus using UCSC genome browser.

6.3 Functional genome analysis: in silico validation

Functional annotation analysis investigated the involvement of non-coding RNA in several cellular processes. There are two way to understand and evaluate the previously insights. First, it is necessary explicated which are the target genes of the specific miRNAs differentially expressed, over and under-regulated in cancer samples. Second, it is possible used the target genes of the miRNAs (reverse expression analysis) that characterized the disjoint network previously analyzed in the section 5. Consequently, we used DIANA-miRpath in order to find target mRNAs. After, the union of the target mRNAs with a score above 3 was used as an input to DAVID EASE and ClueGO to investigate the pathway. Finally, we obtained a KEGG pie-chart corresponding to the functional effect of differentially-expressed miRNAs on cellular pathways, simultaneously targeted by both up-regulated and down-regulated miRNAs (Figure 27). Most of the affected pathways are related to cancer or signal transduction (i.e. Wnt, VEGF, TGF-beta, insulin, and phosphatidylinositol signaling, focal adhesion and colorectal cancer).

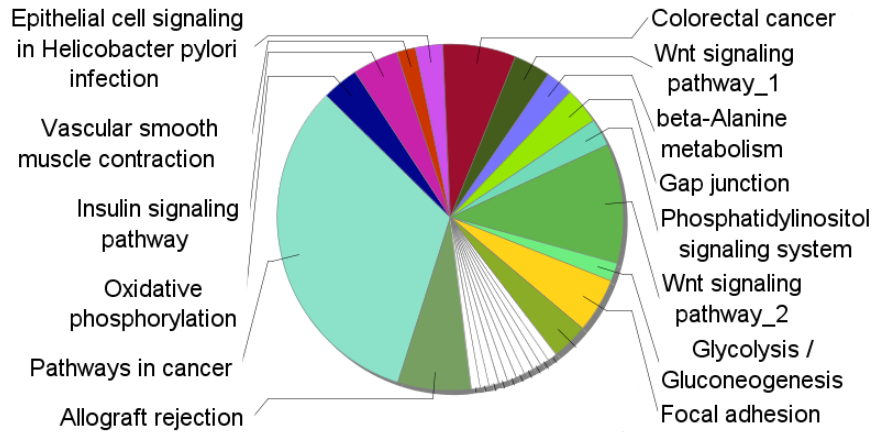


Figure 27 The KEGG pie-chart shows the functional effect of differentially-expressed miRNAs on cellular pathways in cancer.

For lung cancer we had identified 8 minor sub-networks unconnected in cancer, so we inferred a KEGG functional analysis of the miRNAs in the 8 minor sub-networks unconnected. The results showed that the miRNA which are present in the unconnected cliques target genes which are involved in many cancer-related terms, such as focal adhesion, small cell lung cancer, calcium signaling. (Figure 28).

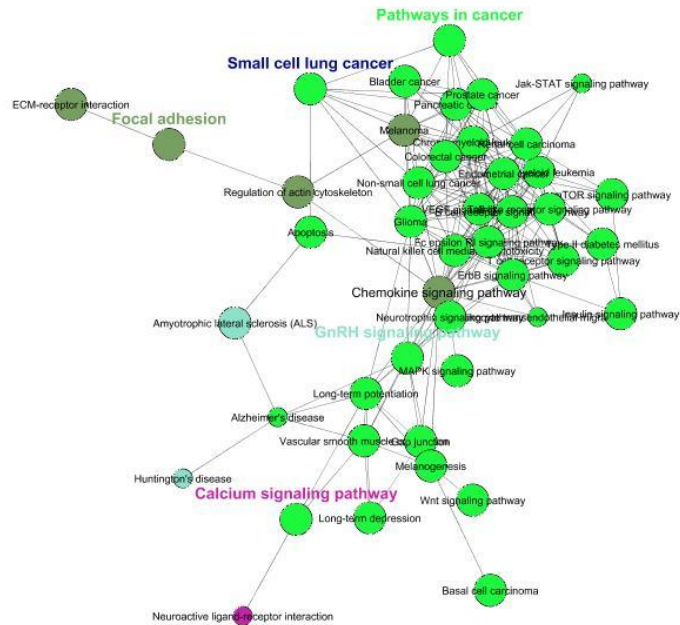


Figure 28 The KEGG functional analysis of 8 disjointed minor miRNA networks in lung adenocarcinoma.

Additionally, we used data from array comparative genomic hybridization (aCGH) to verify where the over or under-expression of miRNA differentially expressed in cancer is connected to genomic region amplified or deleted. We calculated, for each of 20,000 different chromosomal locations, two p-values, one for deletion and one for amplification. To measure miRNA copy number alterations we used their respective host genes or, when unavailable, their two flanking genes. In the Figure 18 we showed amplified and deleted miRNA families. The results from aCGH were overlaid on the expression network in solid cancers. The node labels, for which expression and physical alteration were concordant (i.e. over-expression and amplification), were embolden and visually reinforced with a hexagonally shaped border. The

detection of amplified hsa-miR-17-5p/20/93/106 family using flanking genes was a successful validation of our approach. Genomic amplification and elevated expression of mir-17-92 were both found in several human B-cell lymphomas, and its enforced expression exhibits strong tumorigenic activity in multiple mouse tumor models. In the Figure 18 the hsa-mir-17-92 cluster was in the midst of the cancer network. The top deleted miRNA family was hsa-miR-204/211, followed by other families including hsa-miR-200b/c/429, hsa-miR-141/200a, hsa-miR-125/351, and hsa-miR-218.

6.4 *In vivo* validation

Finally, we experimentally validated the miRNA network with acute lymphocytic leukemia originated in *Mir155* transgenic mice. We compared the miRNA profiles of three leukemia samples from these *Mir155* transgenes to controls from wild type mice (Table A). Then we located the positions in the network for the miRNAs regulated in the transgene's leukemias, most of miRNAs deregulated in these transgenic mice were located close to hsa-miR-155 in the cancer network. The yellow nodes appeared concentrated around the hsa-miR-155 node (black). We separated with a diagonal the hsa-miR-155 half from the other one and we compared the two sides: the difference in yellow nodes concentrations was significant (14 vs. 43, 4 vs. 57, Fisher exact test, 2-tail p-value<0.009). Interestingly the topological distribution of hsa-miR-29s and hsa-miR-181s were not considered as hsa-miR-155 regulated. In fact, hsa-miR-181 over-expression and hsa-miR-29 down-regulation are hallmarks miRNAs in leukemia, and this finding confirmed the independency of these events in cellular transformation.

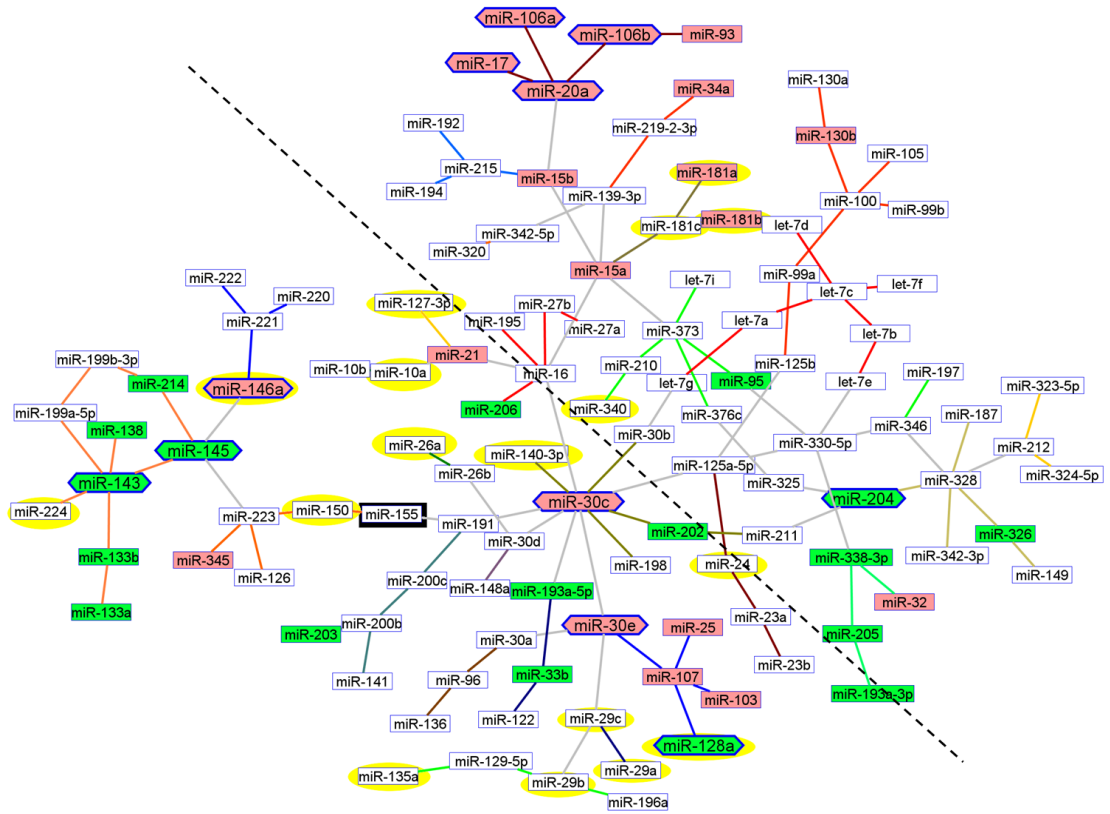


Figure 29 Deregulated miRNAs in leukemia from miR-155 transgenic mice are preferentially located close to miR-155 in the cancer network.

7 FINAL DISCUSSION

Non-coding RNAs were considered the most important novel arms of human post-transcriptional regulatory tools. They can specifically target different genes, often in a one-to-many way. Therefore fine-tuning the level of a single non-coding RNA might affect many pathways in a pleiotropic mode. In fact the effects of microRNAs on cell pathology and physiology are likely to be complex, each miRNA can control translation of tens or even hundreds of different coding messengers and a single mRNA can be controlled by more than one microRNA.

Thus, we have proposed a study of non-coding RNAs in cancer by applying a systems biology approach.

First, we presented a thorough analysis of miRNAs tissue specificity in 50 different normal tissues grouped by 17 systems, corresponding to 1070 human samples. We have found that a small set of miRNAs were tissue-specific while many others were broadly expressed. In the same way, we tested the expression of T-UCRs in 374 normal samples from 46 different tissues, grouped by 16 systems. We studied the global changes of T-UCRs expression, and only few T-UCRs were tissue-specific. We focused our study on the unique T-UCR specific for the embryonic condition, uc.283plus. We confirmed this specificity by Real-time PCR and we proposed this conserved transcript as a potential novel molecular marker of stemness. Furthermore we considered as a possible future investigation the uc.283 plus expression in cancer stem cells sub-populations.

We also investigated oncologic or hemato-oncologic disorders and identified cancer type-specific miRNAs and T-UCRs. Then, we used the differential expressed miRNAs to classify solid cancers. We tested and compared two different approaches: J48graft and JRip, that respectively were a decision-tree and a rule induction algorithm. By the computational optimization of these two classifiers we can correctly defined the type of solid cancer. We suggested, on the basis of the results, that a further implementation can be useful to increase the percentage of

corrected classification, in order to apply this tool for the investigation of cancer of unknown primary origin.

Afterwards, we inferred genetic networks directly from miRNA expression data for normal tissues and for several solid tumors and leukemia. We found that the normal tissues were represented by single complete miRNA networks, while the cancers were characterized by separate and unlinked miRNA sub-networks.

In fact, we analyzed in depth the sub-networks of the cancers and we found miRNAs independent from the general transcriptional program involved in specific cancer-related pathways. We confirmed our finding by several functional genome validation methods.

We investigated also two hematological malignancies: acute myeloid leukemia and chronic lymphocytic leukemia. MiRNAs related to AML and CLL pathogenesis presented a different behavior respect solid cancers and there were only one and two sub-networks respectively. The different observance of solid cancers and leukemia might be due to the diverging pathogenesis mechanisms that include differing oncogenic miRNA networks.

Finally, all these finding, taken together, leads us to conclude that the bioinformatics tools and analysis of “big data” with a system biology approach were fundamentals to try to understand the complexity of the biology of the cancer.

We have analyzed only a part of a huge human genomic landscape, two classes of the non-coding RNA side. Our findings permit us to have contributed to a better analysis of the functions of miRNAs and T-UCRs and their global impact that can effect on the pathogenesis of cancer.

References

1. Darnell JE, Jr.: **Variety in the level of gene control in eukaryotic cells.** *Nature* 1982, **297**(5865):365-371.
2. Galasso M, Elena Sana M, Volinia S: **Non-coding RNAs: a key to future personalized molecular therapy?** *Genome medicine* 2010, **2**(2):12.
3. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ: **Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila.** *Cell* 2007, **128**(6):1089-1103.
4. Grimson A, Srivastava M, Fahey B, Woodcroft BJ, Chiang HR, King N, Degan BM, Rokhsar DS, Bartel DP: **Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals.** *Nature* 2008, **455**(7217):1193-1197.
5. Malone CD, Hannon GJ: **Small RNAs as guardians of the genome.** *Cell* 2009, **136**(4):656-668.
6. Peters L, Meister G: **Argonaute proteins: mediators of RNA silencing.** *Molecular cell* 2007, **26**(5):611-623.
7. Yu W, Gius D, Onyango P, Muldoon-Jacobs K, Karp J, Feinberg AP, Cui H: **Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA.** *Nature* 2008, **451**(7175):202-206.
8. Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Philippov DV, Blaser H, Raz E, Moens CB *et al*: **A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish.** *Cell* 2007, **129**(1):69-82.
9. Kiss T: **Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions.** *Cell* 2002, **109**(2):145-148.
10. Bachellerie JP, Cavaille J, Huttenhofer A: **The expanding snoRNA world.** *Biochimie* 2002, **84**(8):775-790.
11. Balakin AG, Smith L, Fournier MJ: **The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions.** *Cell* 1996, **86**(5):823-834.
12. Luo Y, Li S: **Genome-wide analyses of retrogenes derived from the human box H/ACA snoRNAs.** *Nucleic acids research* 2007, **35**(2):559-571.
13. Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, Pfeffer S, Rajewsky N, Meister G: **A human snoRNA with microRNA-like functions.** *Molecular cell* 2008, **32**(4):519-528.
14. Taft RJ, Kaplan CD, Simons C, Mattick JS: **Evolution, biogenesis and function of promoter-associated RNAs.** *Cell Cycle* 2009, **8**(15):2332-2338.
15. Cao F, Li X, Hiew S, Brady H, Liu Y, Dou Y: **Dicer independent small RNAs associate with telomeric heterochromatin.** *RNA* 2009, **15**(7):1274-1281.

16. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C *et al*: **The transcriptional landscape of the mammalian genome.** *Science* 2005, **309**(5740):1559-1563.
17. **Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs.** *Nature* 2009, **457**(7232):1028-1032.
18. Mao YS, Sunwoo H, Zhang B, Spector DL: **Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs.** *Nat Cell Biol* 2011, **13**(1):95-101.
19. Saxena A, Carninci P: **Long non-coding RNA modifies chromatin: epigenetic silencing by long non-coding RNAs.** *BioEssays : news and reviews in molecular, cellular and developmental biology* 2011, **33**(11):830-839.
20. Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, Haussler D: **Ultraconserved elements in the human genome.** *Science* 2004, **304**(5675):1321-1325.
21. Bejerano G, Lowe CB, Ahituv N, King B, Siepel A, Salama SR, Rubin EM, Kent WJ, Haussler D: **A distal enhancer and an ultraconserved exon are derived from a novel retroposon.** *Nature* 2006, **441**(7089):87-90.
22. Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD *et al*: **In vivo enhancer analysis of human conserved non-coding sequences.** *Nature* 2006, **444**(7118):499-502.
23. Poulin F, Nobrega MA, Plajzer-Frick I, Holt A, Afzal V, Rubin EM, Pennacchio LA: **In vivo characterization of a vertebrate ultraconserved enhancer.** *Genomics* 2005, **85**(6):774-781.
24. Baira E, Greshock J, Coukos G, Zhang L: **Ultraconserved elements: genomics, function and disease.** *RNA biology* 2008, **5**(3):132-134.
25. Katzman S, Kern AD, Bejerano G, Fewell G, Fulton L, Wilson RK, Salama SR, Haussler D: **Human genome ultraconserved elements are ultraselected.** *Science* 2007, **317**(5840):915.
26. Licastro D, Gennarino VA, Petrera F, Sanges R, Banfi S, Stupka E: **Promiscuity of enhancer, coding and non-coding transcription functions in ultraconserved elements.** *BMC Genomics* 2010, **11**:151.
27. Fabbri M, Garzon R, Andreeff M, Kantarjian HM, Garcia-Manero G, Calin GA: **MicroRNAs and noncoding RNAs in hematological malignancies: molecular, clinical and therapeutic implications.** *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK* 2008, **22**(6):1095-1105.
28. Calin GA, Liu CG, Ferracin M, Hyslop T, Spizzo R, Sevignani C, Fabbri M, Cimmino A, Lee EJ, Wojcik SE *et al*: **Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas.** *Cancer Cell* 2007, **12**(3):215-229.
29. Rossi S, Sevignani C, Nnadi SC, Siracusa LD, Calin GA: **Cancer-associated genomic regions (CAGRs) and noncoding RNAs: bioinformatics and therapeutic implications.** *Mamm Genome* 2008, **19**(7-8):526-540.

30. Scaruffi P, Stigliani S, Moretti S, Coco S, De Vecchi C, Valdora F, Garaventa A, Bonassi S, Tonini GP: **Transcribed-Ultra Conserved Region expression is associated with outcome in high-risk neuroblastoma.** *BMC Cancer* 2009, **9**:441.
31. Mestdagh P, Fredlund E, Pattyn F, Rihani A, Van Maerken T, Vermeulen J, Kumps C, Menten B, De Preter K, Schramm A *et al*: **An integrative genomics screen uncovers ncRNA T-UCR functions in neuroblastoma tumours.** *Oncogene* 2010, **29**(24):3583-3592.
32. Filipowicz W, Bhattacharyya SN, Sonenberg N: **Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?** *Nature reviews Genetics* 2008, **9**(2):102-114.
33. Lee RC, Feinbaum RL, Ambros V: **The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*.** *Cell* 1993, **75**(5):843-854.
34. Rana TM: **Illuminating the silence: understanding the structure and function of small RNAs.** *Nat Rev Mol Cell Biol* 2007, **8**(1):23-36.
35. Doench JG, Sharp PA: **Specificity of microRNA target selection in translational repression.** *Genes & development* 2004, **18**(5):504-511.
36. Eis PS, Tam W, Sun L, Chadburn A, Li Z, Gomez MF, Lund E, Dahlberg JE: **Accumulation of miR-155 and BIC RNA in human B cell lymphomas.** *Proc Natl Acad Sci U S A* 2005, **102**(10):3627-3632.
37. Christensen BC, Moyer BJ, Avissar M, Ouellet LG, Plaza SL, McClean MD, Marsit CJ, Kelsey KT: **A *let-7* microRNA-binding site polymorphism in the *KRAS* 3' UTR is associated with reduced survival in oral cancers.** *Carcinogenesis* 2009, **30**(6):1003-1007.
38. Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M, Wojcik SE, Aqeilan RI, Zupo S, Dono M *et al*: **miR-15 and miR-16 induce apoptosis by targeting *BCL2*.** *Proc Natl Acad Sci U S A* 2005, **102**(39):13944-13949.
39. Bonci D, Coppola V, Musumeci M, Addario A, Giuffrida R, Memeo L, D'Urso L, Pagliuca A, Biffoni M, Labbaye C *et al*: **The miR-15a-miR-16-1 cluster controls prostate cancer by targeting multiple oncogenic activities.** *Nat Med* 2008, **14**(11):1271-1277.
40. Garofalo M, Croce CM: **microRNAs: Master regulators as potential therapeutics in cancer.** *Annu Rev Pharmacol Toxicol* 2011, **51**:25-43.
41. Lebanony D, Benjamin H, Gilad S, Ezagouri M, Dov A, Ashkenazi K, Gefen N, Izraeli S, Rechavi G, Pass H *et al*: **Diagnostic assay based on hsa-miR-205 expression distinguishes squamous from nonsquamous non-small-cell lung carcinoma.** *J Clin Oncol* 2009, **27**(12):2030-2037.
42. Hatley ME, Patrick DM, Garcia MR, Richardson JA, Bassel-Duby R, van Rooij E, Olson EN: **Modulation of K-Ras-dependent lung tumorigenesis by MicroRNA-21.** *Cancer Cell* 2010, **18**(3):282-293.

43. Johnson CD, Esquela-Kerscher A, Stefani G, Byrom M, Kelnar K, Ovcharenko D, Wilson M, Wang X, Shelton J, Shingara J *et al*: **The let-7 microRNA represses cell proliferation pathways in human cells.** *Cancer Res* 2007, **67**(16):7713-7722.
44. Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T: **Impaired microRNA processing enhances cellular transformation and tumorigenesis.** *Nat Genet* 2007, **39**(5):673-677.
45. Takamizawa J, Konishi H, Yanagisawa K, Tomida S, Osada H, Endoh H, Harano T, Yatabe Y, Nagino M, Nimura Y *et al*: **Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival.** *Cancer Res* 2004, **64**(11):3753-3756.
46. Yanaihara N, Caplen N, Bowman E, Seike M, Kumamoto K, Yi M, Stephens RM, Okamoto A, Yokota J, Tanaka T *et al*: **Unique microRNA molecular profiles in lung cancer diagnosis and prognosis.** *Cancer Cell* 2006, **9**(3):189-198.
47. Chin LJ, Ratner E, Leng S, Zhai R, Nallur S, Babar I, Muller RU, Straka E, Su L, Burki EA *et al*: **A SNP in a let-7 microRNA complementary site in the KRAS 3' untranslated region increases non-small cell lung cancer risk.** *Cancer Res* 2008, **68**(20):8535-8540.
48. Lee EJ, Gusev Y, Jiang J, Nuovo GJ, Lerner MR, Frankel WL, Morgan DL, Postier RG, Brackett DJ, Schmittgen TD: **Expression profiling identifies microRNA signature in pancreatic cancer.** *Int J Cancer* 2007, **120**(5):1046-1054.
49. Bloomston M, Frankel WL, Petrocca F, Volinia S, Alder H, Hagan JP, Liu CG, Bhatt D, Taccioli C, Croce CM: **MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis.** *Jama* 2007, **297**(17):1901-1908.
50. Zhang Y, Li M, Wang H, Fisher WE, Lin PH, Yao Q, Chen C: **Profiling of 95 microRNAs in pancreatic cancer cell lines and surgical specimens by real-time PCR analysis.** *World J Surg* 2009, **33**(4):698-709.
51. Dillhoff M, Liu J, Frankel W, Croce C, Bloomston M: **MicroRNA-21 is Overexpressed in Pancreatic Cancer and a Potential Predictor of Survival.** *J Gastrointest Surg* 2008.
52. Habbe N, Koorstra JB, Mendell JT, Offerhaus GJ, Ryu JK, Feldmann G, Mullendore ME, Goggins MG, Hong SM, Maitra A: **MicroRNA miR-155 is a biomarker of early pancreatic neoplasia.** *Cancer Biol Ther* 2009, **8**(4):340-346.
53. Ma Y, Yu S, Zhao W, Lu Z, Chen J: **miR-27a regulates the growth, colony formation and migration of pancreatic cancer cells by targeting Sprouty2.** *Cancer Lett*, **298**(2):150-158.
54. Huang X, Ding L, Bennewith KL, Tong RT, Welford SM, Ang KK, Story M, Le QT, Giaccia AJ: **Hypoxia-inducible mir-210 regulates normoxic gene expression involved in tumor initiation.** *Mol Cell* 2009, **35**(6):856-867.
55. Nakata K, Ohuchida K, Mizumoto K, Kayashima T, Ikenaga N, Sakai H, Lin C, Fujita H, Otsuka T, Aishima S *et al*: **MicroRNA-10b is overexpressed in pancreatic cancer,**

- promotes its invasiveness, and correlates with a poor prognosis.** *Surgery*, **150**(5):916-922.
56. Li Y, Vandenboom TG, 2nd, Wang Z, Kong D, Ali S, Philip PA, Sarkar FH: **miR-146a suppresses invasion of pancreatic cancer cells.** *Cancer Res*, **70**(4):1486-1495.
 57. Yan H, Wu J, Liu W, Zuo Y, Chen S, Zhang S, Zeng M, Huang W: **MicroRNA-20a overexpression inhibited proliferation and metastasis of pancreatic carcinoma cells.** *Hum Gene Ther*, **21**(12):1723-1734.
 58. Yu S, Lu Z, Liu C, Meng Y, Ma Y, Zhao W, Liu J, Yu J, Chen J: **miRNA-96 suppresses KRAS and functions as a tumor suppressor gene in pancreatic cancer.** *Cancer Res*, **70**(14):6015-6025.
 59. Wang J, Chen J, Chang P, LeBlanc A, Li D, Abbruzzese JL, Frazier ML, Killary AM, Sen S: **MicroRNAs in plasma of pancreatic ductal adenocarcinoma patients as novel blood-based biomarkers of disease.** *Cancer Prev Res (Phila)* 2009, **2**(9):807-813.
 60. Liu R, Chen X, Du Y, Yao W, Shen L, Wang C, Hu Z, Zhuang R, Ning G, Zhang C *et al*: **Serum MicroRNA Expression Profile as a Biomarker in the Diagnosis and Prognosis of Pancreatic Cancer.** *Clin Chem*.
 61. Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, Magri E, Pedriali M, Fabbri M, Campiglio M *et al*: **MicroRNA gene expression deregulation in human breast cancer.** *Cancer Res* 2005, **65**(16):7065-7070.
 62. Yan LX, Huang XF, Shao Q, Huang MY, Deng L, Wu QL, Zeng YX, Shao JY: **MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis.** *RNA* 2008, **14**(11):2348-2360.
 63. Qian B, Katsaros D, Lu L, Preti M, Durando A, Arisio R, Mu L, Yu H: **High miR-21 expression in breast cancer associated with poor disease-free survival in early stage disease and high TGF-beta1.** *Breast Cancer Res Treat* 2009, **117**(1):131-140.
 64. Farazi TA, Horlings HM, Ten Hoeve JJ, Mihailovic A, Halfwerk H, Morozov P, Brown M, Hafner M, Reyat F, van Kouwenhove M *et al*: **MicroRNA sequence and expression analysis in breast tumors by deep sequencing.** *Cancer Res*, **71**(13):4443-4453.
 65. Volinia S, Galasso M, Sana ME, Wise TF, Palatini J, Huebner K, Croce CM: **Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA.** *Proc Natl Acad Sci U S A*.
 66. Lee JW, Park YA, Choi JJ, Lee YY, Kim CJ, Choi C, Kim TJ, Lee NW, Kim BG, Bae DS: **The expression of the miRNA-200 family in endometrial endometrioid carcinoma.** *Gynecologic oncology* 2011, **120**(1):56-62.
 67. Ladeiro Y, Couchy G, Balabaud C, Bioulac-Sage P, Pelletier L, Rebouissou S, Zucman-Rossi J: **MicroRNA profiling in hepatocellular tumors is associated with clinical features and oncogene/tumor suppressor gene mutations.** *Hepatology* 2008, **47**(6):1955-1963.

68. Pineau P, Volinia S, McJunkin K, Marchio A, Battiston C, Terris B, Mazzaferro V, Lowe SW, Croce CM, Dejean A: **miR-221 overexpression contributes to liver tumorigenesis.** *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**(1):264-269.
69. Schetter AJ, Leung SY, Sohn JJ, Zanetti KA, Bowman ED, Yanaihara N, Yuen ST, Chan TL, Kwong DL, Au GK *et al*: **MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma.** *Jama* 2008, **299**(4):425-436.
70. Schetter AJ, Nguyen GH, Bowman ED, Mathe EA, Yuen ST, Hawkes JE, Croce CM, Leung SY, Harris CC: **Association of inflammation-related and microRNA gene expression with cancer-specific mortality of colon adenocarcinoma.** *Clin Cancer Res* 2009, **15**(18):5878-5887.
71. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, Aldler H, Rattan S, Keating M, Rai K *et al*: **Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia.** *Proceedings of the National Academy of Sciences of the United States of America* 2002, **99**(24):15524-15529.
72. Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M, Wojcik SE, Aqeilan RI, Zupo S, Dono M *et al*: **miR-15 and miR-16 induce apoptosis by targeting BCL2.** *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102**(39):13944-13949.
73. Mi S, Lu J, Sun M, Li Z, Zhang H, Neilly MB, Wang Y, Qian Z, Jin J, Zhang Y *et al*: **MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia.** *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104**(50):19971-19976.
74. Spizzo R, Nicoloso MS, Croce CM, Calin GA: **SnapShot: MicroRNAs in Cancer.** *Cell* 2009, **137**(3):586-586 e581.
75. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB: **Prediction of mammalian microRNA targets.** *Cell* 2003, **115**(7):787-798.
76. Lewis BP, Burge CB, Bartel DP: **Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.** *Cell* 2005, **120**(1):15-20.
77. Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, Lim B, Rigoutsos I: **A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes.** *Cell* 2006, **126**(6):1203-1217.
78. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS: **Human MicroRNA targets.** *PLoS Biol* 2004, **2**(11):e363.
79. Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E: **The role of site accessibility in microRNA target recognition.** *Nat Genet* 2007, **39**(10):1278-1284.
80. Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M *et al*: **Combinatorial microRNA target predictions.** *Nat Genet* 2005, **37**(5):495-500.

81. Kiriakidou M, Nelson PT, Kouranov A, Fitziev P, Bouyioukos C, Mourelatos Z, Hatzigeorgiou A: **A combined computational-experimental approach predicts human microRNA targets.** *Genes & development* 2004, **18**(10):1165-1178.
82. Maragkakis M, Alexiou P, Papadopoulos GL, Reczko M, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, Simossis VA *et al*: **Accurate microRNA target prediction correlates with protein repression levels.** *BMC Bioinformatics* 2009, **10**:295.
83. Papadopoulos GL, Alexiou P, Maragkakis M, Reczko M, Hatzigeorgiou AG: **DIANA-mirPath: Integrating human and mouse microRNAs in pathways.** *Bioinformatics* 2009, **25**(15):1991-1993.
84. Beerenwinkel N, Siebourg J: **Probability, statistics, and computational science.** *Methods in molecular biology* 2012, **855**:77-110.
85. Sakhanenko NA, Galas DJ: **Probabilistic logic methods and some applications to biology and medicine.** *J Comput Biol* 2012, **19**(3):316-336.
86. Liu CG, Calin GA, Volinia S, Croce CM: **MicroRNA expression profiling using microarrays.** *Nature protocols* 2008, **3**(4):563-578.
87. Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M *et al*: **A microRNA expression signature of human solid tumors defines cancer gene targets.** *Proceedings of the National Academy of Sciences of the United States of America* 2006, **103**(7):2257-2261.
88. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ: **miRBase: tools for microRNA genomics.** *Nucleic acids research* 2008, **36**(Database issue):D154-158.
89. Bolstad BM, Irizarry RA, Astrand M, Speed TP: **A comparison of normalization methods for high density oligonucleotide array data based on variance and bias.** *Bioinformatics* 2003, **19**(2):185-193.
90. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A, Kamphorst AO, Landthaler M *et al*: **A mammalian microRNA expression atlas based on small RNA library sequencing.** *Cell* 2007, **129**(7):1401-1414.
91. Frank E, Hall M, Trigg L, Holmes G, Witten IH: **Data mining in bioinformatics using Weka.** *Bioinformatics* 2004, **20**(15):2479-2481.
92. Visel A, Minovitsky S, Dubchak I, Pennacchio LA: **VISTA Enhancer Browser--a database of tissue-specific human enhancers.** *Nucleic Acids Res* 2007, **35**(Database issue):D88-92.
93. Kim TM, Huang W, Park R, Park PJ, Johnson MD: **A developmental taxonomy of glioblastoma defined and maintained by MicroRNAs.** *Cancer research* 2011, **71**(9):3387-3399.
94. Jopling C: **Liver-specific microRNA-122: Biogenesis and function.** *RNA biology* 2012, **9**(2):137-142.
95. Ferracin M, Pedriali M, Veronese A, Zagatti B, Gafa R, Magri E, Lunardi M, Munerato G, Querzoli G, Maestri I *et al*: **MicroRNA profiling for the identification of cancers**

- with unknown primary tissue-of-origin.** *The Journal of pathology* 2011, **225**(1):43-53.
96. Rosenwald S, Gilad S, Benjamin S, Lebanony D, Dromi N, Faerman A, Benjamin H, Tamir R, Ezagouri M, Goren E *et al*: **Validation of a microRNA-based qRT-PCR test for accurate identification of tumor tissue origin.** *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2010, **23**(6):814-823.
 97. Ivanovska I, Ball AS, Diaz RL, Magnus JF, Kibukawa M, Schelter JM, Kobayashi SV, Lim L, Burchard J, Jackson AL *et al*: **MicroRNAs in the miR-106b family regulate p21/CDKN1A and promote cell cycle progression.** *Mol Cell Biol* 2008, **28**(7):2167-2174.
 98. Boucher JM, Peterson SM, Urs S, Zhang C, Liaw L: **The miR-143/145 cluster is a novel transcriptional target of Jagged-1/Notch signaling in vascular smooth muscle cells.** *The Journal of biological chemistry* 2011, **286**(32):28312-28321.
 99. Suh EJ, Remillard MY, Legesse-Miller A, Johnson EL, Lemons JM, Chapman TR, Forman JJ, Kojima M, Silberman ES, Collier HA: **A microRNA network regulates proliferative timing and extracellular matrix synthesis during cellular quiescence in fibroblasts.** *Genome biology* 2012, **13**(12):R121.
 100. Mongroo PS, Rustgi AK: **The role of the miR-200 family in epithelial-mesenchymal transition.** *Cancer biology & therapy* 2010, **10**(3):219-222.
 101. Vasilatou D, Papageorgiou S, Pappa V, Papageorgiou E, Dervenoulas J: **The role of microRNAs in normal and malignant hematopoiesis.** *Eur J Haematol* 2010, **84**(1):1-16.
 102. Aqeilan RI, Calin GA, Croce CM: **miR-15a and miR-16-1 in cancer: discovery, function and future perspectives.** *Cell death and differentiation* 2010, **17**(2):215-220.
 103. Zhu Y, Lu Y, Zhang Q, Liu JJ, Li TJ, Yang JR, Zeng C, Zhuang SM: **MicroRNA-26a/b and their host genes cooperate to inhibit the G1/S transition by activating the pRb protein.** *Nucleic acids research* 2012, **40**(10):4615-4625.
 104. Gao J, Liu QG: **The role of miR-26 in tumors and normal tissues (Review).** *Oncol Lett* 2011, **2**(6):1019-1023.
 105. Wang Y, Zhang X, Li H, Yu J, Ren X: **The role of miRNA-29 family in cancer.** *Eur J Cell Biol* 2013.
 106. Korpala M, Kang Y: **The emerging role of miR-200 family of microRNAs in epithelial-mesenchymal transition and cancer metastasis.** *RNA biology* 2008, **5**(3):115-119.
 107. Karaayvaz M, Pal T, Song B, Zhang C, Georgakopoulos P, Mehmood S, Burke S, Shroyer K, Ju J: **Prognostic significance of miR-215 in colon cancer.** *Clin Colorectal Cancer* 2011, **10**(4):340-347.
 108. Yuan Y, Kasar S, Underbayev C, Prakash S, Raveche E: **MicroRNAs in Acute Myeloid Leukemia and Other Blood Disorders.** *Leuk Res Treatment* 2012, **2012**:603830.
 109. O'Connell RM, Zhao JL, Rao DS: **MicroRNA function in myeloid biology.** *Blood* 2011, **118**(11):2960-2969.

110. Li Z, Huang H, Li Y, Jiang X, Chen P, Arnovitz S, Radmacher MD, Maharry K, Elkahloun A, Yang X *et al*: **Up-regulation of a HOXA-PBX3 homeobox-gene signature following down-regulation of miR-181 is associated with adverse prognosis in patients with cytogenetically abnormal AML.** *Blood* 2012, **119**(10):2314-2324.
111. Havelange V, Stauffer N, Heaphy CC, Volinia S, Andreeff M, Marcucci G, Croce CM, Garzon R: **Functional implications of microRNAs in acute myeloid leukemia by integrating microRNA and messenger RNA expression profiling.** *Cancer* 2011, **117**(20):4696-4706.
112. Garzon R, Garofalo M, Martelli MP, Briesewitz R, Wang L, Fernandez-Cymering C, Volinia S, Liu CG, Schnittger S, Haferlach T *et al*: **Distinctive microRNA signature of acute myeloid leukemia bearing cytoplasmic mutated nucleophosmin.** *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**(10):3945-3950.
113. Gao P, Tchernyshyov I, Chang TC, Lee YS, Kita K, Ochi T, Zeller KI, De Marzo AM, Van Eyk JE, Mendell JT *et al*: **c-Myc suppression of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism.** *Nature* 2009, **458**(7239):762-765.
114. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z, Galon J: **ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks.** *Bioinformatics* 2009, **25**(8):1091-1093.

APPENDIX A

Table A1 Relevant miRNAs associated to solid cancers.

Solid Cancers	miRNA	Functional Role
Breast cancer	miR-21, miR-125b	oncomiR
Breast cancer metastasis	miR-335, miR-206, miR-126	Metastasis suppressor
Lung adenocarcinoma	let-7a, miR-143, miR-145	Tumor suppressor
	miR-17-92 cluster, miR-106b/93/25 cluster	oncomiR
Pancreatic ductal carcinoma	miR-196a, miR-196b	oncomiR
Ovarian carcinoma	miR-199a/b, miR-140, miR-145, miR-204, miR-125a/b	Tumor suppressor
	miR-141, miR-200a/b/c	oncomiR
Hepatocellular carcinoma	miR-21, miR-224, miR-34a, miR-221/222, miR-106a, miR-203	oncomiR
	miR-122a, miR-422b, miR-145, miR-199a	Tumor suppressor
Thyroid papillary cancer	miR-146b, miR-221, miR-222, miR-181b, miR-155, miR-224	oncomiR

Table A2 Prominent miRNAs associated to Luekemia.

Blood malignancies	miRNA	Regulation
CLL	miR-15a, miR-16-1, miR-181a, let-7a, miR-30d, miR-150, miR-92	Downregulation
Pediatric Burkitt's lymphoma, Hodgkin's lymphoma, diffuse large B cell lymphoma	miR-155, miR-17-92	Upregulation
Hodgkin's disease, Burkitt lymphoma cells	miR-9, let-7a	Upregulation
B cell malignancies	miR-143, miR-145	Downregulation
AML	miR-127, miR-154, miR-299, miR-323, miR-368, miR-370	Upregulation
AML	miR-221, miR-222	Downregulation
Hematopoietic malignancies	miR-203	Downregulation

Table A3 . Expression levels of tissue specific microRNAs in normal tissues.
Sorted by information content (IC).

Name	Tissue Specificity	Expression Value	Information Content	Chromosomal Location
miR-302a	EMBRYO	4380	4.00	4:113569339-113569407
miR-302c	EMBRYO	1709	3.99	4: 113569519-113569586
miR-302b	EMBRYO	4929	3.73	4: 113569641-113569713
miR-338-5p	EPIDERMIS / NERVOUS SYSTEM	573 / 451	2.51	17: 79099683-79099749
miR-323-3p	NERVOUS SYSTEM	167	2.47	14: 101492069-101492154
miR-335	BREAST	322	2.37	7: 130135952-130136045
miR-124	NERVOUS SYSTEM	533	2.33	8:9,798,308-9,798,392 8:65,454,260-65,454,368
miR-139-5p	NERVOUS SYSTEM	245	2.27	11: 72326107-72326174
miR-371-5p	EMBRYO	443	1.93	19: 54290929-54290995
miR-325	EPIDERMIS	129	1.91	X: 76225829-76225926
miR-133a	CARDIOVASCULAR SYSTEM / SKELETAL MUSCLE	464/ 661	1.88	18:17,659,657-17,659,744 20:60,572,564-60,572,665
miR-184	RESPIRATORY SYSTEM	259	1.86	15: 79502130-79502213
miR-133b	CARDIOVASCULAR SYSTEM / SKELETAL MUSCLE	394 / 371	1.83	6: 52013721-52013839
miR-211	NERVOUS SYSTEM	180	1.78	15: 31357235-31357344
miR-9	NERVOUS SYSTEM	1346	1.73	1:154656757-154656845 5:87998427-87998513
miR-370	NERVOUS SYSTEM	146	1.72	14: 101377476-101377550
miR-376c	EPIDERMIS	121	1.61	14: 101506027-101506092
miR-132	NERVOUS SYSTEM	552	1.59	17: 1953202-1953302
miR-1	SKELETAL MUSCLE / CARDIOVASCULAR SYSTEM	500 / 350	1.56	20:60,561,958-60,562,028 18:17,662,963-17,663,047
miR-215	KIDNEY	2630	1.55	1: 220291195-220291304
miR-339-5p	LIVER	107	1.54	7: 1062569-1062662
miR-330-3p	LIVER	125	1.53	19: 46142252-46142345
miR-342-5p	EPIDERMIS	326	1.49	14: 100575992-100576090
miR-199b-5p	RESPIRATORY SYSTEM	144	1.39	9: 131007000-131007109
miR-129-3p	NERVOUS SYSTEM	266	1.36	11: 43602944-43603033

Table A4 Expression levels of tissue specific microRNAs during differentiation of embryonic stem cells (ES) . Sorted by information content (IC).

Name	Embryonic tissue Specificity	Expression Value	IC
miR-211	EB 14 day	165	2.65
miR-10b	Monolayer	909	2.20
miR-138	EB 7 day	181	1.88
miR-218	Monolayer	665	1.65
miR-122	Monolayer	3660	1.57
miR-99a	Trophoblast	197	1.54
miR-215	Monolayer /Definitive Endoderm	1673 / 325	1.42
miR-338-3p	EB 7 day	432	1.26
let-7e	iPS / Embryonic Stem Cells	284 / 181	1.23
miR-371-5p	EB 14 day / EB 7 day / ES	1424 / 6274 / 848	1.20
miR-148a	Monolayer	158	1.18
miR-181c	Trophoblast / Monolayer	1500 / 927	1.07
miR-203	EB 14 day / Monolayer / Trophoblast	974 / 750 / 934	1.05
miR-181a	Trophoblast / Monolayer	2943 / 1106	0.99
miR-181b	Trophoblast / Monolayer	8290 / 3397	0.93
miR-192	Monolayer / Definitive Endoderm	1804 / 474	0.84
miR-150	EB 7 day / iPS	501 / 347	0.83
miR-139-3p	iPS / ES	477 / 250	0.82
miR-27b	Trophoblast / Monolayer	554 / 317	0.77
miR-10a	Monolayer	2547	0.75

Table A5 T-UCRs bidirectional in normal cluster tissues.

Chromosoma	Quantity of T-UCRs	T-UCRs bidirectional
1	22	5
2	21	6
3	9	2
4	7	2
5	21	5
6	6	1
7	13	3
8	5	1
9	17	2
10	22	8
11	10	3
12	8	3
13	7	1
14	10	1
15	5	1
16	4	1
17	5	2
18	3	0
19	8	4
20	2	1
22	1	0
X	18	5
TOT	224	57

Table A6 Up-regulated and Down-regulated microRNAs in 31 types of solid cancers (2532 cancers samples vs. 806 corresponding normal samples).

Parametric p-value	FDR	intensities in Solid Cancers	intensities in Normal Tissues	Fold change	microRNA	Chromosomal location
< 1e-07	< 1e-07	967	617.9	1.57	miR-21	17q23.1
< 1e-07	< 1e-07	1378.2	917.7	1.5	miR-25	7q22.1
< 1e-07	< 1e-07	902.7	626.3	1.44	miR-20a	13q31.3
< 1e-07	< 1e-07	925.7	646.9	1.43	miR-17	13q31.3
< 1e-07	< 1e-07	652.3	469	1.39	miR-106a	Xq26.2
< 1e-07	< 1e-07	410	297.8	1.38	miR-106b	7q22.1
1.30E-06	1.08E-05	918.8	697.9	1.32	miR-146a	5q34
< 1e-07	< 1e-07	11893.3	9370.9	1.27	miR-92a	13q31.3,Xq26.2
1.60E-06	1.29E-05	2354.9	1919.4	1.23	miR-103	5q35.1,20p13
< 1e-07	< 1e-07	289.4	237.8	1.22	miR-130b	22q11.21
6.81E-05	0.000387	750.4	615.7	1.22	miR-30c	1p34.2,6q13
< 1e-07	< 1e-07	452.7	372	1.22	miR-93	7q22.1
3.90E-06	2.74E-05	2116.4	1743.6	1.21	miR-107	10q23.31
< 1e-07	< 1e-07	297.7	248	1.2	miR-30e	1p34.2
< 1e-07	< 1e-07	227	191.9	1.18	miR-15a	13q14.2
0.000433	0.002	1209.2	1031.6	1.17	miR-181b	1q32.1,9q33.3
< 1e-07	< 1e-07	191.3	164.8	1.16	miR-15b	3q25.33

0.000605	0.002648	508.2	438.5	1.16	miR-181a	1q32.1,9q33.3
0.000442	0.002002	333.8	287	1.16	miR-32	9q31.3
5.10E-06	3.40E-05	160	147.3	1.09	miR-345	14q32.2
0.000494	0.0022	208.1	193.5	1.08	miR-34a	1p36.22
0.000864	0.00372	154.1	146.2	1.05	miR-374a	Xq13.2
< 1e-07	< 1e-07	275.7	403	0.68	miR-203	14q32.33
< 1e-07	< 1e-07	714.9	1015.1	0.7	miR-145	5q32
0.000003	2.23E-05	273.7	349.6	0.78	miR-205	1q32.2
2.00E-07	1.80E-06	382.5	471.2	0.81	miR-206	6p12.2
< 1e-07	< 1e-07	204	247.9	0.82	miR-33b	17p11.2
3.70E-06	2.67E-05	419.9	507.8	0.83	miR-193a	17q11.2
0.000042	0.000249	175.1	208.5	0.84	miR-204	9q21.12
< 1e-07	< 1e-07	173.9	204.2	0.85	miR-143	5q32
1.00E-07	0.000001	220.5	260.2	0.85	miR-326	11q13.4
< 1e-07	< 1e-07	150.2	176.9	0.85	miR-338	17q25.3
< 1e-07	< 1e-07	164.8	193.6	0.85	miR-9	1q22.5q14.3, 15q26.1
0.000355	0.001789	190.6	220.4	0.86	miR-95	4p16.1
7.40E-06	4.82E-05	225.6	259.4	0.87	miR-138	16q13,3p21.33
0.000435	0.002	160.4	183.8	0.87	miR-183	7q32.2
< 1e-07	< 1e-07	150.4	172.9	0.87	miR-202	10q26.3
8.24E-05	0.000458	529.1	603.1	0.88	miR-128a	2q21.3,3p22.3
0.000921	0.003905	617	694.2	0.89	miR-214	1q24.3

< 1e-07	< 1e-07	139	153.9	0.9	miR-132	17p13.3
< 1e-07	< 1e-07	140.7	156.7	0.9	miR-299	14q32.31
< 1e-07	< 1e-07	140.7	154.8	0.91	miR-129	7q32.1,11p11.2
2.00E-07	1.80E-06	139.6	152.7	0.91	miR-133a	18q11.2,20q13.33
< 1e-07	< 1e-07	129.8	141.1	0.92	miR-139	11q13.4
8.80E-06	5.46E-05	153.4	165	0.93	miR-339	7p22.3
4.50E-06	3.08E-05	140	149.6	0.94	miR-1	20q13.33,18q11.2
9.68E-05	0.000528	140.4	149.6	0.94	miR-133b	6p12.2
< 1e-07	< 1e-07	129.7	137.8	0.94	miR-323	14q32.31
0.000394	0.001946	144.7	152.7	0.95	miR-218	4p15.31,5q34
0.000193	0.00101	130.2	133	0.98	miR-335	7q32.2

Table A7 Expression levels of tissue specific microRNAs in cancer and leukemia.

Name	Cancer Specificity	Expression	IC	Chromosomal location
miR-369-3p	OVARIAN ADENOCARCINOMA	2737	3.95	14:101531935- 101532004
miR-325	PANCREAS NON- FUNCTIONING ENDOCRINE CANCER / INSULINOMA	1625 / 365	3.67	X: 76225829-76225926
miR-208	CC / BCC / EPIDERMIS SCC	256/ 244 / 418	3.36	14:22,927,645- 22,927,715 14:22,957,036- 22,957,112
miR-190	TPC / HEPATOBLASTOMA	211 / 220	3.04	15: 63116156-63116240
miR-184	AMoL / HEPATOBLASTOMA /TPC / MM	448 / 324 / 214 / 171	3.01	15: 79502130-79502213
miR-338-5p	EPIDERMIS SCC / BCC / CC	1928 / 808 / 472	2.70	17: 79099683-79099749
miR-105	HGSIL	234	2.57	X:151,311,347- 151,313,620
miR-376c	LYMPHOMA CUTANEOUS / EPIDERMIS SCC / MM	1134 / 626 / 1223 / 525	2.50	14: 101506027- 101506092
miR-302b	EPIDERMIS SCC / LYMPHOMA CUTANEOUS	336 / 110	2.48	4: 113569641- 113569713
miR-299-5p	HEPATOBLASTOMA / AMoL / TPC	287 / 122 / 148	2.03	14: 101490131- 101490193
miR-19a	LYMPHOMA CUTANEOUS / CML	388 /367	1.83	13: 92003145-92003226
miR-200c	AMoL / HEPATOBLASTOMA	4460 / 4138	1.99	12: 7072862-7072929
miR-143	AMoL / HEPATOBLASTOMA	1043 / 1098	1.94	5: 148808481- 148808586
miR-211	EPIDERMIS SCC / NSCLC	129 / 89	1.86	15: 31357235-31357344
miR-374a	BL SPORADIC / CML / APL	619 / 169 / 212	1.69	X: 73507121-73507192
miR-339-5p	BL ENDEMIC / BL SPORADIC	616 / 738	1.62	7: 1062569-1062662
miR-148a	LYMPHOMA CUTANEOUS	1161	1.64	7: 25989539-25989606
miR-370	BREAST, DUCTAL, LOBULAR, MALE CARCINOMA	157, 227, 239	1.69	14: 101377476- 101377550
miR-199a- 5p	HEPATOBLASTOMA	1122	1.52	19:10,789,102- 10,789,172

				1:170,380,298-170,380,407
miR-215	KIDNEY CARCINOMA / HEPATOBLASTOMA	7247 /2461	1.45	1: 220291195-220291304
miR-19b	CML	232	1.33	13:90,801,447-90,801,533 X:133,131,367-133,131,462
miR-148b	LYMPHOMA CUTANEOUS	657	1.39	12: 54731000-54731098
miR-18a	AMoL / FOLLICULAR ADENOMA	166 / 76	1.38	13: 92003005-92003075
miR-202	AMoL/ HEPATOBLASTOMA	362 / 287	1.32	10: 135061015-135061124
miR-205	AMoL / HEPATOBLASTOMA / LUNG SCC	2029 / 2275 / 2665	1.31	1: 209605478-209605587

Table A8 List of 100 miR selected by class prediction method.

MiRNAs
hsa-let-7e
hsa-let-7i
hsa-miR-100
hsa-miR-105*
hsa-miR-106b
hsa-miR-10a
hsa-miR-122
hsa-miR-122*
hsa-miR-125a-3p
hsa-miR-125b
hsa-miR-128a
hsa-miR-132*
hsa-miR-134
hsa-miR-136
hsa-miR-138-1*
hsa-miR-138
hsa-miR-142-5p

hsa-miR-145
hsa-miR-146a
hsa-miR-149
hsa-miR-150
hsa-miR-155
hsa-miR-15b
hsa-miR-16
hsa-miR-181a
hsa-miR-181b
hsa-miR-187
hsa-miR-190b
hsa-miR-192
hsa-miR-193b
hsa-miR-197
hsa-miR-199b-3p
hsa-miR-1
hsa-miR-200b
hsa-miR-200c
hsa-miR-203
hsa-miR-204
hsa-miR-205
hsa-miR-210
hsa-miR-215
hsa-miR-218
hsa-miR-219-2-3p
hsa-miR-223
hsa-miR-224
hsa-miR-22
hsa-miR-23b
hsa-miR-24
hsa-miR-25*
hsa-miR-26a-1*
hsa-miR-26a-2*
hsa-miR-26b*
hsa-miR-27b

hsa-miR-29a*
hsa-miR-29b-1*
hsa-miR-29b-2*
hsa-miR-29c
hsa-miR-30b*
hsa-miR-30c-1*
hsa-miR-324-3p
hsa-miR-32
hsa-miR-330-5p
hsa-miR-33b
hsa-miR-342-3p
hsa-miR-342-5p
hsa-miR-345
hsa-miR-346
hsa-miR-34c-3p
hsa-miR-361-5p
hsa-miR-369-3p
hsa-miR-371-5p
hsa-miR-374a
hsa-miR-374b
hsa-miR-375
hsa-miR-376b
hsa-miR-383
hsa-miR-410
hsa-miR-412
hsa-miR-421
hsa-miR-424
hsa-miR-433
hsa-miR-484
hsa-miR-486-5p
hsa-miR-487a
hsa-miR-487b
hsa-miR-494
hsa-miR-496
hsa-miR-498

hsa-miR-499-3p
hsa-miR-500*
hsa-miR-516a-5p
hsa-miR-517b
hsa-miR-519b-3p
hsa-miR-543
hsa-miR-551b*
hsa-miR-671-5p
hsa-miR-93*
hsa-miR-99a
hsa-miR-99b
hsa-miR-99b*
hsa-miR-9

Table A9 List of 31 miRNAs selected using evaluator: CfsSubsetEval and as searcher: BestFirst -D1 -N5.

UniqueID
hsa-miR-105*
hsa-miR-10a
hsa-miR-122
hsa-miR-122*
hsa-miR-134
hsa-miR-145
hsa-miR-146a
hsa-miR-15b
hsa-miR-181a
hsa-miR-192
hsa-miR-193b
hsa-miR-197
hsa-miR-210
hsa-miR-215
hsa-miR-218
hsa-miR-219-2-3p
hsa-miR-223
hsa-miR-30c-1*
hsa-miR-330-5p
hsa-miR-33b
hsa-miR-342-3p
hsa-miR-346
hsa-miR-34c-3p
hsa-miR-375
hsa-miR-383

hsa-miR-494
hsa-miR-496
hsa-miR-9

Table A10 List of 15 miRNAs selected.

UniqueID
hsa-let-7e
hsa-miR-105*
hsa-miR-10a
hsa-miR-122
hsa-miR-145
hsa-miR-15b
hsa-miR-193b
hsa-miR-215
hsa-miR-218
hsa-miR-30c-1*
hsa-miR-330-5p
hsa-miR-34c-3p
hsa-miR-375
hsa-miR-383
hsa-miR-9

Table A11 Array CGH datasets.

GEO datasets	# of samples	Cancer type	Platform
GSE4659	32	AML	GPL2873
GSE6472	6	Nasopharyngeal carcinoma	GPL2879
GSE7077	4	Osteosarcoma	GPL2879
GSE7344	16	Glioma	GPL2873
GSE7482	25	ACC	GPL2879
GSE7615	298	Pancreas, glioblastoma, T-ALL, melanoma, colon	GPL2879 GPL4091
GSE7822	14	Melanoma	GPL2879
GSE8398	25	Ewing Sarcoma	GPL2879
GSE8804	13	Myelodysplasia	GPL2879
GSE8918	87	Follicular Lymphoma, CLL, Mantle Cell Lymphoma, Nodal Marginal Zone Lymphoma, Lymphoplasmacytic Lymphoma, Splenic Marginal Zone Lymphoma , MALT	GPL2879
GSE9015	7	Breast	GPL4091
GSE9654	10	Osteosarcoma	GPL2879
SMD	207	Lung, Pancreas, Breast, Melanoma, Fibrous Histiocytoma	Stanford

- GPL2873 Agilent- Human Genome CGH Microarray 44A G4410A
- GPL2879 Agilent- Human Genome CGH Microarray 44B G4410B
- GPL4091 Agilent- Human Genome CGH Microarray 244A G4411B

Table A12. Expression levels are calculated as $2^{(-\Delta \Delta CT)}$ using 18S rRNA as a reference. Average of at least three experiments, each PCR performed in triplicates. Standard errors (ERR.ST.) are reported.

Sample Name	uc.283plus	uc.283plus	uc.283plus	AVERAGE	ERR.ST
Spleen	6.91855E-08	1.82593E-07	2.40273E-07	1.64017E-07	2.90144E-08
SV	3.5482E-07	6.30583E-07	7.71458E-07	5.8562E-07	7.06424E-08
Thymus	3.25623E-07	6.1807E-07	4.94822E-07	4.79505E-07	4.89413E-08
Lungs	1.62718E-06	1.08067E-06	3.6858E-07	1.02548E-06	2.1037E-07
BM	6.44552E-08	3.48473E-07	3.75434E-07	2.62787E-07	5.74297E-08
ES	1.63444E-05	2.09864E-05	1.01926E-05	1.58411E-05	1.80482E-06

Sample Name	miR-302	miR-302	miR-302	AVERAGE	ERR.ST
Spleen	9.48487E-07	1.72605E-06	2.02762E-06	1.56739E-06	1.85596E-07
SV	7.29868E-08	2.40353E-09	8.50101E-08	5.34668E-08	1.48763E-08
Thymus	1.00638E-06	7.11832E-07	1.24061E-06	9.86274E-07	8.8321E-08
Lungs	7.85065E-07	2.40986E-07	1.82429E-07	4.02827E-07	1.10773E-07
BM	1.18112E-06	1.83193E-06	1.31968E-06	1.44424E-06	1.14273E-07
ES	3.35274E-05	3.37314E-05	2.50494E-05	3.07694E-05	1.65158E-06

Table A13 Differentially expressed miRNAs in miR-155 leukemia transgenic mouse.

Parametric p-value	FDR	Intensity in miR-155 Leukemia	Intensity in Wt	Fold Change	miRNA
1.00E-06	2.31E-05	832.1	10.08	82.57	miR--224
4.20E-06	5.67E-05	3426.09	132.36	25.89	miR--9
1.00E-07	7.20E-06	21760.18	1005.44	21.64	miR--217
5.70E-06	7.39E-05	3387.97	219.75	15.42	miR--131
1.00E-07	7.20E-06	5334.39	375.03	14.22	miR--128b
1.00E-07	7.20E-06	31460.89	2922.63	10.76	miR--181b
2.00E-07	7.20E-06	17040.98	1615.73	10.55	miR--181c
2.00E-07	7.20E-06	18356.22	1976.64	9.29	miR--128a
1.28E-05	0.000148	759.87	2695.34	0.28	miR--140
0.000659	0.003556	228.42	874.66	0.26	miR--135a
3.69E-05	0.000332	8491.07	34327.47	0.25	miR--26a
0.000128	0.000984	528.06	2433.76	0.22	miR--10a
3.50E-06	5.55E-05	590.74	2779.66	0.21	miR--425
0.000193	0.001351	514.05	2467.2	0.21	miR--340
0.000474	0.002742	81.06	403.52	0.2	miR--24
2.00E-06	3.78E-05	656.68	3304.56	0.2	miR--140
0.000579	0.003178	30.24	196.77	0.15	miR--218
2.55E-05	0.000243	1006.31	6657.6	0.15	miR--29b
6.00E-07	1.62E-05	1007.33	7465.65	0.13	miR--29a
2.54E-05	0.000243	357.49	3183.78	0.11	miR--29c
< 1e-07	< 1e-07	2176.08	37081.81	0.06	miR--146
2.70E-06	4.60E-05	3339.45	59257.53	0.06	miR--150

APPENDIX B

ESC and IPS and Normal POOL

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2

Relation: UCRescIPSnormaldifferpool-weka.filters.supervised.attribute.AttributeSelection-Eweka.attributeSelection.CfsSubsetEval-Sweka.attributeSelection.BestFirst -D 1 -N 5

Instances: 90

Attributes: 16

uc.283Plus

uc.262Minus

uc.153Plus

uc.192Plus

uc.478Minus

uc.240Plus

uc.246Minus

uc.445Plus

uc.420Minus

uc.20Plus

uc.117Minus

uc.263Plus

uc.341Plus

uc.185Plus

uc.8Plus

Class

Test mode: 10-fold cross-validation

J48 pruned tree

uc.283Plus <= 7.589: Normal differ (49.0/1.0)

uc.283Plus > 7.589

| uc.192Plus <= 7.623

| | uc.20Plus <= 7.878: INDUCED PLURIPOTENT STEM CELLS (5.0)

| | uc.20Plus > 7.878

| | | uc.262Minus <= 10.494: EMBRYONIC STEM CELLS (24.0)

| | | uc.262Minus > 10.494

| | | | uc.246Minus <= 7.307: EMBRYONIC STEM CELLS (2.0)

| | | | uc.246Minus > 7.307: INDUCED PLURIPOTENT STEM CELLS (2.0)

| uc.192Plus > 7.623: INDUCED PLURIPOTENT STEM CELLS (8.0)

Number of Leaves : 6

Size of the tree : 11

Time taken to build model: 0.05 seconds

Correctly Classified Instances	77	85.5556 %
--------------------------------	----	-----------

Incorrectly Classified Instances	13	14.4444 %
----------------------------------	----	-----------

Kappa statistic	0.759
-----------------	-------

Mean absolute error	0.1024
---------------------	--------

Root mean squared error 0.2978
 Relative absolute error 25.5589 %
 Root relative squared error 66.6062 %
 Total Number of Instances 90

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.741	0.095	0.769	0.741	0.755	0.858	ESC
	0.667	0.08	0.625	0.667	0.645	0.888	IPSC
	0.979	0.024	0.979	0.979	0.979	0.967	Normal differ
Weighted Avg.	0.856	0.055	0.857	0.856	0.856	0.921	

=== Confusion Matrix ===

a b c <-- classified as
 20 6 1 | a = EMBRYONIC STEM CELLS
 5 10 0 | b = INDUCED PLURIPOTENT STEM CELLS
 1 0 47 | c = Normal differ

ESC and NORMAL POOL

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2

Relation: UCRescnormaldifferpool-weka.filters.supervised.attribute.AttributeSelection-Eweka.attributeSelection.CfsSubsetEval-Sweka.attributeSelection.BestFirst -D 1 -N 5

Instances: 75

Attributes: 20

uc.10Minus

uc.283Plus

uc.305Minus

uc.20Minus

uc.142Minus

uc.453Minus

uc.172Minus

uc.44Plus

uc.230Plus

uc.309Minus

uc.362Minus

uc.473Plus

uc.229Minus

uc.300Minus

uc.20Plus

uc.34Minus

uc.73Plus

uc.8Plus

uc.426Plus

Class

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree

uc.283Plus <= 7.589: Normal differ (49.0/1.0)

uc.283Plus > 7.589: EMBRYONIC STEM CELLS (26.0)

Number of Leaves : 2

Size of the tree : 3

Time taken to build model: 0.03 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	73	97.3333 %
Incorrectly Classified Instances	2	2.6667 %
Kappa statistic	0.9421	
Mean absolute error	0.0393	
Root mean squared error	0.1642	
Relative absolute error	8.5077 %	
Root relative squared error	34.1622 %	

Total Number of Instances 75

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.963	0.021	0.963	0.963	0.963	0.955	EMBRYONIC STEM CELLS
	0.979	0.037	0.979	0.979	0.979	0.955	Normal differ
Weighted Avg.	0.973	0.031	0.973	0.973	0.973	0.955	

=== Confusion Matrix ===

a b <-- classified as

26 1 | a = EMBRYONIC STEM CELLS

1 47 | b = Normal differ

=== Run information ===

Scheme:weka.classifiers.trees.J48graft -C 0.25 -M 2

Relation: TrainingData323SOLIDnoADENO48miR-weka.filters.unsupervised.attribute

Instances: 140

Attributes: 32

hsa-let-7e

hsa-let-7i

hsa-miR-100

hsa-miR-105*

hsa-miR-10a

hsa-miR-122

hsa-miR-122*

hsa-miR-134
hsa-miR-145
hsa-miR-146a
hsa-miR-15b
hsa-miR-181a
hsa-miR-192
hsa-miR-193b
hsa-miR-197
hsa-miR-210
hsa-miR-215
hsa-miR-218
hsa-miR-219-2-3p
hsa-miR-223
hsa-miR-30c-1*
hsa-miR-330-5p
hsa-miR-33b
hsa-miR-342-3p
hsa-miR-346
hsa-miR-34c-3p
hsa-miR-375
hsa-miR-383
hsa-miR-494
hsa-miR-496
hsa-miR-9
Class

Test mode: evaluate on training data

=== Classifier model (full training set) ===

J48graft pruned tree

```
hsa-miR-218 <= 6.834
| hsa-miR-197 <= 8.748
| | hsa-miR-346 <= 10.961: OVARIAN CARCINOMA (5.0/1.0)
| | hsa-miR-346 > 10.961: LUNG SQUAMOUS CELL CARCINOMA (11.0/1.0)
| hsa-miR-197 > 8.748
| | hsa-miR-192 <= 10.442
| | | hsa-miR-100 <= 10.586
| | | | hsa-miR-383 <= 9.043
| | | | | hsa-miR-496 <= 6.339
| | | | | | hsa-miR-30c-1* <= 6.94
| | | | | | | hsa-miR-181a <= 9.372
| | | | | | | | hsa-miR-346 <= 6.0355: PROSTATE CARCINOMA (0.0|6.0)
| | | | | | | | hsa-miR-346 > 6.0355
| | | | | | | | | hsa-miR-146a <= 3.944: PROSTATE CARCINOMA (0.0|6.0)
| | | | | | | | | hsa-miR-146a > 3.944
| | | | | | | | | | hsa-miR-193b <= 9.3575
| | | | | | | | | | | hsa-miR-122* <= 5.8315
| | | | | | | | | | | | hsa-miR-122 <= 11.4875
```



```

| | | hsa-miR-330-5p > 4.55: COLON CARCINOMA (6.0)
| | | hsa-miR-122* > 5.432: HEPATOCELLULAR CARCINOMA (9.0)
hsa-miR-218 > 6.834
| hsa-miR-375 <= 8.633
| | hsa-miR-145 <= 9.173
| | | hsa-miR-375 <= 5.946: GLIOMA (4.0)
| | | hsa-miR-375 > 5.946
| | | | hsa-miR-100 <= 11.794
| | | | | hsa-miR-223 <= 3.528
| | | | | | hsa-miR-181a <= 9.7665: OVARIAN CARCINOMA (0.0|5.0)
| | | | | | hsa-miR-181a > 9.7665: GLIOMA (2.0/1.0)
| | | | | hsa-miR-223 > 3.528: OVARIAN CARCINOMA (5.0)
| | | | hsa-miR-100 > 11.794: NSCLC (4.0)
| | | hsa-miR-145 > 9.173
| | | hsa-miR-219-2-3p <= 8.017
| | | | hsa-let-7e <= 9.8775: GLIOMA (2.0/1.0)
| | | | hsa-let-7e > 9.8775: KIDNEY CARCINOMA (0.0|5.0)
| | | hsa-miR-219-2-3p > 8.017: KIDNEY CARCINOMA (10.0/1.0)
| hsa-miR-375 > 8.633: PANCREAS CARCINOMA (11.0/1.0)

```

Number of Leaves : 32

Size of the tree : 63

Time taken to build model: 0.08 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	129	92.1429 %
Incorrectly Classified Instances	11	7.8571 %
Kappa statistic	0.9154	
Mean absolute error	0.0163	
Root mean squared error	0.0903	
Relative absolute error	12.2844 %	
Root relative squared error	35.0491 %	
Total Number of Instances	140	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.9	0.008	0.9	0.9	0.9	0.992	BLADDER CANCER
0.8	0.008	0.889	0.8	0.842	0.992	BREAST CARCINOMA
0.8	0.008	0.889	0.8	0.842	0.993	COLON CARCINOMA
1	0.008	0.909	1	0.952	1	ESOPHAGUS CARCINOMA
0.9	0	1	0.9	0.947	1	GASTRIC CANCER
1	0.015	0.833	1	0.909	0.998	GLIOMA
0.9	0	1	0.9	0.947	1	HEPATOCELLULAR CARCINOMA
0.9	0.008	0.9	0.9	0.9	0.992	KIDNEY CARCINOMA

1	0.008	0.909	1	0.952	0.996	LUNG SQUAMOUS CELL CARCINOMA
0.9	0.008	0.9	0.9	0.9	0.994	MELANOMA
0.9	0	1	0.9	0.947	1	NSCLC
0.9	0.008	0.9	0.9	0.9	0.994	OVARIAN CARCINOMA
1	0.008	0.909	1	0.952	0.996	PANCREAS CARCINOMA
1	0	1	1	1	1	PROSTATE CARCINOMA
Weighted Avg.	0.921	0.006	0.924	0.921	0.921	0.996

=== Confusion Matrix ===

```

a b c d e f g h i j k l m n <-- classified as
9 0 0 0 0 0 0 0 0 0 1 0 0 0 0 | a = BLADDER CANCER
1 8 1 0 0 0 0 0 0 0 0 0 0 0 0 | b = BREAST CARCINOMA
0 1 8 0 0 0 0 0 1 0 0 0 0 0 0 | c = COLON CARCINOMA
0 0 0 10 0 0 0 0 0 0 0 0 0 0 0 | d = ESOPHAGUS CARCINOMA
0 0 0 1 9 0 0 0 0 0 0 0 0 0 0 | e = GASTRIC CANCER
0 0 0 0 0 10 0 0 0 0 0 0 0 0 0 | f = GLIOMA
0 0 0 0 0 1 9 0 0 0 0 0 0 0 0 | g = HEPATOCELLULAR CARCINOMA

```

0 0 0 0 0 0 9 0 0 0 0 1 0 | h = KIDNEY CARCINOMA
0 0 0 0 0 0 0 10 0 0 0 0 0 | i = LUNG SQUAMOUS CELL CARCINOMA
0 0 0 0 0 0 0 0 9 0 1 0 0 | j = MELANOMA
0 0 0 0 0 1 0 0 0 0 9 0 0 | k = NSCLC
0 0 0 0 0 0 1 0 0 0 9 0 0 | l = OVARIAN CARCINOMA
0 0 0 0 0 0 0 0 0 0 10 0 | m = PANCREAS CARCINOMA
0 0 0 0 0 0 0 0 0 0 0 10 | n = PROSTATE CARCINOMA

=== Run information ===

Scheme: weka.classifiers.trees.J48graft -C 0.25 -M 2

Relation: TrainingData323SOLIDnoADENO48miR-weka.filters.unsupervised.attribute.

Instances: 140

Attributes: 32

hsa-let-7e
hsa-let-7i
hsa-miR-100
hsa-miR-105*
hsa-miR-10a
hsa-miR-122
hsa-miR-122*
hsa-miR-134
hsa-miR-145
hsa-miR-146a
hsa-miR-15b
hsa-miR-181a
hsa-miR-192
hsa-miR-193b
hsa-miR-197
hsa-miR-210
hsa-miR-215
hsa-miR-218
hsa-miR-219-2-3p

hsa-miR-223
hsa-miR-30c-1*
hsa-miR-330-5p
hsa-miR-33b
hsa-miR-342-3p
hsa-miR-346
hsa-miR-34c-3p
hsa-miR-375
hsa-miR-383
hsa-miR-494
hsa-miR-496
hsa-miR-9
Class

Test mode:user supplied test set: size unknown (reading incrementally)

=== Classifier model (full training set) ===

J48graft pruned tree

```
hsa-miR-218 <= 6.834
| hsa-miR-197 <= 8.748
| | hsa-miR-346 <= 10.961: OVARIAN CARCINOMA (5.0/1.0)
| | hsa-miR-346 > 10.961: LUNG SQUAMOUS CELL CARCINOMA (11.0/1.0)
| hsa-miR-197 > 8.748
| | hsa-miR-192 <= 10.442
| | | hsa-miR-100 <= 10.586
| | | | hsa-miR-383 <= 9.043
| | | | | hsa-miR-496 <= 6.339
| | | | | | hsa-miR-30c-1* <= 6.94
| | | | | | | hsa-miR-181a <= 9.372
| | | | | | | | hsa-miR-346 <= 6.0355: PROSTATE CARCINOMA (0.0|6.0)
| | | | | | | | hsa-miR-346 > 6.0355
| | | | | | | | | hsa-miR-146a <= 3.944: PROSTATE CARCINOMA (0.0|6.0)
| | | | | | | | | hsa-miR-146a > 3.944
| | | | | | | | | | hsa-miR-193b <= 9.3575
```


| | | | | | | | | | hsa-miR-122* <= 5.8315
| | | | | | | | | | hsa-miR-122 <= 11.4875
| | | | | | | | | | | hsa-miR-122 <= 3.944: PROSTATE CARCINOMA (0.0|5.0)
| | | | | | | | | | | hsa-miR-122 > 3.944: ESOPHAGUS CARCINOMA (2.0/1.0)
| | | | | | | | | | | hsa-miR-122 > 11.4875: HEPATOCELLULAR CARCINOMA (0.0|5.0)
| | | | | | | | | | | hsa-miR-122* > 5.8315: HEPATOCELLULAR CARCINOMA (0.0|5.0)
| | | | | | | | | | | hsa-miR-193b > 9.3575: PROSTATE CARCINOMA (0.0|5.0)
| | | | | | | hsa-miR-181a > 9.372: BREAST CARCINOMA (7.0)
| | | | | | hsa-miR-30c-1* > 6.94: ESOPHAGUS CARCINOMA (9.0)
| | | | | hsa-miR-496 > 6.339
| | | | | | hsa-miR-105* <= 5.412: COLON CARCINOMA (3.0/1.0)
| | | | | | hsa-miR-105* > 5.412: BLADDER CANCER (10.0/1.0)
| | | | hsa-miR-383 > 9.043: MELANOMA (10.0/1.0)
| | | | hsa-miR-100 > 10.586
| | | | | hsa-miR-342-3p <= 11.704
| | | | | | hsa-miR-100 <= 13.359
| | | | | | hsa-miR-146a <= 4.872: NSCLC (5.0)
| | | | | | hsa-miR-146a > 4.872
| | | | | | | hsa-miR-383 <= 9.003
| | | | | | | | hsa-miR-375 <= 8.849: BREAST CARCINOMA (2.0/1.0)
| | | | | | | | hsa-miR-375 > 8.849: MELANOMA (0.0|5.0)
| | | | | | | | hsa-miR-383 > 9.003: MELANOMA (0.0|9.0)
| | | | | | | hsa-miR-100 > 13.359: GLIOMA (4.0)
| | | | | | | hsa-miR-342-3p > 11.704: PROSTATE CARCINOMA (10.0)
| | | hsa-miR-192 > 10.442
| | | | hsa-miR-122* <= 5.432
| | | | | hsa-miR-330-5p <= 4.55: GASTRIC CANCER (9.0)
| | | | | hsa-miR-330-5p > 4.55: COLON CARCINOMA (6.0)
| | | | hsa-miR-122* > 5.432: HEPATOCELLULAR CARCINOMA (9.0)
hsa-miR-218 > 6.834
| | hsa-miR-375 <= 8.633
| | | hsa-miR-145 <= 9.173
| | | | hsa-miR-375 <= 5.946: GLIOMA (4.0)
| | | | hsa-miR-375 > 5.946
| | | | | hsa-miR-100 <= 11.794
| | | | | | hsa-miR-223 <= 3.528

```

| | | | | hsa-miR-181a <= 9.7665: OVARIAN CARCINOMA (0.0|5.0)
| | | | | hsa-miR-181a > 9.7665: GLIOMA (2.0/1.0)
| | | | | hsa-miR-223 > 3.528: OVARIAN CARCINOMA (5.0)
| | | | hsa-miR-100 > 11.794: NSCLC (4.0)
| | hsa-miR-145 > 9.173
| | | hsa-miR-219-2-3p <= 8.017
| | | | hsa-let-7e <= 9.8775: GLIOMA (2.0/1.0)
| | | | hsa-let-7e > 9.8775: KIDNEY CARCINOMA (0.0|5.0)
| | | hsa-miR-219-2-3p > 8.017: KIDNEY CARCINOMA (10.0/1.0)
| hsa-miR-375 > 8.633: PANCREAS CARCINOMA (11.0/1.0)

```

Number of Leaves : 32

Size of the tree : 63

Time taken to build model: 0.09 seconds

=== Predictions on test split ===

```

inst#, actual, predicted, error, probability distribution
  1 1:BLADDER 1:BLADDER *0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0
  2 1:BLADDER 1:BLADDER *0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0
  3 1:BLADDER 10:MELANOM + 0.1 0 0 0 0 0 0 0 0 0 *0.9 0 0 0
0
  4 1:BLADDER 1:BLADDER *0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0
  5 1:BLADDER 1:BLADDER *0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0
  6 1:BLADDER 1:BLADDER *0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0
  7 1:BLADDER 1:BLADDER *0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0

```

8 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
9 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
10 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
11 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
12 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
13 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
14 2:BREAST C 3:COLON CA	+ 0	0.333	*0.667	0	0	0	0	0	0	0	0	0	0	0
0 0														
15 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
16 2:BREAST C 2:BREAST C	0	*0.5	0.5	0	0	0	0	0	0	0	0	0	0	0
0														
17 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
18 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
19 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
20 2:BREAST C 1:BLADDER	+ *0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
21 3:COLON CA 3:COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
22 3:COLON CA 3:COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
23 3:COLON CA 3:COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
24 3:COLON CA 2:BREAST C	+ 0	*0.5	0.5	0	0	0	0	0	0	0	0	0	0	0
0														
25 3:COLON CA 3:COLON CA	0	0.333	*0.667	0	0	0	0	0	0	0	0	0	0	0
0 0														

26 3:COLON CA 9:LUNG SQU	+	0	0	0.091	0	0	0	0	0	0	*0.909	0	0	0
0 0														
27 3:COLON CA 3:COLON CA		0	0.333	*0.667	0	0	0	0	0	0	0	0	0	0
0 0														
28 3:COLON CA 3:COLON CA		0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
29 3:COLON CA 3:COLON CA		0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
30 3:COLON CA 3:COLON CA		0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
31 4:ESOPHAGU 4:ESOPHAGU		0	0	0	*1	0	0	0	0	0	0	0	0	0
0														
32 4:ESOPHAGU 4:ESOPHAGU		0	0	0	*1	0	0	0	0	0	0	0	0	0
0														
33 4:ESOPHAGU 4:ESOPHAGU		0	0	0	*1	0	0	0	0	0	0	0	0	0
0														
34 4:ESOPHAGU 4:ESOPHAGU		0	0	0	*1	0	0	0	0	0	0	0	0	0
0														
35 4:ESOPHAGU 4:ESOPHAGU		0	0	0	*1	0	0	0	0	0	0	0	0	0
0														
36 4:ESOPHAGU 4:ESOPHAGU		0	0	0	*0.5	0.5	0	0	0	0	0	0	0	0
0 0														
37 4:ESOPHAGU 4:ESOPHAGU		0	0	0	*1	0	0	0	0	0	0	0	0	0
0														
38 4:ESOPHAGU 4:ESOPHAGU		0	0	0	*1	0	0	0	0	0	0	0	0	0
0														
39 4:ESOPHAGU 4:ESOPHAGU		0	0	0	*1	0	0	0	0	0	0	0	0	0
0														
40 4:ESOPHAGU 4:ESOPHAGU		0	0	0	*1	0	0	0	0	0	0	0	0	0
0														
41 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0
0														
42 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0
0														
43 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0
0														

44	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0																	
45	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0																	
46	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0																	
47	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0																	
48	5:GASTRIC	4:ESOPHAGU	+	0	0	0	*0.5	0.5	0	0	0	0	0	0	0	0	0
0																	
49	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0																	
50	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0																	
51	6:GLIOMA	6:GLIOMA	0	0	0	0	0	*0.5	0.5	0	0	0	0	0	0	0	0
0																	
52	6:GLIOMA	6:GLIOMA	0	0	0	0	0	*0.5	0	0	0	0	0	0.5	0	0	0
0																	
53	6:GLIOMA	6:GLIOMA	0	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
54	6:GLIOMA	6:GLIOMA	0	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
55	6:GLIOMA	6:GLIOMA	0	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
56	6:GLIOMA	6:GLIOMA	0	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
57	6:GLIOMA	6:GLIOMA	0	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
58	6:GLIOMA	6:GLIOMA	0	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
59	6:GLIOMA	6:GLIOMA	0	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
60	6:GLIOMA	6:GLIOMA	0	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
61	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																	
62	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																	
63	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																	
64	7:HEPATOCE	6:GLIOMA	+	0	0	0	0	0	*0.5	0.5	0	0	0	0	0	0	0
0																	
65	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																	

66 7:HEPATOCE 7:HEPATOCE	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
67 7:HEPATOCE 7:HEPATOCE	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
68 7:HEPATOCE 7:HEPATOCE	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
69 7:HEPATOCE 7:HEPATOCE	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
70 7:HEPATOCE 7:HEPATOCE	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
71 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	
0														
72 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	
0														
73 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	
0														
74 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	
0														
75 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	
0														
76 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	
0														
77 8:KIDNEY C 13:PANCREA	+ 0	0	0	0	0	0	0	0	0.091	0	0	0	0	
*0.909 0														
78 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	
0														
79 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	
0														
80 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	
0														
81 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	*0.909	0	0	0	0	
0 0														
82 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	*0.909	0	0	0	0	
0 0														
83 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	*0.909	0	0	0	0	
0 0														

84 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	0	0	*0.909	0	0	0
0 0														
85 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	0	0	*0.909	0	0	0
0 0														
86 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	0	0	*0.909	0	0	0
0 0														
87 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	0	0	*0.909	0	0	0
0 0														
88 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	0	0	*0.909	0	0	0
0 0														
89 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	0	0	*0.909	0	0	0
0 0														
90 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	0	0	*0.909	0	0	0
0 0														
91 10:MELANOM 10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	*0.9	0	0	0
0 0														
92 10:MELANOM 10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	*0.9	0	0	0
0 0														
93 10:MELANOM 12:OVARIAN	+ 0	0	0	0	0	0	0	0	0	0	0.2	0	*0.8	0
0														
94 10:MELANOM 10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	*0.9	0	0	0
0 0														
95 10:MELANOM 10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	*0.9	0	0	0
0 0														
96 10:MELANOM 10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	*0.9	0	0	0
0 0														
97 10:MELANOM 10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	*0.9	0	0	0
0 0														
98 10:MELANOM 10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	*0.9	0	0	0
0 0														
99 10:MELANOM 10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	*0.9	0	0	0
0 0														
100 10:MELANOM 10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	*0.9	0	0	0
0 0														
101 11:NSCLC 6:GLIOMA	+ 0	0	0	0	0	0	*0.5	0	0	0	0	0.5	0	0
0														

124 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
125 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
126 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
127 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
128 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
129 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
130 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
131 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
132 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
133 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
134 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
135 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
136 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
137 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
138 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
139 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
140 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
141 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0
0													

142 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
143 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
144 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
145 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
146 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
147 1:BLADDER 4:ESOPHAGU	+	0	0	0	*0.5	0.5	0	0	0	0	0	0	0	0
0 0														
148 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
149 1:BLADDER 4:ESOPHAGU	+	0	0	0	*0.5	0.5	0	0	0	0	0	0	0	0
0 0														
150 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
151 2:BREAST C 14:PROSTAT	+	0	0	0	0	0	0	0	0	0	0	0	0	0
*1														
152 2:BREAST C 12:OVARIAN	+	0	0	0	0	0	0	0	0	0	0.2	0	*0.8	0
0														
153 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
154 2:BREAST C 12:OVARIAN	+	0	0	0	0	0	0	0	0	0	0.2	0	*0.8	0
0														
155 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
156 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
157 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
158 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
159 2:BREAST C 4:ESOPHAGU	+	0	0	0	*1	0	0	0	0	0	0	0	0	0
0														

160 2:BREAST C 14:PROSTAT	+	0	0	0	0	0	0	0	0	0	0	0	0	0
*1														
161 3:COLON CA 3:COLON CA		0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
162 3:COLON CA 3:COLON CA		0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
163 3:COLON CA 12:OVARIAN	+	0	0	0	0	0	0	0	0	0	0.2	0	*0.8	0
0														
164 3:COLON CA 3:COLON CA		0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
165 3:COLON CA 3:COLON CA		0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
166 3:COLON CA 3:COLON CA		0	0.333	*0.667	0	0	0	0	0	0	0	0	0	0
0 0														
167 3:COLON CA 3:COLON CA		0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
168 3:COLON CA 3:COLON CA		0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
169 3:COLON CA 3:COLON CA		0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
170 3:COLON CA 3:COLON CA		0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
171 4:ESOPHAGU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
172 4:ESOPHAGU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
173 4:ESOPHAGU 1:BLADDER	+	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0
0 0														
174 4:ESOPHAGU 1:BLADDER	+	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0
0 0														
175 4:ESOPHAGU 1:BLADDER	+	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0
0 0														
176 4:ESOPHAGU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
177 4:ESOPHAGU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														

178 4:ESOPHAGU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0															
179 4:ESOPHAGU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0															
180 4:ESOPHAGU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0															
181 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0															
182 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0															
183 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0															
184 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0															
185 5:GASTRIC 7:HEPATOCE	+	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0															
186 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0															
187 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0															
188 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0															
189 5:GASTRIC 6:GLIOMA	+	0	0	0	0	0	*0.5	0.5	0	0	0	0	0	0	0
0															
190 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0															
191 6:GLIOMA 2:BREAST C	+	0	*0.5	0.5	0	0	0	0	0	0	0	0	0	0	0
0															
192 6:GLIOMA 2:BREAST C	+	0	*0.5	0.5	0	0	0	0	0	0	0	0	0	0	0
0															
193 6:GLIOMA 2:BREAST C	+	0	*0.5	0.5	0	0	0	0	0	0	0	0	0	0	0
0															
194 6:GLIOMA 2:BREAST C	+	0	*0.5	0.5	0	0	0	0	0	0	0	0	0	0	0
0															
195 6:GLIOMA 2:BREAST C	+	0	*0.5	0.5	0	0	0	0	0	0	0	0	0	0	0
0															

196	6:GLIOMA 2:BREAST C	+	0	*0.5	0.5	0	0	0	0	0	0	0	0	0	0	0
0																
197	6:GLIOMA 6:GLIOMA		0	0	0	0	0	*1	0	0	0	0	0	0	0	0
198	6:GLIOMA 6:GLIOMA		0	0	0	0	0	*1	0	0	0	0	0	0	0	0
199	6:GLIOMA 6:GLIOMA		0	0	0	0	0	*1	0	0	0	0	0	0	0	0
200	6:GLIOMA 6:GLIOMA		0	0	0	0	0	*1	0	0	0	0	0	0	0	0
201	7:HEPATOCE 7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
202	7:HEPATOCE 5:GASTRIC	+	0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																
203	7:HEPATOCE 7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
204	7:HEPATOCE 7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
205	7:HEPATOCE 7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
206	7:HEPATOCE 7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
207	7:HEPATOCE 7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
208	7:HEPATOCE 5:GASTRIC	+	0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																
209	7:HEPATOCE 13:PANCREA	+	0	0	0	0	0	0	0	0.091	0	0	0	0	0	0
*0.909	0															
210	7:HEPATOCE 5:GASTRIC	+	0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																
211	8:KIDNEY C 8:KIDNEY C		0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	0
0																
212	8:KIDNEY C 8:KIDNEY C		0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	0
0																
213	8:KIDNEY C 8:KIDNEY C		0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	0
0																
214	8:KIDNEY C 8:KIDNEY C		0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	0
0																
215	8:KIDNEY C 8:KIDNEY C		0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	0
0																

216 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
217 8:KIDNEY C 6:GLIOMA	+	0	0	0	0	0	0	*0.5	0.5	0	0	0	0	0
0														
218 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
219 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*1	0	0	0	0	0
0														
220 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
221 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	0	*0.909	0	0	0	0
0 0														
222 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	0	*0.909	0	0	0	0
0 0														
223 9:LUNG SQU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
224 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	0	*0.909	0	0	0	0
0 0														
225 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	0	*0.909	0	0	0	0
0 0														
226 9:LUNG SQU 9:LUNG SQU	0	0	0.091	0	0	0	0	0	0	*0.909	0	0	0	0
0 0														
227 9:LUNG SQU 3:COLON CA	+	0	0.333	*0.667	0	0	0	0	0	0	0	0	0	0
0 0														
228 9:LUNG SQU 3:COLON CA	+	0	0.333	*0.667	0	0	0	0	0	0	0	0	0	0
0 0														
229 9:LUNG SQU 1:BLADDER	+	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0
0														
230 9:LUNG SQU 1:BLADDER	+	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0
0														
231 10:MELANOM 4:ESOPHAGU	+	0	0	0	*1	0	0	0	0	0	0	0	0	0
0 0														
232 10:MELANOM 5:GASTRIC	+	0	0	0	0	*1	0	0	0	0	0	0	0	0
0														
233 10:MELANOM 10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	*0.9	0	0	0
0 0														

234	10:MELANOM	10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	0	0	*0.9	0	0	
0	0																	
235	10:MELANOM	10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	0	0	*0.9	0	0	
0	0																	
236	10:MELANOM	4:ESOPHAGU	+	0	0	0	*1	0	0	0	0	0	0	0	0	0	0	
0	0																	
237	10:MELANOM	10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	0	0	*0.9	0	0	
0	0																	
238	10:MELANOM	10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	0	0	*0.9	0	0	
0	0																	
239	10:MELANOM	10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	0	0	*0.9	0	0	
0	0																	
240	10:MELANOM	10:MELANOM	0.1	0	0	0	0	0	0	0	0	0	0	0	*0.9	0	0	
0	0																	
241	11:NSCLC	12:OVARIAN	+	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	
0																		
242	11:NSCLC	12:OVARIAN	+	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	
0																		
243	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	0
244	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	0
245	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	0
246	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	0
247	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	0
248	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	0
249	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	0
250	11:NSCLC	2:BREAST C	+	0	*0.5	0.5	0	0	0	0	0	0	0	0	0	0	0	0
0																		
251	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	0.2	0	*0.8	0	
0																		
252	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	0.2	0	*0.8	0	
0																		
253	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	
0																		
254	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	
0																		

255 12:OVARIAN 13:PANCREA	+	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
256 12:OVARIAN 13:PANCREA	+	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
257 12:OVARIAN 8:KIDNEY C	+	0	0	0	0	0	0	0	*0.9	0	0	0	0.1
0													
258 12:OVARIAN 8:KIDNEY C	+	0	0	0	0	0	0	0	*0.9	0	0	0	0.1
0													
259 12:OVARIAN 14:PROSTAT	+	0	0	0	0	0	0	0	0	0	0	0	0
*1													
260 12:OVARIAN 12:OVARIAN		0	0	0	0	0	0	0	0	0	0.2	0	*0.8
0													
261 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
262 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
263 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
264 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
265 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
266 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
267 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
268 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
269 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
270 13:PANCREA 8:KIDNEY C	+	0	0	0	0	0	0	0	*0.9	0	0	0	0.1
0													
271 14:PROSTAT 14:PROSTAT		0	0	0	0	0	0	0	0	0	0	0	0
*1													
272 14:PROSTAT 14:PROSTAT		0	0	0	0	0	0	0	0	0	0	0	0
*1													


```

273 14:PROSTAT 14:PROSTAT    0  0  0  0  0  0  0  0  0  0  0  0  0  0
*1
274 14:PROSTAT 13:PANCREA    + 0  0  0  0  0  0  0  0  0.091 0  0  0  0
*0.909 0
275 14:PROSTAT  6:GLIOMA    + 0  0  0  0  0  0  *1  0  0  0  0  0  0
0
276 14:PROSTAT 14:PROSTAT    0  0  0  0  0  0  0  0  0  0  0  0  0  0
*1
277 14:PROSTAT 14:PROSTAT    0  0  0  0  0  0  0  0  0  0  0  0  0  0
*1
278 14:PROSTAT 14:PROSTAT    0  0  0  0  0  0  0  0  0  0  0  0  0  0
*1
279 14:PROSTAT 12:OVARIAN    + 0  0  0  0  0  0  0  0  0  0  0.2  0  *0.8  0
0
280 14:PROSTAT 12:OVARIAN    + 0  0  0  0  0  0  0  0  0  0  0.2  0  *0.8  0
0

```

=== Evaluation on test set ===

=== Summary ===

Correctly Classified Instances	217	77.5 %
Incorrectly Classified Instances	63	22.5 %
Kappa statistic	0.7577	
Mean absolute error	0.0369	
Root mean squared error	0.168	
Relative absolute error	27.7984 %	
Root relative squared error	65.2313 %	
Total Number of Instances	280	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.85	0.023	0.739	0.85	0.791	0.933	BLADDER CANCER
0.65	0.031	0.619	0.65	0.634	0.851	BREAST CARCINOMA
0.85	0.042	0.607	0.85	0.708	0.948	COLON CARCINOMA
0.5	0.023	0.625	0.5	0.556	0.741	ESOPHAGUS CARCINOMA

0.85	0.015	0.81	0.85	0.829	0.941	GASTRIC CANCER
0.7	0.019	0.737	0.7	0.718	0.845	GLIOMA
0.75	0.004	0.938	0.75	0.833	0.897	HEPATOCELLULAR CARCINOMA
0.9	0.015	0.818	0.9	0.857	0.963	KIDNEY CARCINOMA
0.75	0.004	0.938	0.75	0.833	0.873	LUNG SQUAMOUS CELL CARCINOMA
0.8	0.004	0.941	0.8	0.865	0.918	MELANOMA
0.8	0	1	0.8	0.889	0.925	NSCLC
0.7	0.031	0.636	0.7	0.667	0.9	OVARIAN CARCINOMA
0.95	0.019	0.792	0.95	0.864	0.965	PANCREAS CARCINOMA
0.8	0.012	0.842	0.8	0.821	0.894	PROSTATE CARCINOMA
Weighted Avg.	0.775	0.017	0.789	0.775	0.776	0.899

=== Confusion Matrix ===

```

a b c d e f g h i j k l m n <-- classified as
17 0 0 2 0 0 0 0 0 1 0 0 0 0 | a = BLADDER CANCER
1 13 1 1 0 0 0 0 0 0 0 0 2 0 2 | b = BREAST CARCINOMA
0 1 17 0 0 0 0 0 0 1 0 0 1 0 0 | c = COLON CARCINOMA
3 0 7 10 0 0 0 0 0 0 0 0 0 0 0 | d = ESOPHAGUS CARCINOMA
0 0 0 1 17 1 1 0 0 0 0 0 0 0 0 | e = GASTRIC CANCER
0 6 0 0 0 14 0 0 0 0 0 0 0 0 0 | f = GLIOMA
0 0 0 0 3 1 15 0 0 0 0 0 0 1 0 | g = HEPATOCELLULAR CARCINOMA
0 0 0 0 0 1 0 18 0 0 0 0 0 1 0 | h = KIDNEY CARCINOMA
2 0 3 0 0 0 0 0 15 0 0 0 0 0 0 | i = LUNG SQUAMOUS CELL CARCINOMA
0 0 0 2 1 0 0 0 0 16 0 1 0 0 0 | j = MELANOMA
0 1 0 0 0 1 0 0 0 0 16 2 0 0 0 | k = NSCLC
0 0 0 0 0 0 0 3 0 0 0 14 2 1 0 | l = OVARIAN CARCINOMA
0 0 0 0 0 0 0 1 0 0 0 0 19 0 0 | m = PANCREAS CARCINOMA
0 0 0 0 0 1 0 0 0 0 0 2 1 16 0 | n = PROSTATE CARCINOMA

```

=== Run information ===

Scheme: weka.classifiers.trees.J48graft -C 0.25 -M 2

Relation: TrainingData323SOLIDnoADENO48miR-weka.filters.unsupervised.attribute.Remove-R2-3,7-11,13,15-16,18-23,26-29,31-32,34-36,40-45,47

Instances: 140

Attributes: 16

hsa-let-7e
hsa-miR-105*
hsa-miR-10a
hsa-miR-122
hsa-miR-145
hsa-miR-15b
hsa-miR-193b
hsa-miR-215
hsa-miR-218
hsa-miR-30c-1*
hsa-miR-330-5p
hsa-miR-34c-3p
hsa-miR-375
hsa-miR-383
hsa-miR-9
Class

Test mode:evaluate on training data

=== Classifier model (full training set) ===

J48graft pruned tree

```
hsa-miR-218 <= 6.834
| hsa-miR-215 <= 10.152
| | hsa-miR-383 <= 9.043
| | | hsa-miR-34c-3p <= 5.821
| | | | hsa-miR-193b <= 8.367
| | | | | hsa-miR-383 <= 6.914
| | | | | hsa-let-7e <= 9.765: LUNG SQUAMOUS CELL CARCINOMA (8.0)
| | | | | hsa-let-7e > 9.765: GLIOMA (3.0)
| | | | | hsa-miR-383 > 6.914: ESOPHAGUS CARCINOMA (6.0/1.0)
| | | | hsa-miR-193b > 8.367: PROSTATE CARCINOMA (11.0/1.0)
| | | hsa-miR-34c-3p > 5.821
| | | | hsa-miR-30c-1* <= 9.066
```

| | | | | hsa-miR-10a <= 9.198
| | | | | | hsa-miR-122 <= 8.518: BLADDER CANCER (10.0/1.0)
| | | | | | hsa-miR-122 > 8.518
| | | | | | | hsa-let-7e <= 7.8325: BLADDER CANCER (0.0|7.0)
| | | | | | | hsa-let-7e > 7.8325: BREAST CARCINOMA (3.0/1.0)
| | | | | | hsa-miR-10a > 9.198
| | | | | | | hsa-miR-218 <= 3.565
| | | | | | | | hsa-miR-30c-1* <= 7.097
| | | | | | | | | hsa-miR-15b <= 5.272: COLON CARCINOMA (2.0)
| | | | | | | | | hsa-miR-15b > 5.272: BREAST CARCINOMA (7.0)
| | | | | | | | | hsa-miR-30c-1* > 7.097
| | | | | | | | | | hsa-miR-105* <= 6.507: ESOPHAGUS CARCINOMA (2.0)
| | | | | | | | | | hsa-miR-105* > 6.507: LUNG SQUAMOUS CELL CARCINOMA (2.0)
| | | | | | | | | | hsa-miR-218 > 3.565: NSCLC (6.0/1.0)
| | | | | | | | | | hsa-miR-30c-1* > 9.066
| | | | | | | | | | | hsa-miR-105* <= 8.394: ESOPHAGUS CARCINOMA (2.0)
| | | | | | | | | | | hsa-miR-105* > 8.394: OVARIAN CARCINOMA (4.0)
| | | | | | | | | | | | hsa-miR-383 > 9.043: MELANOMA (11.0/1.0)
| | | | | | | | | | | | | hsa-miR-215 > 10.152
| | | | | | | | | | | | | | hsa-miR-330-5p <= 4.55
| | | | | | | | | | | | | | | hsa-miR-34c-3p <= 5.432: GASTRIC CANCER (9.0)
| | | | | | | | | | | | | | | hsa-miR-34c-3p > 5.432: HEPATOCELLULAR CARCINOMA (9.0)
| | | | | | | | | | | | | | | | hsa-miR-330-5p > 4.55: COLON CARCINOMA (7.0)
| | | | | | | | | | | | | | | | | hsa-miR-218 > 6.834
| | | | | | | | | | | | | | | | | | hsa-miR-375 <= 8.633
| | | | | | | | | | | | | | | | | | | hsa-miR-145 <= 9.173
| | | | | | | | | | | | | | | | | | | | hsa-miR-375 <= 5.946: GLIOMA (4.0)
| | | | | | | | | | | | | | | | | | | | hsa-miR-375 > 5.946
| | | | | | | | | | | | | | | | | | | | | hsa-let-7e <= 9.638
| | | | | | | | | | | | | | | | | | | | | | hsa-miR-122 <= 7.79: OVARIAN CARCINOMA (2.0)
| | | | | | | | | | | | | | | | | | | | | | hsa-miR-122 > 7.79: NSCLC (6.0/1.0)
| | | | | | | | | | | | | | | | | | | | | | | hsa-let-7e > 9.638: OVARIAN CARCINOMA (3.0)
| | | | | | | | | | | | | | | | | | | | | | | | hsa-miR-145 > 9.173
| | | | | | | | | | | | | | | | | | | | | | | | | hsa-miR-193b <= 6.339: KIDNEY CARCINOMA (10.0/1.0)
| | | | | | | | | | | | | | | | | | | | | | | | | hsa-miR-193b > 6.339
| | | | | | | | | | | | | | | | | | | | | | | | | | hsa-miR-330-5p <= 3.538: KIDNEY CARCINOMA (0.0|5.0)

| | | | hsa-miR-330-5p > 3.538: GLIOMA (2.0/1.0)
 | hsa-miR-375 > 8.633: PANCREAS CARCINOMA (11.0/1.0)

Number of Leaves : 26

Size of the tree : 51

Time taken to build model: 0.05 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	130	92.8571 %
Incorrectly Classified Instances	10	7.1429 %
Kappa statistic	0.9231	
Mean absolute error	0.0167	
Root mean squared error	0.0914	
Relative absolute error	12.6061 %	
Root relative squared error	35.505 %	
Total Number of Instances	140	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.9	0.008	0.9	0.9	0.9	0.992	BLADDER CANCER
0.9	0.008	0.9	0.9	0.9	0.995	BREAST CARCINOMA
0.9	0	1	0.9	0.947	0.999	COLON CARCINOMA
0.9	0.008	0.9	0.9	0.9	0.995	ESOPHAGUS CARCINOMA
0.9	0	1	0.9	0.947	0.998	GASTRIC CANCER
0.8	0.008	0.889	0.8	0.842	0.988	GLIOMA
0.9	0	1	0.9	0.947	1	HEPATOCELLULAR CARCINOMA
0.9	0.008	0.9	0.9	0.9	0.992	KIDNEY CARCINOMA
1	0	1	1	1	1	LUNG SQUAMOUS CELL CARCINOMA
1	0.008	0.909	1	0.952	0.996	MELANOMA
1	0.015	0.833	1	0.909	0.992	NSCLC
0.9	0	1	0.9	0.947	0.997	OVARIAN CARCINOMA

1	0.008	0.909	1	0.952	0.996	PANCREAS CARCINOMA
1	0.008	0.909	1	0.952	0.996	PROSTATE CARCINOMA
Weighted Avg.	0.929	0.005	0.932	0.929	0.928	0.995

=== Confusion Matrix ===

```

a b c d e f g h i j k l m n <-- classified as
9 0 0 0 0 0 0 0 0 1 0 0 0 0 | a = BLADDER CANCER
1 9 0 0 0 0 0 0 0 0 0 0 0 0 | b = BREAST CARCINOMA
0 1 9 0 0 0 0 0 0 0 0 0 0 0 | c = COLON CARCINOMA
0 0 0 9 0 0 0 0 0 0 1 0 0 0 | d = ESOPHAGUS CARCINOMA
0 0 0 1 9 0 0 0 0 0 0 0 0 0 | e = GASTRIC CANCER
0 0 0 0 0 8 0 0 0 0 1 0 0 1 | f = GLIOMA
0 0 0 0 0 1 9 0 0 0 0 0 0 0 | g = HEPATOCELLULAR CARCINOMA
0 0 0 0 0 0 0 9 0 0 0 0 1 0 | h = KIDNEY CARCINOMA
0 0 0 0 0 0 0 0 10 0 0 0 0 0 | i = LUNG SQUAMOUS CELL CARCINOMA
0 0 0 0 0 0 0 0 0 10 0 0 0 0 | j = MELANOMA
0 0 0 0 0 0 0 0 0 0 10 0 0 0 | k = NSCLC
0 0 0 0 0 0 0 1 0 0 0 9 0 0 | l = OVARIAN CARCINOMA
0 0 0 0 0 0 0 0 0 0 0 0 10 0 | m = PANCREAS CARCINOMA
0 0 0 0 0 0 0 0 0 0 0 0 0 10 | n = PROSTATE CARCINOMA

```

== Run information ==

```

Scheme:weka.classifiers.trees.J48graft -C 0.25 -M 2
Relation: TrainingData323SOLIDnoADENO48miR-weka.filters.unsupervised.attribute.Remove-R2-
3,7-11,13,15-16,18-23,26-29,31-32,34-36,40-45,47
Instances: 140
Attributes: 16
    hsa-let-7e
    hsa-miR-105*
    hsa-miR-10a
    hsa-miR-122
    hsa-miR-145
    hsa-miR-15b
    hsa-miR-193b

```

hsa-miR-215
hsa-miR-218
hsa-miR-30c-1*
hsa-miR-330-5p
hsa-miR-34c-3p
hsa-miR-375
hsa-miR-383
hsa-miR-9
Class

Test mode:user supplied test set: size unknown (reading incrementally)

=== Classifier model (full training set) ===

J48graft pruned tree

```
hsa-miR-218 <= 6.834
| hsa-miR-215 <= 10.152
| | hsa-miR-383 <= 9.043
| | | hsa-miR-34c-3p <= 5.821
| | | | hsa-miR-193b <= 8.367
| | | | | hsa-miR-383 <= 6.914
| | | | | | hsa-let-7e <= 9.765: LUNG SQUAMOUS CELL CARCINOMA (8.0)
| | | | | | hsa-let-7e > 9.765: GLIOMA (3.0)
| | | | | hsa-miR-383 > 6.914: ESOPHAGUS CARCINOMA (6.0/1.0)
| | | | hsa-miR-193b > 8.367: PROSTATE CARCINOMA (11.0/1.0)
| | | hsa-miR-34c-3p > 5.821
| | | | hsa-miR-30c-1* <= 9.066
| | | | | hsa-miR-10a <= 9.198
| | | | | | hsa-miR-122 <= 8.518: BLADDER CANCER (10.0/1.0)
| | | | | | hsa-miR-122 > 8.518
| | | | | | | hsa-let-7e <= 7.8325: BLADDER CANCER (0.0|7.0)
| | | | | | | hsa-let-7e > 7.8325: BREAST CARCINOMA (3.0/1.0)
| | | | | hsa-miR-10a > 9.198
| | | | | | hsa-miR-218 <= 3.565
| | | | | | | hsa-miR-30c-1* <= 7.097
```

```

| | | | | | | hsa-miR-15b <= 5.272: COLON CARCINOMA (2.0)
| | | | | | | hsa-miR-15b > 5.272: BREAST CARCINOMA (7.0)
| | | | | | | hsa-miR-30c-1* > 7.097
| | | | | | | hsa-miR-105* <= 6.507: ESOPHAGUS CARCINOMA (2.0)
| | | | | | | hsa-miR-105* > 6.507: LUNG SQUAMOUS CELL CARCINOMA (2.0)
| | | | | | hsa-miR-218 > 3.565: NSCLC (6.0/1.0)
| | | | hsa-miR-30c-1* > 9.066
| | | | hsa-miR-105* <= 8.394: ESOPHAGUS CARCINOMA (2.0)
| | | | hsa-miR-105* > 8.394: OVARIAN CARCINOMA (4.0)
| | hsa-miR-383 > 9.043: MELANOMA (11.0/1.0)
| hsa-miR-215 > 10.152
| | hsa-miR-330-5p <= 4.55
| | | hsa-miR-34c-3p <= 5.432: GASTRIC CANCER (9.0)
| | | hsa-miR-34c-3p > 5.432: HEPATOCELLULAR CARCINOMA (9.0)
| | hsa-miR-330-5p > 4.55: COLON CARCINOMA (7.0)
hsa-miR-218 > 6.834
| hsa-miR-375 <= 8.633
| | hsa-miR-145 <= 9.173
| | | hsa-miR-375 <= 5.946: GLIOMA (4.0)
| | | hsa-miR-375 > 5.946
| | | | hsa-let-7e <= 9.638
| | | | | hsa-miR-122 <= 7.79: OVARIAN CARCINOMA (2.0)
| | | | | hsa-miR-122 > 7.79: NSCLC (6.0/1.0)
| | | | hsa-let-7e > 9.638: OVARIAN CARCINOMA (3.0)
| | hsa-miR-145 > 9.173
| | | hsa-miR-193b <= 6.339: KIDNEY CARCINOMA (10.0/1.0)
| | | hsa-miR-193b > 6.339
| | | | hsa-miR-330-5p <= 3.538: KIDNEY CARCINOMA (0.0|5.0)
| | | | hsa-miR-330-5p > 3.538: GLIOMA (2.0/1.0)
| hsa-miR-375 > 8.633: PANCREAS CARCINOMA (11.0/1.0)

```

Number of Leaves : 26

Size of the tree : 51

Time taken to build model: 0.05 seconds

=== Predictions on test split ===

inst#,	actual,	predicted,	error,	probability distribution
1	1:BLADDER	1:BLADDER		*0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0				
2	1:BLADDER	1:BLADDER		*0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0				
3	1:BLADDER	10:MELANOM	+	0.091 0 0 0 0 0 0 0 0 0 *0.909 0 0
0	0			
4	1:BLADDER	1:BLADDER		*0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0				
5	1:BLADDER	1:BLADDER		*0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0				
6	1:BLADDER	1:BLADDER		*0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0				
7	1:BLADDER	1:BLADDER		*0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0				
8	1:BLADDER	1:BLADDER		*0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0				
9	1:BLADDER	1:BLADDER		*0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0				
10	1:BLADDER	1:BLADDER		*0.9 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0				
11	2:BREAST C	2:BREAST C		0 *1 0 0 0 0 0 0 0 0 0 0 0 0
0				
12	2:BREAST C	2:BREAST C		0 *1 0 0 0 0 0 0 0 0 0 0 0 0
0				
13	2:BREAST C	2:BREAST C		0 *1 0 0 0 0 0 0 0 0 0 0 0 0
0				
14	2:BREAST C	2:BREAST C		0 *1 0 0 0 0 0 0 0 0 0 0 0 0
0				
15	2:BREAST C	2:BREAST C		0 *1 0 0 0 0 0 0 0 0 0 0 0 0
0				

16 2: BREAST C 2: BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
17 2: BREAST C 2: BREAST C	0	*0.667	0.333	0	0	0	0	0	0	0	0	0	0	0
0 0														
18 2: BREAST C 2: BREAST C	0	*0.667	0.333	0	0	0	0	0	0	0	0	0	0	0
0 0														
19 2: BREAST C 2: BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
20 2: BREAST C 1: BLADDER	+	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0
0														
21 3: COLON CA 3: COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
22 3: COLON CA 3: COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
23 3: COLON CA 3: COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
24 3: COLON CA 3: COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
25 3: COLON CA 3: COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
26 3: COLON CA 3: COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
27 3: COLON CA 2: BREAST C	+	0	*0.667	0.333	0	0	0	0	0	0	0	0	0	0
0 0														
28 3: COLON CA 3: COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
29 3: COLON CA 3: COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
30 3: COLON CA 3: COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
31 4: ESOPHAGU 4: ESOPHAGU	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
32 4: ESOPHAGU 4: ESOPHAGU	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
33 4: ESOPHAGU 4: ESOPHAGU	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														

34	4:ESOPHAGU	11:NSCLC	+	0	0	0	0.167	0	0	0	0	0	0	0	0	*0.833	0
0	0																
35	4:ESOPHAGU	4:ESOPHAGU		0	0	0	*0.833	0.167	0	0	0	0	0	0	0	0	0
0	0																
36	4:ESOPHAGU	4:ESOPHAGU		0	0	0	*0.833	0.167	0	0	0	0	0	0	0	0	0
0	0																
37	4:ESOPHAGU	4:ESOPHAGU		0	0	0	*0.833	0.167	0	0	0	0	0	0	0	0	0
0	0																
38	4:ESOPHAGU	4:ESOPHAGU		0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0	0																
39	4:ESOPHAGU	4:ESOPHAGU		0	0	0	*0.833	0.167	0	0	0	0	0	0	0	0	0
0	0																
40	4:ESOPHAGU	4:ESOPHAGU		0	0	0	*0.833	0.167	0	0	0	0	0	0	0	0	0
0	0																
41	5:GASTRIC	5:GASTRIC		0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0	0																
42	5:GASTRIC	5:GASTRIC		0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0	0																
43	5:GASTRIC	5:GASTRIC		0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0	0																
44	5:GASTRIC	5:GASTRIC		0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0	0																
45	5:GASTRIC	5:GASTRIC		0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0	0																
46	5:GASTRIC	5:GASTRIC		0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0	0																
47	5:GASTRIC	5:GASTRIC		0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0	0																
48	5:GASTRIC	4:ESOPHAGU	+	0	0	0	*0.833	0.167	0	0	0	0	0	0	0	0	0
0	0																
49	5:GASTRIC	5:GASTRIC		0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0	0																
50	5:GASTRIC	5:GASTRIC		0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0	0																
51	6:GLIOMA	6:GLIOMA		0	0	0	0	*0.5	0.5	0	0	0	0	0	0	0	0
0	0																

52	6:GLIOMA	11:NSCLC	+	0	0	0	0	0	0.167	0	0	0	0	*0.833	0	0
0																
53	6:GLIOMA	6:GLIOMA		0	0	0	0	0	*1	0	0	0	0	0	0	0
54	6:GLIOMA	6:GLIOMA		0	0	0	0	0	*1	0	0	0	0	0	0	0
55	6:GLIOMA	6:GLIOMA		0	0	0	0	0	*1	0	0	0	0	0	0	0
56	6:GLIOMA	14:PROSTAT	+	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909																
57	6:GLIOMA	6:GLIOMA		0	0	0	0	0	*1	0	0	0	0	0	0	0
58	6:GLIOMA	6:GLIOMA		0	0	0	0	0	*1	0	0	0	0	0	0	0
59	6:GLIOMA	6:GLIOMA		0	0	0	0	0	*1	0	0	0	0	0	0	0
60	6:GLIOMA	6:GLIOMA		0	0	0	0	0	*1	0	0	0	0	0	0	0
61	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0
0																
62	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0
0																
63	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0
0																
64	7:HEPATOCE	6:GLIOMA	+	0	0	0	0	0	*0.5	0.5	0	0	0	0	0	0
0																
65	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0
0																
66	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0
0																
67	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0
0																
68	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0
0																
69	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0
0																
70	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	0	*1	0	0	0	0	0	0
0																
71	8:KIDNEY C	8:KIDNEY C		0	0	0	0	0	0	*0.9	0	0	0	0.1	0	0
0																
72	8:KIDNEY C	8:KIDNEY C		0	0	0	0	0	0	*0.9	0	0	0	0.1	0	0
0																

73 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
74 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
75 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
76 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
77 8:KIDNEY C 13:PANCREA	+ 0	0	0	0	0	0	0	0	0.091	0	0	0	0	0
*0.909 0														
78 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
79 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
80 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
81 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
82 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
83 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
84 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
85 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
86 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
87 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
88 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
89 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
90 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														

91	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
92	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
93	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
94	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
95	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
96	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
97	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
98	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
99	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
100	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
101	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
102	11:NSCLC	11:NSCLC	0	0	0	0.167	0	0	0	0	0	0	0	0	*0.833	0	0
0																	
103	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
104	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
105	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
106	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
107	11:NSCLC	11:NSCLC	0	0	0	0.167	0	0	0	0	0	0	0	0	*0.833	0	0
0																	
108	11:NSCLC	11:NSCLC	0	0	0	0.167	0	0	0	0	0	0	0	0	*0.833	0	0
0																	

109	11:NSCLC	11:NSCLC	0	0	0	0.167	0	0	0	0	0	0	0	*0.833	0	0	
0																	
110	11:NSCLC	11:NSCLC	0	0	0	0.167	0	0	0	0	0	0	0	*0.833	0	0	
0																	
111	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	
0																	
112	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	
0																	
113	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	
0																	
114	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	
0																	
115	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	
0																	
116	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	
0																	
117	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	
0																	
118	12:OVARIAN	8:KIDNEY C	+	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0																	
119	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	
0																	
120	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	
0																	
121	13:PANCREA	13:PANCREA	0	0	0	0	0	0	0	0.091	0	0	0	0	0	0	
*0.909	0																
122	13:PANCREA	13:PANCREA	0	0	0	0	0	0	0	0.091	0	0	0	0	0	0	
*0.909	0																
123	13:PANCREA	13:PANCREA	0	0	0	0	0	0	0	0.091	0	0	0	0	0	0	
*0.909	0																
124	13:PANCREA	13:PANCREA	0	0	0	0	0	0	0	0.091	0	0	0	0	0	0	
*0.909	0																
125	13:PANCREA	13:PANCREA	0	0	0	0	0	0	0	0.091	0	0	0	0	0	0	
*0.909	0																
126	13:PANCREA	13:PANCREA	0	0	0	0	0	0	0	0.091	0	0	0	0	0	0	
*0.909	0																

127 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
128 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
129 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
130 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909 0													
131 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
132 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
133 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
134 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
135 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
136 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
137 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
138 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
139 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
140 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
141 1:BLADDER 9:LUNG SQU	+	0	0	0	0	0	0	0	0	*1	0	0	0
0													
142 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0
0													
143 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0
0													
144 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0
0													

145 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
146 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
147 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
148 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
149 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
150 1:BLADDER 1:BLADDER	*0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
151 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
152 2:BREAST C 11:NSCLC	+ 0	0	0	0.167	0	0	0	0	0	0	0	*0.833	0	0
0 0														
153 2:BREAST C 3:COLON CA	+ 0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
154 2:BREAST C 3:COLON CA	+ 0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
155 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
156 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
157 2:BREAST C 2:BREAST C	0	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
158 2:BREAST C 2:BREAST C	0	*0.667	0.333	0	0	0	0	0	0	0	0	0	0	0
0 0														
159 2:BREAST C 2:BREAST C	0	*0.667	0.333	0	0	0	0	0	0	0	0	0	0	0
0 0														
160 2:BREAST C 1:BLADDER	+ *0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
161 3:COLON CA 11:NSCLC	+ 0	0	0	0.167	0	0	0	0	0	0	0	*0.833	0	0
0 0														
162 3:COLON CA 3:COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														

163 3:COLON CA 3:COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
164 3:COLON CA 3:COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
165 3:COLON CA 3:COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
166 3:COLON CA 4:ESOPHAGU	+	0	0	0	*1	0	0	0	0	0	0	0	0	0
0														
167 3:COLON CA 3:COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
168 3:COLON CA 3:COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
169 3:COLON CA 3:COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
170 3:COLON CA 3:COLON CA	0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
171 4:ESOPHAGU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
172 4:ESOPHAGU 4:ESOPHAGU	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0 0														
173 4:ESOPHAGU 4:ESOPHAGU	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0 0														
174 4:ESOPHAGU 11:NSCLC	+	0	0	0	0.167	0	0	0	0	0	0	0	*0.833	0
0 0														
175 4:ESOPHAGU 9:LUNG SQU	+	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
176 4:ESOPHAGU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
177 4:ESOPHAGU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
178 4:ESOPHAGU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
179 4:ESOPHAGU 3:COLON CA	+	0	0	*1	0	0	0	0	0	0	0	0	0	0
0														
180 4:ESOPHAGU 4:ESOPHAGU	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
0 0														

181	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0																
182	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0																
183	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0																
184	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0																
185	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0																
186	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0																
187	5:GASTRIC	7:HEPATOCE	+	0	0	0	0	0	0	*1	0	0	0	0	0	0
0																
188	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0																
189	5:GASTRIC	8:KIDNEY C	+	0	0	0	0	0	0	*1	0	0	0	0	0	0
0																
190	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0																
191	6:GLIOMA	11:NSCLC	+	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																
192	6:GLIOMA	11:NSCLC	+	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																
193	6:GLIOMA	11:NSCLC	+	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																
194	6:GLIOMA	14:PROSTAT	+	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909																
195	6:GLIOMA	11:NSCLC	+	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																
196	6:GLIOMA	14:PROSTAT	+	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909																
197	6:GLIOMA	11:NSCLC	+	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																
198	6:GLIOMA	14:PROSTAT	+	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909																

199	6:GLIOMA	6:GLIOMA	0	0	0	0	0	*1	0	0	0	0	0	0	0	0
200	6:GLIOMA	2:BREAST C	+	0	*1	0	0	0	0	0	0	0	0	0	0	0
0																
201	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
202	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
203	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
204	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
205	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
206	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
207	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
208	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
209	7:HEPATOCE	13:PANCREA	+	0	0	0	0	0	0	0	0.091	0	0	0	0	0
*0.909	0															
210	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
211	8:KIDNEY C	8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	0
0																
212	8:KIDNEY C	8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	0
0																
213	8:KIDNEY C	6:GLIOMA	+	0	0	0	0	0	*0.5	0.5	0	0	0	0	0	0
0																
214	8:KIDNEY C	8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	0
0																
215	8:KIDNEY C	8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	0
0																
216	8:KIDNEY C	8:KIDNEY C	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0	0
0																

217 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
218 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
219 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
220 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*0.9	0	0	0	0.1	0
0														
221 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
222 9:LUNG SQU 11:NSCLC	+ 0	0	0	0.167	0	0	0	0	0	0	0	0	*0.833	0
0 0														
223 9:LUNG SQU 3:COLON CA	+ 0	0	*1	0	0	0	0	0	0	0	0	0	0	0
0														
224 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
225 9:LUNG SQU 6:GLIOMA	+ 0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0														
226 9:LUNG SQU 9:LUNG SQU	0	0	0	0	0	0	0	0	0	*1	0	0	0	0
0														
227 9:LUNG SQU 1:BLADDER	+ *0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
228 9:LUNG SQU 1:BLADDER	+ *0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
229 9:LUNG SQU 1:BLADDER	+ *0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
230 9:LUNG SQU 1:BLADDER	+ *0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0														
231 10:MELANOM 1:BLADDER	+ *0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0
0 0														
232 10:MELANOM 10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0 0														
233 10:MELANOM 10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0 0														
234 10:MELANOM 10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0 0														

235	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
236	10:MELANOM	11:NSCLC	+	0	0	0	0.167	0	0	0	0	0	0	0	*0.833	0	0
0	0																
237	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
238	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
239	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
240	10:MELANOM	10:MELANOM	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.909	0	0
0	0																
241	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
242	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
243	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
244	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
245	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
246	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
247	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
248	11:NSCLC	11:NSCLC	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833	0	0
0																	
249	11:NSCLC	11:NSCLC	0	0	0	0.167	0	0	0	0	0	0	0	0	*0.833	0	0
0																	
250	11:NSCLC	11:NSCLC	0	0	0	0.167	0	0	0	0	0	0	0	0	*0.833	0	0
0																	
251	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	0
0																	
252	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0	0	0	0	0	0	0	*1	0	0
0																	

253	12:OVARIAN	11:NSCLC	+	0	0	0	0	0	0.167	0	0	0	0	*0.833	0
0	0														
254	12:OVARIAN	11:NSCLC	+	0	0	0	0	0	0.167	0	0	0	0	*0.833	0
0	0														
255	12:OVARIAN	13:PANCREA	+	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909	0														
256	12:OVARIAN	13:PANCREA	+	0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909	0														
257	12:OVARIAN	8:KIDNEY C	+	0	0	0	0	0	0	0	*0.9	0	0	0	0.1
0															
258	12:OVARIAN	8:KIDNEY C	+	0	0	0	0	0	0	0	*0.9	0	0	0	0.1
0															
259	12:OVARIAN	11:NSCLC	+	0	0	0	0.167	0	0	0	0	0	0	*0.833	0
0	0														
260	12:OVARIAN	11:NSCLC	+	0	0	0	0.167	0	0	0	0	0	0	*0.833	0
0	0														
261	13:PANCREA	13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909	0														
262	13:PANCREA	13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909	0														
263	13:PANCREA	13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909	0														
264	13:PANCREA	13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909	0														
265	13:PANCREA	13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909	0														
266	13:PANCREA	13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909	0														
267	13:PANCREA	13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909	0														
268	13:PANCREA	13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909	0														
269	13:PANCREA	13:PANCREA		0	0	0	0	0	0	0	0.091	0	0	0	0
*0.909	0														
270	13:PANCREA	8:KIDNEY C	+	0	0	0	0	0	0	0	*0.9	0	0	0	0.1
0															

271 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
272 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
273 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
274 14:PROSTAT 13:PANCREA	+	0	0	0	0	0	0	0.091	0	0	0	0	0
*0.909 0													
275 14:PROSTAT 6:GLIOMA	+	0	0	0	0	0	*1	0	0	0	0	0	0
0													
276 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
277 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
278 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													
279 14:PROSTAT 1:BLADDER	+	*0.9	0.1	0	0	0	0	0	0	0	0	0	0
0													
280 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0.091	0	0	0	0	0	0	0
*0.909													

=== Evaluation on test set ===

=== Summary ===

Correctly Classified Instances	222	79.2857 %
Incorrectly Classified Instances	58	20.7143 %
Kappa statistic	0.7769	
Mean absolute error	0.0357	
Root mean squared error	0.1639	
Relative absolute error	26.9091 %	
Root relative squared error	63.6456 %	
Total Number of Instances	280	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
---------	---------	-----------	--------	-----------	----------	-------

0.9	0.031	0.692	0.9	0.783	0.955	BLADDER CANCER
0.75	0.008	0.882	0.75	0.811	0.91	BREAST CARCINOMA
0.85	0.031	0.68	0.85	0.756	0.933	COLON CARCINOMA
0.6	0.008	0.857	0.6	0.706	0.832	ESOPHAGUS CARCINOMA
0.85	0	1	0.85	0.919	0.949	GASTRIC CANCER
0.45	0.015	0.692	0.45	0.545	0.804	GLIOMA
0.9	0.004	0.947	0.9	0.923	0.973	HEPATOCELLULAR CARCINOMA
0.9	0.019	0.783	0.9	0.837	0.959	KIDNEY CARCINOMA
0.65	0.008	0.867	0.65	0.743	0.821	LUNG SQUAMOUS CELL CARCINOMA
0.9	0.004	0.947	0.9	0.923	0.948	MELANOMA
1	0.062	0.556	1	0.714	0.969	NSCLC
0.55	0	1	0.55	0.71	0.834	OVARIAN CARCINOMA
0.95	0.019	0.792	0.95	0.864	0.965	PANCREAS CARCINOMA
0.85	0.015	0.81	0.85	0.829	0.917	PROSTATE CARCINOMA
Weighted Avg.	0.793	0.016	0.822	0.793	0.79	0.912

=== Confusion Matrix ===

```

a b c d e f g h i j k l m n <-- classified as
18 0 0 0 0 0 0 0 1 1 0 0 0 0 | a = BLADDER CANCER
2 15 2 0 0 0 0 0 0 0 1 0 0 0 | b = BREAST CARCINOMA
0 1 17 1 0 0 0 0 0 0 1 0 0 0 | c = COLON CARCINOMA
0 0 5 12 0 0 0 0 1 0 2 0 0 0 | d = ESOPHAGUS CARCINOMA
0 0 0 1 17 0 1 1 0 0 0 0 0 0 | e = GASTRIC CANCER
0 1 0 0 0 9 0 0 0 0 6 0 0 4 | f = GLIOMA
0 0 0 0 0 1 18 0 0 0 0 0 1 0 | g = HEPATOCELLULAR CARCINOMA
0 0 0 0 0 1 0 18 0 0 0 0 1 0 | h = KIDNEY CARCINOMA
4 0 1 0 0 1 0 0 13 0 1 0 0 0 | i = LUNG SQUAMOUS CELL CARCINOMA
1 0 0 0 0 0 0 0 0 18 1 0 0 0 | j = MELANOMA
0 0 0 0 0 0 0 0 0 0 20 0 0 0 | k = NSCLC
0 0 0 0 0 0 0 3 0 0 4 11 2 0 | l = OVARIAN CARCINOMA
0 0 0 0 0 0 0 1 0 0 0 0 19 0 | m = PANCREAS CARCINOMA
1 0 0 0 0 1 0 0 0 0 0 0 1 17 | n = PROSTATE CARCINOMA

```

=== Run information ===

Scheme:weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1

Relation: TrainingData323SOLIDnoADENO48miR-weka.filters.unsupervised.attribute.Remove-R2-3,7-11,13,15-16,18-23,26-29,31-32,34-36,40-45,47

Instances: 140

Attributes: 16

hsa-let-7e
hsa-miR-105*
hsa-miR-10a
hsa-miR-122
hsa-miR-145
hsa-miR-15b
hsa-miR-193b
hsa-miR-215
hsa-miR-218
hsa-miR-30c-1*
hsa-miR-330-5p
hsa-miR-34c-3p
hsa-miR-375
hsa-miR-383
hsa-miR-9
Class

Test mode:evaluate on training data

=== Classifier model (full training set) ===

JRIP rules:

=====

(hsa-miR-383 <= 4.417) and (hsa-miR-30c-1* >= 4.024) and (hsa-miR-215 <= 5.322) => Class=LUNG SQUAMOUS CELL CARCINOMA (11.0/3.0)

(hsa-miR-9 >= 8.487) => Class=GLIOMA (10.0/2.0)

(hsa-miR-122 >= 11.89) => Class=HEPATOCELLULAR CARCINOMA (9.0/0.0)

(hsa-miR-145 <= 3.852) and (hsa-let-7e >= 9.894) => Class=PROSTATE CARCINOMA (9.0/0.0)

(hsa-miR-383 >= 9.519) => Class=MELANOMA (9.0/1.0)

(hsa-miR-15b >= 8.367) and (hsa-miR-218 <= 4.893) => Class=BREAST CARCINOMA (8.0/1.0)

(hsa-miR-145 >= 11.285) and (hsa-miR-383 >= 6.18) => Class=GASTRIC CANCER (8.0/0.0)

(hsa-miR-145 <= 7.731) and (hsa-miR-193b <= 8.717) => Class=NSCLC (9.0/1.0)
 (hsa-miR-330-5p >= 8.863) => Class=NSCLC (2.0/0.0)
 (hsa-miR-330-5p >= 7.761) and (hsa-miR-122 <= 7.731) => Class=BLADDER CANCER (4.0/0.0)
 (hsa-miR-218 <= 3.513) and (hsa-miR-10a <= 9.198) => Class=BLADDER CANCER (4.0/0.0)
 (hsa-miR-145 <= 9.173) and (hsa-miR-375 <= 8.103) => Class=OVARIAN CARCINOMA (8.0/1.0)
 (hsa-miR-330-5p <= 3.533) => Class=KIDNEY CARCINOMA (5.0/0.0)
 (hsa-miR-218 >= 7.41) => Class=PANCREAS CARCINOMA (12.0/2.0)
 (hsa-miR-383 >= 7.298) => Class=ESOPHAGUS CARCINOMA (11.0/3.0)
 => Class=COLON CARCINOMA (21.0/12.0)

Number of Rules : 16

Time taken to build model: 0.11 seconds

=== Predictions on training set ===

inst#	actual	predicted	error	probability distribution
1	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
2	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
3	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
4	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
5	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
6	1:BLADDER	3:COLON CA	+	0.048 0.095 *0.429 0.048 0.095 0.048 0 0.095 0.095 0
7	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
8	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
9	1:BLADDER	9:LUNG SQU	+	0.091 0.091 0 0 0 0 0 0 0 *0.727 0 0 0

10 1:BLADDER 1:BLADDER	*1	0	0	0	0	0	0	0	0	0	0	0	0	0
0														
11 2:BREAST C 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
12 2:BREAST C 2:BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0	0	0
0 0														
13 2:BREAST C 9:LUNG SQU	+	0.091	0.091	0	0	0	0	0	0	*0.727	0	0	0	0
0 0.091														
14 2:BREAST C 2:BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0	0	0
0 0														
15 2:BREAST C 2:BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0	0	0
0 0														
16 2:BREAST C 2:BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0	0	0
0 0														
17 2:BREAST C 2:BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0	0	0
0 0														
18 2:BREAST C 2:BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0	0	0
0 0														
19 2:BREAST C 2:BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0	0	0
0 0														
20 2:BREAST C 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
21 3:COLON CA 3:COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
22 3:COLON CA 3:COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
23 3:COLON CA 3:COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
24 3:COLON CA 3:COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
25 3:COLON CA 3:COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
26 3:COLON CA 3:COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
27 3:COLON CA 4:ESOPHAGU	+	0	0	0.091	*0.727	0	0	0	0	0	0.091	0		
0.091 0 0														

28 3:COLON CA 3:COLON CA	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0
0 0.048 0 0										
29 3:COLON CA 3:COLON CA	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0
0 0.048 0 0										
30 3:COLON CA 3:COLON CA	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0
0 0.048 0 0										
31 4:ESOPHAGU 4:ESOPHAGU	0	0	0.091	*0.727	0	0	0	0	0.091	0
0.091 0 0										
32 4:ESOPHAGU 4:ESOPHAGU	0	0	0.091	*0.727	0	0	0	0	0.091	0
0.091 0 0										
33 4:ESOPHAGU 4:ESOPHAGU	0	0	0.091	*0.727	0	0	0	0	0.091	0
0.091 0 0										
34 4:ESOPHAGU 3:COLON CA	+ 0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0
0 0.048 0 0										
35 4:ESOPHAGU 4:ESOPHAGU	0	0	0.091	*0.727	0	0	0	0	0.091	0
0.091 0 0										
36 4:ESOPHAGU 4:ESOPHAGU	0	0	0.091	*0.727	0	0	0	0	0.091	0
0.091 0 0										
37 4:ESOPHAGU 4:ESOPHAGU	0	0	0.091	*0.727	0	0	0	0	0.091	0
0.091 0 0										
38 4:ESOPHAGU 4:ESOPHAGU	0	0	0.091	*0.727	0	0	0	0	0.091	0
0.091 0 0										
39 4:ESOPHAGU 6:GLIOMA	+ 0	0	0	0.1	0	*0.8	0	0	0	0.1
0 0 0 0										
40 4:ESOPHAGU 4:ESOPHAGU	0	0	0.091	*0.727	0	0	0	0	0.091	0
0.091 0 0										
41 5:GASTRIC 5:GASTRIC	0	0	0	0	*1	0	0	0	0	0
0 0 0 0										
42 5:GASTRIC 5:GASTRIC	0	0	0	0	*1	0	0	0	0	0
0 0 0 0										
43 5:GASTRIC 5:GASTRIC	0	0	0	0	*1	0	0	0	0	0
0 0 0 0										
44 5:GASTRIC 5:GASTRIC	0	0	0	0	*1	0	0	0	0	0
0 0 0 0										
45 5:GASTRIC 3:COLON CA	+ 0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0
0 0.048 0 0										

46	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0																
47	5:GASTRIC	3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0	0.048	0	0													
48	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0																
49	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0																
50	5:GASTRIC	5:GASTRIC	0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0																
51	6:GLIOMA	3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0	0.048	0	0													
52	6:GLIOMA	12:OVARIAN	+	0	0	0	0	0	0.125	0	0	0	0	0	0	*0.875
0																
53	6:GLIOMA	6:GLIOMA	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																
54	6:GLIOMA	6:GLIOMA	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																
55	6:GLIOMA	6:GLIOMA	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																
56	6:GLIOMA	6:GLIOMA	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																
57	6:GLIOMA	6:GLIOMA	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																
58	6:GLIOMA	6:GLIOMA	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																
59	6:GLIOMA	6:GLIOMA	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																
60	6:GLIOMA	6:GLIOMA	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																
61	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
62	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																
63	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
0																

64 7:HEPATOCE 7:HEPATOCE	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
65 7:HEPATOCE 2:BREAST C	+	0	*0.875	0	0	0	0	0	0.125	0	0	0	0	0
0 0														
66 7:HEPATOCE 7:HEPATOCE	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
67 7:HEPATOCE 7:HEPATOCE	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
68 7:HEPATOCE 7:HEPATOCE	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
69 7:HEPATOCE 7:HEPATOCE	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
70 7:HEPATOCE 7:HEPATOCE	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
71 8:KIDNEY C 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0	0
0 0.048 0 0														
72 8:KIDNEY C 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0	0
0 0.048 0 0														
73 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
74 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
75 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
76 8:KIDNEY C 13:PANCREA	+	0	0	0	0	0	0	0	0.167	0	0	0	0	0
*0.833 0														
77 8:KIDNEY C 13:PANCREA	+	0	0	0	0	0	0	0	0.167	0	0	0	0	0
*0.833 0														
78 8:KIDNEY C 10:MELANOM	+	0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0
0 0														
79 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
80 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
81 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	0	*0.727	0	0	0
0 0.091														

82 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0
0 0.091													
83 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0
0 0.091													
84 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0
0 0.091													
85 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0
0 0.091													
86 9:LUNG SQU 3:COLON CA	+ 0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0	0
0 0.048 0 0													
87 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0
0 0.091													
88 9:LUNG SQU 3:COLON CA	+ 0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0	0
0 0.048 0 0													
89 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0
0 0.091													
90 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0
0 0.091													
91 10:MELANOM 10:MELANOM	0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0
0 0													
92 10:MELANOM 10:MELANOM	0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0
0 0													
93 10:MELANOM 10:MELANOM	0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0
0 0													
94 10:MELANOM 4:ESOPHAGU	+ 0	0	0.091	*0.727	0	0	0	0	0	0	0.091	0	0
0.091 0 0													
95 10:MELANOM 10:MELANOM	0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0
0 0													
96 10:MELANOM 10:MELANOM	0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0
0 0													
97 10:MELANOM 6:GLIOMA	+ 0	0	0	0.1	0	*0.8	0	0	0	0	0.1	0	0 0
0													
98 10:MELANOM 10:MELANOM	0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0
0 0													
99 10:MELANOM 10:MELANOM	0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0
0 0													

100	10:MELANOM	10:MELANOM	0	0	0	0	0	0	0	0	0	0.111	0	*0.889	0	0	
0	0																
101	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	*0.889	0.111	0	
0																	
102	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	*0.889	0.111	0	
0																	
103	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	*0.889	0.111	0	
0																	
104	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	*0.889	0.111	0	
0																	
105	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	*0.889	0.111	0	
0																	
106	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	*0.889	0.111	0	
0																	
107	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	0
108	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	*0.889	0.111	0	
0																	
109	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	*0.889	0.111	0	
0																	
110	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	0	0	*1	0	0	0
111	12:OVARIAN	4:ESOPHAGU	+	0	0	0.091	*0.727	0	0	0	0	0	0	0.091	0		
0.091	0	0															
112	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0.125	0	0	0	0	0	*0.875			
0	0																
113	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0.125	0	0	0	0	0	*0.875			
0	0																
114	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0.125	0	0	0	0	0	*0.875			
0	0																
115	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0.125	0	0	0	0	0	*0.875			
0	0																
116	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0.125	0	0	0	0	0	*0.875			
0	0																
117	12:OVARIAN	12:OVARIAN	0	0	0	0	0	0.125	0	0	0	0	0	*0.875			
0	0																
118	12:OVARIAN	3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0				
0	0.048	0	0														

119 12:OVARIAN 11:NSCLC	+	0	0	0	0	0	0	0	0	0	0	0	0	0	*0.889	0.111
0 0																
120 12:OVARIAN 12:OVARIAN		0	0	0	0	0	0.125	0	0	0	0	0	0	0	*0.875	
0 0																
121 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0		
*0.833 0																
122 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0		
*0.833 0																
123 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0		
*0.833 0																
124 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0		
*0.833 0																
125 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0		
*0.833 0																
126 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0		
*0.833 0																
127 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0		
*0.833 0																
128 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0		
*0.833 0																
129 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0		
*0.833 0																
130 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0		
*0.833 0																
131 14:PROSTAT 14:PROSTAT		0	0	0	0	0	0	0	0	0	0	0	0	0		
*1																
132 14:PROSTAT 14:PROSTAT		0	0	0	0	0	0	0	0	0	0	0	0	0		
*1																
133 14:PROSTAT 14:PROSTAT		0	0	0	0	0	0	0	0	0	0	0	0	0		
*1																
134 14:PROSTAT 14:PROSTAT		0	0	0	0	0	0	0	0	0	0	0	0	0		
*1																
135 14:PROSTAT 14:PROSTAT		0	0	0	0	0	0	0	0	0	0	0	0	0		
*1																
136 14:PROSTAT 9:LUNG SQU	+	0.091	0.091	0	0	0	0	0	0	0	0	0	0	0	*0.727	0 0 0
0 0.091																

```

137 14:PROSTAT 14:PROSTAT    0  0  0  0  0  0  0  0  0  0  0  0  0
*1
138 14:PROSTAT 14:PROSTAT    0  0  0  0  0  0  0  0  0  0  0  0  0
*1
139 14:PROSTAT 14:PROSTAT    0  0  0  0  0  0  0  0  0  0  0  0  0
*1
140 14:PROSTAT 14:PROSTAT    0  0  0  0  0  0  0  0  0  0  0  0  0
*1

```

=== Evaluation on training set ===

=== Summary ===

```

Correctly Classified Instances    114      81.4286 %
Incorrectly Classified Instances   26      18.5714 %
Kappa statistic                   0.8
Mean absolute error               0.0406
Root mean squared error           0.1425
Relative absolute error           30.6135 %
Root relative squared error       55.3295 %
Total Number of Instances        140

```

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.8	0	1	0.8	0.889	0.981	BLADDER CANCER
0.7	0.008	0.875	0.7	0.778	0.962	BREAST CARCINOMA
0.9	0.092	0.429	0.9	0.581	0.945	COLON CARCINOMA
0.8	0.023	0.727	0.8	0.762	0.968	ESOPHAGUS CARCINOMA
0.8	0	1	0.8	0.889	0.985	GASTRIC CANCER
0.8	0.015	0.8	0.8	0.8	0.975	GLIOMA
0.9	0	1	0.9	0.947	0.997	HEPATOCELLULAR CARCINOMA
0.5	0	1	0.5	0.667	0.939	KIDNEY CARCINOMA
0.8	0.023	0.727	0.8	0.762	0.972	LUNG SQUAMOUS CELL CARCINOMA
0.8	0.008	0.889	0.8	0.842	0.981	MELANOMA
1	0.008	0.909	1	0.952	0.997	NSCLC
0.7	0.008	0.875	0.7	0.778	0.96	OVARIAN CARCINOMA

	1	0.015	0.833	1	0.909	0.992	PANCREAS CARCINOMA
	0.9	0	1	0.9	0.947	0.996	PROSTATE CARCINOMA
Weighted Avg.	0.814	0.014	0.862	0.814	0.822	0.975	

=== Confusion Matrix ===

```

a b c d e f g h i j k l m n <-- classified as
8 0 1 0 0 0 0 0 1 0 0 0 0 0 | a = BLADDER CANCER
0 7 2 0 0 0 0 0 1 0 0 0 0 0 | b = BREAST CARCINOMA
0 0 9 1 0 0 0 0 0 0 0 0 0 0 | c = COLON CARCINOMA
0 0 1 8 0 1 0 0 0 0 0 0 0 0 | d = ESOPHAGUS CARCINOMA
0 0 2 0 8 0 0 0 0 0 0 0 0 0 | e = GASTRIC CANCER
0 0 1 0 0 8 0 0 0 0 0 1 0 0 | f = GLIOMA
0 1 0 0 0 0 9 0 0 0 0 0 0 0 | g = HEPATOCELLULAR CARCINOMA
0 0 2 0 0 0 0 5 0 1 0 0 2 0 | h = KIDNEY CARCINOMA
0 0 2 0 0 0 0 0 8 0 0 0 0 0 | i = LUNG SQUAMOUS CELL CARCINOMA
0 0 0 1 0 1 0 0 0 8 0 0 0 0 | j = MELANOMA
0 0 0 0 0 0 0 0 0 0 10 0 0 0 | k = NSCLC
0 0 1 1 0 0 0 0 0 0 1 7 0 0 | l = OVARIAN CARCINOMA
0 0 0 0 0 0 0 0 0 0 0 0 10 0 | m = PANCREAS CARCINOMA
0 0 0 0 0 0 0 0 0 1 0 0 0 0 9 | n = PROSTATE CARCINOMA

```

=== Run information ===

Scheme:weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1

Relation: TrainingData323SOLIDnoADENO48miR-weka.filters.unsupervised.attribute.Remove-R2-3,7-11,13,15-16,18-23,26-29,31-32,34-36,40-45,47

Instances: 140

Attributes: 16

- hsa-let-7e
- hsa-miR-105*
- hsa-miR-10a
- hsa-miR-122
- hsa-miR-145
- hsa-miR-15b
- hsa-miR-193b
- hsa-miR-215

hsa-miR-218
hsa-miR-30c-1*
hsa-miR-330-5p
hsa-miR-34c-3p
hsa-miR-375
hsa-miR-383
hsa-miR-9
Class

Test mode:user supplied test set: size unknown (reading incrementally)

=== Classifier model (full training set) ===

JRIP rules:

=====

(hsa-miR-383 <= 4.417) and (hsa-miR-30c-1* >= 4.024) and (hsa-miR-215 <= 5.322) => Class=LUNG
SQUAMOUS CELL CARCINOMA (11.0/3.0)

(hsa-miR-9 >= 8.487) => Class=GLIOMA (10.0/2.0)

(hsa-miR-122 >= 11.89) => Class=HEPATOCELLULAR CARCINOMA (9.0/0.0)

(hsa-miR-145 <= 3.852) and (hsa-let-7e >= 9.894) => Class=PROSTATE CARCINOMA (9.0/0.0)

(hsa-miR-383 >= 9.519) => Class=MELANOMA (9.0/1.0)

(hsa-miR-15b >= 8.367) and (hsa-miR-218 <= 4.893) => Class=BREAST CARCINOMA (8.0/1.0)

(hsa-miR-145 >= 11.285) and (hsa-miR-383 >= 6.18) => Class=GASTRIC CANCER (8.0/0.0)

(hsa-miR-145 <= 7.731) and (hsa-miR-193b <= 8.717) => Class=NSCLC (9.0/1.0)

(hsa-miR-330-5p >= 8.863) => Class=NSCLC (2.0/0.0)

(hsa-miR-330-5p >= 7.761) and (hsa-miR-122 <= 7.731) => Class=BLADDER CANCER (4.0/0.0)

(hsa-miR-218 <= 3.513) and (hsa-miR-10a <= 9.198) => Class=BLADDER CANCER (4.0/0.0)

(hsa-miR-145 <= 9.173) and (hsa-miR-375 <= 8.103) => Class=OVARIAN CARCINOMA (8.0/1.0)

(hsa-miR-330-5p <= 3.533) => Class=KIDNEY CARCINOMA (5.0/0.0)

(hsa-miR-218 >= 7.41) => Class=PANCREAS CARCINOMA (12.0/2.0)

(hsa-miR-383 >= 7.298) => Class=ESOPHAGUS CARCINOMA (11.0/3.0)

=> Class=COLON CARCINOMA (21.0/12.0)

Number of Rules : 16

Time taken to build model: 0.09 seconds

=== Predictions on test split ===

inst#,	actual,	predicted,	error,	probability distribution
1	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
2	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
3	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
4	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
5	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
6	1:BLADDER	3:COLON CA	+	0.048 0.095 *0.429 0.048 0.095 0.048 0 0.095 0.095 0 0 0.048 0 0
7	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
8	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
9	1:BLADDER	9:LUNG SQU	+	0.091 0.091 0 0 0 0 0 0 *0.727 0 0 0 0 0.091
10	1:BLADDER	1:BLADDER	*1	0 0 0 0 0 0 0 0 0 0 0 0 0 0
11	2:BREAST C	3:COLON CA	+	0.048 0.095 *0.429 0.048 0.095 0.048 0 0.095 0.095 0 0 0.048 0 0
12	2:BREAST C	2:BREAST C	0	*0.875 0 0 0 0 0.125 0 0 0 0 0 0 0
13	2:BREAST C	9:LUNG SQU	+	0.091 0.091 0 0 0 0 0 0 *0.727 0 0 0 0 0.091
14	2:BREAST C	2:BREAST C	0	*0.875 0 0 0 0 0.125 0 0 0 0 0 0 0
15	2:BREAST C	2:BREAST C	0	*0.875 0 0 0 0 0.125 0 0 0 0 0 0 0

16 2: BREAST C 2: BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0
0 0												
17 2: BREAST C 2: BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0
0 0												
18 2: BREAST C 2: BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0
0 0												
19 2: BREAST C 2: BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0
0 0												
20 2: BREAST C 3: COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	
0 0.048 0 0												
21 3: COLON CA 3: COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	
0 0.048 0 0												
22 3: COLON CA 3: COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	
0 0.048 0 0												
23 3: COLON CA 3: COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	
0 0.048 0 0												
24 3: COLON CA 3: COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	
0 0.048 0 0												
25 3: COLON CA 3: COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	
0 0.048 0 0												
26 3: COLON CA 3: COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	
0 0.048 0 0												
27 3: COLON CA 4: ESOPHAGU	+	0	0	0.091	*0.727	0	0	0	0	0	0.091	0
0.091 0 0												
28 3: COLON CA 3: COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	
0 0.048 0 0												
29 3: COLON CA 3: COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	
0 0.048 0 0												
30 3: COLON CA 3: COLON CA		0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	
0 0.048 0 0												
31 4: ESOPHAGU 4: ESOPHAGU		0	0	0.091	*0.727	0	0	0	0	0	0.091	0
0.091 0 0												
32 4: ESOPHAGU 4: ESOPHAGU		0	0	0.091	*0.727	0	0	0	0	0	0.091	0
0.091 0 0												
33 4: ESOPHAGU 4: ESOPHAGU		0	0	0.091	*0.727	0	0	0	0	0	0.091	0
0.091 0 0												

34 4:ESOPHAGU 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0.048	0	0
35 4:ESOPHAGU 4:ESOPHAGU		0	0	0.091	*0.727	0	0	0	0	0	0	0.091	0	0.091	0
36 4:ESOPHAGU 4:ESOPHAGU		0	0	0.091	*0.727	0	0	0	0	0	0	0.091	0	0.091	0
37 4:ESOPHAGU 4:ESOPHAGU		0	0	0.091	*0.727	0	0	0	0	0	0	0.091	0	0.091	0
38 4:ESOPHAGU 4:ESOPHAGU		0	0	0.091	*0.727	0	0	0	0	0	0	0.091	0	0.091	0
39 4:ESOPHAGU 6:GLIOMA	+	0	0	0	0.1	0	*0.8	0	0	0	0	0.1	0	0	0
40 4:ESOPHAGU 4:ESOPHAGU		0	0	0.091	*0.727	0	0	0	0	0	0	0.091	0	0.091	0
41 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
42 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
43 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
44 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
45 5:GASTRIC 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0.048	0	0
46 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
47 5:GASTRIC 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0.048	0	0
48 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
49 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
50 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
51 6:GLIOMA 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0.048	0	0

52	6:GLIOMA	12:OVARIAN	+	0	0	0	0	0	0.125	0	0	0	0	0	0	*0.875	0
0																	
53	6:GLIOMA	6:GLIOMA		0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																	
54	6:GLIOMA	6:GLIOMA		0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																	
55	6:GLIOMA	6:GLIOMA		0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																	
56	6:GLIOMA	6:GLIOMA		0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																	
57	6:GLIOMA	6:GLIOMA		0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																	
58	6:GLIOMA	6:GLIOMA		0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																	
59	6:GLIOMA	6:GLIOMA		0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																	
60	6:GLIOMA	6:GLIOMA		0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0																	
61	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																	
62	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																	
63	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																	
64	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																	
65	7:HEPATOCE	2:BREAST C	+	0	*0.875	0	0	0	0	0.125	0	0	0	0	0	0	0
0	0																
66	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																	
67	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																	
68	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																	
69	7:HEPATOCE	7:HEPATOCE		0	0	0	0	0	*1	0	0	0	0	0	0	0	0
0																	

70 7:HEPATOCE 7:HEPATOCE	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
71 8:KIDNEY C 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
72 8:KIDNEY C 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
73 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
74 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
75 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
76 8:KIDNEY C 13:PANCREA	+	0	0	0	0	0	0	0	0.167	0	0	0	0	0
*0.833 0														
77 8:KIDNEY C 13:PANCREA	+	0	0	0	0	0	0	0	0.167	0	0	0	0	0
*0.833 0														
78 8:KIDNEY C 10:MELANOM	+	0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0
0 0														
79 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
80 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
81 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0	0
0 0.091														
82 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0	0
0 0.091														
83 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0	0
0 0.091														
84 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0	0
0 0.091														
85 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0	0
0 0.091														
86 9:LUNG SQU 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
87 9:LUNG SQU 9:LUNG SQU	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0	0	0
0 0.091														

88 9:LUNG SQU 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0.048	0	0
0															
89 9:LUNG SQU 9:LUNG SQU		0.091	0.091	0	0	0	0	0	0	*0.727	0	0	0	0	0
0		0.091													
90 9:LUNG SQU 9:LUNG SQU		0.091	0.091	0	0	0	0	0	0	*0.727	0	0	0	0	0
0		0.091													
91 10:MELANOM 10:MELANOM		0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0	0
0		0													
92 10:MELANOM 10:MELANOM		0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0	0
0		0													
93 10:MELANOM 10:MELANOM		0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0	0
0		0													
94 10:MELANOM 4:ESOPHAGU	+	0	0	0.091	*0.727	0	0	0	0	0	0.091	0	0.091	0	0
0.091		0		0											
95 10:MELANOM 10:MELANOM		0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0	0
0		0													
96 10:MELANOM 10:MELANOM		0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0	0
0		0													
97 10:MELANOM 6:GLIOMA	+	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
0															
98 10:MELANOM 10:MELANOM		0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0	0
0		0													
99 10:MELANOM 10:MELANOM		0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0	0
0		0													
100 10:MELANOM 10:MELANOM		0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0	0
0		0													
101 11:NSCLC 11:NSCLC		0	0	0	0	0	0	0	0	0	*0.889	0.111	0	0	0
0															
102 11:NSCLC 11:NSCLC		0	0	0	0	0	0	0	0	0	*0.889	0.111	0	0	0
0															
103 11:NSCLC 11:NSCLC		0	0	0	0	0	0	0	0	0	*0.889	0.111	0	0	0
0															
104 11:NSCLC 11:NSCLC		0	0	0	0	0	0	0	0	0	*0.889	0.111	0	0	0
0															
105 11:NSCLC 11:NSCLC		0	0	0	0	0	0	0	0	0	*0.889	0.111	0	0	0
0															

125 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.167	0	0	0	0
*0.833 0													
126 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.167	0	0	0	0
*0.833 0													
127 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.167	0	0	0	0
*0.833 0													
128 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.167	0	0	0	0
*0.833 0													
129 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.167	0	0	0	0
*0.833 0													
130 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.167	0	0	0	0
*0.833 0													
131 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
132 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
133 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
134 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
135 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
136 14:PROSTAT 9:LUNG SQU	+	0.091	0.091	0	0	0	0	0	0	0	*0.727	0	0
0 0.091													
137 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
138 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
139 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
140 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
141 1:BLADDER 11:NSCLC	+	0	0	0	0	0	0	0	0	0	*1	0	0
0													
142 1:BLADDER 1:BLADDER	*1	0	0	0	0	0	0	0	0	0	0	0	0
0													

143 1:BLADDER 1:BLADDER	*1	0	0	0	0	0	0	0	0	0	0	0	0	0
0														
144 1:BLADDER 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
145 1:BLADDER 1:BLADDER	*1	0	0	0	0	0	0	0	0	0	0	0	0	0
0														
146 1:BLADDER 1:BLADDER	*1	0	0	0	0	0	0	0	0	0	0	0	0	0
0														
147 1:BLADDER 1:BLADDER	*1	0	0	0	0	0	0	0	0	0	0	0	0	0
0														
148 1:BLADDER 1:BLADDER	*1	0	0	0	0	0	0	0	0	0	0	0	0	0
0														
149 1:BLADDER 1:BLADDER	*1	0	0	0	0	0	0	0	0	0	0	0	0	0
0														
150 1:BLADDER 1:BLADDER	*1	0	0	0	0	0	0	0	0	0	0	0	0	0
0														
151 2:BREAST C 1:BLADDER	+ *1	0	0	0	0	0	0	0	0	0	0	0	0	0
0														
152 2:BREAST C 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
153 2:BREAST C 9:LUNG SQU	+	0.091	0.091	0	0	0	0	0	0	*0.727	0	0	0	0
0 0.091														
154 2:BREAST C 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
155 2:BREAST C 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0			
0 0.048 0 0														
156 2:BREAST C 2:BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0	0	0
0 0														
157 2:BREAST C 2:BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0	0	0
0 0														
158 2:BREAST C 2:BREAST C	0	*0.875	0	0	0	0	0.125	0	0	0	0	0	0	0
0 0														
159 2:BREAST C 12:OVARIAN	+	0	0	0	0	0	0.125	0	0	0	0	0	0	*0.875
0 0														
160 2:BREAST C 12:OVARIAN	+	0	0	0	0	0	0.125	0	0	0	0	0	0	*0.875
0 0														

161 3:COLON CA 3:COLON CA	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0				
0 0.048 0 0														
162 3:COLON CA 3:COLON CA	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0				
0 0.048 0 0														
163 3:COLON CA 1:BLADDER	+ *1	0	0	0	0	0	0	0	0	0	0	0	0	0
0														
164 3:COLON CA 3:COLON CA	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0				
0 0.048 0 0														
165 3:COLON CA 3:COLON CA	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0				
0 0.048 0 0														
166 3:COLON CA 11:NSCLC	+ 0	0	0	0	0	0	0	0	0	0	0	*1	0	0
0														
167 3:COLON CA 7:HEPATOCE	+ 0	0	0	0	0	0	0	*1	0	0	0	0	0	0
0														
168 3:COLON CA 3:COLON CA	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0				
0 0.048 0 0														
169 3:COLON CA 3:COLON CA	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0				
0 0.048 0 0														
170 3:COLON CA 3:COLON CA	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0				
0 0.048 0 0														
171 4:ESOPHAGU 4:ESOPHAGU	0	0	0.091	*0.727	0	0	0	0	0	0	0.091	0		
0.091 0 0														
172 4:ESOPHAGU 4:ESOPHAGU	0	0	0.091	*0.727	0	0	0	0	0	0	0.091	0		
0.091 0 0														
173 4:ESOPHAGU 1:BLADDER	+ *1	0	0	0	0	0	0	0	0	0	0	0	0	0
0														
174 4:ESOPHAGU 5:GASTRIC	+ 0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0														
175 4:ESOPHAGU 5:GASTRIC	+ 0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0														
176 4:ESOPHAGU 3:COLON CA	+ 0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0				
0 0.048 0 0														
177 4:ESOPHAGU 4:ESOPHAGU	0	0	0.091	*0.727	0	0	0	0	0	0	0.091	0		
0.091 0 0														
178 4:ESOPHAGU 5:GASTRIC	+ 0	0	0	0	*1	0	0	0	0	0	0	0	0	0
0														

179 4:ESOPHAGU 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0.048	0	0
180 4:ESOPHAGU 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0.048	0	0
181 5:GASTRIC 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0.048	0	0
182 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
183 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
184 5:GASTRIC 10:MELANOM	+	0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0	0
185 5:GASTRIC 4:ESOPHAGU	+	0	0	0.091	*0.727	0	0	0	0	0	0.091	0	0.091	0	0
186 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
187 5:GASTRIC 8:KIDNEY C	+	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
188 5:GASTRIC 8:KIDNEY C	+	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
189 5:GASTRIC 8:KIDNEY C	+	0	0	0	0	0	0	0	*1	0	0	0	0	0	0
190 5:GASTRIC 5:GASTRIC		0	0	0	0	*1	0	0	0	0	0	0	0	0	0
191 6:GLIOMA 11:NSCLC	+	0	0	0	0	0	0	0	0	0	*0.889	0.111	0	0	0
192 6:GLIOMA 6:GLIOMA		0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
193 6:GLIOMA 6:GLIOMA		0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
194 6:GLIOMA 6:GLIOMA		0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
195 6:GLIOMA 6:GLIOMA		0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0	0
196 6:GLIOMA 2:BREAST C	+	0	*0.875	0	0	0	0	0.125	0	0	0	0	0	0	0

197	6:GLIOMA	6:GLIOMA	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0
0															
198	6:GLIOMA	6:GLIOMA	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0
0															
199	6:GLIOMA	6:GLIOMA	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0
0															
200	6:GLIOMA	2:BREAST C	+	0	*0.875	0	0	0	0	0.125	0	0	0	0	0
0	0														
201	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0
0															
202	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0
0															
203	7:HEPATOCE	3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0
0	0.048	0	0												
204	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0
0															
205	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0
0															
206	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0
0															
207	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0
0															
208	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0
0															
209	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0
0															
210	7:HEPATOCE	7:HEPATOCE	0	0	0	0	0	0	*1	0	0	0	0	0	0
0															
211	8:KIDNEY C	5:GASTRIC	+	0	0	0	0	*1	0	0	0	0	0	0	0
0															
212	8:KIDNEY C	13:PANCREA	+	0	0	0	0	0	0	0.167	0	0	0	0	0
0	*0.833	0													
213	8:KIDNEY C	3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0	0
0	0.048	0	0												
214	8:KIDNEY C	6:GLIOMA	+	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0
0															

215 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*1	0	0	0	0	0
0														
216 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*1	0	0	0	0	0
0														
217 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*1	0	0	0	0	0
0														
218 8:KIDNEY C 6:GLIOMA	+	0	0	0	0.1	0	*0.8	0	0	0	0.1	0	0	0
0														
219 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*1	0	0	0	0	0
0														
220 8:KIDNEY C 8:KIDNEY C	0	0	0	0	0	0	0	0	*1	0	0	0	0	0
0														
221 9:LUNG SQU 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0.048	0	0
0 0.048 0 0														
222 9:LUNG SQU 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0.048	0	0
0 0.048 0 0														
223 9:LUNG SQU 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0.048	0	0
0 0.048 0 0														
224 9:LUNG SQU 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0.048	0	0
0 0.048 0 0														
225 9:LUNG SQU 9:LUNG SQU		0.091	0.091	0	0	0	0	0	0	*0.727	0	0	0	0
0 0.091														
226 9:LUNG SQU 9:LUNG SQU		0.091	0.091	0	0	0	0	0	0	*0.727	0	0	0	0
0 0.091														
227 9:LUNG SQU 1:BLADDER	+	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
228 9:LUNG SQU 1:BLADDER	+	*1	0	0	0	0	0	0	0	0	0	0	0	0
0														
229 9:LUNG SQU 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0.048	0	0
0 0.048 0 0														
230 9:LUNG SQU 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0.095	0.095	0	0.048	0	0
0 0.048 0 0														
231 10:MELANOM 12:OVARIAN	+	0	0	0	0	0	0.125	0	0	0	0	0	0	0
*0.875 0 0														
232 10:MELANOM 10:MELANOM		0	0	0	0	0	0	0	0.111	0	*0.889	0	0	0
0 0														

233	10:MELANOM	10:MELANOM	0	0	0	0	0	0	0	0.111	0	*0.889	0	0
0	0													
234	10:MELANOM	10:MELANOM	0	0	0	0	0	0	0	0.111	0	*0.889	0	0
0	0													
235	10:MELANOM	10:MELANOM	0	0	0	0	0	0	0	0.111	0	*0.889	0	0
0	0													
236	10:MELANOM	4:ESOPHAGU	+	0	0	0.091	*0.727	0	0	0	0	0	0.091	0
0.091	0	0												
237	10:MELANOM	10:MELANOM	0	0	0	0	0	0	0	0.111	0	*0.889	0	0
0	0													
238	10:MELANOM	11:NSCLC	+	0	0	0	0	0	0	0	0	*0.889	0.111	0
0	0													
239	10:MELANOM	10:MELANOM	0	0	0	0	0	0	0	0.111	0	*0.889	0	0
0	0													
240	10:MELANOM	4:ESOPHAGU	+	0	0	0.091	*0.727	0	0	0	0	0	0.091	0
0.091	0	0												
241	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	*0.889	0.111	0
0														
242	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	*0.889	0.111	0
0														
243	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	*0.889	0.111	0
0														
244	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	*0.889	0.111	0
0														
245	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	*0.889	0.111	0
0														
246	11:NSCLC	13:PANCREA	+	0	0	0	0	0	0	0.167	0	0	0	0
*0.833	0													
247	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	*0.889	0.111	0
0														
248	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	*0.889	0.111	0
0														
249	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	*1	0	0
0														
250	11:NSCLC	11:NSCLC	0	0	0	0	0	0	0	0	0	*0.889	0.111	0
0														

251 12:OVARIAN 4:ESOPHAGU	+	0	0	0.091	*0.727	0	0	0	0	0	0	0.091	0	0.091	0	0
252 12:OVARIAN 9:LUNG SQU	+	0.091	0.091	0	0	0	0	0	0	0	0	*0.727	0	0	0	0
253 12:OVARIAN 9:LUNG SQU	+	0.091	0.091	0	0	0	0	0	0	0	0	*0.727	0	0	0	0
254 12:OVARIAN 12:OVARIAN		0	0	0	0	0	0.125	0	0	0	0	0	0	0	*0.875	0
255 12:OVARIAN 13:PANCREA	+	0	0	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833
256 12:OVARIAN 8:KIDNEY C	+	0	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
257 12:OVARIAN 8:KIDNEY C	+	0	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
258 12:OVARIAN 5:GASTRIC	+	0	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
259 12:OVARIAN 5:GASTRIC	+	0	0	0	0	*1	0	0	0	0	0	0	0	0	0	0
260 12:OVARIAN 8:KIDNEY C	+	0	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
261 13:PANCREA 11:NSCLC	+	0	0	0	0	0	0	0	0	0	0	0	0	*0.889	0.111	0
262 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833
263 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833
264 13:PANCREA 8:KIDNEY C	+	0	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
265 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833
266 13:PANCREA 8:KIDNEY C	+	0	0	0	0	0	0	0	*1	0	0	0	0	0	0	0
267 13:PANCREA 13:PANCREA		0	0	0	0	0	0	0	0.167	0	0	0	0	0	0	*0.833
268 13:PANCREA 3:COLON CA	+	0.048	0.095	*0.429	0.048	0.095	0.048	0	0	0.095	0.095	0	0	0.048	0	0

269 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.167	0	0	0	0
*0.833	0												
270 13:PANCREA 13:PANCREA	0	0	0	0	0	0	0	0	0.167	0	0	0	0
*0.833	0												
271 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
272 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
273 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
274 14:PROSTAT 11:NSCLC	+ 0	0	0	0	0	0	0	0	0	0	0	*0.889	0.111
0 0													
275 14:PROSTAT 2:BREAST C	+ 0	*0.875	0	0	0	0	0	0.125	0	0	0	0	0
0 0													
276 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
277 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
278 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
279 14:PROSTAT 14:PROSTAT	0	0	0	0	0	0	0	0	0	0	0	0	0
*1													
280 14:PROSTAT 2:BREAST C	+ 0	*0.875	0	0	0	0	0	0.125	0	0	0	0	0
0 0													

=== Evaluation on test set ===

=== Summary ===

Correctly Classified Instances	191	68.2143 %
Incorrectly Classified Instances	89	31.7857 %
Kappa statistic	0.6577	
Mean absolute error	0.0566	
Root mean squared error	0.1932	
Relative absolute error	42.634 %	
Root relative squared error	75.0182 %	
Total Number of Instances	280	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.8	0.019	0.762	0.8	0.78	0.943	BLADDER CANCER
0.5	0.019	0.667	0.5	0.571	0.859	BREAST CARCINOMA
0.8	0.112	0.356	0.8	0.492	0.86	COLON CARCINOMA
0.55	0.027	0.611	0.55	0.579	0.829	ESOPHAGUS CARCINOMA
0.6	0.023	0.667	0.6	0.632	0.829	GASTRIC CANCER
0.75	0.015	0.789	0.75	0.769	0.893	GLIOMA
0.9	0.004	0.947	0.9	0.923	0.97	HEPATOCELLULAR CARCINOMA
0.5	0.031	0.556	0.5	0.526	0.841	KIDNEY CARCINOMA
0.5	0.023	0.625	0.5	0.556	0.898	LUNG SQUAMOUS CELL CARCINOMA
0.7	0.008	0.875	0.7	0.778	0.923	MELANOMA
0.95	0.027	0.731	0.95	0.826	0.96	NSCLC
0.4	0.015	0.667	0.4	0.5	0.705	OVARIAN CARCINOMA
0.8	0.019	0.762	0.8	0.78	0.89	PANCREAS CARCINOMA
0.8	0	1	0.8	0.889	0.919	PROSTATE CARCINOMA
Weighted Avg.	0.682	0.024	0.715	0.682	0.686	0.88

=== Confusion Matrix ===

```

a b c d e f g h i j k l m n <-- classified as
16 0 2 0 0 0 0 0 1 0 1 0 0 0 | a = BLADDER CANCER
1 10 5 0 0 0 0 0 2 0 0 2 0 0 | b = BREAST CARCINOMA
1 0 16 1 0 0 1 0 0 0 1 0 0 0 | c = COLON CARCINOMA
1 0 4 11 3 1 0 0 0 0 0 0 0 0 | d = ESOPHAGUS CARCINOMA
0 0 3 1 12 0 0 3 0 1 0 0 0 0 | e = GASTRIC CANCER
0 2 1 0 0 15 0 0 0 0 1 1 0 0 | f = GLIOMA
0 1 1 0 0 0 18 0 0 0 0 0 0 0 | g = HEPATOCELLULAR CARCINOMA
0 0 3 0 1 2 0 10 0 1 0 0 3 0 | h = KIDNEY CARCINOMA
2 0 8 0 0 0 0 0 10 0 0 0 0 0 | i = LUNG SQUAMOUS CELL CARCINOMA
0 0 0 3 0 1 0 0 0 14 1 1 0 0 | j = MELANOMA
0 0 0 0 0 0 0 0 0 0 19 0 1 0 | k = NSCLC
0 0 1 2 2 0 0 3 2 0 1 8 1 0 | l = OVARIAN CARCINOMA
0 0 1 0 0 0 0 2 0 0 1 0 16 0 | m = PANCREAS CARCINOMA
0 2 0 0 0 0 0 0 1 0 1 0 0 16 | n = PROSTATE CARCINOMA

```