Journal of Biomedical Informatics 44 (2011) 615–620

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Predicting microRNA modulation in human prostate cancer using a simple String IDentifier (SID1.0)

Maria C. Albertini [a,*], Fabiola Olivieri [b], Raffaella Lazzarini [b], Francesca Pilolli [c], Francesco Galli [c], Giorgio Spada [d], Augusto Accorsi [a], Maria R. Rippo [b], Antonio D. Procopio [b]

[a] Dipartimento di Scienze Biomolecolari, Sezione di Biochimica e Biologia molecolare, Università degli Studi di Urbino "Carlo Bo", Urbino, Italy
[b] Dipartimento di Patologia Molecolare e Terapie Innovative, Università Politecnica delle Marche, Ancona, Italy
[c] Dipartimento di Medicina Interna, Università degli Studi di Perugia, Italy
[d] Dipartimento di Scienze di Base e Fondamenti, Università degli Studi di Urbino "Carlo Bo", Urbino, Italy

ARTICLE INFO

Article history:
Received 6 August 2010
Available online 18 February 2011

Keywords:
Computational methods
Prostate cancer (CaP)
SID1.0
PicTar
DIANA-MicroT 3.0
microRNA prediction
Exhaustive search algorithm
Fortran language

ABSTRACT

To make faster and efficient the identification of mRNA targets common to more than one miRNA, and to identify new miRNAs modulated in specific pathways, a computer program identified as SID1.0 (simple String IDentifier) was developed and successfully applied in the identification of deregulated miRNAs in prostate cancer cells. This computationally inexpensive Fortran program is based on the strategy of exhaustive search and specifically designed to screen shared data (target genes, miRNAs and pathways) available from PicTar and DIANA-MicroT 3.0 databases. As far as we know this is the first software designed to filter data retrieved from available miRNA databases. SID1.0 takes advantage of the standard Fortran intrinsic functions for manipulating text strings and requires ASCII input files. In order to demonstrate SID1.0 applicability, some miRNAs expected from the literature to associate with cancerogenesis (miR-125b, miR-148a and miR-141), were randomly identified as main entries for SID1.0 to explore matching sequences of mRNA targets and also to explore KEGG pathways for the presence of ID codes of targeted genes. Besides genes and pathways already described in the literature, SID1.0 has proven to useful for predicting other genes involved in prostate carcinoma. These latter were used to identify new deregulated miRNAs: miR-141, miR-148a, miR-19a and miR-19b. Prediction data were preliminary confirmed by expression analysis of the identified miRNAs in androgen-dependent (LNCaP) and independent (PC3) prostate carcinoma cell lines and in normal prostatic epithelial cells (PrEC).
© 2011 Elsevier Inc. All rights reserved.

1. Introduction

MicroRNAs (miRNAs) are small non-coding RNAs that take part in post-transcriptional regulation either by arresting the translation or by cleavage (degradation) of mRNA targets [1]. MicroRNA regulation is performed through pairing of the miRNA to sites in the messenger RNA of protein coding genes. MicroRNAs have been thought to be involved in many biological processes (i.e. cell proliferation, death and differentiation) and are believed to regulate the expression of approximately one-third of all human genes [2]. Mature miRNAs bind to their target mRNAs by complete or incomplete complementation of their 5′-end nucleotides 1–8 (seed sequences) with a binding site in the 3′- or 5′-untranslated regions (UTRs) of target transcripts or in the coding sequences [3]. This process results in direct cleavage of the targeted mRNAs or inhibition of translation. Currently, nearly 800 human miRNAs have been identified. Target-prediction computer programs often predict that a miRNA could target tens to hundreds of genes, and that a gene can be targeted by many miRNAs. Thus, the expectation that miRNAs play important roles in coordinating many cellular processes, particularly those involved in development and disease, and even in various cancers [4,5]. Prediction servers such as TargetScan [6], miRanda [7], RNAhybrid [8], PicTar [9] and the recent DIANA-MicroT 3.0 [10] provide information about the miRNA-target interactions. This latter, in particular, has been designed to cover a quick and easy interpretation of the involvement of a series of genes in biological processes. DIANA-MicroT 3.0 actually measures the enrichment of genes involved in pathways within a given list [10].

It has already been demonstrated that miRNAs are aberrantly expressed in prostate cancer (CaP) [11]. In androgen-dependent (AD) and androgen-independent (AI) CaP cell lines differential expression of miR-148a was found [11]. MiR-125b expression was reported to be increased in benign prostate lines compared with CaP lines and androgen receptor positive (AR-positive) cell lines expressed more miR-125b than androgen receptor negative (AR-negative) cell lines [5]. It has been recently reported that even

* Corresponding author. Fax: +39 (0) 722 305260.
E-mail address: maria.albertini@uniurb.it (M.C. Albertini).

1532-0464/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2011.02.006

miR-148a expression levels are increased in AR-positive cell lines (PrEC and LNCaP) in comparison to AR-negative cell lines (PC3 and DU145) [11]. Furthermore, in humans, a member of the miRNA-200 family (miR-141) has been reported to distinguish patients with prostate cancer from healthy controls [12]. Down-regulation of miRNA-200 family members has been identified to be involved in tumor metastasis [13]. These studies showed that miR-125b, miR-148a and miR-141 may have a key role in prostate tumorigenesis.

We developed a simple String IDentifier (SID1.0) to identify common mRNA targets of miR-125b, miR-148a and miR-141 and the associated pathways. In order to identify common strings, SID1.0 performs an exhaustive search within each individual input file and reports the result (i.e. the number of common targets) on an ASCII output file. More sophisticated search strategies have not been explored to date, since the size of the input files (hence the number of IDs to be scrutinized) is, in the experiments performed so far, limited to a few hundreds or thousands. This makes the whole procedure computationally inexpensive. SID1.0 has been developed in a Mac OSX environment and compiled with the Intel Fortran compiler. As far as we know this is the first software designed to filter data retrieved from available databases (e.g.: TargetScan, miRanda, RNAhybrid, PicTar, DIANA-MicroT 3.0), allowing us for retrieving new information that is not directly available from them. The analysis performed by SID1.0, i.e. the identification of common strings, is not directly available through the user interface of the databases above. It is difficult to evaluate the strengths or the limitations of alternative computational approaches because similar tools, possibly developed by other investigators, have not been described or made available to date. Moreover, despite its simplicity, this technique can allow to predict new miRNAs targeting genes modulated in these pathways. To validate the reliability of predictions made by SID1.0, we ascertained that SID1.0 predicted miRNAs were modulated in CaP (LNCaP androgen-dependent and PC3 androgen-independent) cells that were compared with normal prostatic epithelial (PrEC) cells.

2. Materials and methods

2.1. Data acquisition

We obtained from PicTar database (see http://pictar.mdc-berlin.de/) the predicted target genes of the miRNAs we were interested in. The targets of a miRNA are indicated with a specific gene ID system (Refseq ID). For each miRNA a dataset (i.e., a group list of Refseq IDs) of the predicted targeted genes was created [9]. Since a visual inspection of the IDs would be unpractical due to their large number (even thousands IDs), they have been automatically indexed using a simple program written in Fortran (SID1.0; String IDentifier, see http://www.fis.uniurb.it/spada/SID_minipage.html) that looks for Refseq IDs shared by the predicted target genes of the different datasets [14]. In Fig. 1 we show the experimental workflow. SID1.0 is in fact based on an algorithm of sequential exhaustive search that has been implemented in Fortran 90 using very elementary methods. SID1.0 performs an exhaustive search within each individual one-column ASCII input file and reports the result (i.e. the number of common targets) on an ASCII output file in the form of a table that summarizes the common IDs. Thus, the main advantage of SID1.0, which works as a filter on the information provided by the web pages hosting the miRNA databases, is that it is completely independent from the algorithms on which the databases rely upon. In this way, our procedure builds upon the prediction algorithms used in the databases, whose outputs are scrutinized by SID1.0. SID1.0 has been developed and tested in an Mac OS X environment and is currently compiled using the Intel Fortran 90 compiler. A version of SID1.0 based on the free, open source g95 compiler is under development.

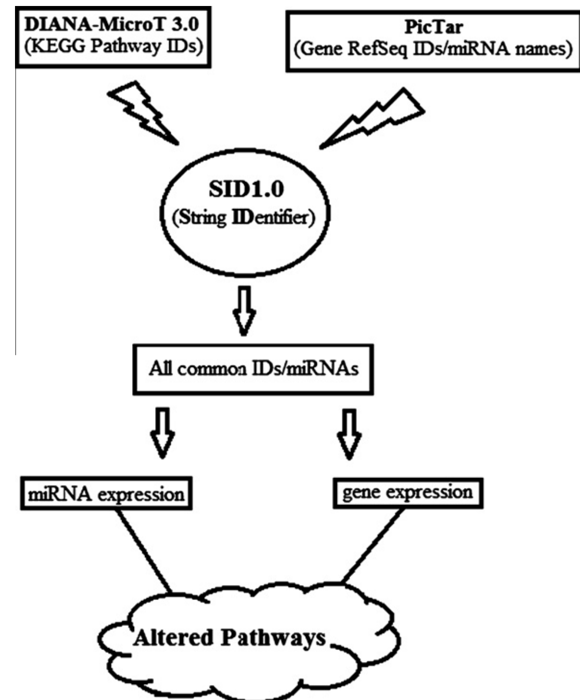

Fig. 1. Workflow illustrating the main steps of the informatics and the experimental procedure. Data from any other available databases can be used in addition to those indicated here.

Each gene in the group list has the related information in NCBI's Entrez Nucleotide database. It is possible to perform a reverse search by obtaining from PicTar database the miRNAs predicted to target a gene. For each gene a dataset of the miRNAs predicted to target a gene was created. The names of the miRNAs were indexed using SID1.0, which looks for miRNA names shared by the predicted targeting miRNAs of the different datasets. Furthermore, for a defined miRNA name, target genes can be automatically retrieved from the DIANA-microT 3.0 database. A list of gene names or a list of RefSeq IDs are provided and the program translates them into Ensembl IDs. The list of genes is compared to the Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathways Database [10] and IDs are indexed using SID1.0 that looks for KEGGs pathway IDs shared by the predicted target genes of the different datasets. In this way, we were able to obtain the common target genes of specific miRNAs, the common targeting miRNAs of specific genes and the common pathways of specific miRNAs (see Fig. 1). PicTar database is trained to identify both binding sites targeted by a single miRNA and those that are co-regulate by several miRNAs in a coordinated manner.

2.2. Cell lines and culture

Prostate epithelial cells (PrEC) were kindly provided by Dr. Antonella Farsetti from the Department of Experimental Oncology, Molecular Oncogenesis Laboratory, Experimental Research Center, Regina Elena (Regina Elena Cancer Institute of Rome, Italy) [15]. The cells were cultured according to the manufacturer's protocols (Cambrex BioScience, Verviers, Belgium). Human prostate cancer cell lines (LNCaP and PC3) were cultured in RPMI 1640 supplemented with 10% fetal bovine serum (Invitrogen), 4.5 g/L glucose, and 0.1 mol/L HEPES, penicillin (100 units/ml) and streptomycin (100 μg/ml).

2.3. Expression levels of predicted MicroRNAs

Extraction of miRNAs from cell culture was performed from $1x10^6$ cells per samples by Total RNA Purification Kit according to manufacturer's protocol (NORGEN). Predicted miRNA expression was carried out by using TaqMan MicroRNA Assays (Applied Biosystem). Expression data were normalized according to the *Ct* of a miRNA that preliminary experiments showed not to be modulated by CaP pathogenesis, such as U6 snRNA [16]. Relative amounts of miRNA were calculated by the comparative threshold cycle (*Ct*) method as: $2^{-\Delta Ct}$, where $\Delta Ct = Ct_{\mathrm{miRNA}} - Ct_{\mathrm{U6}}$.

2.4. Statistical analysis

Statistical analyses were performed using Student's *t*-test for unpaired data and *P*-values < 0.05 were considered significant. Data are presented as mean ± SEM.

3. Results

3.1. MicroRNAs and prostate cancer (CaP)

It has already been demonstrated that miRNAs are aberrantly expressed in prostate cancer (CaP) [11]. Considering that a given miRNA may have up to 100 mRNA targets, the issue is how to decide which CaP-related targets of the miRNAs may be biologically active and important. To date computational algorithms have played a central role in the discovery of the majority of miRNAs known, as well as in prediction of their targets (and associated pathways).

Available data in literature suggest that miR-125b, miR-148a and miR-141 may be related to prostatic tumorigenesis, AI growth (AR may even regulate their expression) and tumor metastasis [4,13]. We used these three miRNAs to extract their potential targets in the searchable database of published miRNA sequences provided by the miRNA registry. Using SID1.0, we found the common nucleic acid codes (RefSeq ID datasets) shared by miR-125b, miR-148a and miR-141 targets. This approach identified the following seven common target genes: ATXN1 (ataxin variant 1); RNF44 (ring finger protein 44); AKAP1 (A kinase anchor protein 1); TRPS1 (trichorhinophalangeal syndrome 1); E2F3 (E2F transcription factor 3); CPEB4 (cytoplasmic polyadenylation element binding protein 4) and ESRRG (estrogen-related receptor gamma transcript variant 1, 2 and 3) (Table 1).

Between them, E2F3 has been already demonstrated to play a crucial role in the control of cell cycle, in the action of tumor suppressor proteins and it has been even shown to increase with the down-regulation of miR-125b in human prostate cancer tissues [17]. It has also been demonstrated that ESRRG may perform an antiproliferative or tumor-suppressing function in prostate cancer cell lines [18]. It has been observed that TRPS1 is possibly involved in oxidative stress and apoptosis in androgen-independent DU145 prostate cancer cells [19], while the remaining predicted target genes have still not been considered in prostate cancer. By these results, we can deduce that SID1.0 distinguished some target genes (E2F3, ESRRG and TRPS1) previously demonstrated to be involved in prostate tumorigenesis, and that the remaining ones are potentially involved.

3.2. Common pathways potentially altered by the expression of miR-125b, miR-148a and miR-141

We used the web-based computational database DIANA-MicroT 3.0 to extract the KEGG pathway IDs potentially altered by the expression of the following miRNAs: miR-125b, miR-148a and miR-141 (obtaining 3 ID datasets). By using SID1.0 we found the common KEGG pathway IDs (obtained from the 3 KEGG datasets of each miRNA) shared by miR-125b, miR-148a and miR-141. The output of DIANA-MicroT 3.0 is a table that illustrates the number and name of the genes involved in a determined pathway (i.e. O-Glycan biosynthesis) and compares it with the expected number of genes, depending on how many genes compose the specific pathway. As shown in Table 2, the name of the common pathways indentified by the correspondent KEGG Pathway IDs were only 5: O-Glycan biosynthesis; mTOR (mammalian Target of Rapamycin) signaling pathway; glycan structures-biosynthesis 1; Axon guidance and Neuroactive ligand-receptor interaction.

The O-Glycan biosynthesis and glycan structures-biosynthesis 1 pathways are responsible of O-glycosylated glycoproteins biosynthesis which have antigenic properties. mTOR is a key protein kinase controlling signal transduction from various growth factors and upstream proteins to the level of mRNA translation and ribosome biogenesis, with pivotal regulatory effects on cell cycle progression, cellular proliferation and growth, autophagy and angiogenesis. Axon guidance pathway represents a key stage in the formation of neuronal network. Neuroactive ligand-receptor interaction is important for neuron function. To evaluate the involvement of these pathways in prostate carcinogenesis we found support from the literature, where two of these pathways (O-Glycan biosynthesis and mTOR) have already been described to be modulated (see Section 4).

3.3. Prediction of modulated miRNAs in CaP cell lines

We used the seven common target genes identified in the first analysis (Table 1) to evaluate common miRNAs across them. The number of possible combinations obtained by SID1.0 across the 7 dataset of targets were 41 in total, representing the common genes across the 7, 4, 3 and 2 datasets (the missing combinations indicate that any common miRNAs between 6 and 5 datasets were found). As shown in Table 3, only miR-141 was common across all the 7 target datasets.

Table 1
Common targets (RefSeq ID) of hsa-miR-125b (531 elements), hsa-miR-141 (422 elements) and hsa-miR-148a (434 elements). We found nine common genes on a total of 95 possible combinations (targets common either to 2 or all of the 3 miRNAs). We reported only the common target RefSeq IDs (in the first column) to all the miRNAs and their corresponding annotation (second column). The database used for this analysis was PicTar.

Common RefSeq ID	Annotation
NM_000332	Homo sapiens ataxin 1 (ATXN1)
NM_014901	Homo sapiens ring finger protein 44 (RNF44)
NM_139275	Homo sapiens A kinase (PRKA) anchor protein 1 (AKAP1), nuclear gene encoding mitochondrial protein, transcript variant 2
NM_014112	Homo sapiens trichorhinophalangeal syndrome I (TRPS1)
NM_001949	Homo sapiens E2F transcription factor 3 (E2F3)
NM_030627	Homo sapiens cytoplasmic polyadenylation element binding protein 4 (CPEB4)
NM_001438	Homo sapiens estrogen-related receptor gamma (ESRRG), transcript variant 1
NM_206594	Homo sapiens estrogen-related receptor gamma (ESRRG), transcript variant 2
NM_206595	Homo sapiens estrogen-related receptor gamma (ESRRG), transcript variant 3

Table 2
Common pathways (KEGG Pathway IDs) of hsa-miR-125b (92 elements), hsa-miR-141 (111 elements) and hsa-miR-148a (5 elements). We found five common pathways (to the 3 miRNAs) on a total of 74 total possible combinations (common pathways to either to 2 or all of the 3 miRNAs). In the table are reported only the pathways common to the 3 miRNAs. We indicated the KEGG Pathway name (first column), the gene names involved in each pathway (second column) and the identification ID of the pathway used by the KEGG database (third column). The database used for this analysis was DIANA-MicroT 3.0.

KEGG pathway name	Gene name	KEGG pathway ID
O-Glycan biosynthesis	OGT	hsa00512
mTOR signaling pathway	TSC1, RHEB	hsa04150
Glycan structures – biosynthesis 1	ST3GAL3, OGT	hsa01030
Axon guidance	SEMA6A, CXCL12, EPHA7, NRP1, EPHA2, CDC42	hsa04360
Neuroactive ligand-receptor interaction	GRIN2D, GRIN3A, ADRB1, THRB, CALCR, DRD2	hsa04080

Table 3
Common miRNAs predicted to target the genes: ATXN1 (five elements), RNF44 (27 elements), AKAP1 (25 elements), TRPS1 (50 elements), E2F3 (49 elements), CPEB4 (eight elements) and ESRRG (six elements which were the same in the 3 variants). We reported only the miRNAs common (column 1) to 7, 4 and 3 different datasets (the others are not shown) with the corresponding RefSeq ID (second column) and the annotation (third column). We found 1 common miRNAs (hsa-miR-141) on a total of 41 possible combinations (common miRNAs to 7, 4, 3 and 2 targets). The database used for this analysis was PicTar.

Common miRNAs	RefSeq ID	Annotation
hsa-miR-141	NM_000332	ATXN1
	NM_014901	RNF44
	NM_139275	AKAP1
	NM_014112	TRPS1
	NM_030627	CPEB4
	NM_001438	ESRRG
	NM_001949	E2F3
hsa-miR-200a	NM_000332	ATXN1
	NM_014901	RNF44
	NM_001438	ESRRG
	NM_001949	E2F3
hsa-miR-301, hsa-miR-152, hsa-miR-148a, hsa-miR-148b, hsa-miR-125a, hsa-miR-125b	NM_001949	E2F3
	NM_014901	RNF44
	NM_139275	AKAP1
	NM_014112	TRPS1
hsa-miR-130a, hsa-miR-130b	NM_014901	RNF44
	NM_139275	AKAP1
	NM_014112	TRPS1
	NM_001949	E2F3
hsa-miR-96	NM_000332	ATXN1
	NM_139275	AKAP1
	NM_001438	ESRRG
hsa-miR-23a	NM_000332	ATXN1
	NM_030627	CPEB4
	NM_001438	ESRRG
hsa-miR-19a, hsa-miR-19b	NM_014901	RNF44
	NM_139275	AKAP1
	NM_014112	TRPS1

These data suggest that these miRNAs could be potentially modulated in CaP. Then, we decided to evaluate the expression of some of these miRNAs (hsa-miR-19a; hsa-miR-19b; hsa-miR-148a; hsa-miR-141) in CaP cell lines: LNCaP androgen-dependent, PC3 androgen-independent and normal prostatic epithelial (PrEC) cells. We remark that it is not possible to rank the miRNAs obtained in the analysis of Table 3, since the biological relevance cannot be estimated by our informatic procedure. Each prediction requires an experimental validation. As shown in Fig. 2, all the miRNAs quantified were modulated (up-regulated and/or down-regulated) in the CaP cell lines tested. Some of them, hsa-miR-141 and hsa-miR-148a, are down-regulated (Fig. 2A and B): hsa-miR-141 is down-expressed in LNCaP and PC3, while hsa-miR-148a is down-expressed only in LNCaP compared to PrEC cells. Conversely, hsa-miR-19a and hsa-miR-19b are up-regulated (Fig. 2C and D): both of them are highly expressed in PC3 but down-expressed in LNCaP compared to PrEC cells.

We can assume that all the miRNAs analyzed are differentially expressed in LNCaP androgen-dependent and PC3 androgen-independent compared to normal prostatic epithelial (PrEC) cells. Surprisingly, when we extracted the miRNAs targeting the genes involved in each pathway, (see Table 4), we found most of the miRNAs predicted in the previous analysis (Table 3).

4. Discussion

Animals encode miRNAs in their genomes which attack protein coding genes altering protein expression by cleavage or translational repression. In human malignancies, including prostate cancer, miRNAs have shown a capability of distinguishing the different tissue developmental and differentiation states. In this study, using miRNAs and related target predictions documented in PicTar and DIANA-MicroT 3.0 databases, we identified the miRNAs and target genes that potentially cause differences in prostate cancer. Taking advantage of the above prediction servers we obtained the information for the miRNA-target interactions desired. By our bioinformatic strategy based on SID1.0, a computationally inexpensive Fortran 90 program, we identified shared data (target genes, miRNA and pathways). SID1.0 performs an exhaustive search based on the manipulation of characters strings in Fortran. Given its versatility, SID1.0 can manage different IDs associated with gene expression, miRNA and pathways (even in large number and from any available databases). SID1.0 software has been designed to filter data retrieved from available databases, thus allowing us for retrieving new information that is not directly available from those databases. Our proposed approach has successfully identified some deregulated miRNA in LNCaP androgen-dependent and PC3 androgen-independent compared to normal prostatic epithelial (PrEC) cells.

By identifying from literature modulated miRNAs in cancerogenesis and in particular in prostate cancer, we extracted the genes modulated by miR-125b, miR-148a and miR-141 in PicTar database. The predicted genes were the following: ATXN1 (ataxin variant 1); RNF44 (ring finger protein 44); AKAP1 (A kinase anchor protein 1); TRPS1 (trichorhinophalangeal syndrome 1); E2F3 (E2F transcription factor 3); CPEB4 (cytoplasmic polyadenylation element binding protein 4) and ESRRG (estrogen-related receptor gamma transcript variant 1, 2 and 3) (Table 1).

From the literature it has already been shown that E2F3, ESRRG and TRPS1 are modulated in prostate cancer: E2F3 controls cell cycle, acts on tumor suppressor proteins and increases with the down-regulation of miR-125b in human prostate cancer tissues [17]; ESRRG functions as antiproliferative or tumor-suppressor in prostate cancer cell lines [18]; TRPS1 is involved in oxidative stress, and possibly in apoptosis in androgen-independent DU145 prostate cancer cells [19]. The remaining target genes have not been described in prostate cancerogenesis but seem to be involved in other pathogenesis mechanisms, ATXN1 expression can be modulated by miR-19a affecting neuropathology [20].

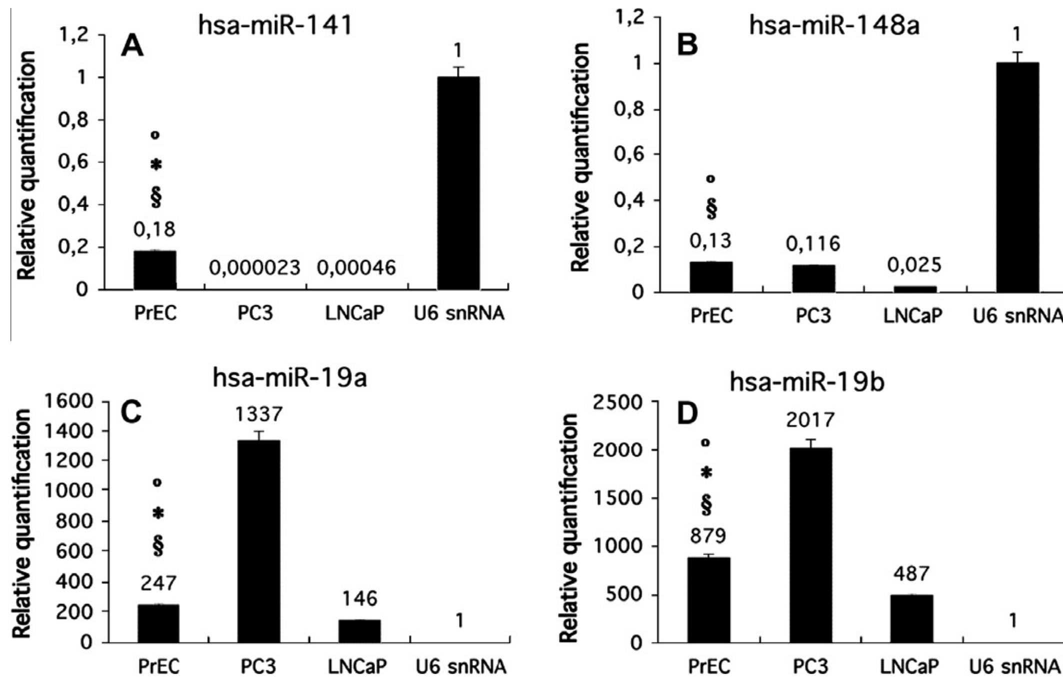

Fig. 2. Relative quantification of down-regulated (A and B) and up-regulated (C–D) miRNAs in CaP cell lines (LNCaP androgen-dependent, PC3 androgen-independent and PrEC normal prostatic epithelial cells). Relative quantification was normalized by U6 snRNA endogenous control. Data are presented as mean ± SEM and *P*-values < 0.05 were considered significant: °PrEC vs. U6 snRNA; *PrEC vs. PC3; §PrEC vs. LNCaP.

We have also used the DIANA-MicroT 3.0 database to extract the pathways altered by miR-125b, miR-148a and miR-141. SID1.0 was able to predict the common pathways shared by miR-125b, miR-148a and miR-141: O-Glycan biosynthesis (and glycan structures-biosynthesis 1); mTOR signaling pathway; Axon guidance and Neuroactive ligand-receptor interaction (Table 2). These pathways are known to be involved in antigen expression, in cell proliferation, growth, autophagy, angiogenesis and in the formation of neuronal network [21,22]. The O-Glycan biosynthesis and mTOR pathways have already been described to be involved in prostate cancer. Actually, the functions of carbohydrates present in normal cells decrease in prostate cancer cells [23]. Furthermore, it has been demonstrated that the mTOR pathway is dominant over AR signaling in prostate cancer cells [24] and its inhibition can lead to suppressed proliferation of aggressive prostate cancer cell lines in vitro [25]. SID1.0 was able to correctly predict the involvement of these two pathways and suggests the possible role of the other pathways in prostate cancerogenesis.

Next, the genes predicted in our previous analysis (Table 1) were used in the PicTar database to extract the targeting miRNAs. In this way we were able to predict the modulation of different miRNAs (Table 3). Between them, we randomly decided to test hsa-miR-19a; hsa-miR-19b; hsa-miR-148a and hsa-miR-141 in LNCaP androgen-dependent and PC3 androgen-independent to be compared to normal prostatic epithelial (PrEC) cells. As shown in the results section, all these miRNAs were modulated in our CaP cells (Fig. 1) indicating not only a correct prediction but also differentially expressed miRNAs in androgen-dependent and -independent compared to normal prostatic cells. Furthermore, as shown in Table 4, extracting the miRNAs targeting the genes involved in each pathway we found most of the miRNAs predicted in the previous analysis (Table 3). We did not identify any other genes, but another analysis could be performed with the genes targeting the miRNAs above. This will be the subject of future investigations.

It is worth to note that our data collectively may predict the involvement of miRNAs in prostate cancerogenesis. Serum levels of hsa-miR-141 have been reported to discriminate patients with prostate cancer from healthy controls [12]. In renal cell carcinoma and colorectal cancer cells, it has been observed a down-regulation of hsa-miR-141 [26,27] as we observed in our CaP cell lines. As demonstrated by other authors [11], we observed a down-regula-

Table 4
MiRNAs targeting the genes belonging to the previously identified common pathways. The genes without our predicted miRNAs are omitted. We indicated the KEGG Pathway name (first column), the gene names involved in each pathway (second column) and the miRNAs predicted in the previous analysis (third column). The database used for this analysis was PicTar.

KEGG pathway name	Gene name	Predicted miRNAs
O-Glycan biosynthesis	OGT	hsa-miR-141, hsa-miR-301, hsa-miR-200a, hsa-miR-130a, hsa-miR-130b, hsa-miR-125a, hsa-miR-125b.
mTOR signaling pathway	TSC1	hsa-miR-141, hsa-miR-200a.
Glycan structures – biosynthesis 1	OGT	hsa-miR-141.
Axon guidance	SEMA6A, CXCL12, EPHA2	hsa-miR-141, hsa-miR-200a.
	EPHA7	hsa-miR-141, hsa-miR-200a,
	NRP1	hsa-miR-23a
	CDC42	hsa-miR-141, hsa-miR-200a, hsa-miR-148a, hsa-miR-148b, hsa-miR-130a, hsa-miR-130b. hsa-miR-200a.
Neuroactive ligand-receptor interaction	GRIN3A	hsa-miR-141, hsa-miR-200a.
	CALCR	hsa-miR-141, hsa-miR-200a

tion of hsa-miR-148a in PC3 cells compared to PrEC, but we did not find higher levels in LNCaP (on the contrary, its expression was even more decreased compared to PC3 and PrEC). However, the functional roles and target genes of miR-148a in prostate cancer have not yet been documented. The over-expression of hsa-miR-19a in PC3 androgen-independent cells may decrease cytotoxic ATXN1 expression while its down-expression in LNCaP androgen-dependent cells may increase the cytotoxic effect of ATXN1 as observed in neurodegenerative pathology [20]. Furthermore, hsa-miR-19a showed increased expression upon glioma progression [28] which could be similarly investigated in the progression from androgen-dependent to -independent CaP cells. In line with our results on PC3 androgen-independent cells, over-expression of hsa-miR-19b has been demonstrated to be involved in carcinogenesis [29]. Interestingly, our results demonstrated that this form is down-expressed in LNCaP androgen-dependent cells compared to PrEC (and PC3 androgen-independent) cells.

5. Conclusions

In view of the relatively large number of data involved in the analyses described in this study, we conclude that even a simple program of string identification as SID1.0 has proven to be a very useful system to predict miRNAs involved in tumor progression thereby representing a potential tool in diagnostics and clinical monitoring of cancer. From a general stand point SID1.0 is simply an implementation of a classical and well known search strategy in computer science. Its merit relies upon the biological application that is described in this paper.

Comparison of human cancers with the corresponding normal tissues revealed that some miRNAs are consistently deregulated in cancer, which may also suggest a direct role of miRNAs in tumourigenesis. Beside to deregulated miRNAs, because of its broad applicability, the proposed prediction tool can help to identify gene targets and biological pathways relevant to cancer development and treatment.

Acknowledgments

We thank Dr. Antonella Farsetti and Dr. Annalisa Grasselli, from the Department of Experimental Oncology, Molecular Oncogenesis Laboratory, Experimental Research Center, Regina Elena (Regina Elena Cancer Institute of Rome, Italy) for providing the PrEC cell line. We are grateful to Dr. Laura Teodori, from the BAS-BIOTEC-MED, Italian Agency for New Technology, Energy and the Environment (Rome, Italy) for the kind support.

References

[1] Yan X, Chao T, Tu K, Zhang Y, Xie L, Gong Y, et al. Improving the prediction of human microRNA target genes by using ensemble algorithm. FEBS Lett 2007;581:1587–93.
[2] Inui M, Martello G, Piccolo S. MicroRNA control of signal transduction. Nat Rev Mol Cell Biol 2010;11:252–63.
[3] Duursma AM, Kedde M, Schrier M, le Sage C, Agami R. miR-148 targets human DNMT3b protein coding region. RNA 2008;14:872–7.
[4] Aguda BD, Kim Y, Piper-Hunter MG, Friedman A, Marsh CB. MicroRNA regulation of a cancer network: consequences of the feedback loops involving miR-17-92, E2F, and Myc. Proc Natl Acad USA 2008;105:19678–83.
[5] DeVere White RW, Vinall RL, Tepper CG, Shi XB. MicroRNAs and their potential for translation in prostate cancer. Urol Oncol 2009;27:307–11.
[6] Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 2005;120:15–20.
[7] John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human microRNA targets. PLoS Biol 2004;2:e363.
[8] Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R. Fast and effective prediction of microRNA/target duplexes. RNA 2004;10:1507–17.
[9] Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, et al. Combinatorial microRNA target predictions. Nat Genet 2005;37:495–500.
[10] Papadopoulos GL, Alexiou P, Maragkakis M, Reczko M, Hatzigeorgiou AG. DIANA-mirPath: integrating human and mouse microRNAs in pathways. Bioinformatics 2009;25:1991–3.
[11] Fujita Y, Kojima K, Ohhashi R, Hamada N, Nozawa Y, Kitamoto A, et al. MiR-148a attenuates paclitaxel-resistance of hormone-refractory, drug-resistant prostate cancer PC3 cells by regulating MSK1 expression. J Biol Chem 2010;285:19076–84.
[12] Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, et al. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci USA 2008;105:10513–8.
[13] Baffa R, Fassan M, Volinia S, O'Hara B, Liu CG, Palazzo JP, et al. MicroRNA expression profiling of human metastatic cancers identifies cancer gene targets. J Pathol 2009;219:214–21.
[14] Chapman SJ. Fortran 95/2003 for scientists and engineers. 3rd ed. ISBN. McGraw-Hill; 2007 (978-0-07-319157-7).
[15] Nanni S, Priolo C, Grasselli A, D'Eletto M, Merola R, Moretti F, et al. Epithelial-restricted gene profile of primary cultures from human prostate tumors: a molecular approach to predict clinical behavior of prostate cancer. Mol Cancer Res 2006;4:79–92.
[16] Tsuchiya Y, Nakajima M, Takagi S, Taniya T, Yokoi T. MicroRNA regulates the expression of human cytochrome P450 1B1. Cancer Res. 2006;66:9090–8.
[17] Ozen M, Creighton CJ, Ozdemir M, Ittmann M. Widespread deregulation of microRNA expression in human prostate cancer. Oncogene 2008;27:1788–93.
[18] Yu S, Wang X, Ng CF, Chen S, Chan FL. ERRgamma suppresses cell proliferation and tumor growth of androgen-sensitive and androgen-insensitive prostate cancer cells and its implication as a therapeutic target for prostate cancer. Cancer Res. 2007;67:4904–14.
[19] Chang GTG, Gamble SC, Jhamai M, Wait R, Bevan CL, Brinkmann AO. Proteomic analysis of proteins regulated by TRPS1 transcription factor in DU145 prostate cancer cells. Biochim Biophys Acta 2007;1774:575–82.
[20] Lee Y, Samaco RC, Gatchel JR, Thaller C, Orr HT, Zoghbi HY, et al. MiR-19, miR-101, and miR-130 co-regulate ATXN1 levels to potentially modulate SCA1 pathogenesis. Nat Neurosci 2008;11:1137–9.
[21] Ciuffreda L, Di Sanza C, Incani UC, Milella M. The mTOR pathway: a new target in cancer therapy. Curr Cancer Drug Targets 2010;10:484–95.
[22] Rzeszowska-Wolny J, Herok R, Widel M, Hancock R. X-irradiation and bystander effects induce similar changes of transcript profiles in most functional pathways in human melanoma cells. DNA Repair 2009;8:732–8.
[23] Lee SH, Hatakeyama S, Yu SY, Bao X, Ohyama C, Khoo KH, et al. Core3 O-glycan synthase suppresses tumor formation and metastasis of prostate carcinoma PC3 and LNCaP cells through down-regulation of alpha2beta1 integrin complex. J Biol Chem 2009;284:17157–69.
[24] Kaarbø M, Mikkelsen OL, Malerød L, Qu S, Lobert VH, Akgul G, et al. PI3K-AKT-mTOR pathway is dominant over androgen receptor signaling in prostate cancer cells. Cell Oncol 2010;32:11–27.
[25] Kobayashi T, Shimizu Y, Terada N, Yamasaki T, Nakamura E, Toda Y, et al. Regulation of androgen receptor transactivity and mTOR-S6 kinase pathway by Rheb in prostate cancer cell proliferation. Prostate 2010;70:866–74.
[26] Nakada C, Matsuura K, Tsukamoto Y, Tanigawa M, Yoshimoto T, Narimatsu T, et al. Genome-wide microRNA expression profiling in renal cell carcinoma: significant down-regulation of miR-141 and miR-200c. J Pathol 2008;216:418–27.
[27] Minghua M, Xia MG, Chen X, Lin Z, Xu Y, Ma Y, et al. MicroRNA-141 regulates smad interacting protein 1 (SIP1) and inhibits migration and invasion of colorectal cancer cells. Dig Dis Sci 2009:2365–72.
[28] Malzkorn B, Wolter M, Liesenberg F, Grzendowski M, Stühler K, Meyer HE, et al. Identification and functional characterization of microRNAs involved in the malignant progression of gliomas. Brain Pathol 2009;20:539–50.
[29] Castellano L, Giamasa G, Jacoba J, Coombesa RC, Lucchesib W, Thiruchelvama P, Bartonc G, Jiaod LR, Waite R, Waxmana J, Hannonf GJ, Stebbinga J. The estrogen receptor-alpha-induced microRNA signature regulates itself and its transcriptional response. Proc Natl Acad 2009;106:15732–37.