MicroRNA Regulation by Epigenetic Mechanisms

Inaugural-Dissertation

zur

Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Rui Pedro Lousa das Neves

Coimbra (Portugal)

Dezember 2008

Aus dem Institut für Transplatationsdiagnostik und Zelltherapeutika
Der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

Referent:	PD. Dr. Markus Uhrberg
Koreferent:	Prof. Dr. Rüdiger Simon
Tag der mündlichen Prüfung:	02-12-2008

Zusammenfassung

MiRNAs haben eine Schlüsselstellung in der Molekular- und Zellbiologie eingenommen, da sie in mehreren wichtigen biologischen Prozessen, so die embryonale Entwicklung, die Zelldifferenzierung, die Apoptose und in malignen Erkrankungen involviert sind. Nachdem die biologische Funktion der miRNAs erkannt wurde, stieg das Interesse an den Regulationsmechanismen der miRNA Expression. Die vorliegende Studie konzentriert sich auf die Rolle der DNA-Methylierung in der Regulation der miRNA Expression. In einem ersten Schritt wurde die Expression von 234 miRNAs in natürlichen Killerzellen (NK) Zellen untersucht. Auf der Basis einer globalen miRNA Expressionsanalysetechnik konnten 16 miRNAs identifiziert werden, die nach Kultivierung der NK-Zellen in Anwesenheit von 5-AZA-CdR, ein DNA demethylierendes Agens, verstärkt exprimiert wurden. Hierbei wurde gezeigt, dass die Transkription der Schlüsselfaktoren der miRNA Prozessierungsmaschinerie von der DNA Demethylierung unbeeinträchtigt bleibt. Desweiteren erwies sich die Expressionszunahme derjenigen miRNAs, die intragenisch lokalisiert sind, unabhängig von den sie umgebenden Transkriptionseinheiten. Zusammengenommen liefern diese Ergebnisse Evidenz dafür, dass die DNA Methylierung eine direkte Auswirkung auf die miRNA Expression haben kann. In der Tat konnte für zwei miRNAs (miRNA-200c und 141) eine eindeutige Korrelation zwischen dem Methylierungsstatus des experimentell identifizierten miRNA Promotors und der miRNA Expression in verschiedenen Zelltypen gezeigt werden.

Die Bedeutung dieser differentiellen DNA-Methylierung im Promotorbereich der genannten miRNAs wurde evaluiert. DNA Methylierung führte zur Inaktivierung des Promotors *in-vitro*. Zusätzlich wurde gezeigt, dass in Brustkrebs-Zelllinien dieser Promotor durch DNA-Methylierung stillgelegt ist.

Die Beobachtung, dass die DNA Methylierung einen entscheidenden Einfluss auf die miRNA Expression haben kann, liefert eine Erklärung für die differentielle Expression von überlappenden miRNA Genen, wie sie von uns und anderen Arbeitsgruppen beobachtet worden ist.

Erst vor kurzem wurde berichtet, dass die miRNA200c und die miRNA142 eine grundlegende Rolle im Metastasierungsprozess des Mammakarzinoms spielen. Die vorliegende Arbeit legt nahe, dass hierbei der DNA Methylierungsstatus ausschlaggebend sein könnte.

Summary

miRNAs have became a major focus of molecular and cellular biology due to their involvement in several critical processes that cover embryonic development, cell differentiation, apoptosis and several malignancies including cancer. At the same time as their biologic function is being unravelled, there has been also an increasing interest on the mechanisms regulating their expression. The present study focused on the potential role of DNA methylation as epigenetic regulator of miRNA gene expression. In a first step, the expression of 234 miRNA was analyzed in natural killer (NK) cells, which served as a model system for expression of miRNAs in the haematopoietic compartment. Based on global miRNA expression analysis, 16 miRNAs were identified as being up-regulated after culturing the cells in the presence of 5-AZA-dcR, a DNA-demethylating drug with clinical applicability on cancer treatment. Epigenetic regulation of miRNA expression levels was not based on expression changes of enzymes involved in the miRNA processing machinery as demonstrated by mRNA expression analysis. Furthermore, the expression of miRNAs that were located intragenically, was frequently found to be regulated independently from the surrounding transcripts. Together, experimental evidence pointed to a direct effect of DNA methylation on the regulation of miRNAs loci. Indeed, for two clustered miRNAs (miRNA-200c and 141), a correlation was found between the methylation pattern of the putative promoter region, which was identified by reporter gene analysis, and their corresponding expression levels in different cell models. The importance of the differentially methylated area was assessed and evidence provided showing silencing of the identified promoter region upon DNA methylation in breast cancer cells. Additionally, based on our findings demonstrating the existence of several overlapping transcriptional units (by 5`RACE experiments), the present work provides a logical explanation for differential expression of these two miRNAs in certain cell types. The work gains additional clinical relevance through the recent observation that miR200c and miR141 play crucial roles in the epithelial-to-mesenchymal transition process that leads to metastasis formation in several tumour cell types such as breast cancer. The experimental data of the present thesis implicate DNA methylation of the miR200c/141 locus as one possible cause for those events.

Index

1 - Introduction

1.1 - Epigenetics

In the PubMed registry the oldest publication on epigenetics dates from 1964 but the concept of *epigenetics* seems to have emerged two decades earlier. This term is attributed to Conrad Waddington in 1942[1] and, literally meaning "outside conventional genetics", it had been used to name all the events that, at that time, could not be explained by the current knowledge in genetics or by the Mendelian form of inheritance. According to the original definition, epigenetics would involve the study of "the causal interactions between genes and their products, which brings phenotype into being". This broad concept accommodated several apparently unrelated events, in which phenotypic differences were not associated with genetic differences, and genetic variations did not lead to phenotypic variations. As time passed and knowledge increased, these phenomena appeared to be related which justified the classification as *epigenetic*.

More recent definitions of epigenetics can be ascribed to several authors[2-5]. As definitions, they show slight differences, but they all converge in the idea that epigenetics should be considered as something stably heritable and able to change the expression of a gene, its final outcome, or a cellular phenotype, without interfering with its DNA sequence.

A good example illustrating the role of epigenetics is the cascade of events occurring during development and cell differentiation. In such events, where many different cell types originate from the same one cell, DNA sequence is always constant, and so it cannot justify the different phenotypes that cells present. The cell fate and the complex sequence of events that take place, seems to be defined by the way the cell, in a given moment, interprets the DNA-coded information and the way the cell decides which information to read.

The last decades have contributed with pieces of evidence supporting the idea that the reasons leading to different interpretation of the same genome are postsynthetic modifications of DNA or posttranscriptional modifications of DNA-associated proteins. These modifications seem to be recognized by proteins that facilitate the appropriate downstream biological effects.

As there is no change in the genetic information, the processes leading to development and cellular differentiation can be seen as a consequence of epigenetic events and the modifications on the genome can be accepted as epigenetic elements.

1.1.1 - Histone modifications

Chromatin, as initially described, was considered as an octamer of the four core histones (H3, H4, H2A, H2B) with a passive packaging function, stabilizing and organizing the DNA[6]. With time, different types of chromatin were discovered and their different role in transcription was also recognized. The differences between the chromatin forms result

from at least 8 types of distinct covalent and non-covalent modifications on histones (acetylation, methylation, phosphorylation, ubiquitylation, sumoylation, ADP ribosylation, deimination and proline isomerization)[7], occurring mainly on the unstructured N-terminal tails of the histones that contrasts with the globular structure of the rest of the protein that accommodates the DNA.

Acetylation of lysine 9 and methylation on lysine 4 of Histone 3 are well characterized examples of modifications contributing to a relaxed chromatin structure that confer transcriptional activity. Those modifications also promote the binding of transcriptional factors and other regulatory proteins, ensuring the correct transcriptional and splicing processes. The relaxed spatial organization that allows the transcriptional activity is characteristic of euchromatin which DNA content consists mainly of coding sequences representing only less than 4% of the genome in mammals.
In contrast, there are also histone modifications like the methylation of the lysine 9 and lysine 27 of Histone 3 that are able to remodel the chromatin, creating local compact structures and environments more prone to remain silent and that constitute the heterochromatin.

The local physical alterations leading to a more or less compact DNA structure, and driven by histone modifications can be due to two main causes: one is called "cis-effect" and are alterations on local electrostatic charges. For example the addition of a positive charge caused by acetylation and, in contrast, the addition of a negative charge caused by phosphorylation. Cis-effect comprehends also modulations of histone tail structure as verified upon ubiquitination. The other effect to be considered leading to the local physical alterations of DNA structure is called "trans-effect", and it is caused the binding of proteins to some modifications. These modification-binding partners can facilitate downstream chromatin-modulating events.

The chromatin structure of a certain genomic region is not static, altering between relaxed or compact states. Histone modifications contribute to that. They are placed, removed and maintained by a large group of enzymes, whose number is continuously increasing. Among them are: histone acetyltransferases[8] (HAT) that acetylate specific lysine residues of histones; histone deacetylases[9] (HDAC) that are able to remove the acetylation marks; histone methytransferases[10]; histone kinases[11] that phosphorylate specific serine or threonine residues; and phosphatases (PPTases) that remove the same phosphate marks. Many of these enzymes catalyze the respective reactions with remarkable specificity in terms of target residues and act depending on the cellular context (e.g. dependent on external or internal stimuli).

There are hundreds of histone modifications described so far and their role in gene expression regulation is being revealed in more detail at a fast speed. In what extent some

modifications affect the chromatin structure and coordinate to regulate gene expression, it is not completely understood. The knowledge gained from the characterized modifications, and the big number of modifications yet to be characterized, represent an enormous potential for regulating cell functional responses.

1.1.2 - DNA methylation

Chromatin structure affecting gene expression can also be modulated by modifications on the DNA itself. DNA methylation is the best characterized modification affecting chromatin structure. Initially described in 1950[12], only 25 years later, in 1975, it was associated with the stable maintenance of a particular gene expression pattern during mitotic cell division[13,14]. The current knowledge on this phenomenon is steadily growing and its biological significance has been demonstrated by several studies [15,16].

Chemically, DNA methylation consists on the covalent addition of a methyl group to the 5th position carbon (m^5C) of the cytosine carbon ring. *In vivo*, this reaction is catalyzed by DNA methyltransferases (DMNT1, DNMT3a and DNMT3b) and the methyl group donor is s-adenosylmethionine (SAM). Methylation seems to occur in a symmetric manner, as the two DNA strands are equally methylated[17].

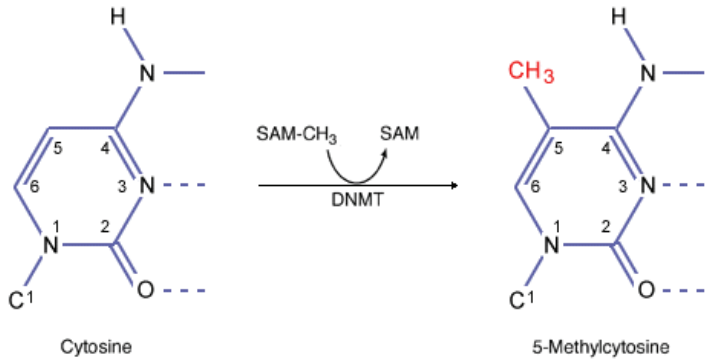


Fig.1 - Methylation of cytosine residues by DNMTs

Methylation of cytosines occurs in all eukaryotes except yeast. The extent of methylation as well as the sequence motif in which it occurs is variable among organisms. The worm *C. elegans* is almost devoid of m^5C [18] while the insect *D. melanogaster* presents very low levels of m^5C mostly on CpT dinucleotides (with less extend also in CpA and CpG)[19,20]. Plants show a high level of methylation occurring mostly on CpApG and CpTpG motifs but also in CpG with less extent[21]. Methylation on plants can also occur on adenine residues (N6-methyladenine, m^6A) of their mitochondrial DNA[22]. Bacteria, in its turn, show three types of methylation: m^5C, m^6A and also N4-methylcytosine (m^4C) on several sequence motifs[23-25].

Human somatic cells present a relatively high methylation level and the addition of a methyl group to cytosine residues involves almost exclusively the cytosines of CpG dinucleotides as for all other mammals[26]. These CpG dinucleotides are not randomly distributed on the human genome, with some regions lacking any and some regions presenting an extremely high CG relative density. These last regions are denominated as CpG islands and, by convention, are 0,4-3 Kb in length, have a G+C content higher than 55% and are enriched in the CpG dinucleotides relative to the rest of genome (the observed/expected ratio is >0,6). Computational analysis predicts the existence of ca. ~29,000 of CpG islands in the human genome[27,28], preferentially located at the 5' ends of many genes[29]. Estimations point that between 60% and 80% of the human genes are associated with CpG islands[30-32].
Although the methylation status can vary depending on the differentiation stage and the cell type, the majority of CpG islands stay unmethylated in all stages of development and in all tissues[33].

1.1.3 - The role of DNA methylation

In bacterial DNA, methylation has an important protective function. It is included in a restriction-modification system that allows the bacteria to distinguish its own DNA (methylated) from foreign DNA (unmethylated) and, subsequently, selectively digest this last by the action of restriction enzymes sensitive to methylation[34].
DNA methylation in mammals has been associated with other functions[35]. It has been involved in some crucial processes in mammalian development such as genomic imprinting and X chromosome inactivation. The influence of DNA methylation on gene expression regulation is now widely accepted as a powerful mechanism to regulate gene activity either at a local level through the effects on single gene promoters or at a more global level through mechanisms influencing many genes and even entire chromosomes.

Methylation of CG dinucleotides, located in the regulatory region of a certain transcript, is able to directly repress its expression[36]. The extent of this effect depends on the position and density of m^5CpG relative to the promoter[37-39]. Two mechanisms may lead to gene silencing by promoter methylation: Many transcription factors are known to be able to bind CG-containing motifs where the presence of a methylated cytosine can interfere with that binding. This is not applicable to all genes and transcription factors, as some do not contain or are not able to bind to any CpG-containing motifs[40,41]. But some examples, as is the case for CTCF protein, support this mechanism. CTCF is a protein able to insulate the promoter of the *Igf2* gene in mice from the influence of remote enhancers. The CTCF DNA-binding-motif contains CpG dinucleotides and the binding of CTCF relies on a demethylated status of the area[42]; Considering an alternative mechanism, methyl groups can, on the other hand, promote the binding of proteins to DNA. This is the case of the

m^5CpG-binding proteins. As m^5CpG binding proteins known so far all act as transcriptional repressors, the final outcome of this in general is gene silencing. Examples of this are the MeCP2 and its relatives, the m^5CpG-containing-motifs-binding-domain proteins MBD1-MBD4 as well as the unrelated protein Kaiso. These enzymes were shown to target specifically m^5CpG with no other apparent sequence specificity[43-45]. MBD2 and MeCP2 are thought to mediate transcriptional repression via interactions with co-repressor complexes containing HDAC. This suggests a coordinated action between epigenetic elements and introduces a new level of complexity.
Evidence has been provided showing that the influence of DNA methylation on gene silencing is also dependent on the action of other, yet uncharacterized, players. Methylation would appear as a secondary event contributing to the stability of inactivation as seen for the phosphoglycerate gene on the inactive X chromosome that is already silent when its promoter is methylated[46,47].

There are several examples showing the impact of DNA methylation on gene expression regulation, with physiologic impact for the cell and for the organism. An example of how DNA methylation controls gene expression at a cell-type level is given by the Maspin gene that shows a clear demethylated promoter in gene-expressing epithelial cells and a highly methylated one in cells where the gene is silent[48]. DNA methylation has also been involved in gene expression regulation at clonotypical level. The killer cell immunoglobulin-like receptors (KIR) genes coding for membrane receptors, involved in key steps of immunologic responses, are characteristic of NK cells or closely related NK-T cells[49]. Although within a same individual all cells have the same genetic background, these receptors are differentially expressed in the NK cell population (clonotypical expression)[50] as a consequence of differential methylation of the KIR loci[51,52].
Similarly, the MAGE genes expressed only on germ and tumour cells where low levels of genomic methylation are observed, are silent in somatic cells where the promoter is methylated[53,54]. This suggests also a putative role of MAGE gene product in cell differentiation and tumorigenesis.
Although DNA methylation can contribute to keep genes silent, it can also indirectly contribute to maintain other genes or even entire chromosomes transcriptionally active. That is the case for the Xist gene which initiates the widespread methylation process culminating in the inactivation of one X chromosome. The active X chromosome so, has to stay protected from inactivation by Xist action and that is ensured by keeping the Xist gene silent by gene locus methylation[55].

1.1.4 - DNA methylation during development

Several experiments suggest the involvement of DNA methylation in development and cell differentiation processes. DNA methylation pattern of a cell, changes in a highly

orchestrated way during development involving genome-wide as well as gene specific demethylation and de novo methylation.

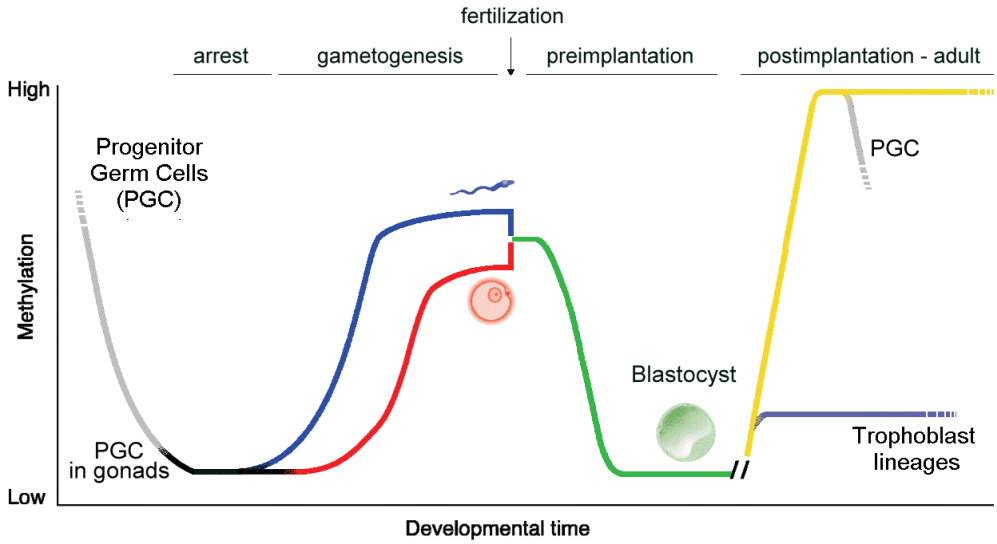


Fig. 2 - Methylation reprogramming during development (adapted from ref.[56])

Genome-wide DNA methylation in primordial germ cells is low when they enter the gonads and start the mitotic (male) and meiotic (female) arrest. It increases during the proliferative gametogenesis stages. During spermatogenesis, the degree of methylation achieved is higher than in oogenesis and that can contribute to the differential genomic imprinting observed in the parental genomes of the fertilized oocyte. Demethylation starts on both genomes after fertilization reaching a very low genomic methylation grade during the pre-implantation stages[57,58]. Demethylation of paternal genome seems to occur by a fast and active process while the maternal genome seems to be passively demethylated during subsequent cell divisions.

The biological significance of this methylation cycle is not completely understood. Experimental evidence show that spontaneous deamination can affect both cytosine or 5-methylcytosine but, while in the first case it generates a uracil residue that can be recognized and excised from DNA without further consequences, on the last case, the deamination of 5-methylcytosine leads to the formation of a thymine residue that can not be excised constituting so a mutation by trasition[59]. So, the demethylation observed in the germ cells seems to have some benefits of reducing the mutation rate protecting against that transition. There are also lines of evidence showing that the functions of centromeres are affected when they are demethylated[60] and so, especially during the gametogenesis, as it is the phase in which cell proliferation and DNA replication happens, a higher degree of methylation can contribute to a lower frequency of structural abnormalities in chromosomes[61]. Indeed, genome methylation is a pre-requisite for the massive phenotypical and functional changes occurring during the gametogenesis, and especially spermatogenesis.

The methylation occurrence is dependent also on the individual and environmental factors, and the following erasure of the parental methylated pattern, can constitute an opportunity for resetting the imprints removing any acquired epigenetic marks[62-65]. Although this demethylation observed for paternal genome is the best example suggesting the existence of an active demethylating process, clear evidence for the enzymes responsible for this process *in vivo* do not exist so far[66-68].
In the early embryo, just after implantation, *de novo* methylation sets a new pattern[69-73]. The main factors responsible for the process are DNMT3 enzymes[74,75] which act mainly on embryonic stem cells and early embryos and have a crucial importance for mammalian development[76]. Although the mechanism of pattern establishment is still not understood and the biological function of this methylation event is still under debate, DNA methylation contributes to the compartmentalization of the genome and to the formation of gene expression patterns that lead to cell differentiation.

DNMT1, another enzyme from DNMT family, has no/low de novo methylation capacity, but it rather recognizes hemi-methylated DNA and serves for the maintenance of methylation patterns during late stages of embryo development and adult life. DNMT1 is also responsible for the transmission of those same patterns in dividing cells, acting just after DNA replication take place. The role of DNMT1 in the embryo is crucial since deletion of DNMT1 results in global demethylation and embryonic lethality[77-79]. According to the semi-conservative mode of DNA inheritance, the DNA molecules upon replication contain one parental strand and one newly synthesized strand. The parental strand does not loose its methylation pattern and the m^5C are recognized by DNMT1 enzyme that copies those patterns into the complementary strand, thereby ensuring the transmission of methylation patterns during cell division.

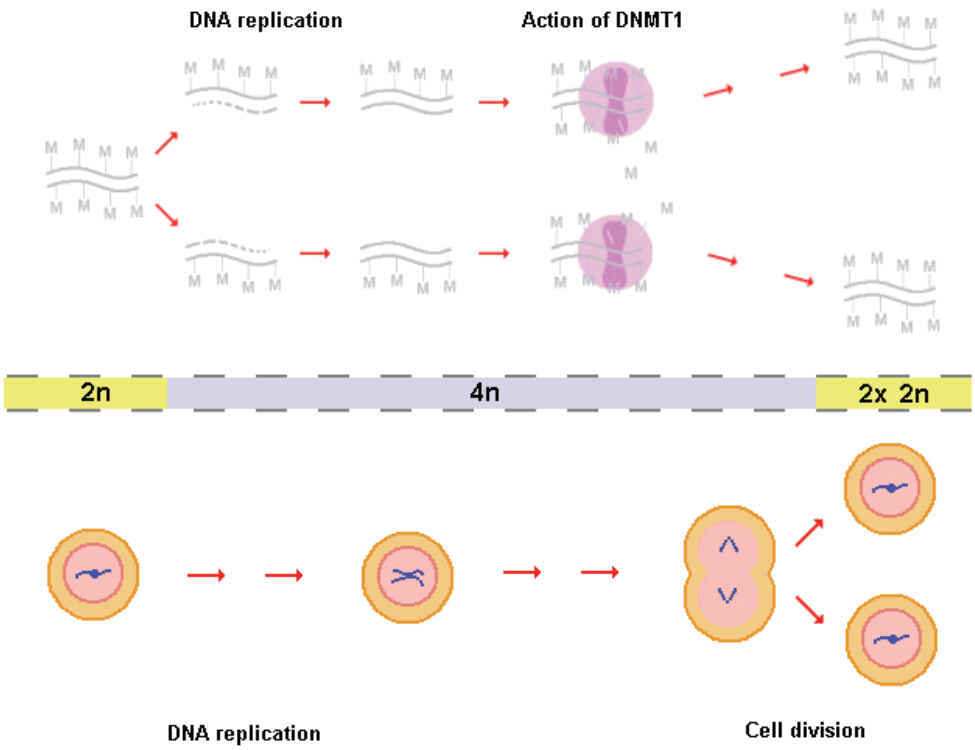


Fig. 3 – Action of DNMT1 during mitosis. DNMT1 acts on newly synthesized DNA after DNA replication and before cell division.

Although methylation maintenance is highly efficient, rare errors might occur during this process in both directions of a loss or gain of sporadic punctual methylation[80]. However, these errors occur at low frequency and as they cause only punctual changes they have limited, if any, consequences on the methylation status of the region.

1.1.5 - DNA methylation in cancer and in its therapy

Some epigenetic marks, like DNA methylation, are not conserved through the life time of the individuals. They can be reversible and temporary and, therefore, they can affect chromatin structure transiently in response to internal or external stimuli with the subsequent consequences on the gene outcome. Increasing evidence has been supporting the idea that phenotypic changes observed on cancer cells resulting from chromosomal instability, activation of oncogenes, silencing of tumour suppressor genes as well as inactivation of DNA repair systems are associated with epigenetic abnormalities including DNA methylation[81,82]. In cancer genome-wide hypomethylation is generally observed while, at the same time, DNA hypermethylation of promoter regions of selective genes, especially tumour suppressor genes, occurs. These epigenetic changes have been associated with development and maintenance of malignancies[83-88], for examples leukemias[89,90].
Hypermethylation can also be the cause of genetic changes characteristic of cancer. The dinucleotides m^5CpG are prompt to suffer point mutations and so, in hypermethylated

regions, the probability of mutation occurrence is higher[91]. Also methylation silencing of MLH1 mismatch repair genes can cause genome-wide microsatellite instability[92]. Hypomethylation, on the other hand, is also associated with a higher rate of mutations caused by chromosomal rearrangements and loss of heterozygosity[93,94].

These observations have increased the interest of the scientific community in the anti-tumour role that epigenetic drugs may have. Decitabine (SuperGen, Ireland) is the most widely used DNA demethylating drug. Decitabine is the commercial patented name for 5-aza-2'deoxycytidine (5-AZA-CdR) a deoxycytidine analogue that incorporate with DNA forming irreversible covalent bonds with DNMT at cytosine sites targeted for methylation[95]. These bonds inactivate DNMTs and when DNA replication takes place without active DNMTs it leads to a global genome-wide hypomethylation in subsequent cell generations. This effect is mainly observed when low doses of 5-AZA-CdR are used. At high dosage, the effect of the interaction of 5-AZA-CdR with DNA gets predominant causing programmed cell dead by apoptosis due to DNA synthesis arrest[96-98].

First clinical trial regarding the usage of 5-AZA-CdR started in early 80ths. Results of phase 1 and phase 2 clinical studies helped to define some clinical-biological parameters of the compound and showed encouraging results obtained mainly for haematological tumours. Acute myeloid leukaemia (AML), myeloid dysplastic syndrome (MDS) and chronic myeloid leukaemia (CML) patients have been included in those studies. Several drug regimens have been tested so far and studies are still in progress. Again, a dosage effect is observed and a higher number of clinical responses are observed for milder dosages[99]. Authors speculate that higher doses may suppress clinical responses by abrogation of an immune-mediated effect[100,101]. Effects of the drug are gradual but long lasting even after treatment cessation.
Multiple mechanisms may be involved in achievement of complete remissions, including apoptosis or senescence following gene expression normalization, immune-mediated effects related to differentiation of leukaemic cells into dendritic cells or changes in the cell-surface markers of malignant cells, or even a direct cytotoxic effect on leukaemic cells. Although is likely the involvement of DNA methylation on some of these processes, it is still not known in which of them is in fact involved.

1.2 - miRNAs

The so called "junk DNA fraction", comprising approximately 90% of the mammalian genome, has been receiving increasing attention in the last two decades. Although initially thought to have no functional relevance, new approaches have been revealing important biological roles for non coding DNA. Within this genomic portion, many novel non-protein-coding RNAs have been identified with relevant biologic functions in the cell.

Among the novel species of RNA molecules discovered are the micro RNAs (miRNAs). Due to its size and genomic characteristics, these small (~22nt) RNA species stayed hidden from the conventional genetic approaches for long time.

The first miRNA was described in 1993[102]. It was characterized as the final product of lin-4 gene known to interfere with the timing of developmental events in *C. elegans*[103,104]. These publications innovate also because, the authors were already able to provide the first hints on the way how miRNAs could regulate its target genes. These observations are still valid nowadays.
From these first findings until today where a large number of publications appears every week, there was a time gap. Eventually the second miRNA (let-7) was characterized only in 2000 in the same model organism, showing some common features with the lin-4 namely its potential mode of action and relevance for development[105,106]. The year 2000 also represents the great breakthrough for miRNAs when homologous were identified in thirteen other species of animals including humans[107]. The evidence pointing to a role of miRNAs in differentiation and development, plus the idea of a generalized presence among several organisms, opened the door for a world wide interest in these small RNAs.

In 2007 miRNAs were also found in much simpler organisms like unicellular algae[108,109] contradicting a previous idea that only multicellular organisms would have miRNAs and that its acquirement could have been an evolutionary step in order to a more complex organism design.
At the present time, miRNAs are thought to be present in all animals, plants and alga [110]. They are evolutionarily conserved, being common to many species. There are 6396 miRNAs identified so far, 678 of them encoded on the human genome which is thought to encode a total of about 1000 miRNAs, approximately 3% of the total number of genes[111-113].

The existing knowledge about their biogenesis has been increasing including main aspects of the maturation processes leading to the formation of effectors miRNA species. In contrast, the processes governing miRNAs action, including selection of target genes, are yet poorly understood. It is currently hypothesised that miRNAs are regulating 30% of the human genome-encoded genes[114,115].

1.2.1 - Biogenesis and localization of miRNAs

miRNAs are either encoded isolated on the genome or grouped in clusters. miRNAs are transcribed as part of long transcripts, the primary micro RNA (pri-miRNA). These pri-miRNAs are Pol II transcripts[116], and like messenger RNAs (mRNAs), they contain a 5' 7-

methyl guanosine cap as well as a 3' poly-A tail, they can be several Kb long and they can undergo splicing and even show alternative splicing forms[117-119].

Initially pri-miRNAs were thought to constitute specific transcriptional units coding only for miRNAs[120-123]. Later, better informatics analysis tools allowed an enhanced characterization of the genome and it was realized that pri-miRNAs transcriptional units can also be responsible for the synthesis of other transcripts. Indeed, a vast majority of miRNAs are now considered to be located in intronic regions of protein coding transcripts[124]. From the rest, some are located in introns or exons of transcripts with unknown function[125] while others remain without a defined transcriptional unit leaving room for the idea of special and dedicated transcriptional units.

In the cell nucleus, the newly transcribed pri-miRNA acquires a specific secondary structure characterized by loops and stems. miRNAs are contained in an arm of imperfect hairpin structures formed by approximately 80nt that can be recognized in the nucleus by the so called microprocessor complex, formed (in the mammals) by DiGeorge critical region 8 protein (DGCR8) and Drosha (also known as RN3)[126-129], one of the three RNase type III expressed in human cells. RNase III enzymes cleave double stranded RNA introducing a 2nt 3' overhang at the cleavage site. The action of Drosha on the hairpin structure containing the miRNA leads to the liberation of a shorter hairpin of approximately 70nt containing the 3' 2nt overhang, which is called precursor miRNA (pre-miRNA).

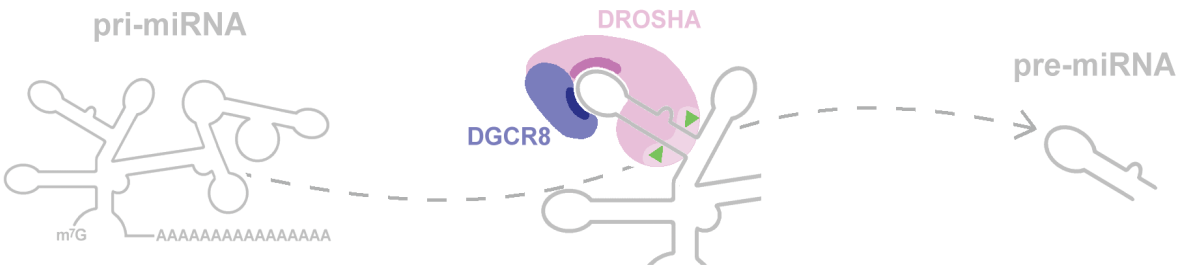


Fig. 4 - Schematic representation of a pri-miRNA secondary structure and the action of Drosha cropping an hairpin with production of a pre-miRNA. Drosha, represented in pink associates with DGCR8 (in blue) and the dsRNA binding domains of both (in darker pink and blue colours respectively) recognize the terminal loop of the stem and position the RNase domain (in lighter pink) to cleave the two RNA strands. Cleavage sites are represented by the green triangles. After cleavage, pre-miRNA is liberated.

The cropping by Drosha is governed by the structure and conformation of the microprocessor complex. The dsRNA binding domain of Drosha recognize the terminal loop of the hairpin while the two RNase domains cleave the RNA at a fixed distance from that terminal loop. The DGCR8 proteins present in the same complex also contain a dsRNA binding domain that helps the right positioning of the hairpin inside the complex. The recognition and accommodation of the pri-miRNA in the microprocessor complex is

sequence independent but structure dependent. The sequences of the several miRNAs hairpins are different and consequently its structure is also slightly different varying in the length as well as in the number of unpaired nucleotides. But despite that, the imperfect hairpin structure is a common feature in every pri-miRNA known so far. Disruption of either the stem part, inside or outside the miRNA, or the terminal loop by deletion mutations can block Drosha processing while changes on the sequences do not have any effect[130,131].

The formed pre-miRNA is then exported to the cytoplasm by Exportin-5 (Exp5), a GTPase specific nucleo-cytoplasmic transporter dependent on Ran[132,133]. On the nucleus, Exp5 forms a heterodimer with Ran-GTP that recognizes the pre-miRNA. Once again, the binding is dependent on the RNA structure and not on its nucleotide sequence. This heterodimer recognizes small RNAs with a stem of more than 16nt and a short 3' overhang[134-136]. When the complex passes the nuclear pore and reaches the cytoplasm, the Ran-GTP is hydrolysed to Ran-GDP and consequently the pre-miRNA is released.
Some experimental evidence suggest that Exp5 might have also a miRNA stabilizing role besides the transporting function. Inhibition of Exp5 expression results in the expected loss of the cytoplasmic miRNA fraction but does not result in an accumulation of its nuclear precursor [137].

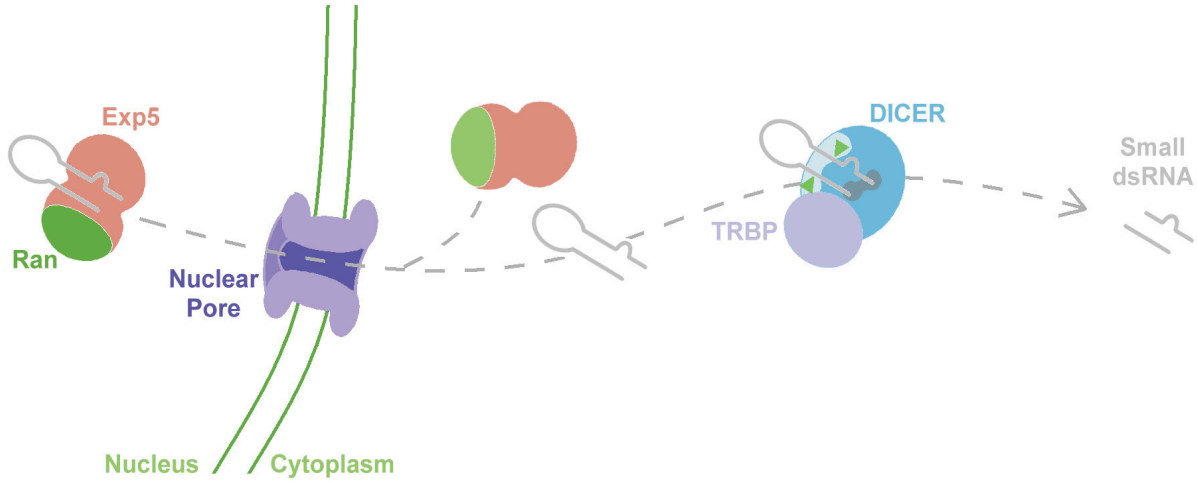


Fig. 5 - Schematic representation of pre-miRNA exporting mediated by Exp5. Once in the cytoplasm, free pre-miRNA is recognized by a heterodimer of Dicer and TRBP. The PAZ domain of dicer (dark blue on the fig) recognizes the 3' end overhang of the pre-miRNA and, together with TRBP, place the hairpin in a position that the RNase domains (lighter blue on the fig) can cleave it generating a small double stranded RNA of approximately 22nt long, with a 3' end overhang in both terminuses.

Once in the cytoplasm, the pre-miRNA is recognized by Dicer, the second RNase III intervening in the maturation process[138-141]. Dicer also acts in a heterodimer complex together with TAR RNA binding protein (TRBP). Unlike the nuclear Drosha that plays a role in the processing of human rRNAs[142], the cytoplasmic RNase III Dicer seems to participate exclusively in the maturation of miRNAs.
Dicer has high affinity for the 2nt 3' overhang of the pre-miRNAs conferred by a PAZ domain[143,144]. The dsRNA binding domains from Dicer and the TRBP protein, position the

hairpin in a way that the cleavage by the two RNase III domains from Dicer occurs approximately 22nt from the existing overhang[145]. This new cleavage also creates a 2nt 3' overhang and so, the resulting 22nt long dsRNA contains one in each end.

The two strands of the small dsRNA are then separated by a helicase and often, one of them is incorporated into the miRNA induced silencing complex (miRISC) while the other is likely to suffer rapid degradation. The mechanisms behind the choice of which strand will act as a miRNA are not completely understood but some thermodynamic factors seem to play a role. Generally, the strand with its 5' end located at the less stable end of the small duplex is the chosen one[146,147]. Occasionally, both arms of the pre-miRNA can originate two mature miRNA forms as is the case for miRNA-22[148].
For a miRNA to be effective on repressing translational processes, it has to be assembled into a miRISC complex. This assembly seems to be dynamic and coupled to the Dicer processing, but its details are not completely understood.
In mammals, the complex is mainly composed of proteins from the Argonaute family (AGO1 to AGO4) but also includes other proteins like the fragile x mental retardation protein (FMRP) and some processing-bodies (P-bodies) components such as GW182 and RCK/p54, that are essential for the translation modulation and induction of repression respectively[149-151].

Fig. 6 - Schematic representation of final steps of the miRNA maturation. The two strands of the small dsRNA are separated by a helicase and one of them, the mature miRNA, is incorporated into RISC complex, while the other, the miRNA*, is most often degraded.

Where the assembly of miRISC complex and the repression of target mRNAs occur, is still not known. Experimental evidence colocalize mature miRNA with P-bodies[152], a structure rich in enzymes involved in mRNA decay that potentially might also play a role in miRNA mediated mRNA degradation. The involvement of P-bodies is supported by the colocalization of repressed mRNAs and small amounts of the four AGO proteins also in the same structures[153-157]. But the same elements (miRNAs, target mRNA and AGO proteins) also colocalize with other cellular structures such as stress granules[158] so the results are not yet conclusive.

1.2.2 - miRNA's mode of action

Interestingly, the first clues about the miRNA mode of action were published even before the description of the first miRNA[159]. A paper from 1991 already referred to the 3' UTR as an element able to facilitate the down-regulation of lin-14 mRNA, another gene involved in time control during *C. elegans* development [160]. When miRNA lin-4 was described in 1993, authors realised immediately that it shows complementarity to 7 conserved sequence motifs in the 3' UTR of the lin-14 mRNA[161,162]. Later other works showed that over-expressing miRNA lin-4, a strong reduction of 10 fold could be observed at the lin-14 protein level but the transcriptional rate as well as the polyadenylation status were not affected, and only slight reduction at the mRNA level could be seen[163]. Together, these data suggested the action of miRNAs as a translational repressor.
Since those initial contributions, many other authors have analysed the way by which miRNAs are able to down-regulate gene expression, but the mechanisms behind the action of miRNAs on target mRNAs are still not completely understood.

With few exceptions, miRNAs interact with the 3' UTR of mRNA transcripts by nucleotide complementarity driving the miRISC to that portion of the mRNA. Contrarily to plants, where miRNAs pair almost perfectly with their target mRNA, in animals, for most of the cases, the pairing is imperfect being governed by some rules that started to be uncovered in the early work by V. Ambros and his team[164]. Mutational analysis showed that a perfect binding of the region between nucleotides 2 and 8 of the miRNAs (known as seed) and their targets is an absolute requirement. Higher repression efficiency is obtained when several binding sites are present in the same UTR.

How is the binding of an element in the 3' end able to repress an event occurring along several kilo bases of a normal RNA strand? A simplistic view would point to a repression occurring in the end of the translation process when the ribosome involved in the translation, would get closer to the miRISC positioned in the 3'UTR. However, that seems not to be the case.
Although some authors show and defend the idea that repression occurs during the elongation phase of the translation (post-initiation) by a "ribosome drop-off"[165], the most consistent experimental evidence points to a repression at the initiation step. According to recent publications, the repression seems to occur when the ribosome is not even completely assembled at the opposite mRNA terminus[166,167]. For that, spatial rearrangements of the mRNA, and the very far end m^7G cap play an important role. The translation initiation and the positioning of the ribosome subunits on the AUG codon are governed by a complex containing some translation factors. One is eIF4E that recognize the m^7G cap and position the whole complex. The same complex also integrates another factor, eIF4G, with affinity for proteins, like the polyadenylate-binding protein 1 (PABP1), that in turn bind to the poly-A tail in the 3'UTR[168]. This binding brings the two terminuses

to a close proximity, influencing positively the translation initiation processes[169] when RISC complex is absent.

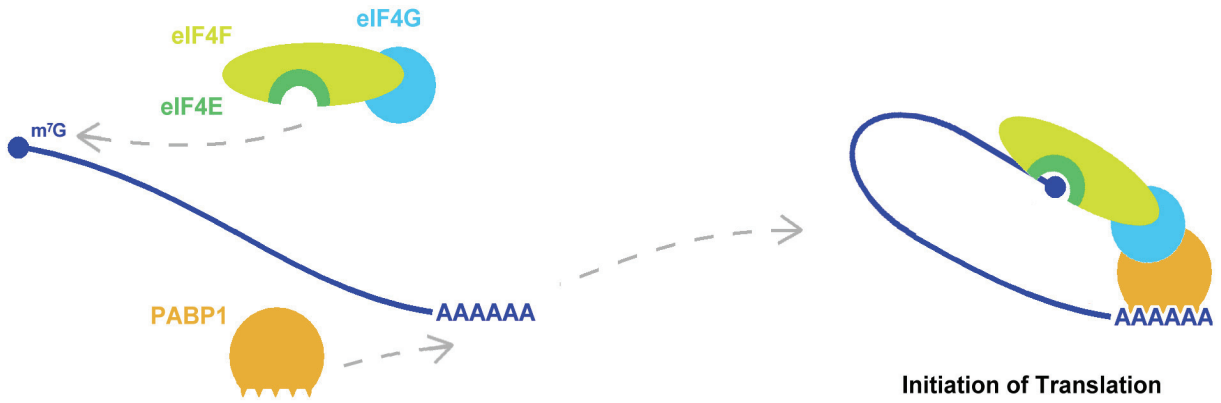


Fig. 7 - Schematic representation of the major events leading to the recruitment of ribosomal subunits and subsequent translation initiation of an mRNA.

When RISC complex is present at the 3' end, this circularization brings also the miRISC closer to the 5' end, and recently, it was indeed shown that the AGO proteins of miRISC have a domain characterized by two aromatic residues with affinity for the m^{7}G cap[170]. The binding of AGO2 to the m^{7}G cap is able to repress translation and the proposed mechanism is that this binding competes with the one of the eIF4E preventing the positioning of the translation complex and the subsequent recruitment of the ribosome subunits.

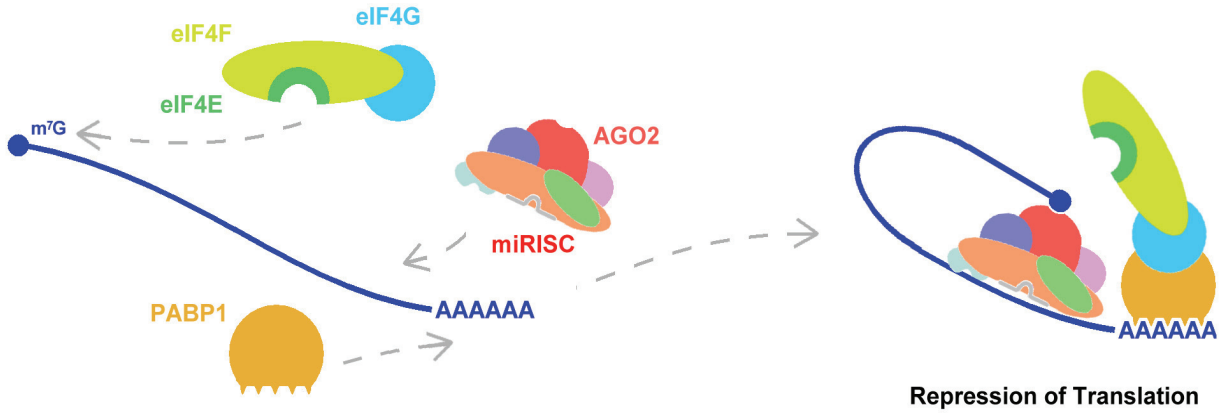


Fig. 8 - Schematic representation of the major events of translation initiation inhibition by the presence of miRISC bound to the 3'UTR of an mRNA.

The proposed model provides also an explanation for the fact that the presence of more miRNA binding motifs in the 3'UTR of an mRNA makes the repression more effective as more AGOs are present. As more AGOs proteins are present, more chances exist for inhibition of the eIF4E factor binding to the m^{7}G cap.

For most of the cases the miRNA-dependent repression of translation does not lead to a degradation of the mRNA. But an increasing number of examples show that a moderate

(1,2 to 2 fold) target mRNA degradation can also occur via miRNA action by destabilization and deadenylation[171-175] possibly in a cell specific manner[176],.
Presently 3 different possible degradation mechanisms are considered[177]: One is called "endonucleolytic cleavage" and may occur when a full complementarity between the miRNA and its target mRNA is verified (as between miRNA-196 and HoxB8 mRNA)[178,179]. In this mechanism mRNA is supposedly cleaved by the AGO2 protein, the only Argonaute protein with endonuclease activity although it normally do not integrate the miRISC (it is present in the siRISC); Another mechanism is the one mediated by the seed[180]. It is not really understood but evidences show that it differs from the siRNA mediated effect probably by involving different proteins; A third mechanism is called "AU-rich element (ARE)-mediated decay". This is the one observed for miRNA-16[181] which mediates degradation by binding of its central motif UAAAUAUU (not of the seed) fully complementary to typical ARE sites present on its target transcripts.

1.2.3 - Regulation of miRNA expression

In contrast to the significant advances in the last years in terms of our understanding on the miRNA biogenesis, nowadays not many things are understood about the way how miRNAs themselves are regulated. The mechanisms that govern the expression of miRNAs and facilitate the tissue/cell specific expression are not known yet. "What regulates miRNA expression?" is one of main questions currently under investigation and debate.
Possible checkpoints for expression of mature and active miRNAs may reside in the transcription, in the processing, in the sub-cellular localization and/or in the stability.
The enzymes involved in the processing and on the cellular localization of miRNAs like Drosha, DGCR8, Exp5, Dicer and TRBP1 are required for all miRNAs. At present evidence is lacking that of any of these enzymes has differential affinity for any special group of miRNA. These facts make the processing steps, as well as the sub-cellular localization, unlikely checkpoints and it attributes to those key enzymes a constitutive rather than a regulatory role[182]. But some few examples exist of a posttranscriptional regulation of miRNAs as the case of let-7 in sea urchins in which the miRNA mature form is only found in the adult while its precursors are found already in embryonic stages[183].
Experimental evidence suggest that miRNAs stability is high[184,185]. That might be due to the physical characteristics of the molecule and, as they are common to all miRNAs, it is also unlikely to constitute a regulatory step.

Regulation of miRNA expression is, so, more likely to occur at transcription level. The available experimental data suggest that a large fraction of pri-miRNAs, or even probably all, are Pol II transcripts, and as such structurally similar to normal mRNAs[186-189]. By analogy, it is also likely that transcription levels of pri-miRNAs might be dependent on a

complex balance between transcription factors. Importantly, a better understanding of the regulation of miRNA expression requires the definition and localization of the miRNA promoter. On that respect, several authors have given some insights in the identification of miRNA promoters and in the identification of transcription factors regulating miRNA expression.

In 2003 two independent works identified two regions of 521bps and 116bps on the *C. elegans* genome essential for the expression of miRNAs lys-6 and let-7, respectively [190,191]. The authors of the second study did not identify any known promoter sequence or transcription binding site in that area. They instead attributed the promoter activity to a novel element they called temporal regulatory element (TRE) characterized by two inverted sequences located approximately at position -1,2Kbps and able to drive basal expression of the miRNAs. The authors also verified experimentally that the sequence element can bind to an uncharacterized protein that could act as transcriptional factor[192]. In 2004, work done in human cells identified a region able to drive the transcription of the cluster containing miRNAs-23a,27a,24-2[193]. The region lacked any known promoter elements commonly required for Pol II transcription initiation, including TATA box, the initiator element, the downstream promoter element (DPE), the TFIIB recognition element (BRE), and neither the less common elements like downstream core element (DCE), the DPE found in the human glial fibrillary acidic protein, the multiple start site element downstream 1 (MED-1) and the proximal sequence element (PSE)[194]. The authors identified a CG-rich region in the area and point it as possible player on the promoter activity[195].

In 2005 a binding site was identified for the muscle specific transcription factor serum response factor (SRF). Located upstream of miRNA-1, that is conserved among species and able to regulate the expression of that miRNA[196]. Another study provided evidence for C-Myc activation of the miRNA-17 cluster expression by binding to the intron 1 of the pri-miRNA[197]. Also in 2005 another study done in human haematopoietic cells lead to the identification of two contiguous C/EBPα binding elements between positions -730 and -709 of 5' end of miRNA-223 able to drive expression of this miRNA[198]. In 2006 the RE1 silencing transcript factor (REST) was shown to repress the expression of miRNA-124a[199] . A more recent study used an algorithm to predict miRNA promoters in several species including *H. sapiens*[200]. This algorithm identified some sequence motifs and predicted that most of the miRNA promoters are located in a relative close proximity of the pre-miRNA hairpins. Interestingly, this work also points to the existence of a small number of miRNAs containing conserved CpG islands located in close proximity. This fact might be interpreted as indicative of evolutionary conservation and suggests some role in expression regulation.

Taken together, the above mentioned publications strongly suggest that regulation does not occur by a single mechanism common to all miRNAs and that they even might involve novel transcription factors or other similar elements.

1.2.4 - The role of epigenetics in miRNA expression

As previously described, some experimental evidence show similarities between the regulation of miRNA expression and mRNA transcription. Assuming this, DNA methylation occurring on CpG islands could also likely play an important role in regulating miRNA expression. Although likely, until 2005, at the time this work was started, there was no experimental evidence for any direct influence of epigenetics on miRNA expression.

The first studies focusing epigenetic regulation of miRNA expression were published in 2006. The impact of histone acetylation in miRNA expression was the aim of the first study. The work reported that HDAC inhibition by HDACi LAQ824 drug led to rapid alteration of the expression of some miRNAs in the human breast cancer cell line, SKBr3[201].

Once that the known epigenetic factors can act on a coordinate manner, the influence of histone modifications was also investigated together with the influence of DNA methylation. The first evidence for the involvement of DNA methylation on regulation of miRNAs came in a publication from the group of the "epigeneticist" P.A Jones documenting up-regulation of some miRNAs on T24 bladder cancer cells after simultaneous treatment with 5-AZA-CdR and 4-phenylbutyric acid (PBA), a HDAC inhibitor[202].
After that initial study, some other publications appeared focusing only on the effect of DNA methylation on miRNA expression, however showing contradictory results. Two independent studies reported no induction of miRNAs after 5-AZA-CdR treatment on the human lung cancer cell lines A549 and NCI-H157[203,204]. In contrast, another study reported in 2007 up-regulation of miRNAs after 5-AZA-CdR treatment of malignant cholangiocytes MzChA-1 cells[205]. In the same year, DNMT1 and DNMT3b were knocked out on HCT-116 cells, a colon cancer cell line, and an up-regulation of some miRNAs was observed when compared with the control cell line[206].
Contrary to the above cited publications, some others exist that included primary cells on their study. In one, hypermethylation was observed in the loci of miRNA-9-1, 124a3, 148, 152 and 663 in a significant percentage of primary breast cancers and breast cancer cell lines correlating with low expression levels[207]. The same authors also observed that, treatment of some cell lines with 5-AZA-CdR lead to up-regulation of miRNA-9. A partly similar working procedure was used by others resulting in the identification of a different set of miRNAs (miRNA-34b, 132, 137, 193a, 203) as being silent via aberrant DNA methylation in oral squamous carcinoma cell lines as well as in primary cells of the same malignancy[208]. Higher expression levels were also obtained after treatment of cell lines with 5-AZA-CdR. A third publication reports that the down-regulation of miRNA-223 in acute myeloid leukaemia (AML) patients and cell line samples, find an explanation on the locus

hypermethylation verified[209]. Additionally, treatment of AML cell lines with 5-AZA-CdR lead to the restoration of miRNA-223 expression levels and blasts differentiation.

All the publications reporting the effect of DNA methylation on miRNA expression suggest that only a modest number of miRNAs is affected. Interestingly, each study observed a different pool of miRNAs being up-regulated. This suggests that the control of miRNAs expression by epigenetic mechanisms, in particular DNA methylation, might be tissue specific. If that is the case, the cell condition is also likely to be important and, an apart, that was not discussed in these studies.
One other common feature to the cited studies is that the effect of the DNA demethylating drug 5-AZA-CdR was always investigated using cell lines. Although natural, this aspect might be of high relevance and should also be considered. In general, DNA methylation status found in cell lines differs from the status of the corresponding primary cells (is less methylated). The global hypomethylation and the sporadic local hypermethylation observed in cell lines is more similar to the one of a more undifferentiated or tumour cell[210]. Data still do not exist about the impact of DNA demethylating drugs on primary healthy and untransformed somatic cells.

2 - Objectives

As pointed out in the introductory chapter, the role of miRNAs on cell differentiation, proliferation and in tumorigenesis has been widely documented. A crucial regulatory role in these processes was also postulated for epigenetic factors. Notably, the two mechanisms are able to interfere with gene expression but at different time points. While DNA methylation acts at pre-transcriptional level, miRNA regulation acts at post-transcriptional level.

DNA methylation occurs at CpG islands. There are around 29000 CpG islands and 678 miRNA genes distributed in the genome. It is so far unclear if an evolutionary bias exists towards either separating or superimposing these two genomic elements and to which extend this might lead to epigenetic control of miRNA expression.

The present work was designed to fill a gap in the knowledge about the influence of DNA methylation on miRNA expression in human primary, healthy and differentiated cells. Employing novel molecular approaches for global analysis of miRNA expression, the major objective of the study was to define the role of epigenetic regulation for miRNA expression. Starting from global analyses of miRNA expression (in a human haematopoietic cell model: natural killer cells), this work aimed to directly identify epigenetically regulated miRNAs and subsequently define their regulatory role and clinically relevance including association with tumorigenesis.

3 - Materials and Methods

3.1 - Cell culture

3.1.1 - Collection of blood samples

Blood samples were collected after informed consent, from healthy donors using standard procedures, by the technical staff of the Institute of Transplantation Diagnostics and Cell Therapeutics. Approximately 120ml of peripheral blood was drawn in 10ml BD Vacutainer™ EDTA Tubes (Becton Dickinson Labware,U.S.A.).

3.1.2 - Isolation of peripheral blood mononuclear cells (PBMCs)

Isolation of the PBMCs from peripheral blood was done according standard protocol using Ficoll-Hypaque density gradient. For that, blood was initially diluted with PBS pH7.3 (Serac, Germany) at room temperature. The proportion was 70ml of blood for 30ml of PBS. Each 30ml of Blood+PBS suspension were gently transferred to a polypropylene conical (falcon) tube (Becton Dickinson Labware, U.S.A) containing 15ml of Ficoll Separating Solution (Biochrom AG, Germany) at room temperature. Subsequently, a centrifugation was performed for 35min at 2000rpm with a maximum acceleration and without braking force.

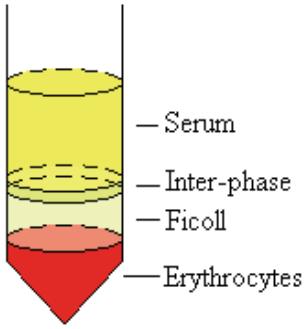


Fig. 9 – Schematic representation of falcon tube containing the peripheral blood and Ficoll after centrifugation.

With a plastic and non reusable transfer-pipette (Sarstedt, Germany), the inter-phase ring composed by mononuclear cells was recovered for another falcon tube. Recovered inter-phases from different falcon tubes belonging to a same donor were pooled at this step and the available empty volume on the falcon was filled up with 4ºC PBS. After a short mix, cell suspensions were centrifuged for 7min at 1900rpm at 4ºC. The supernatant was inverted to another falcon and centrifuged again in order to maximize the cell recovery while pellet was kept on ice. After this second centrifugation, supernatant was discarded and both pellets were resuspended first in 4ºC PBS, pooled in one same tube and centrifuged again for 7min at 1900rpm at 4ºC. In order to lyse any remaining erythrocytes not completely separated from the mononuclear cells, the new pellet was resuspend in 20ml NH_4Cl solution at 4ºC and the suspension was kept at room temperature for 10 min. Subsequently, falcon tube was filled with 4ºC PBS/EDTA and centrifuged for 7min at

1900 rpm at 4ºC. The supernatant was discarded and pellet containing the isolated PBMCs was resuspend in 4ºC PBS/EDTA. 300µl of this suspension were used for cell counting and analysis.

3.1.3 - Isolation of NK cells (CD56+ CD3-) from PBMCs

NK cells, defined as CD56+ CD3- cells, were isolated from PBMCs using magnetic activated cell sorting (MACS) technology (Miltenyi Biotec, Germany) in a two-step procedure illustrated by the following figure.

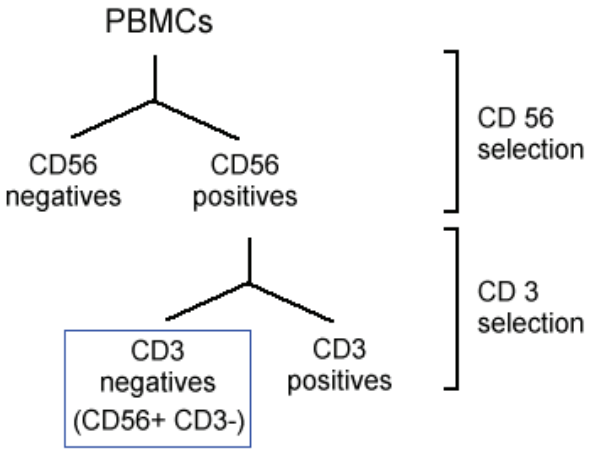


Fig. 10 – Schematic representation of the two-step procedure used to isolate NK cells. A first step consisted on a positive selection of CD56+ cells and a second step consisted on a negative selection eliminating CD3+ from the previously CD56+ population.

CD 56+ cells isolation

The isolated PBMCs were pelleted and resuspended in the residual volume. 5-10µl/10^7 cells of CD56+ magnetic beads suspension were added to the cell suspension and incubated for 30 min at 4ºC. This mixture was washed with 4ºC PBS/EDTA, resuspended in 2ml of 4ºC PBS/EDTA and passed through a 40µm strainer in order to eliminate cell clumps. This clump-free cell suspension was applied on a large size column (LS) previously equilibrated with 5ml of 4ºC PBS/EDTA and placed on the magnet. The CD56+ cells (retained on the column) were recovered releasing the column from the magnet and eluting its content with 5ml of 4ºC PBS/EDTA. 500µl of this elution solution was used to count the cells.

In order to remove the CD56 magnetic beads from the cells, 20µl of *Release Reagent* was added per 1ml of cell suspension and this was incubated for 10min at 4ºC. Subsequently, the cells were pelleted, ressuspended in 50µl of 4ºC PBS/EDTA per 10^7 cells and incubated with a *Stop Reagent* (30µl/10^7 cells).

CD3- cells isolation from CD56+ cell population

This isolation step is needed when working with peripheral blood samples. On this samples there are cells CD56+ CD3+ (also called NK-like-T-Cells) that should be removed.

20µl of the CD3 beads/10^7 of cells were added to the cell suspension containing the stop reagent from the previous selection and this was incubated for 15 min at 4ºC. After incubation, cells were washed and ressuspended in 1ml of 4ºC PBS/EDTA.

This was applied on a mini column equilibrated with 3ml of 4ºC PBS/EDTA and the flow-through solution containing the CD56+ CD3- cell population was collected and washed. All collected fractions were analysed by flow cytometry (FACS technology (Becton Dickinson, USA)) using antibodies specific for CD8, CD3 and CD56.

3.1.4 - Cell culture conditions and transformation

Primary NK cells

Culture medium was RPMI-1640 with L-glutamine and 25mM Hepes (Lonza, Belgium) supplemented with 10% Fetal Calf Serum (Biochrom AG, Germany), 1% of Penicillin/Streptomycin (Gibco, U.S.A.), 5% human-albumin (5% octalbin, Octapharma, Switzerland) and 1000 units/ml of interleukine-2 (Novartis, Switzerland).

Cultures were initially established on 96 well plate (flat bottom) containing $0{,}5x10^6$ cells in 200µl medium per well. As cell number increased, cultures were been transferred into bigger wells in culture plates. When density of living cells reached $1{,}5x10^6$cells/ml in a total volume of 2ml, cells were transferred into culture flasks (Greiner Bio-one, Germany) and the cell density was kept around $1{,}5x10^6$cells/ml in increasing volumes.

Transfection of primary NK cells was done using a calculated amount of DNA (Tab. 5) referenced to 5µg, and nucleofection technology (Amaxa, Germany) according to the manufacturer's protocol. $5x10^6$ cells were used on each assay. Briefly: $5x10^6$ cells per test were washed in PBS and resuspended in 90 µl Buffer 2A. 5µg of DNA was added to the cell suspension, mixed and transferred to a nucleofection cuvette (Invitrogen, U.S.A.). Nucleofection was performed using program U-01 and immediately, 500 µl medium were added and the whole volume was transferred to a well, of a 6 well plate, containing 1,5 ml of pre-warmed culture medium. Culture was maintained for 4h before harvesting.

NK 3.3 cell line

(Human natural killer cell line)

Culture medium was RPMI 1640 with L-glutamine and 25mM Hepes (Lonza, Belgium) supplemented with 10% Fetal Calf Serum (Biochrom AG, Germany), 1% Penicillin/Streptomycin (Gibco, U.S.A.) and 1% lymphocyte supernatant[211]. Cultures were initially established with $5x10^6$ unfrozen cells on 25cm^2 culture flasks (Greiner Bio-one,

Germany) in 10ml medium culture. After stabilization, cell density was kept around 0,5 - 0,6 x10^6/ml in increasing volumes.

Transfection of NK 3.3 cells was done using nucleofection technology (Amaxa, Germany) according to the manufacturer's protocol and briefly described above for primary NK cells.

NK L cell line

(Human natural killer cell line)

Culture medium was RPMI 1640 with L-glutamine and 25mM Hepes (Lonza, Belgium) supplemented with 10% Fetal Calf Serum (Biochrom AG, Germany), 1% Penicillin/Streptomycin (Gibco, U.S.A.) and 1% lymphocyte supernatant[212]. Cultures were initially established with 5x10^6 unfrozen cells on 25cm^2 culture flasks (Greiner Bio-one, Germany) in 10ml medium culture. After stabilization, cell density was kept around 0,5 - 0,6 x10^6/ml in increasing volumes.

Transfection of NK L cells was done using nucleofection technology (Amaxa, Germany) according to the manufacturer's protocol and briefly described above for primary NK cells.

Hela cell line

(Human Negroid cervix epitheloid carcinoma)

Culture medium was DMEM Glutamax™ I (1g/L D-glucose) (Gibco, U.S.A.) supplemented with 10% Fetal Calf Serum (Biochrom AG, Germany) and 1% Penicillin/Streptomycin (Gibco, U.S.A.). Cultures were initially established with 5x106 unfrozen cells on 25cm^2 culture flasks (Greiner Bio-one, Germany) in 10ml medium culture. Cell passages involved trypsin (trypsin EDTA 1x solution (Gibco, U.S.A)) treatment and were done near confluence diluting the cells in 1:10 ratio.

Transfection of Hela cells was performed using plasmid DNA and FuGENE® HD Transfection Reagent (Roche Applied Science, Germany). This is a multi-component reagent that forms a complex with DNA and then, is able to interact with cell membranes allowing the entry of the DNA into the cells. It has minimal cell cytotoxicity. Transfections were performed according to the manufacturer's protocol. Briefly: on the day prior the transfection, 3x10^5 Hela cells were plated per well in a 6-well plate in 2ml of culture medium. For the transfection, the amount of plasmid calculated (Tab. 5) referenced to 2µg, was diluted in 100µl H_2O NaCl 150mM solution and added to 6µl of transfection reagent at room temperature. This mixture was added to the cells and the culture was maintained for 24h, 48h or 72h before harvesting. FuGENE® HD Transfection Reagent

HepG2 cell line

(Human hepatocellular liver carcinoma cell line)

Culture medium was RPMI 1640 with L-glutamine and 25mM Hepes (Lonza, Belgium) supplemented with 15-20% Fetal Calf Serum (Biochrom AG, Germany), 1% of Penicillin/Streptomycin (Gibco, U.S.A.). Cultures were initially established with $5x10^6$ unfrozen cells on $25cm^2$ culture flasks (Greiner Bio-one, Germany) in 10ml medium culture. Cell passages involved trypsin (trypsin EDTA 1x solution from Gibco, U.S.A) treatment and were done before confluence (twice per week) diluting the cells in 1:4 – 1:6 ratios.

Transfection of HepG2 cells was performed as described above for Hela cells. For HepG2 cells, were used 4 $x10^5$ cells, the same calculated amounts of plasmids (Tab. 5) referenced to 2μg, and 5μl of FuGENE® HD Transfection Reagent (Roche Applied Science, Germany).

MCF7 cell line

(Human adduct breast cancer cell line)

Culture medium was RPMI 1640 with L-glutamine and 25mM Hepes (Lonza, Belgium) supplemented with 15-20% Fetal Calf Serum (Biochrom AG, Germany), 1% of Penicillin/Streptomycin (Gibco, U.S.A.). Cultures were initially established with $5x10^6$ unfrozen cells on $25cm^2$ culture flasks (Greiner Bio-one, Germany) in 10ml medium culture. Cell passages involved trypsin (trypsin EDTA 1x solution from Gibco, U.S.A) treatment and were done before confluence (two or three times per week) diluting the cells in 1:3 - 1:6 ratios.

Transfection of MCF7 cells was performed as described above for Hela cells. For MCF7 cells, were used 3 $x10^5$ cells, the same calculated amounts of plasmids (Tab. 5) referenced to 2μg, and 6μl of FuGENE® HD Transfection Reagent (Roche Applied Science, Germany).

3.1.5 - Incubation with 5-aza-2'-deoxycytidine (5-AZA-CdR) and Trichostatin A (TSA)

On day 20 of culture, primary NK cells were incubated with the above described culture medium supplemented with 2 μM 5-aza-2´-deoxycytidine (5-AZA-CdR) (Sigma, U.S.A.) for 6 days, with 5-AZA-CdR and 0,5μM of the histone deacetylases (HDAC) inhibitor Trichostatin A (TSA) (Sigma, U.S.A.) also for 6 days or were kept under the control culturing conditions (without 5-AZA-CdR or TSA) for a same period of time.

To keep the concentration of 5-AZA-CdR constant during the 6 day period, 5-AZA-CdR was renewed every 24 hours due to compound instability.

3.2 - Molecular Biology

3.2.1 - Extraction and purification of nucleic acids

RNA Isolation

Total RNA, as well as the fraction of RNA enriched in species smaller that 200nt long, was extracted from cultured cells using mirVana™ miRNA isolation kit (Ambion, U.S.A.) according to the manufacturer's protocol. To perform Northern blotting experiments, total RNA was isolated using Trizol® reagent (Invitrogen, U.S.A) according to the manufacturer's protocol.

DNA Isolation

DNA extraction from cultured cells was done using QIAamp® DNA Blood Mini Kit (Qiagen, Germany) according to the manufacturer's protocol.

Nucleic acids purification

PCR products, digestion-reaction products and labelled northern blot probes were purified using QIAquick® PCR Purification Kit (Qiagen, Germany) or, alternatively, using standard phenol-chloroform protocol. Briefly, 1 volume of Phenol/Chloroform/Isoamilic (25:24:1) mixture (all compounds from Roth, Germany) was added to a PCR reaction and agitated well. The mixture was centrifuged 10min at maximum speed in a bench top centrifuge (Hettich, Germany). The upper phase (aqua.) containing the DNA was recovered to other tube. Subsequently, DNA was precipitated by addition of 1/20 volumes of NaCl 5M and 3 volumes of cold EtOH 100%. The suspension was kept at -80ºC for 30min for better precipitation and centrifuged 10 min at 4ºC at maximum speed. The supernatant was carefully removed and the pellet was washed with 300µl of cold EtOH 70% and centrifuged 10 min at 4ºC at maximum speed. Supernatant was removed and the pellet was dried at room temperature being, finally, resuspended in water.

PCR products were also purified from preparative agarose gels. For that, bands containing the expected DNA fragments were cut out from the gel over a U.V. transilluminator using a scalpel. Purification was done using QIAquick® Gel Extraction Kit (Qiagen, Germany) according to the manufacturer's protocol.

Although ligation-reaction products generated during cloning procedures were not run on agarose gels, they were purified using an adapted protocol of the QIAquick® Gel Extraction Kit (Qiagen, Germany).

Plasmids were extracted from transformed bacteria using QIAprep® Miniprep Kit (Qiagen, Germany) or QIAfilter Maxi Kit (Qiagen, Germany) according to the manufacturer's respective protocols.

3.2.2 - miRNA profiling by TaqMan® MicroRNA assay

TaqMan® MicroRNA Assay

The miRNA profiling was performed by Dr. Sandra Weinhold using the TaqMan® MicroRNA Assays, Human panel (Applied Biosystems, U.S.A.) and an adapted protocol. In a bench top centrifuge, 1000 sorted cells were pellet at 3000rpm and resuspended on 5µl of PBS. To promote cell lysis, the pellet was incubated at 95ºC for 5min and transferred immediately into ice. All miRNAs were converted into cDNA in one reverse transcription reaction. For that 7µl of a master mix containing enzymes and buffers necessary for cDNA synthesis as well as 4µl of a pool of specific primers for 48 miRNAs were added. These 8x40µl multiplex reverse transcription reactions took place on a Gene Amp® PCR System 9700 (Applied Biosystems, U.S.A.).

16ºC 30min ; 42ºC 30min ; 85ºC 5min ; 4ºC ∞

The cDNA fragments generated were about 70nts long consisting on the mature miRNA cDNA copy and the primers used for cDNA synthesis.

Subsequently, a real time polymerase chain reaction (PCR) using TaqMan® probe technology was done separately for each of 234 miRNAs. The cDNA generated was diluted 1:10 and split into 234 single wells of 96-well-plates containing a universal primer, a primer specific for each miRNA and a TaqMan® probe also specific for each miRNA (4,5µl of single Taqman MicroRNA Assay). The reaction was carried in an ABI 7900 HT Fast Real Time PCR System (Applied Biosystems, U.S.A.) using the following conditions:

95ºC 10min ; 40x(95ºC 15sec; 60ºC 60sec)

During this PCR reaction, in the annealing step, both primers bind to the respective terminus of the target fragment while the TaqMan® probe binds to an inner sequence motif of it. During elongation step of PCR reaction, the Taq DNA polymerase elongates both primers destroying the TaqMan® probe. Each TaqMan® probe is labelled with a reporter fluorophore and a quencher fluorophore. The destruction of the probe, leads to the separation of both fluorophores and this separation allows the reporter fluorophore to emit fluorescent radiation that can be detected by a sensor. The amount of emitted light will be representative, on a proportional way, of the amount of PCR product already generated after each PCR cycle.

The light signal is read and interpreted by a reader sensor after each PCR cycle and is obtained a graphic representation of the PCR product generated in terms of PCR cycles run.

Graphic representations from the data recorded allow the setup of some parameters for analysis like the interval of cycles to be used for constitution of a base line measurement of light intensity and the light intensity threshold above which it is considered a positive signal. The number of cycles at which the light intensity has the same value as the established threshold is recorded as cycle threshold (Ct). Lower Ct values mean that sooner in the PCR reaction was obtained a certain amount of product and that means that the template was more abundant in the initial population. – It other words, to a higher amount of specific miRNA that was present on the sample corresponds a lower Ct value. For Ct values determination, a base line was calculated based on the cycles interval 3-15 and a Ct value of 0,05 was used for all analysis made.
Technical duplicates or triplicates were made of every sample and, when possible, biological duplicates were also made using a different sample of NK cells from the same donor.

Normalization of the Ct values

Normalization was done by the method of Vandesompele *et. al*[13]. That involved the constitution of a normalization factor based on the geometric mean of the most stably expressed miRNAs across all the tested runs. The choice of the most stable genes is based on the assumption that two perfect reference genes behave equally across different samples maintaining constant their relative expression. The authors developed an algorithm that analyses the Ct values of every combination of two genes and defines the most stably expressed genes across the several samples. The more stable the expression of a gene is, the better it can work as a reference. As there are no perfect normalizer genes, the same method defends that is not the expression of a single gene that allows the best normalization but the expression of several most stable genes. It is suggested to constitute a normalization factor calculated as the geometric mean of the Ct values of the more stably expressed genes.
The expression stability of 234 miRNAs across 11 independent runs (NK cells with and without AZA) from 8 NK cell samples from 2 different donors was calculated using two algorithms named qBase and geNorm (available on-line at http://medgen.ugent.be/~jvdesomp/genorm/). A normalization factor for each sample was calculated as the geometric mean of the Ct values obtained of the 5 most stable miRNAs (see Formula (1) in section 4.3). Subsequently, all the miRNA Ct values of a given sample were normalized on the basis of ΔCt to the calculated normalization factor. In the end, the data obtained from different experiments was compared by the ΔΔCt method.
Up-regulation was considered when a >2 fold (Log2 > 1) difference in the expression levels was observed and down-regulation was considered when a <0,5 fold (Log2 < -1) difference in the expression levels was observed.

Before the method was applied to the TaqMan miRNA dataset, the data from all miRNAs which do not show a detectable expression in all samples was removed from it and was not

considered to constitute the normalization factor. The first reason is naturally the idea that a normalizer gene is chosen due to its stable expression and so has to be expressed. The mathematical conception behind the Vandesompele's method brings another reason. Once that the real-time PCR program has 40 cycles, by default, to transcripts not detected, the analysis software attribute a Ct value of 40. So, this value does not translate a real expression but a situation in which, at cycle 40, signal was still not detected. So, if a certain miRNA is not detected in several samples, by the algorithm that would be interpreted as a mistrustful tremendously stable expression at Ct=40.

3.2.3 - Detection and quantification of RNA transcripts

cDNA was prepared from 1µg of total RNA using Moloney-Murine Leukaemia Virus Reverse Transcriptase (M-MLV RT) (Promega, U.S.A.) and Oligo-dT primers according to the manufacturer's standard protocol in a 20µl or 50µl reaction.
Relative quantification of the transcripts was done by SYBR Green real-time PCR using QuantiTect SYBR Green PCR Kit (Qiagen, Germany) using 12,5µl of Master Mix, 1,5µl of cDNA and 10ρmol of each primer. All the primers used for cDNA amplification were synthesized (Thermo Electron GmbH, Germany) according sequences designed with software Oligo® version 6,71 (Molecular Biology Insights, U.S.A.) and Clone Manager Professional Suite version 6,00 (Scientific & Educational Software, U.S.A.). The exception is the sense primer for detection of Drosha transcript that was according a previous publication[214]. The sequences are listed in the following table.

Tab. 1 – Sequence of the primers used for transcript quantification.

Transcript	Sense Primer	Antisense Primer	Length (bps)	Tm (ºC)
miR-22	GGAAGGTGACAGAAATGGG	GGAAGTAGGGCGGTGAAG	221	56,0
miR-33	GAAAGCAGATGGGCAGCAGAG	GTATGCTGGGCGGAAGCTGT	289	59,0
miR-128	CAAGACAGTCAATTCCAGTCCG	CTGCCTCCAAATCCTTCATTGTT	285	54,0
miR-449	CAGAGGCAGTCGAGCAGAGAGG	GGCTACTCGCAACAGGAACCTCT	268	56,0
	TGCGAGTAGCCCCATTACC	CAGTCTCTGTCCATTGCATGAG	255	56,0
Drosha	CATGCACCAGATTCTCCTGTA	GGGTCTCCTGCATAACTCAAC	270	55,0
TRBP	GGTGGCCCTCAAACACCT	CTGCACCACCAGCTCCTG	244	58,0
Exp5	TCAAAACCGTCTTGGAAAGAAT	TGGAGGTAGGTGAAAAGAGGTC	286	55,0
Dicer	GCTTGAAATGCTTGGCGACTC	CCTCCTCATCCTCCTCCTCGT	329	55,0
DGCR8	GAGACAGAGTGGATGAAGAGG	TCTCTGTGTAGGTACACCGG	289	55,0
β-actin	GAAGATCCTCACCGAGCGC	AGGGTAGATGGTGGTGCCG	352	58,0

SYBR Green real-time PCR amplification was performed in an ABI Prism™ 7700 Sequence Detector (Perkin Elmer, U.S.A), using the following conditions:

95ºC 13min ; 40x(95ºC 60sec; Tm 25sec; 72ºC 25sec)

Subsequently, results were analysed in the original software of the thermal cycler in order to obtain the Ct values, and normalization was done using the ΔΔCt method considering the expression of β-actin transcript.

3.2.4 - Analysis of DNA methylation status

The analysis of the DNA methylation status is based on the distinction of methylated and non methylated cytosine residues. That is achieved by performing a chemical modification (bisulfite conversion) prior to DNA amplification by PCR and sequence analysis.

Bisulfite conversion

DNA was chemically treated with bisulfite reagents using MethylEasy™ DNA Bisulphite Modification Kit (Human Genetics Signatures, Australia) or alternatively using EpiTect® Bisulfite Kit (Epigenomics and Qiagen, Germany) according to the manufacturer's protocols.

Both kits convert non methylated cytosines of single stranded DNA into uracil by incubation with highly concentrated bisulfite salt solutions at high temperature and low pH. After conversion, DNA was purified. That was made by two different methods in the two used kits. The MethylEasy™ uses an ethanol precipitation while the EpiTect® uses a silica column based procedure. After purification, DNA was used for *in vitro* amplification.

DNA *in vitro* amplification

1,5μl of bisulfite-converted DNA was amplified in 25μl PCR using 1,5 units of HotStar Taq® DNA polymerase (Qiagen, Germany), 0,4mM of each dNTPs (PeqLab, Germany) and 10ρmol of each primer. All the primers used were synthesized (Thermo Electron GmbH, Germany) and according sequences designed with previously mentioned primer-designing software. Designing took in consideration the bisulfite conversion made and avoid CpG dinucleotides. The sequences are listed in the two following tables.

Tab. 2 – Sequence of the primers used for methylation status analysis of the putative regulatory region of specific miRNAs surrounding transcripts.

Region	Sense Primer	Antisense Primer	Length (bps)	Tm (ºC)
miR-22	GGAGGGTTGGGTTATATTTGTT	AACTCACACCCACCTAATACTAAATA	391	54,5
miR-33	TGAGATGAGTTTGGTTTTATTG	ATCTCCAAACCACCCAAC	301	55,7
miR-128	GGAGGTAAAGGGGTGGAGAGAAGATA	CTCTCCCCTACCTACCTCCTTCTCC	410	54,7
miR-449	TGGAGAAAGAGAAGTGGGTGTA	CCAAATCCCCCACTCCTC	431	54,2

Tab. 3 – Sequence of the primers used for methylation status analysis of the region surrounding specific miRNAs.

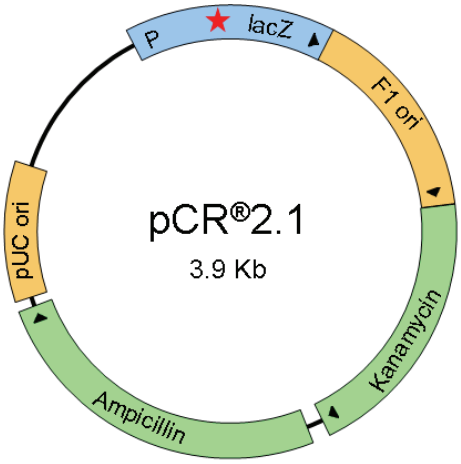
Region	Sense Primer	Antisense Primer	Length (bps)	Tm (ºC)
miR-34a	ATAGGTGTTGGGGAGAGGTAG	AACCCCTATACCTTTTTCCTTC	364	50,5
miR-132	TTGGTTTTAGATTGTTTATTG	CCATAACTATAAACTATTACCTCC	416	51,7
miR-140	TGTTTTGGATAGTAGGGTAGAAG	ACCAATCCCCAACTACTAAAC	409	50,7
miR-148	GAGTGAGTTAGTTGTTGGTAAAAG	CACAACCTCTAAAAACAAAATTC	441	50,7
miR-194-2/192	GTTTAGGAAGGAGTGTTGG	AATTCATAAATCAAAACCCTATAC	411	51,4
miR-200c/141	TTGAGTTTGGGATTGTAG	AAAACCTCCATCATTACC	393	52,5

PCR amplification was performed in a Gene Amp® PCR System 9700 (Applied Biosystems, U.S.A) using the following conditions:

95ºC 14min ; 30x(95ºC 50sec; Tm 30sec; 72ºC 30sec) ; 72ºC 7min ; 4ºC ∞

Sequence analysis

After amplification, PCR products were purified after run on preparative 0,8% agarose gel. Taking advantage of the single adenine 3' overhang present in fresh PCR products, fragments were ligated into pCR®2.1 vector (Invitrogen, U.S.A.) using the Original TA Cloning® Kit (Invitrogen, U.S.A.) according to the manufacturer's protocol in 10µl reactions incubated at 14ºC overnight.



P (LacZ) - promoter of *lacZ*α fragment.

LacZ – LacZα fragment - encodes the first 146 amino acids of *β-galactosidase* gene complemented in trans with the Ω fragment. It allows blue/white colony screening.

F1 ori - Phage F1-derived origin of replication.

Kanamycin - *aph(3')-Ia* gene that confers Kanamycin resistance and allows selection of transformed E. coli carrying the plasmid.

Ampicillin - *β-lactamase* gene that confers Ampicillin resistance and allows selection of transformed E. coli carrying the plasmid.

pUC ori - pUC-vector derived replicon (high copy number).

Fig. 11 - Map and annotation of the pCR®2.1 vector. Direction of gene transcription is represented by the black triangles and the multiple-cloning site is represented by the star.

2µl of ligation products (without purification) were used to transform chemically competent TOP10F' *E. coli* bacteria (Invitrogen, U.S.A.) according instructions of the above mentioned kit.

E. coli TOP10F' strain genotype:

F'[lacI^q Tn10(tet^R)] mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 deoR nupG recA1 araD139 Δ(ara-leu)7697 galU galK rpsL(Str^R) endA1 λ$^-$

50-500µl of transformed-bacteria suspension were spread on Lennox L Broth (LB) Base agar plates containing ampicillin (20 g/L LB medium powder (Invitrogen, USA); 12g/L of Agar (Invitrogen, USA); 100mg/L ampicillin (Sigma, U.S.A.)) and supplemented prior each use with 40µl of 20mg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) (Roth, Germany) and 40µl of 100mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (PeqLab, Germany). After overnight culture at 37°C, approximately 30 white colonies were selected. The clones were incubated overnight in 2 ml of LB medium (20 g/L) with 100mg/L Ampicillin and plasmids were extracted and sequenced using BigDye® Terminator v1.1 cycle sequencing kit (Applied Biosystems, U.S.A) and T7 primers hybridizing in the vector sequence. Reaction was performed in a Gene Amp® PCR System 9700 (Applied Biosystems, U.S.A) using standard reaction conditions. The products were separated by capillary electrophoresis on a 3730 DNA Analyser (Applied Biosystems, U.S.A) by Gaby Tillmann, Gabiele Häger and Damian Marx. Alignment of the obtained sequences was done using Clone Manager Professional Suite version 6,00 (Scientific & Educational Software, U.S.A.) and final images illustrating the results were prepared using Paint Shop Pro version 5,00 (JascSoftware, U.S.A.).

3.2.5 - Cloning of putative promoter sequences

Genomic DNA sequences corresponding to putative regulatory regions of specific miRNAs were amplified by PCR. 20ng of genomic DNA were used as template in 6x 25µl PCR containing 1,5 units of HotStar Taq® DNA polymerase and corresponding buffer (Qiagen, Germany), 0,4mM of each dNTPs (PeqLab, Germany) and 10ρmol of primers.
Primers contained tags with BglII (sense primer) and HindIII (antisense primer) restriction-enzymes-recognition sites. All the primers used for these amplifications were synthesized (Thermo Electron GmbH, Germany) according sequences designed with previously mentioned primer-designing software. The sequences are listed in the following table.

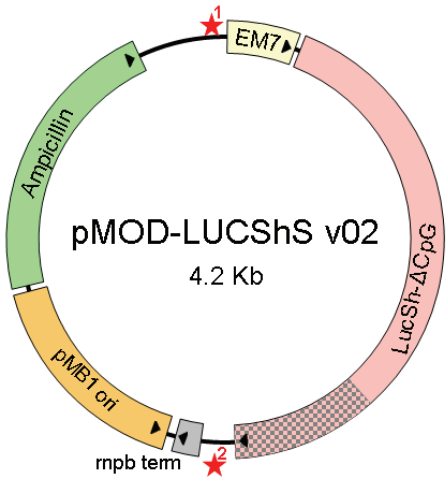
Tab. 4 - sequence of the primers used for genomic amplification

Region	Sense Primer	Antisense Primer	Length (bps)	Tm (ºC)
miR-194-2/192	GAAGATCTTC AGGGCAGAGGTGACAGTGG	CCCAAGCTTGGG TCTCTGCTGACTGCTGGACA	725	58
miR-200c/141	GAAGATCTTC CCGCTTTTTCTACCTCTGGAG	CCCAAGCTTGGG CTGCCGAGAGAACCCACC	1207	58
Control	GAAGATCTTC AGAGCCTTGCCGTAACAGA	CCCAAGCTTGGG CCCTCCCATTTAACCATC	720	58

Amplification was performed using a touchdown PCR strategy in a Gene Amp® PCR System 9700 (Applied Biosystems, U.S.A), using the following conditions:

95ºC 14min ; 5x(95ºC 50sec; 60ºC 30sec; 72ºC 70sec) ; 35x(95ºC 50sec; 58ºC 30sec; 72ºC 70sec) ; 72ºC 7min ; 4ºC ∞

After amplification, PCR products were purified after run on preparative 0,8% agarose gel. 1,5µg of the purified PCR product were digested during 16h at 37ºC, simultaneously with 15 units of BglII and 15 units of HindIII restriction enzymes (Invitrogen, U.S.A.) in a 20µl reaction.
Subsequently, products of the digestion were purified and used in a 10µl ligation reaction with 90ng of BglII and HindIII digested pMOD-LucShS 1 v02 vector (later referred simply as pMOD vector) (Invivogen, U.S.A.), or BglII and HindIII digested pGL3-basic vector (Promega, USA). The amount of insert used was calculated in order to obtain a molar ratio 1:10 (vector:insert). 2 units of T4 DNA ligase (Invitrogen, U.S.A.) were used per reaction.



EM7 - bacterial promoter enabling a low constitutive expression of the synthetic Luc::Sh that confer to the bacteria resistance to zeocin.

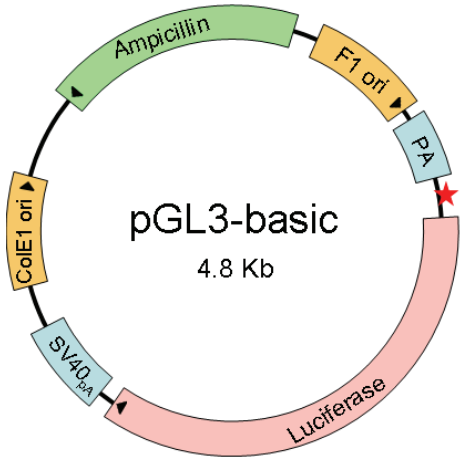
LucSh-ΔCpG - synthetic fusion gene between the firefly luciferase gene and the Sh ble gene (shadowed part) conferring zeocin resistence. Both genes were modified and contain no CpG.

rnpb term - *E. coli* transcription terminator.

pMB1 ori – ColE1-derived origin of replication.

Ampicillin – *β-lactamase* gene that confers Ampicillin resistance and allows selection of transformed E. coli carrying the plasmid.

Fig. 12 – Map and annotation of the pMOD-LucShS v02 vector (later referred simply as pMOD vector). Direction of gene transcription is represented by the black triangles and the multiple-cloning sites are represented by the stars. The multiple cloning site containing BglII and HindIII restriction sites is the number 1.



Ampicillin - *β-lactamase* gene that confers Ampicillin resistance and allows selection of transformed E. coli carrying the plasmid.

F1 ori - Phage F1-derived origin of replication.

PA - Synthetic poly(A) signal to reduce background expression.

Luciferase - *luc*$^{+}$ gene, a modification of firefly luciferase gene.

SV40$_{pA}$ - poly-adenylation site.

ColE1 ori - ColE1-derived origin of replication.

Fig. 13 – Map and annotation of the pGL3-basic vector (later referred simply as pGL3 vector). Direction of gene transcription is represented by the black triangles and the multiple-cloning site is represented by the star.

The result of the ligation reaction was purified and 1/10 of the volume was used to transform electro-competent E. coli SURE strain bacteria (Stratagene, U.S.A.) using an Gene Pulser® electroporator (BioRad, U.S.A.) and standard *E. coli* electroporation settings (Voltage: 2500V; Resistance: 200Ω; Capacitance: 25μF).

E. coli SURE strain genotype:

e14– (McrA–) Δ(mcrCB-hsdSMR-mrr)171 endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC::Tn5 (Kan^R) uvrC [F' proAB lacI^{q}ZΔM15 Tn10 (Tet^R)]

50-500μl of transformed-bacteria suspension were spread on LB agar plates containing ampicillin (see concentrations above). Overnight clones were tested for the presence of insert in 25μl colony PCR using 1 unit of Taq DNA polymerase (Qiagen, Germany), 0,4mM of each dNTPs (PeqLab, Germany) and 10ρmol of the primers used to generate the insert. The positives clones were grown for plasmid preparation and, in the end all sequences were verified by sequencing reaction using Big Dye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, U.S.A.).

The obtained plasmids, on their native form or *in vitro* methylated (see next section) were used to transform cells. The protocol used for transformation is described above. The same molar quantity of all plasmids was used and that was calculated in the basis of the plasmid size. A plasmid was used as a reference for the calculations.

Tab. 5 – Names and sizes of the plasmids constructed, and amounts used in transformation experiments. Grey-highlighted values were the reference considered to calculate the amount of the other plasmids in order to have the same molar quantity in every transformation.

Plasmid name	Plasmid size (bps)	Amount used in NK cells nucleofection (μg)	Amount used in Hela, MepG2 and MCF7 cells transformation (μg)
pMOD	4198	3,88	2,00
pMOD_194-2/192	4923	4,55	2,34
pMOD_200c/141	5405	5,00	2,58
pMOD_control	4931	4,56	2,35
pMOD_GAPDH	4682	4,33	2,23
pGL3-basic	4818		2,30
pGL3_194-2/192	5533		2,64
pGL3_200c/141	6015		2,86
pGL3_control	5539		2,64
pGL3_GAPDH	5318	4,92	2,53

3.2.6 - *In vitro* DNA methylation

DNA methylation was done using the CpG methyltransferase, M.SssI (NEB, U.S.A.). 25µg of vector were methylated using 14U of M.SssI with 160µM of S-adenosylmethionine in a 30µl reaction during 4 hours at 37ºC.
Methylation status of the vector was verified digesting 100ng of methylated and non methylated vector with 2U of BstUI (NEB, U.S.A.) restriction enzyme sensitive to methylation. Digestion was performed at 37ºC during 2h and subsequently, the results were run on agarose gel. Different band pattern was considered indicative of successful *in vitro* methylation.

3.2.7 - Luciferase reporter gene assay

Firefly luciferase activity was measured using the Luciferase Reporter Gene assay, high sensitivity Kit (Roche Applied Science, Germany) according to the manufacturer's protocol. Briefly, for cell suspensions (like NK3.3, NKL and primary NK cells) the cells were washed in PBS, resuspended in 70µl lysis buffer and incubated at 20ºC for 15min to allow lysis. The cell soluble fraction containing proteins was separated from the membranes by high speed centrifugation. The activity of the luciferase present in the soluble fraction was measured by incubation with luciferin (the luciferase substrate). The action of the luciferase on the luciferin results on light emission at 562nm. The activity is proporcional to the emitted light and that was measured on a miniLumat LB 9506 luminometer (Berthold, Germany).
For adherent cells (like Hela, HepG2 and MCF7), before lysis, cells were treated with trypsin EDTA 1x solution (Gibco, U.S.A) and incubated for 5-10min at 37ºC in order to be resuspended. After, cells were washed with PBS and the previously described protocol was followed.
Luciferase activity units are arbitrary and for their normalization the protein content of the samples was considered (method described below).

3.2.8 - Quantification of protein content

Protein content was used as normalization factor for the luciferase reporter gene assay. It was determined by the DC Protein Assay Kit (BioRad, U.S.A.) according to the manufacturer's protocol. Briefly, proteins react with copper in the presence of a catalyst and under alkaline pH conditions (achieved by solution A). Those copper treated proteins are able to reduce a Folin reagent (solution B from the kit) that, when reduced, has a maximum absorbance at 750 nm. Absorbance (measured on a DU® 640

Spectrophotometer from Beckman, U.S.A) is proportional to the amount of Folin reduced which is proportional to the amount of protein on the solution. Absorbance units are relative to a blank sample constituted by lysis buffer from the Luciferase reporter gene assay.

3.2.9 - 5' Rapid amplification of cDNA ends (5'RACE)

RACE experiments were performed using 5'/3' RACE Kit, 2nd Generation (Roche Applied Science, Germany) according to the manufacturer's protocol. Briefly, 2μg of total RNA were used as template in a 20μl reverse transcription reaction using gene specific primer n.1 (SP1) and the other standard components provided with the kit. Reaction was performed using Gene Amp® PCR System 9700 (Applied Biosystems, U.S.A), using the following conditions:

55ºC 60min ; 85ºC 5min

Following this, to the 3' end of the newly synthesized cDNA, a poly-A tail was added using a terminal transferase and dATPs following according the recommendations of the kit. Subsequently, cDNA amplification by PCR was performed. On a first PCR round was used an Oligo-dT-anchor primer and a second gene specific primer (SP2). A second PCR round was performed, this time using the anchor primer and a third gene specific primer (SP3).

Tab. 6 - Sequence of the primers used on the RACE experiment.

Race experiment	Primer name	Primer sequence
miR-200c	SP1	CTCCATCATTACCCGGCAGT
	SP2	ATTAGAGACTCCCAACCGCA
	SP3	AAACACTGCTGGGTAAGACGA
miR-141	SP1	GGGAGCCATCTTTACCAGAC
	SP2	CCAGACAGTGTTAGGAGCTTCA
	SP3	ACTGTACTGGAAGATGGACCC
miR-200c and 141	Oligo-dT-anchor primer	GACCACGCGTATCGATGTCGAC(16T)V
miR-200c and 141	anchor primer	GACCACGCGTATCGATGTCGAC

PCR reactions were performed in the thermal cycler above mentioned, using the following conditions:

94ºC 2min ; 10x(94ºC 20sec; 55ºC 30sec; 72ºC 40sec) ; 25x(94ºC 20sec; 55ºC 30sec; 72ºC 40sec*) ; 72ºC 7min * - 10sec increment per cycle

PCR products were purified and used freshly on a TA cloning strategy as already described above and the clones were sequenced. The sequences obtained had the gene specific primer at their 3' end. The 5' end, corresponding to the transcript starting sequence was variable once that the transcripts containing the primer sequence could have different starting sites.

3.2.10 - Northern Blot

Polyacrylamide Gel electrophoresis of RNA samples

7μg of total RNA from Hela cells isolated using Trizol® (Invitrogen, U.S.A) according to the manufacturer's protocol, were separated on 15% (w/v) Polyacrylamide gel done with 3-morpholinopropane-1-sulfonic acid (MOPS) buffer and containing 8M of urea. Gel was run at a constant voltage of 250V during approximately 1h.

Blotting and membrane cross link

The content of the gel was blotted unto a neutral nylon membrane (Amersham biosciences, U.S.A.) during 1,5h in an electro-blotting device (Biometra, Germany) at a constant voltage of 20V. Membrane cross link was performed using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) reagent (Sigma, U.S.A.) as described in a previously publication[215].

Probes and ^{32}P labelling

Northern probes contained a LNA modified base on every 3rd position (Thermo Electron GmbH, Germany).

Tab. 7 – Sequence of the probes used in the northern blot experiments for detection of miR-141, miR-200c and U6 RNA. LNA modified bases are signed with a * after the respective base letter.

Target	Probe Sequence
miR-141	CCA*TCT*TTA*CCA*GAC*AGT*GTT*A
miR-200c	TCC*ATC*ATT*ACC*CGG*CAG*TAT*TA
U6 RNA	GCC*ATG*CTA*ATC*TTC*TCT*GTA*T

50ρmol of probe were labelled with 2,5μl of freshly acquired radioactive ATP (^{32}P replacing the normal gamma phosphate) in the presence of 10 units of T4 DNA kinase (Invitrogen, U.S.A) and the corresponding enzyme's "*forward kinase buffer*". Reaction took place during 1h at 37ºC. Non incorporated and radioactive ATP was removed using the QIAquick® Nucleotide Removal Kit (Qiagen, Germany). 1μl was used to measure the incorporated radioactivity in terms of counts per minute.

Membrane pre-hybridization and hybridization

Membrane pre-hybridization was done with 6ml of pre-hybridization solution (5xSSC (75mM tri-sodium citrate and 750mM sodium chloride); 20mM Na_2HPO_4 (pH7,2); 1% Sodium Dodecyl Sulfate (SDS); 1x Denhard's solution) at 60ºC during at least 1h in a rotating incubator. During this incubation time, 300μg of salmon sperm DNA denatured at 95ºC for 5min were added to the solution in order to block the membrane reducing possible background signal. Prior hybridization, buffer was renewed and a probe volume equivalent to $3x10^6$ counts per minute was added to the solution. Hybridization took place overnight at 60ºC (hybridization temperature determined specifically for the probes used).

Membrane washing and analysis

Washing of the membrane was done incubating the membrane two times during 10min in 10ml of 5xSSC and 1% SDS solution, and one time, during 10min in 10ml of 1xSSC and 1% SDS solution. All the incubations were done at the same temperature used for hybridization and both solutions were equilibrated at the same temperature before usage. After washing, the membrane was sealed in a plastic foil and was analysed in a phosphoimager.

3.3 - Bioinformatics

All information on miRNA sequences and genomic localization were obtained from miRNA registry[216,217] available on (http://microrna.sanger.ac.uk).

Genomic sequences and transcripts sequences were obtained from the Ensenbl Genome browser (available on: http://www.ensembl.org).

CpG islands detection analysis on genomic sequences was done according to published method[218,219] (available on-line tool on: http://cpgislands.usc.edu) using as CpG island definition: %CG= 50%; ObsCpG/ExpCpG= 0.6; Length= 200bp; Gap between adjacent islands= 100bp.

Analysis of the genomic context in regard of repetitive elements was done using on-line tool RepeatMasker web server (available on: http://www.repeatmasker.org).

Conservation analysis of genomic sequences among species was done using UCSC Genome browser[220] (assembly from Mar.2006 available on: http://genome.ucsc.edu).
Information on DNA Polymerase II binding islands was obtained from a previous publication and using its available on-line data[221] (available on: http://dir.nhlbi.nih.gov/papers/lmi/epigenomes/hgtcell.html).
In silico prediction of Polymerase II promoters was done with an available on-line tool (available on: http://www.fruitfly.org/seq_tools/promoter.html).
In silico prediction of transcription factors binding sites was done with the AliBaba2.1 on-line available tool (http://gene-regulation.com/pub/programs.html).

All this information was compiled in figures and maps using Paint Shop Pro version 5,00 (JascSoftware, U.S.A.) and Clone Manager Professional Suite version 6,00 (Scientific & Educational Software, U.S.A.).

4 - Results

4.1 - Treatment of NK cells with 5-AZA-CdR

NK cells, defined as CD56+ CD3-, were isolated from 120 ml blood of healthy donors by the magnetic activated cell sorting procedure. The number of cells obtained was around 60 $x10^6$.

During the first days of the 20-day-period conceived for culture stabilization, the percentage of dead cells increased while the proliferation was very low. After 5 days in optimal culturing conditions, the population started to increase due to a reduced mortality and an increased proliferation. After 20 days, cell culture was split and 2 μM 5-AZA-CdR was introduced in some of the newly constituted cultures. Renewal of 5-AZA-CdR was done every 24 hours due to compound instability.
The DNA demethylating effect of 5-AZA-CdR on NK cells was verified by FACS analysis using the expression of KIR receptors as an indicator[222].

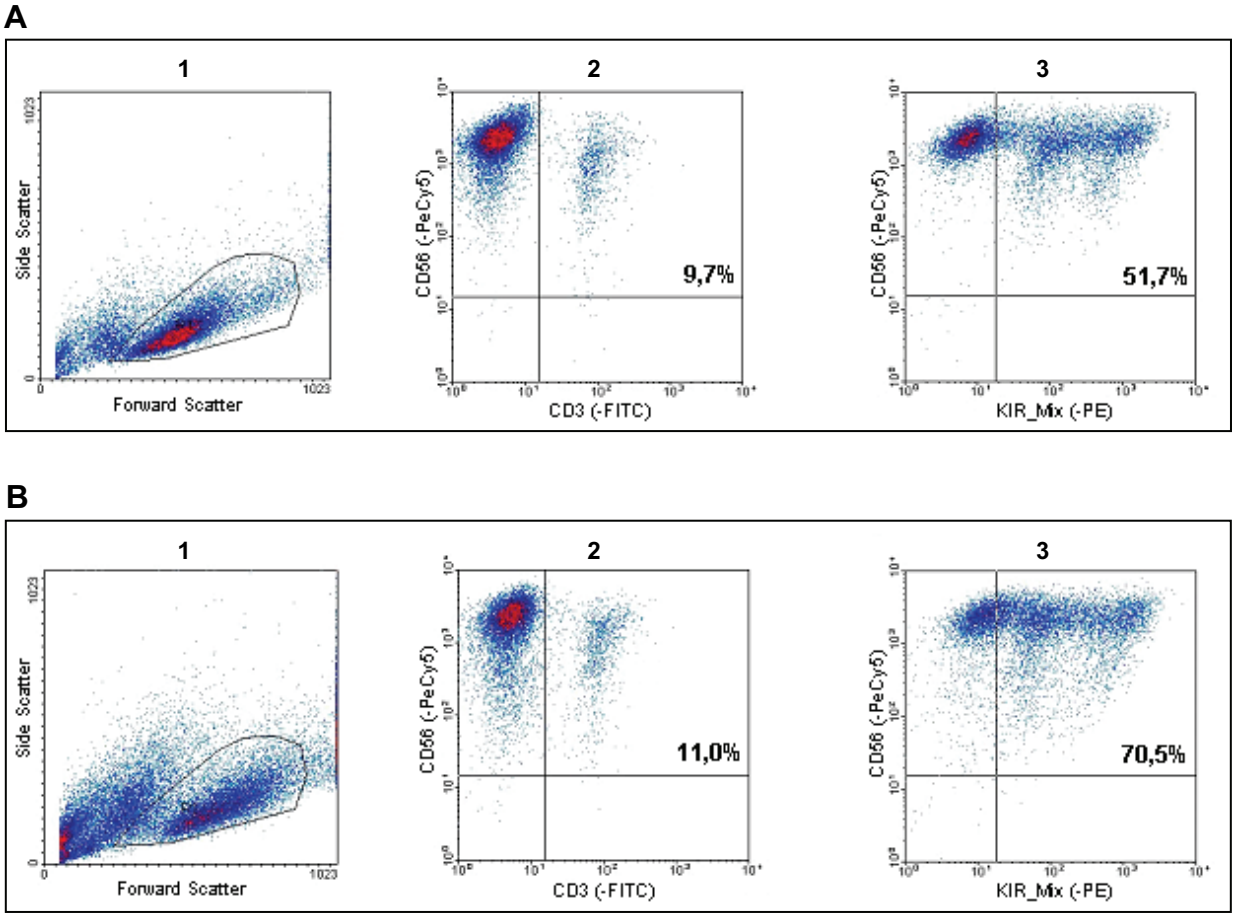


Fig. 14 - Dotplot from flow cytometry analysis of NK cell samples: A) sample after 26 days kept under standard conditions (control sample); B) sample kept under the same conditions for 20 days and treated with 2μM of 5-AZA-CdR for further 6 days. On first dotplots, in A) and B) is represented the side scatter vs forward scatter analysis with the gate defining the living cell population for the subsequent analysis. The second dotplots represent the CD3-FITC vs CD56-PeCγ5 analysis and the third dotplots represent the KIRmix-PE vs CD56-PeCγ5 analysis.

After stabilization of NK cell cultures, when 5-AZA-CdR was introduced in the culture, the mortality started to increase. It is possible to be observed on ungated cell population of the forward vs side scatter dotplot (dotplots n.1) that the mortality reached a considerable level after 6 days of treatment. The gate specifying the living population was defined based on the physical properties of the cells.
The second column of dotplots show that the NK cell populations, defined as CD56+ CD3- cells, was not 100% pure. On both samples a percentage (9,7% and 11,0%) of CD56+,CD3+ cells known to be NK-T cells is detectable.
On the KIRmix vs CD56 dotplot it can be seen that the percentage of KIR expressing cells is 51,7% in the control sample and 70,5% in the 5-AZA-CdR treated one. This difference is due to a shift of the KIR negative population present in the control sample in the direction of KIR expression.

4.2 - Impact of 5-AZA-CdR on key factors of miRNAs maturation machinery in NK cells

Due to the crucial role that the enzymes Drosha, DGCR8, Exp5, Dicer and TRBP2 play in the maturation pathway of miRNA, a possible effect of 5-AZA-CdR in those key enzymes is expected to have consequences in the miRNA expression. Due to the observation of alteration of miRNA expression upon treatment of NK with 5-AZA-CdR, the possible effect of 5-AZA-CdR on those key enzymes was investigated. First an analysis of the genomic organization of the transcripts encoding Drosha, DGCR8, Exp5, Dicer and TRBP2 was done to elucidate any possible relation with CpG islands location.

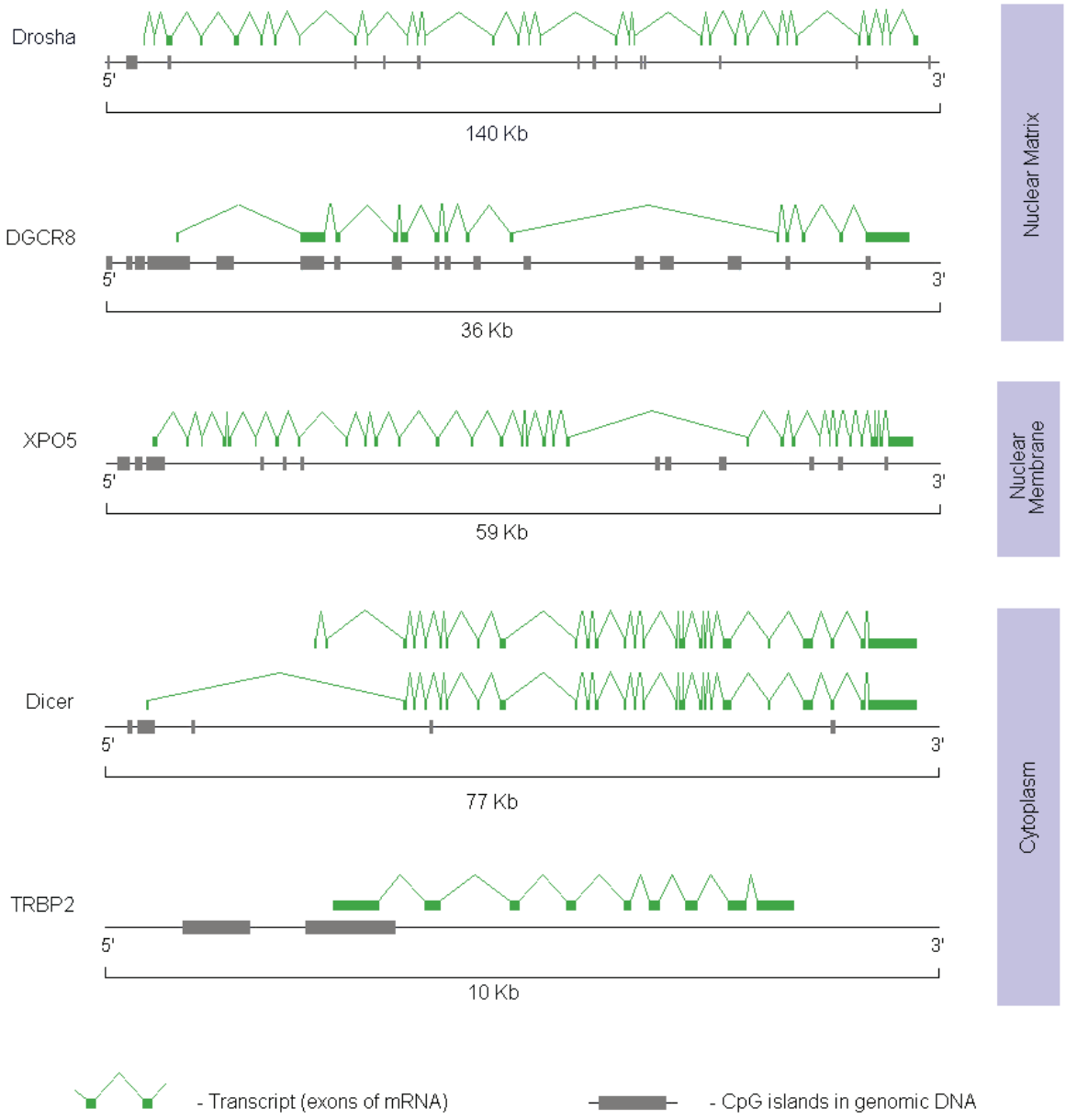


Fig. 15 – Genomic organization of the transcripts encoding the main proteins involved in the miRNA maturation pathway (Drosha, DGCR8, Exp5, Dicer, TRBP2) and CpG island localization on the same genomic area. Information on the subcellular localization of the mature form of the protein is placed on the right side of the figure.

The maps obtained show that the transcripts encoding the main proteins involved in the miRNA maturation pathway are composed of different number of exons (ranging form 9 for TRBP2 to 32 for Drosha and Exp5) as part of genes also significantly different in size (ranging from approximately 6 Kb for TRBP2 to approximately 130Kb for Drosha). Besides showing differences in organization, the results show also that the first exon of DGCR8, Exportin-5, Dicer and TRBP2 transcripts are located within CpG islands. In contrast, the first exon of Drosha transcript is not within a CpG but it is in close proximity of one.

In order to determine if up-regulation of any of these transcripts occurred after 5-AZA-CdR treatment in NK cells, the transcripts abundance was measured by real-time PCR in NK cells treated with 5-AZA-CdR and compared with control NK cells.

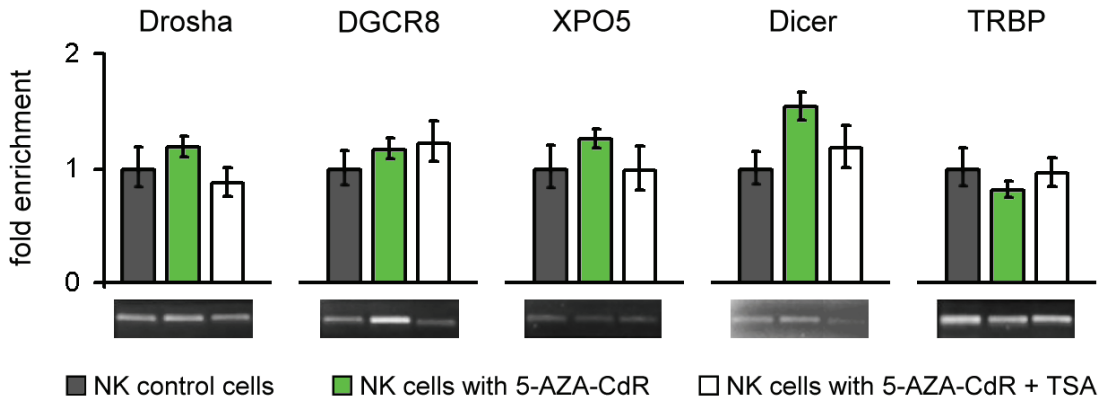


Fig. 16 – Comparison between the expression of the main proteins involved in the miRNA maturation pathway (Drosha, DGCR8, Exp5, Dicer, TRBP2) in a NK control cell sample, in a 5-AZA-CdR treated NK cell sample and in a 5-AZA-CdR+TSA treated NK cell sample. The results represented in the bars were obtained by real-time PCR. Measurements were made in duplicate or triplicate and the values were normalized first by comparison with β-actin expression (ΔCt) and then by comparison with the expression verified for the NK control cell sample (ΔΔCt). The bands in the lower part were obtained in independent end point PCR reactions being the fragments run on an agarose gel stained with Ethidium bromide.

The expression analysis done showed no significant up-regulation of four transcripts when 5-AZA-CdR alone or cumulatively with TSA was introduced in the cultures. For Dicer transcript, a small up-regulation was verified in the 5-AZA-CdR experimental condition.

4.3 - miRNA profiling in NK cells and susceptibility to 5-AZA-CdR

miRNA profiling was done using a real-time PCR semi-quantitative method[223]. All the primers for the specific detection of each miRNA are compiled in a commercially available kit: the TaqMan® MicroRNA Assays Human Panel allowing the detection of 234 miRNAs. A first analysis of the obtained cycle threshold (Ct) values showed a high concordance between technical duplicates and triplicates. Between biological duplicates, the concordance observed was lower. It was also visible that a significant number of miRNAs was not expressed in any of the NK samples under the tested experimental conditions.

The normalization of the real-time PCR results involved the choice of the best normalization factor to use. That was done analyzing the relative expression stability of all miRNAs as described in the accurate normalization method of Vandesompele[224]. The

input values were the Ct values of those miRNA for which expression was detected in all samples. An algorithm created by the authors compared all possible miRNAs pairs in terms of their expression in all samples and determines which miRNAs pairs have a more similar expression profile among the several samples as indication of stability. Finally it organizes the miRNAs in a list sorted by their calculated stability. According to this method, for the 13 TaqMan® runs of material from the 8 NK samples tested and used on this study, the most stably expressed miRNAs were miRNAs-103, 23a, 19b, 191, 30d.

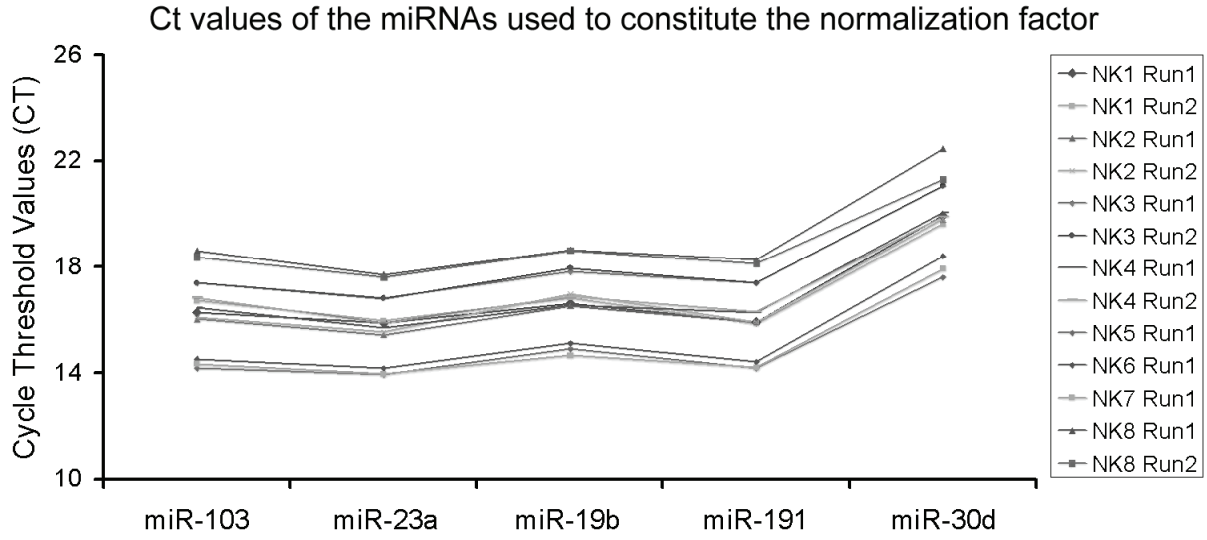


Fig. 17 – Ct values obtained for the 5 most stable miRNAs in the TaqMan® assay, in 8 samples (NK1-8) and a total of 13 runs.

The Ct values obtained for the most stably expressed 5 miRNAs show a visible difference among the tested samples. Both runs of sample NK8 show the highest Ct values approximately 4 cycles higher than the runs of samples NK5, 6 and 7 that show the lowest Ct values. It is possible to observe a common shape of the graphs resulting from the union of the Ct values points of the several miRNAs in a same sample. The overlapping of the same graphs illustrates the reproducibility of the assay between the technical and some biological replicates.

A normalization factor for each sample was calculated as the geometric mean (formula 1) of the Ct values verified for the most stably expressed miRNAs in that sample.

$$\text{N.F. Sample1} = \sqrt[5]{((\text{Ct miR-103}).(\text{Ct miR-23a}).(\text{Ct miR-19b}).(\text{Ct miR-191}).(\text{Ct miR-30d}))} \quad (1)$$

Subsequently, all the miRNA Ct values were normalized for each sample on the basis of ΔCt to this normalization factor. Finally, the data obtained from 5-AZA-CdR treated samples was compared with data obtained from control NK cell sample by the ΔΔCt method.

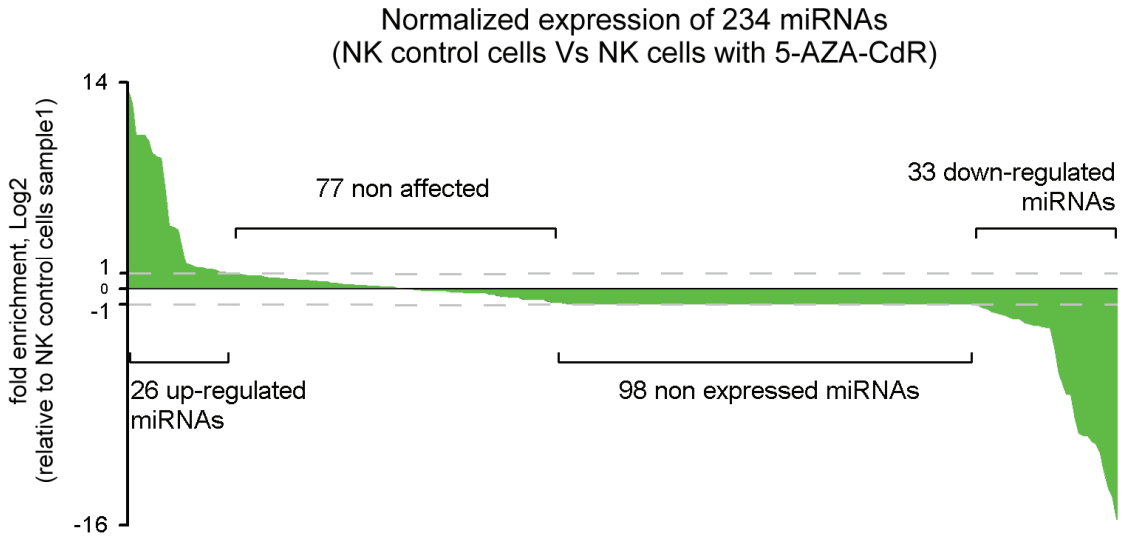


Fig. 18 – Comparison between the normalized miRNA expression data obtained from a NK cell sample treated with 5-AZA-CdR and the normalized expression data from a control NK cell sample. The graphic representation reflects that comparison made for 234 miRNAs on an area plot with a Log2 scale in which, the value 0 translate an equal expression, positive values translate a higher expression of a certain miRNA on the NK sample treated with 5-AZA-CdR vs the NK control sample and, negative values translate the opposite situation.

Considering up-regulation when a >2 fold increase (log2>1) was calculated and down-regulation when a <0,5 fold decrease (log2<-1) was calculated, on the representative experiment shown above, n=26 (26/234=11,1%) miRNAs were up-regulated and n=33 (33/234=14,1%) miRNAs were down-regulated while n=77 (77/234=32,9%) do not show a differential expression. On this particular experiment 98 (98/234=41,9%) miRNAs were not expressed. These last are represented on the long plateau occurring at approximately Log_2=-1. As explained (see section Materials and Methods), not expressed miRNAs are attributed with a Ct of 40, and as the normalization factor assumes a different value for each sample, when a Ct of 40 is compared with those values, a different ΔCt is obtained. Samples are compared on the basis of correlations between ΔCt values (ΔΔCt method) and, when two different ΔCt values are compared, it will result in a Log_2(ΔΔCt) value different of zero.

From the n=26 up-regulated miRNAs presented on the representative experiment, n=16 showed also to be up-regulated when the data from NK cells treated with 5-AZA-CdR was compared simultaneously with two biological replicates from a control NK cells sample from the same donor.

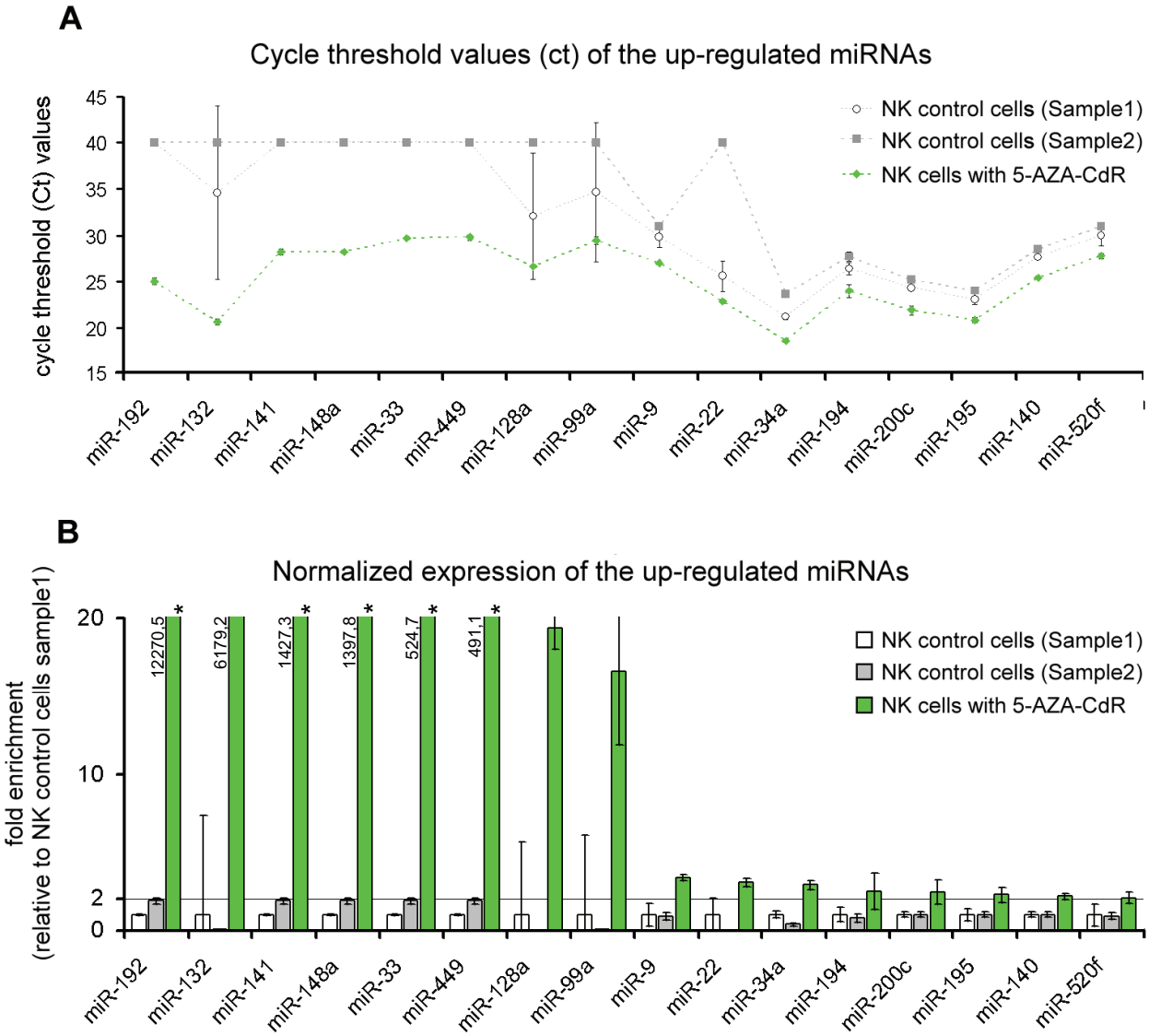


Fig. 19 – Comparison between the original Ct values (A) and also the normalized miRNA expression data (B) from the 16 miRNA showing up-regulation, obtained from a NK cell sample treated with 5-AZA-CdR and from two biological replicates of a same control NK cell sample. **A)** The graphic representation of the Ct values is reflected on a dot and line plot. Lower Ct values correspond to higher expression. **B)** The graphic representation of the normalized data is reflected on a bar plot made for 16 miRNAs for which an over-expression was observed in the 5-AZA-CdR condition when compared with both biological replicates of the control situation. The linear scale used represents a fold enrichment in which, the value 1 translates to an equal expression and values greater than 2 translate to a higher expression when compared with the first biological replicate. TaqMan measurements used for both A) and B) were done in triplicate or duplicate (technical replicates) in order to obtain the error bars. (*) - Fold enrichment is not applicable since expression is not observed on NK control cells (sample 1).

As observed in the bar plot, miRNA-192, 141, 148a, 33 and 449 show up-regulations of, respectively, 12270,5; 1427,3; 1397,8; 524,7 and 491,1 times fold enrichment in the 5-AZA-CdR treated sample. In fact the fold enrichment is not applicable to these cases since these miRNAs were not expressed in any of both control samples (Ct=40). The value is a mathematical calculation resulting from the normalization method used.

miRNA-132, 128a, 99a and 22 were detected in the sample 1 (first biological replicate), were not detected in sample 2 (second biological replicate) and showed a significantly

higher expression in the 5-AZA-CdR treated sample. miRNA-9, 34a, 194, 200c, 195, 140 and 520f showed a detectable expression on both control samples and showed a significant higher expression in the 5-AZA-CdR treated sample.

4.4 - Influence of 5-AZA-CdR on the surrounding transcripts containing the inducible miRNAs

Trying to understand the cause of the observed effect of 5-AZA-CdR on the 16 up-regulated miRNAs of Fig. 19 in NK cells, an analysis of the genomic context of those miRNAs was performed. That analysis revealed 3 different situations:

Tab. 8 – Genomic context of 14 up-regulated miRNAs

		miRNA	N observed	%
Intergenic		miRNA-192, 132, 200c, 141, 148a, 34a, 520f	7	50%
Intragenic	**Intronic**	miRNA-449, 128a, 33, 140, 99a, 195	6	~43%
	Exonic	miRNA-22	1	~7%
		total	14	100%

miR-194 was excluded from these calculations because it has different genomic localizations with different genomic contexts:

194-1 – Chrom 1 – intergenic and intronic antisense

194-2 – Chrom 11 – exonic

miR-9 was also excluded for the same reasons:

9-1 – Chrom 1 – intronic

9-2 – Chrom 5 – intergenic

9-3 – Chrom 15 – intronic

From the 14 miRNAs considered in the analysis, 7 do not show any positional relation with any known or predicted transcript with the same orientation while 6 are located within an intron and 1 is located in an exon of known or predicted transcripts. The analysis was thus concentrated on the transcripts surrounding the 7 intragenic miRNAs (6 intronic + 1 exonic).

A database search allowed the collection of the available information on the surrounding transcripts:

miRNA-449 - Q86Y33 (CDC20B)
The NCBI record of this transcript is provisional and the transcript is not fully characterized. The information available suggests that it codes for a cell division cycle 20 homolog B (S. cerevisiae)-like protein. Expression evidence were only found at transcript level in large human cDNA libraries[225].

miRNA128a - R3HD1/R3HDM1 (R3H domain containing 1)
The NCBI record of this transcript is also provisional and the transcript is not fully characterized. It is predicted as a protein coding gene. Expression evidence were already found at protein level.

miRNA-33 - SREBF2 (Sterol Regulatory element Binding protein-2)
This gene encodes a ubiquitously expressed transcription factor that controls cholesterol homeostasis by stimulating transcription of sterol-regulated genes[226].

miRNA-140 - WWP2 (NEDD-4-Like Ubiquitin-protein ligase)
This gene encodes also a widely expressed protein, member of the NEDD4-like protein family that is involved in the ubiquitination pathways[227].

miRNA-99a - NP_001005733.1 (C21orf34)
In the NCBI registry, it appears as a validated transcript coding for an uncharacterized hypothetical protein LOC388815. Expression evidence, are only found at transcript level ubiquitously expressed at reasonable high levels[228].

miRNA-195 - AC027763.1-001
The only registry is on the ENSEMBL genome browser and refers it as a novel transcript not verified experimentally and with no further description.

miRNA-22 - 116284.2 (C17orf91)
In the NCBI registry, it appears as a validated transcript coding for an uncharacterized hypothetical protein LOC84981. Expression evidence were only found at transcript level in large human cDNA libraries[229].

The collected information on the transcripts does not associate any of them with any specific function in NK cells. Considering the genomic contexts, a genomic mapping of the transcripts surrounding the 7 intragenic miRNAs was performed.

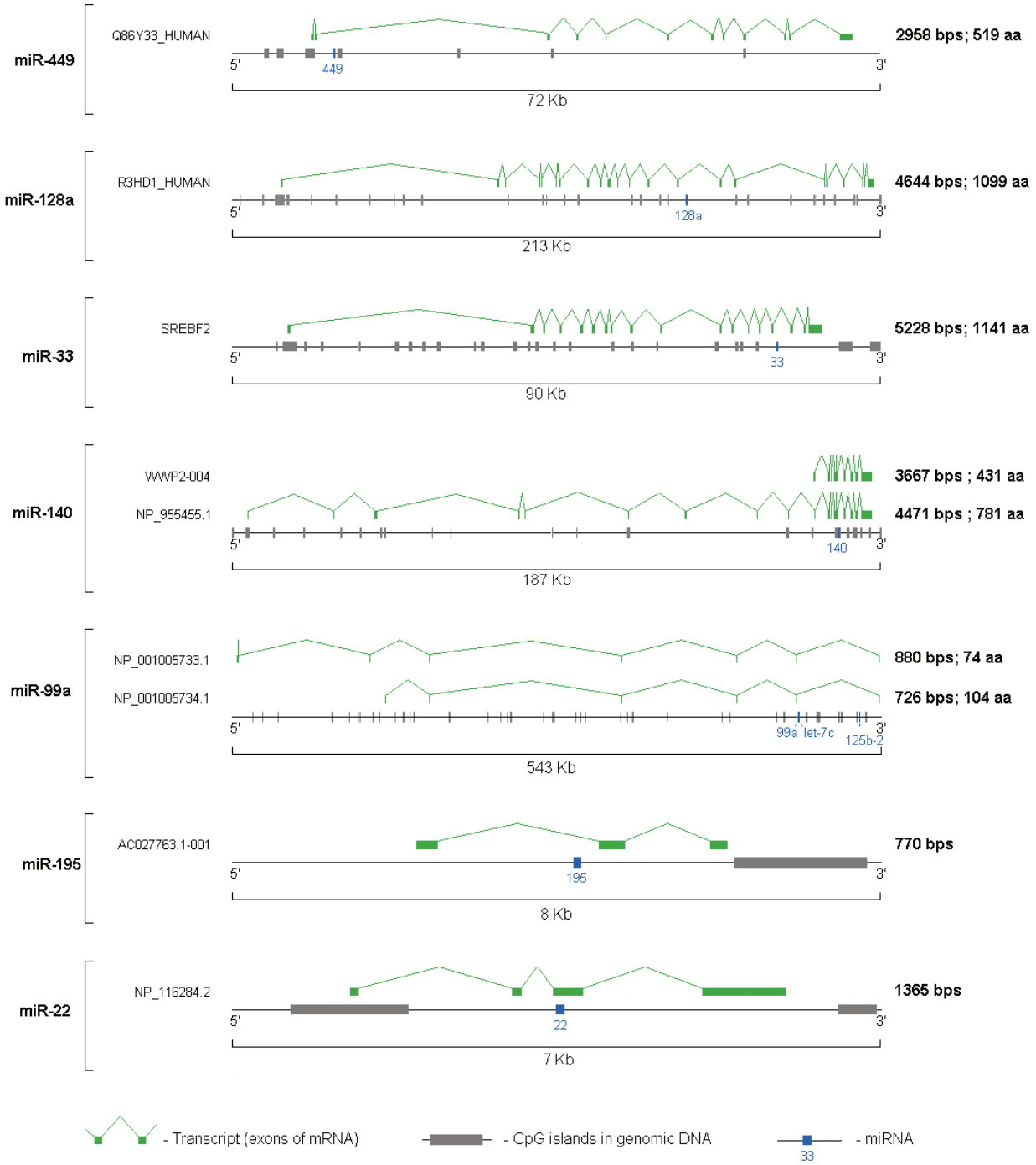


Fig. 20 - Genomic organization of the transcripts that surround the intragenic miRNAs (miRNA-449, 128a, 33, 140, 99a, 195 and 22) and CpG island localization on the same genomic area. miRNA localization is represented in blue.

Similarly to the mapping results obtained for the main transcripts involved in the miRNA maturation pathway, the maps of the transcripts surrounding the intragenic miRNAs showed that these transcripts are composed of widely divergent numbers of exons (ranging form 3 for AC027763.1-001 to 22 for R3HD1_HUMAN) and that they are part of genes also significantly different in size (ranging from approximately 4 Kb for AC027763.1-001 to approximately 530Kb for NP001005733.1).

For 5 of the 7 transcripts analysed, the genomic organization revealed a striking correlation between the position of their first exon and the position of CpG islands. The exceptions to this are the transcripts surrounding miRNAs 99a and 195. From the 5 candidate cases, the transcript surrounding miRNA-140 shows two splicing variants with the first exon located in different CpG islands. For this reason, it was not considered for subsequent analysis that concentrated only on the transcripts surrounding miRNAs-22, -128a, -33 and -449.

As for the case of the transcripts involved in the maturation pathway, the expression of these 4 transcripts was analysed.

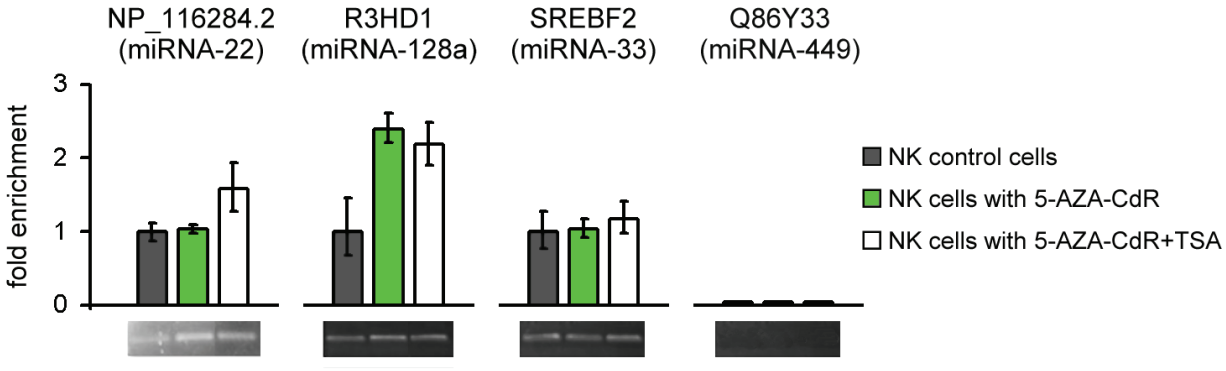


Fig. 21 - Comparison between the expression of the transcripts surrounding miRNA-22, 128a, 33 and 449 in a NK control cell sample, in a 5-AZA-CdR treated NK cell sample and in 5-AZA-CdR+TSA treated NK cell sample. The results represented in the bars were obtained by real-time PCR. Measurements were made in duplicate or triplicate. The results were normalized first by comparison with β-actin expression (ΔCt) and then by comparison with the expression verified for the NK control cell sample (ΔΔCt). The bands in lower part were obtained in independent end point PCR reactions, being the fragments run on an agarose gel stained with ethidium bromide.

The present results of the expression of surrounding transcripts of 4 up-regulated miRNAs, show a significant contrast with the results obtained previously for the expression of the 4 miRNAs themselves (Fig. 19A). The expression of the transcripts NP_116284.2, R3HD1 and SREBF2 was detected in all samples tested. The exception is the transcript surrounding miRNA-449 whose expression was not detected in any sample. Although poor primer design or any other technical problem occurring specifically during the amplification of this transcript cannot be excluded, the transcript seems to be absent in NK cells.

In terms of up-regulation, also a contrast is found when these results are compared with the ones obtained for the miRNAs themselves (Fig. 19B). For the transcript NP_116284.2 no up-regulation was observed when 5-AZA-CdR alone was introduced on the culture. Only when 5-AZA-CdR was introduced cumulatively with TSA, an up-regulation of approximately 1,5 fold was detected. For the transcript R3HD1, on both samples containing 5-AZA-CdR an up-regulation of approximately 2,3 fold was detected. For the SREBF2 transcript, no up-regulation was detected when 5-AZA-CdR was introduced alone or cumulatively with TSA.

Despite the up-regulation verified initially for miRNAs-22, 128a, 33 and 449 (Fig. 19) and the correlation between the position of the first exon of the surrounding transcripts and CpG islands (Fig. 20), only for one of the surrounding transcripts up-regulation was detected. In order to clarify the putative effect of 5-AZA-CdR on the expression of the surrounding transcripts, in special on the one surrounding miRNA-128a, an analysis of the methylation status of the CpG island where the first exon of the same four transcripts is located in, was performed. That analysis was done by genomic sequencing of bisulfite converted DNA from a control NK cell sample.

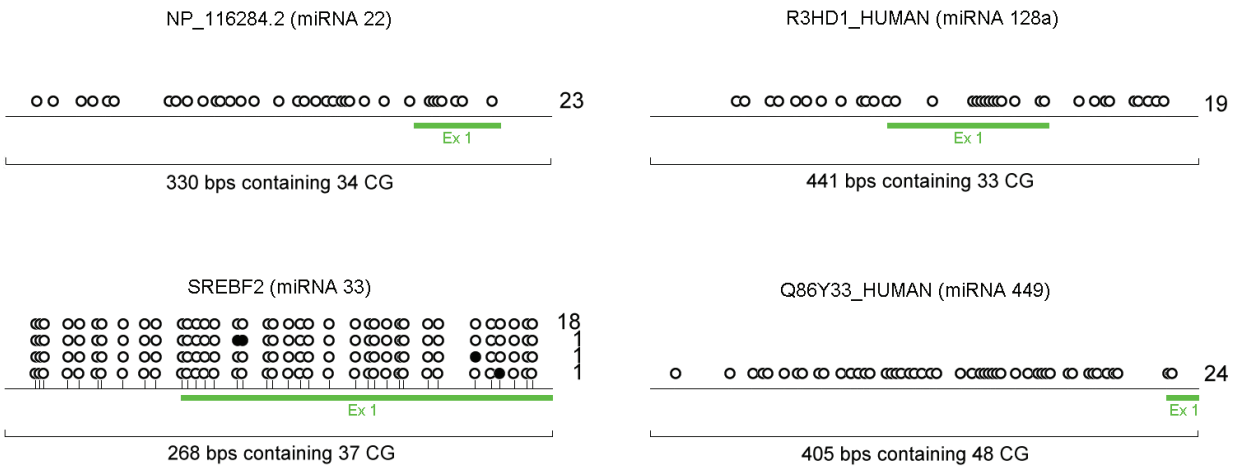


Fig. 22 - Methylation status analysis of DNA from a control NK cell sample for the regions near by the first exon (represented as Ex 1 in the figure) of the transcripts surrounding miRNA-22, 128a, 33 and 499. White dots represent demethylated CpG dinucleotides and black dots represent methylated CG dinucleotides. The number of clones analyzed that showed the represented sequence is indicated on the right side of each scheme.

Approximately 20 clones were analyzed in each experiment. The results showed that three of the four analyzed regions are completely demethylated and one exhibits only marginal DNA methylation in control NK cell samples in agreement with the insensitivity to 5-AZA-CdR observed for the transcripts (Fig. 21).

4.5 - Location of up-regulated miRNAs on CG-rich regions, highly methylated in NK cells

Following the last results, in order to continue the investigation on the cause of the observed effect of 5-AZA-CdR on the 16 up-regulated miRNAs, the analysis turned to concentrate on the miRNA loci itself. In a first step, the miRNA surrounding region was analyzed to elucidate any relation with CpG islands location. For this analysis, this time miRNA-194 was also included due to the close proximity of one of its locations (miRNA-

194-2) with miRNA-192. miRNA-9 was still not considered due to its several genomic locations all independent from any other induced miRNAs.

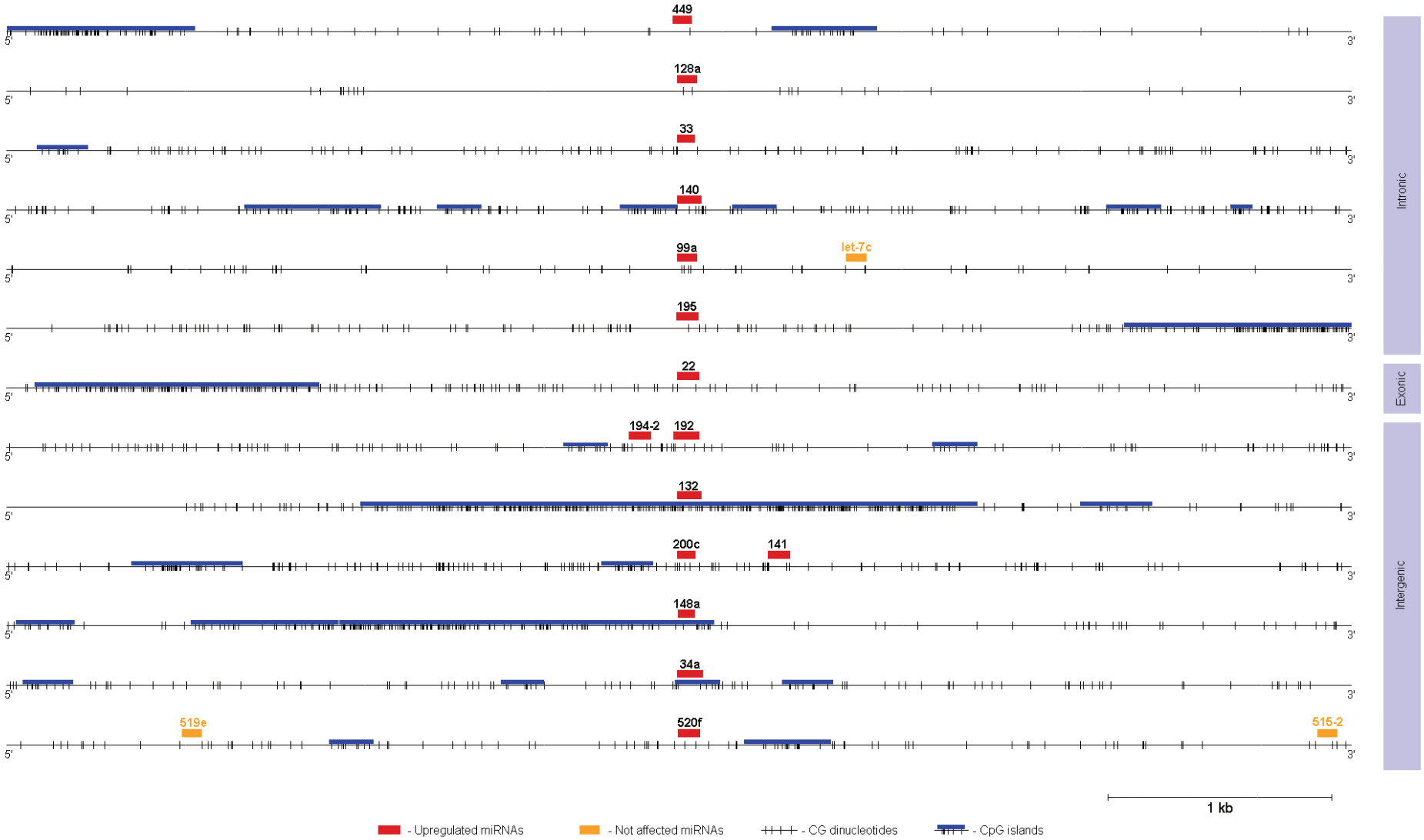


Fig. 23 - Genomic organization of a 6Kb region containing the 16 up-regulated miRNAs of Fig. 19. The analysis focused on the location of CG nucleotides and CpG islands. Information on the genomic context of the miRNA in terms of the correlation with surrounding transcript is presented on the right side.

The analysis of the miRNAs surrounding regions revealed for some cases a colocalization of miRNAs and CG-rich regions, many of them falling into the category of classical CpG islands. Colocalization is more evident for the seven first intergenic miRNAs (localized in 5 loci) all of them within or in close proximity to a CpG island. The intragenic miRNA-140 appears also within a CG-rich region between two well defined CpG islands.

To investigate if the CpG islands co-localizing with miRNAs could have a direct effect on the initially verified expression levels and 5-AZA-CdR mediated inducibility of these same miRNAs (Fig. 19), the methylation status of those islands was analyzed by genomic sequencing of bisulfite converted DNA of control NK cell samples. Analysis was concentrated on the miRNAs presenting highest correlation with CpG island positioning according to the last figure (miRNA-194-2/192, 132, 200c/141, 148a, 34a, 140). Although showing a good correlation, analysis was not done for miRNA-520f because the location of defined CpG island downstream the miRNA location was not considered as supportive of the idea of a regulatory role of that CpG island on the putative promoter of the miRNA. For the analysis were defined fragments of approximately 400bps located upstream the miRNAs.

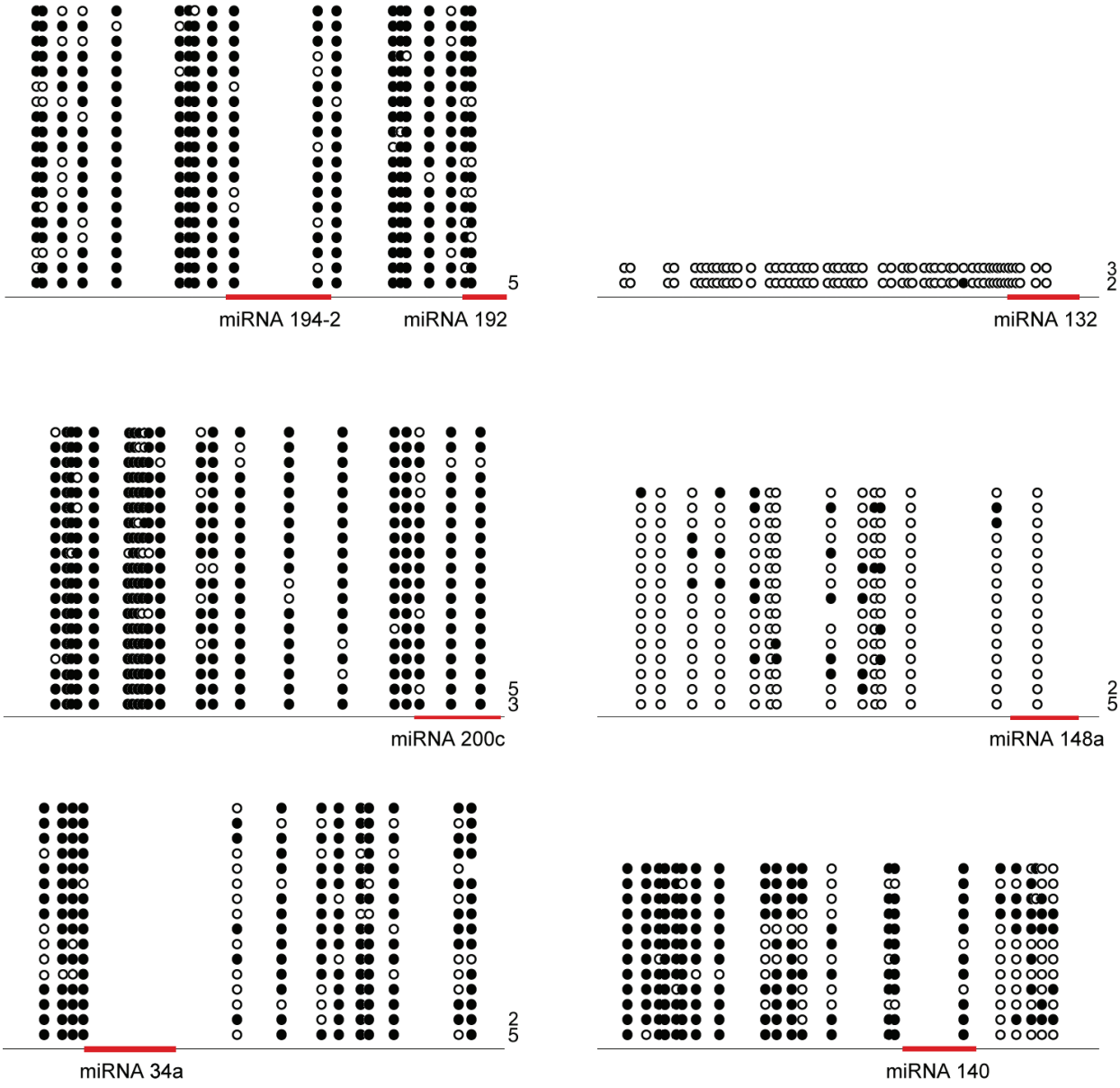


Fig. 24 - Methylation status analysis of DNA from a control NK cell sample done for the regions surrounding miRNA-194-2/192, 132, 200c/141, 148a, 34a and 140. White dots represent demethylated CpG dinucleotides and black dots represent methylated CG dinucleotides. The number of clones analyzed that showed the represented sequence is indicated on the right side of each scheme.

The number of CpG dinucleotides on the analyzed fragments ranged from 13 for miRNA-34a to 54 for miRNA-132. The analysis of the methylation status of those miRNA surrounding regions revealed two different situations. For miRNA-192/194-2, 200c/141 and 34a a very high level of methylation in the control NK cells sample was observed. For miRNA-148a a low methylation was detected and for miRNA-132 one same CG dinucleotide was found to be methylated in two clones. For this last miRNA the number of clones analyzed for this particular sequence is very low (n=5) due to low efficiency in the cloning procedure. For the intragenic miRNA-140, high methylation was observed.

These results showed that the four loci containing a total of 6 miRNAs-192/194-2, 200c/141, 34a and 140, found to be absent or to have a low expression in control NK cells, and to be inducible by 5-AZA-CdR (Fig. 19), are associated with highly methylated CpG in those cells. To explore that correlation also in other cells types, an expression profile database called smiRNAdb (available on http://www.mirz.unibas.ch/smiRNAdb[230]) was utilized to identify cell lines where these 6 candidate miRNAs are present in higher abundance.

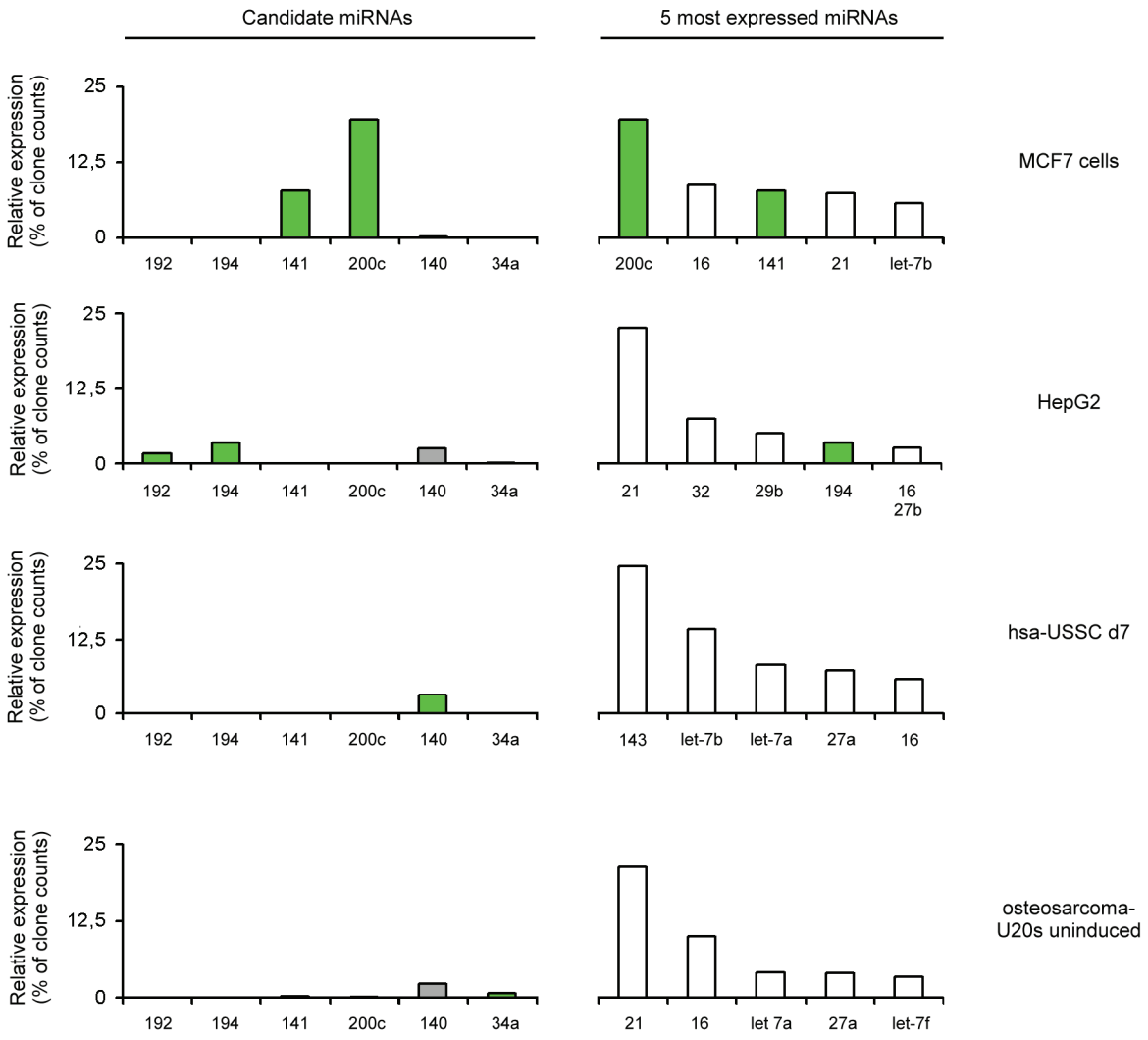


Fig. 25 - Relative expression of the 6 candidate miRNAs from Fig. 24 in several cell lines. The values were obtained by dividing the number of clones containing a specific miRNA by the total number of miRNA containing clones obtained in the cloning procedure. For each cell line the 5 miRNAs with highest relative expression, are presented on the right part of the graphic and on those, highlighted in green, are the ones also belonging to the group of 6 candidates.

On a global view of the database, it was possible to observe that the expression profile of miRNAs 141 and 200c and also 194 and 192 show a significant level of concordance. Among the group of 15 samples showing highest expression of miRNA-200c only 4 do not belong also to group of the 15 samples showing highest abundance on miRNA-141. A similar situation was observed for the other pair miRNA-194-2/192.

miRNA-200c shows its maximum relative expression in the breast cancer cell line MCF7. In that sample, the miRNA-200c is the highest expressed counting for 19,6% (156 clones) of all detected miRNAs (795 clones). miRNA-141 also shows a high expression being the 3rd more abundant with 8,7% relative counts (69 out of 795 clones).
The other candidate miRNAs do not show such levels of abundance in any cell sample present in the database. miRNA-194 shows its highest relative expression in hepatocell-carcinoma-Huh7-HCV cells (data not shown) counting 5% (3 out of 60 clones) being this sample also the 2nd more enriched in miRNA-192 with 2,5% (1 out of 60 clones). Nevertheless, this sample is not very well characterized on the database once that only 60 clones in total were obtained from it. Among the samples in which miRNAs-194 and 192 show the highest relative expression, there are other better characterized hepatocyte samples namely the HepG2 cell line. This is the second richest sample in miRNA-194 (3,5%, 47 out of 1338 clones) and the third in miRNA-192 (1,7%, 23 out of 1338 clones). In those cells, miRNA-194 is the 4th more expressed while 192 is the 16th.
The richest cell type in miRNA-140 is the hsa-USSC-d7 but only with 3,4% of the counts (52 out of 1517 clones). In this cell, miRNA-140 is the 14th in the list of the most expressed ones. In its turn, the cell type with highest expression miRNA-34a is the osteosarcoma-U20s-uninduced but counting just for 0,7% of the total counts (13 out of 1818 clones). In this sample, this miRNA occupies the 35th position in the ranking of the most expressed ones.

Since the two clustered pairs of miRNA-200c/141 and miRNA-194-2/192 were among the top 5 most expressed in a specific cell type (MCF7 and HepG2 respectively), analysis of the methylation status for the loci containing these miRNAs was performed by genomic sequencing of bisulfite converted DNA of samples from the corresponding cell lines. The objective was to better understand if DNA demethylation would be associated with the high expression observed.

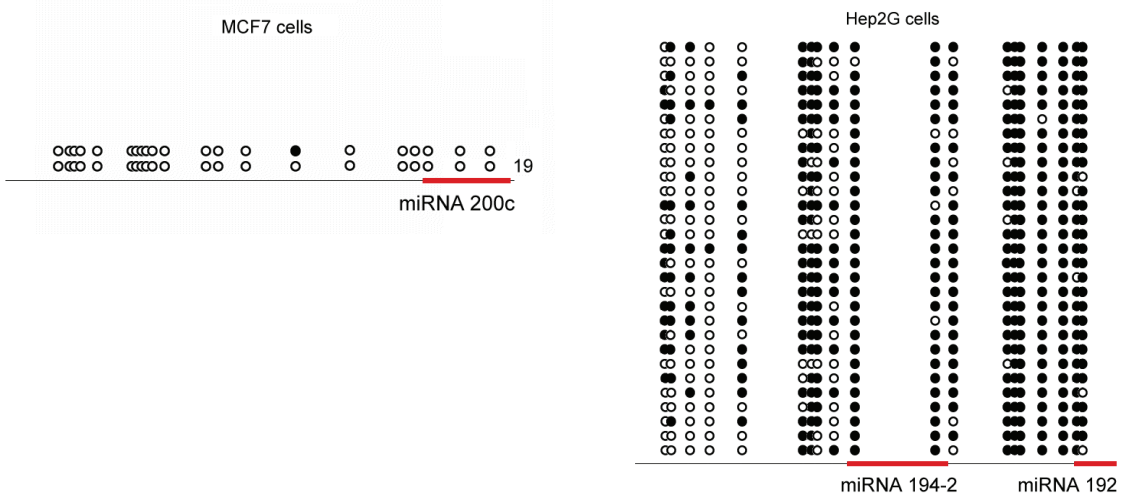


Fig. 26 - Methylation status analysis of MCF7 and HepG2 cells DNA, done for the regions surrounding miRNA-200c/141 and 194-2/192 respectively. White dots represent demethylated CpG dinucleotides and black dots represent methylated CG dinucleotides. The number of clones analyzed that showed the represented sequence is indicated on the right side of each scheme.

20 clones were analyzed in the cloning procedure involving MCF7 DNA and 29 involving HepG2 DNA. They show that the miRNA-200c/141 CpG cluster is fully demethylated in MCF7 cells and that the miRNA-194-2/192 CpG cluster is partially demethylated in HepG2 cells. In this last cluster, the demethylation is only significant in the 5' part of the CpG island especially in the first 5 CG dinucleotides.
The results associate a DNA demethylated status with high expression levels of these miRNAs in the tested cell samples (Fig. 25) and represent a significant contrast with the higher methylation status observed in DNA of control NK cells in which miRNA expression was low (Fig. 19 and Fig. 24).

4.6 - Localization of the transcription starting site for the pri-miRNAs

Although the sequence of the miRNA precursor form of approximately 70 nucleotides (pre-miRNA) is well defined, little is know about the primary RNA molecules from which these pre-miRNAs are originated from, the pri-miRNAs. According to current knowledge on mRNA transcription regulation, the influence of DNA methylation on transcription happens when the methylation occurs at the promoter region often including the first exon of the transcript[231].
In order to better understand how the methylation status of the analyzed regions could influence the miRNA expression, the transcription starting site of the pri-miRNAs was investigated. In a first approach, an *in silico* analysis was done for the four loci containing the 6 candidate miRNAs-192/194-2, 200c/141, 34a and 140. A region of 1,5Kb was considered and analyzed in terms of genomic conservation among 17 mammalian species (Chimp, Rhesus, Bushbaby, TreeShrew, Rat, Mouse, GuineaPig, Rabbit, Shrew, Hedgehog, Dog, Cat, Horse, Cow, Armadillo, Elephant, Tenrec). In parallel the same regions were analyzed *in silico* for presence of predicted Pol II promoters using a Neural Network Promoter Prediction tool (NNPP version 2.2) available on-line (http://www.fruitfly.org/seq_tools/promoter.html). This tool questions the input sequences to a database populated with *D. melanogaster* and *H. sapiens* promoter sequences.

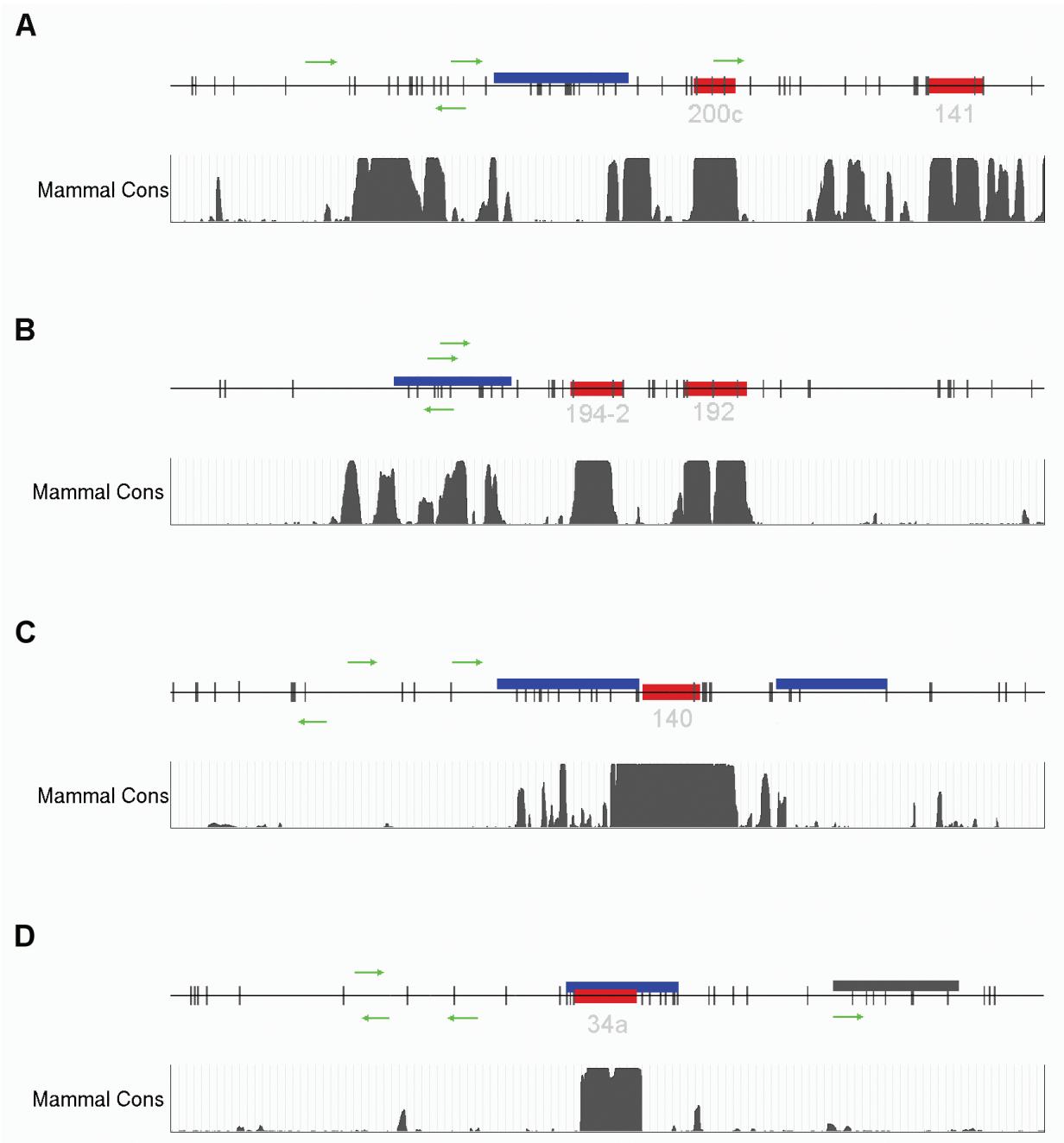


Fig. 27 - *In silico* analysis of the 1,5Kb genomic region surrounding the 6 candidate miRNAs. Analysis was focused on the location of CG nucleotides (short vertical bars), CpG islands (long horizontal blue bars), *in silico* Pol II promoter (green arrows) and conservation among 17 mammal species (area plot). Higher conservation is represented by higher values on the graph. Pre-miRNAs are represented by red boxes.

This *in silico* analysis showed that all miRNAs spots are extremely conserved among the 17 mammal species considered in this analysis. Concerning the surrounding regions, two distinct situations are observed. For miRNAs 200c/141 and 194-2/192 it is visible that the existence of highly conserved motifs among mammalian species not just on the miRNA spot but also shortly upstream the miRNAs location correlating also with predicted Pol II promoters. For miRNA 140 and 34a, although Pol II promoters are also predicted for the upstream region, the conservation analysis does not reveal any conserved motifs among mammals besides the miRNA spots. In addition, the correlation between the predicted

promoters and CpG dinucleotides context is different. On the two first cases, the promoters are located in regions with much higher CG content than on the two last ones. The same fragments were also analyzed for the existence of transcription factors binding sites (TFBS) by the AliBaba2.1 on-line tool (http://gene-regulation.com/pub/programs.html). The analysis revealed 174 TFBS for the fragment containing miRNA-200c/141; 148 TFBS for the fragment containing miRNA-194-2/192; 133 TFBS for the fragment containing miRNA-140; and 156 TFBS for the fragment containing miRNA-34a. The distribution of these TFBS was homogeneous along the 1,5 kb of the fragments.

For all the reasons stated above, attention was focused on the two pairs, miRNAs-200c/141 and 194-2/192, and to gain experimental insights into the transcription initiation site of the primary transcripts containing the miRNAs 200c and 141, a 5' RACE experiment was done. The source of RNA for this experiment was MCF7 cells due to their high expression of these two miRNAs.

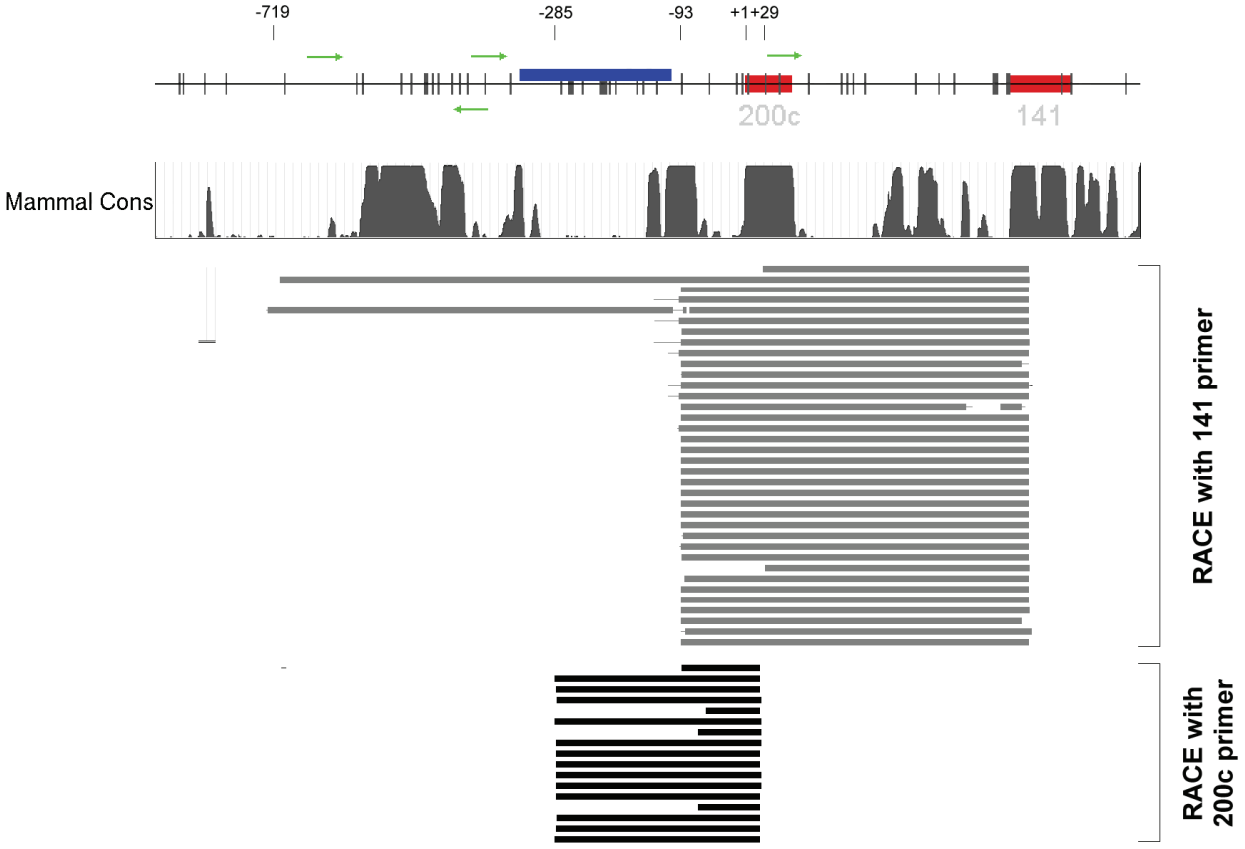


Fig. 28 – Results of the 5' Race experiment. The horizontal bars in the lower part represent the sequence obtained in single clones resulting from the cloning procedure. The bars are in scale and aligned with the mammalian conservation plot, Pol II promoter analysis, CpG dinucleotides and CpG islands of the same 1,5Kb genomic region. Numbers above the plot refer to the relative positions in terms of nucleotide distance to the 1st nucleotide of pre-miRNA-200c.

The results allowed the identification of two main transcription initiation sites and two other "marginal" ones. The main transcriptional unit containing miRNA-141 starts at position -93 and contains also miR-200c. The transcription initiation site of that transcript coincides with the transcription initiation site detected in 4 clones obtained when using the 200c primer. That same transcript seems not to be the most abundant one containing miRNA-200c. Considering the RACE results obtained with 200c primer, a more abundant transcript initiating around position -285 was detected. Two other "marginal" transcript initiation sites were also identified, one longer starting around position -719 and one shorter starting at position +29, both detected only with usage of 141 primer.
All the four sites detected do not correlate with any predicted splicing acceptor site (intron/exon boundary) and that strength the idea that they are transcription starting sites and not intron starting sites belonging to a TU starting further upstream and already partially processed.

Notably the main transcription initiation site identified (-285) is located just inside the particular CpG island from which the methylated status was previously shown to be associated with the expression levels of the corresponding miRNA (Fig. 24 and Fig. 26). Also the second main transcription initiation site is located very close to the same CpG island.

In the pursue of a better characterization of the 1,5Kb region surrounding miRNAs-200c/141, a sequence analysis focusing several aspects was performed. No core TATA-box (TATAAA/T) or CAAT-box (GGNCAATCT) promoter elements were detected. Four core GC boxes (GGGCGG) were discovered at positions -862, -440, -18 and -9. In terms of transcription factor binding sites, analysis identified 174 segments as potential binding sites (http://www.gene-regulation.com/pub/programs.html). There was no visible predominant localization of those sites being homogeneously spread along the sequence. Putative Transcription Starting Site prediction pointed a site between positions -538 and -531 (http://rulai.cshl.edu/software/index1.htm) but it does not seem to correlate with the experimental data obtained on the 5' RACE experiment. Also, no core human poly(A)-sites (AATAAA or ATTAAA) were detected (http://rulai.cshl.edu/software/index1.htm) suggesting that the experimentally verified TU have their 3' end downstream of the 1,5Kb region analyzed.
Noticeable, a predicted splicing donor site (boundary exon/intron) exists between miRNA-200c and miRNA141, at position +366 (gcaactgGTgagcgc), with a very high cut off value (0,98 out of 1) being likely its functional relevance (http://www.fruitfly.org/seq_tools/splice.html).

The binding capacity of Pol ll *in vivo* was checked using the available on-line data from a previous publication in which the high throughput method of Chromatin Immunoprecipitation-on-chip (ChIP-on-chip) was used.

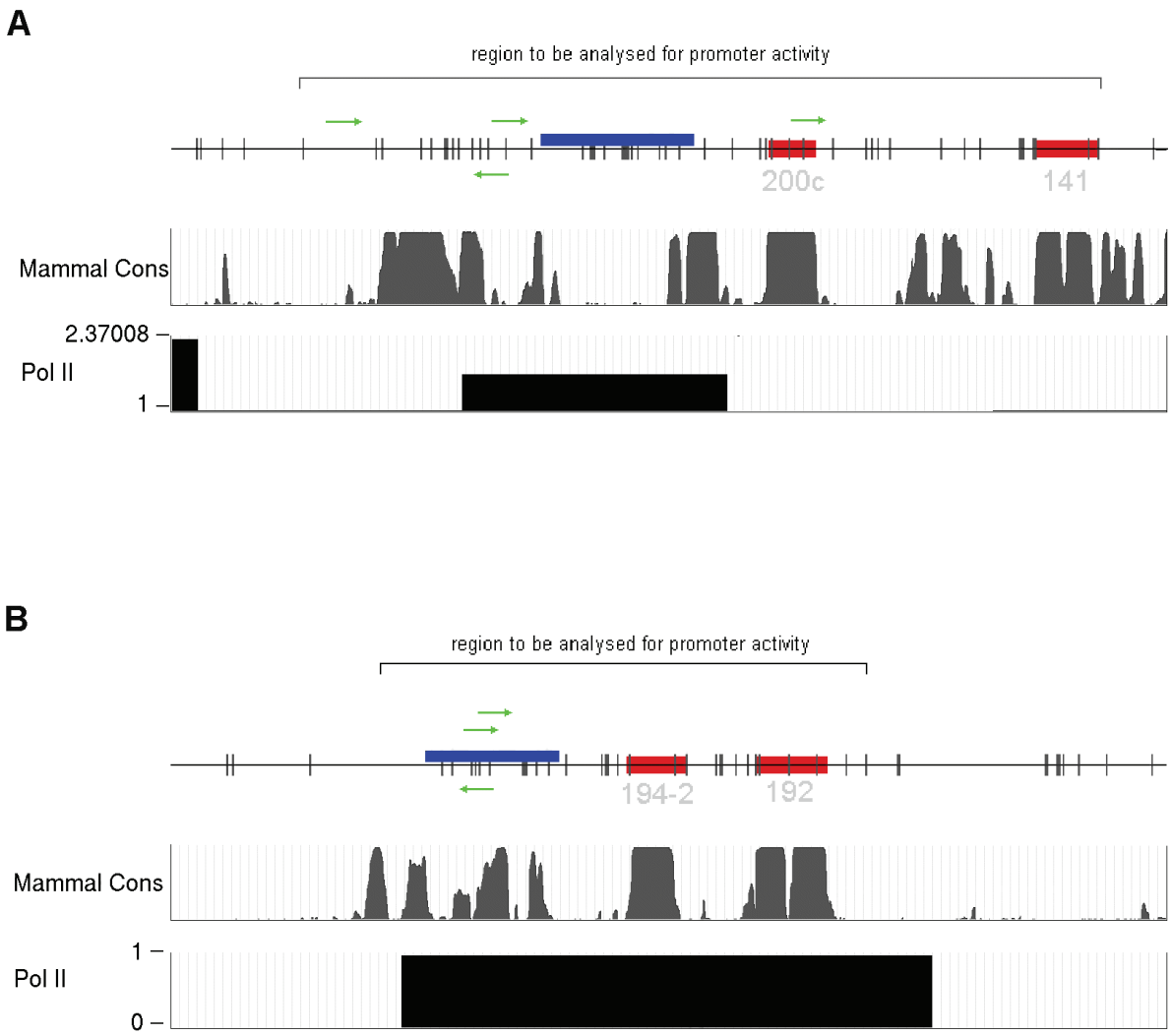


Fig. 29 - Pol II binding capacity data extracted from an on-line databank. On each scheme **A)** and **B)** a 1,5Kb region is represented on scale. Signal at Pol II section reveals binding capacity and the data is aligned with the previously shown results of Fig. 27 focusing on the mammal conservation, Pol II promoter analysis, CpG dinucleotides and CpG islands of the same genomic region. A region defined for subsequent promoter analysis is also represented.

The results show that on the experimental model used on the ChIP-on-chip experiment (CD4 cells) Pol II has *in vivo* capacity to bind a fragment that colocalize with the defined CpG islands, with some conserved motifs, with some predicted Pol II promoter and with the experimentally verified transcription starting sites.

4.7 - Identification of the promoter region of miRNAs 200c/141

The previous results guided the selection of candidate fragments in order to test for *in vivo* promoter activity of these areas. Fragment definition was based on information about the conserved motifs, the predicted Pol II promoters, the CpG islands and also the fragments to which Pol II binding capacity was detected. The idea was to choose a fragment that would include both miRNAs of each pair and their selected surroundings. The defined fragments consisted of 1207bps for the miRNAs-200c/141 and of 725bps for the miRNAs-194-2/192.

The defined fragments were cloned upstream of the luciferase transcription unit of the pMOD vector to use in standard gene reporter assays. A fragment consisting of 730 bps belonging to the 5th intron of KIR 2DL3 and with no predicted promoter activity was used as a negative control. A construct with the GAPDH promoter was used as positive control.

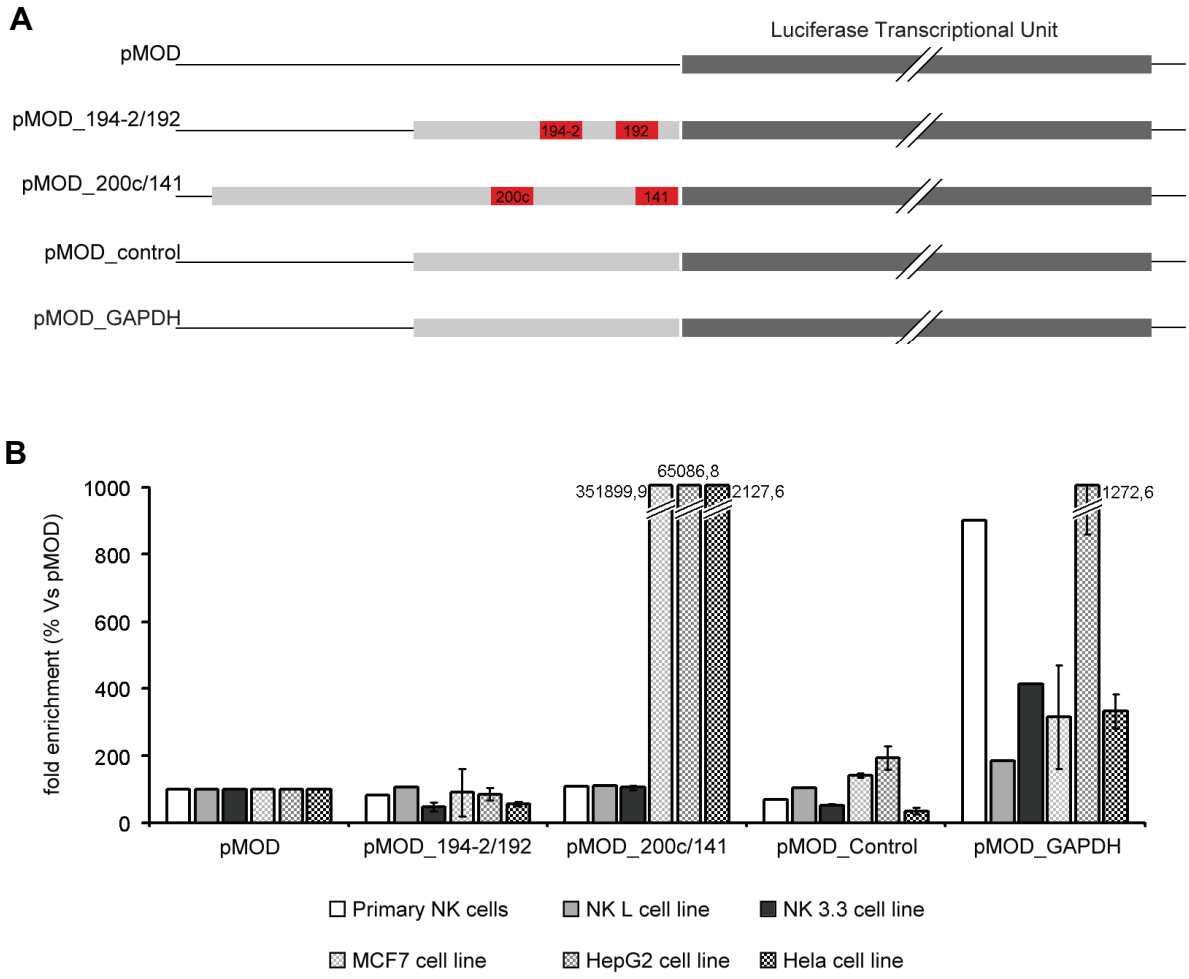


Fig. 30 - A) Schematic representation of the pMOD constructs used for gene reporter assays. B) Results of gene reporter assays obtained from experiments in which three different NK cell samples were nucleofected and three different cell line samples were transfected with the pMOD constructs. The luciferase activity values were normalized first based on the protein content and subsequently compared to the values obtained for the pMOD without any cloned promoter.

The luciferase gene reporter assay showed that the construct containing the miRNA-194-2/192 and its surroundings, did not show any promoter activity beyond the basal promoter activity of the pMOD vector without any cloned promoter. In turn, the 1207bps fragment

containing miRNA-200c/141 and its surroundings, showed a very strong promoter activity in MCF7, HepG2 and Hela cells. This activity was much higher than that of the GAPDH construct normally used as positive control. The results were confirmed by cloning the same fragments into pGL3 vector, another luciferase expressing vector (data not shown). The promoter activity obtained for the fragment containing the miRNA-200c/141 continued to be strongest, although a much stronger promoter activity was observed for the GAPDH promoter sequence using the pGL3 constructs.
For unknown reasons, in the MCF7 and HepG2 cell lines, the construct used as negative control showed a promoter activity higher than the basal promoter activity of the pMOD vector without any cloned promoter.

The results show the existence of a strong promoter located on the fragment that also contains miRNA-200c/141. On the artificial construct used, the promoter lead to an over-expression of luciferase gene but in fact the results do not show that the promoter can also drive the transcription of units that later can generate mature forms of miRNAs. On other words, the results do not show that the promoter detected is the miRNA promoter.

In order to prove that the promoter present in the cloned fragment containing could indeed lead to the expression of mature forms of the miRNAs contained on it, mature miRNAs expression was verified by northern blotting using probes complementary to both miRNAs in pMOD and pMOD_200c/141 transfected cells. In the experiment, Hela cells were used once that they show a low natural expression level of the two miRNAs and that allows more easily the identification of any over-expression effect. As the time taking from expression to complete maturation of miRNAs is not known, in the experiment were considered 3 time points (24h, 48h and 72h).

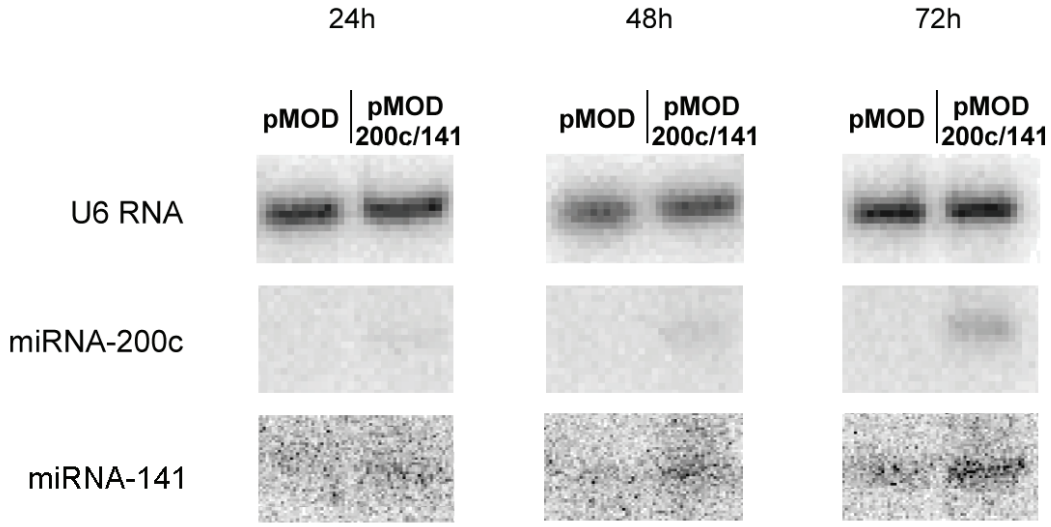


Fig. 31 - Northern blot results using independently a U6 RNA probe, a miRNA-200c probe and a miRNA-141 probe. The bands shown for the miRNAs correspond to the mature form of the miRNAs (22nt for miRNA-141 and 23nt for miRNA-220c) and not to any larger precursor.

In these northern blot results it is possible to observe an increasing expression of the miRNA-200c and 141. The expression shows time dependence, increasing during all the time course of the experiment and so the expression becomes more evident 72h after transfection.
These results show that the promoter cloned is, indeed, able to lead to the formation of mature forms of the present miRNAs.

On the same time points (24h, 48h and 72h), luciferase activity was also measured in order to monitor the promoter activity.

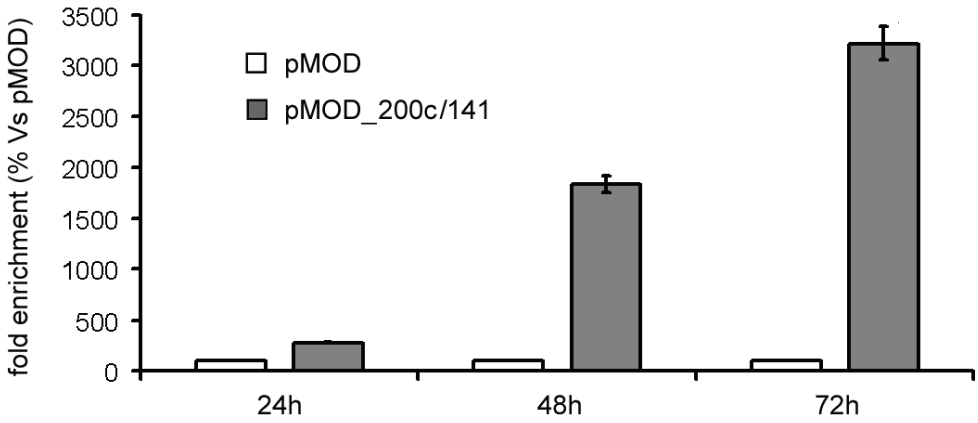


Fig. 32 - Gene reporter assay results done in Hela cell, 24h, 48h and 72h after transfection with pMOD or pMOD_200c/141 construct. Data was normalized against protein content and then pMOD_200c/141 data was normalized with pMOD data so that the scale represents a comparison in terms of fold enrichment.

The results showed again the high promoter activity of the cloned fragment and confirmed the time dependence of that activity already observed in the northern blots presented on the previous figure. Combining the results of this figure and the ones present on figure 29, is possible to state that the promoter identified is able to drive a simultaneous expression of the miRNAs and the luciferase transcription unit located downstream.

4.8 - Functional relevance of the methylation status for the activity of the identified promoter

The pMOD vector (see material and methods section) used in the previously reported gene reporter assays contains a CpG-less luciferase transcriptional unit. Performing an *in vitro* methylation of pMOD constructs, methylation of CpG will occur in other regions of the construct without interfering with the luciferase transcriptional unit. This feature makes this vector suitable to evaluate the direct influence of DNA methylation on the *in vivo*

activity of the promoter present, since it is the only region of the transcriptional unit able to be affected by the mentioned *in vitro* methylation.
In vitro methylation of the vector was performed and, subsequently, the methylation status was verified by digestion of both the native and methylated forms of the construct, with the methylation sensitive restriction enzyme BstUI that recognizes CGCG motifs.

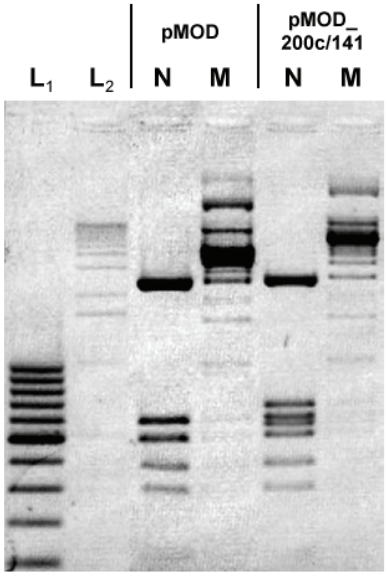


Fig. 33 - Agarose gel electrophoresis of the fragments obtained by digestion of both native (N) and *in vitro* methylated (M) forms of the pMOD vector and pMOD_200c/141 construct with BstUI restriction enzyme. On the left is shown 100bps size ladder (L_1) and 1Kb size ladder (L_2) used to evaluate the sizes of the originated fragments.

The picture taken to the gel shows two distinct band patterns obtained after digestion of the native (N) and the methylated (M) form of the pMOD vector and pMOD_200c/141 construct. The fragments resulting from the digestion of the methylated forms have a higher molecular weight than the fragments obtained from the native ones. This shows that after *in vitro* methylation the vectors were more protected from the BstUI digestion confirming that an extensive construct methylation was achieved with the method used.
Noticeable is the fact that, although the luciferase transcription unit does not contain CpG dinucleotides, these seem to be present in other regions of the vector since the BstUI enzyme is specific for the CGCG sequence motif.

Using the native and the methylated forms of the pMOD-200c/141 construct, a gene reporter assay was performed using again Hela cells.

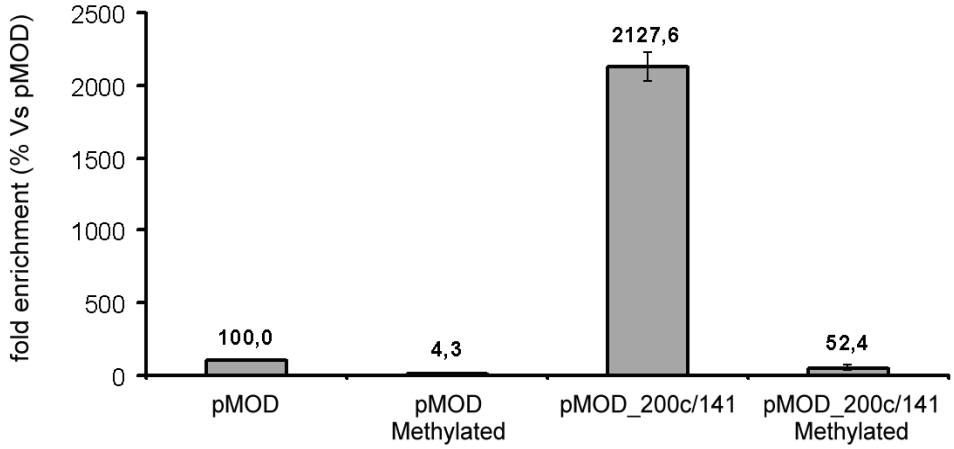


Fig. 34 - Gene reporter assay results done in Hela cells, 48h after transfection with pMOD, pMOD_200c/141 or *in vitro* methylated pMOD_200c/141 construct. Data was normalized against protein content and then pMOD_200c/141 and *in vitro* methylated pMOD_200c/141 data were normalized with pMOD data so that the scale represents a comparison in terms of fold enrichment.

The results show a significant reduction of the basal expression level of pMOD vector when it was methylated but more importantly, these results show also that the strong activity of the promoter contained in the pMOD_200c/141 construct was also almost completely silenced after *in vitro* methylation. These observations prove definitively that methylation has an important role in the activity of the identified promoter and substantiate the previous results of Fig. 24 and Fig. 26 showing an association between the methylation status of that same area and the expression of the miRNA in different cell types.

5 - Discussion

The networks of factors and co-factors leading to a precise gene regulation have been one of the major focuses of cellular biology in the last decades. The complexity of interactions between genes and their known regulators are not able to explain all the observed phenotypes and so there is still the necessity to find other players of this game. miRNAs were recently recognised as important elements of those networks. In the last 4 - 5 years, the knowledge on their capacity to regulate gene expression has increased exponentially and subsequently, other questions started to arise, namely about the mechanisms regulating the expression of the miRNAs. After all, what regulate the regulators? Therefore, this work aimed to contribute to a better understanding of regulation of miRNA expression, exploring the putative role of DNA methylation. In the present work, initially a comparison was performed between the expression of miRNA in primary Natural Killer (NK) cells kept in culture under standard conditions and others kept in culture in the presence of the DNA demethylating drug 5-AZA-CdR. This drug is widely used in epigenetic studies and is clinically approved and used in several therapeutic protocols of patients with haematological tumours. Understanding the molecular basis of 5-AZA-CdR action on the cells has attracted the attention of many researchers due the possible application of this drug for clinical purposes.
The effect of 5-AZA-CdR on NK cells was already reported. A previously published work shows that treatment of NK cells with this drug leads to an increased expression of KIR receptors as a direct consequence of KIR loci DNA demethylation[232]. As the expression of KIR receptors can be monitored by flow cytometry, this constitutes a convenient read-out system to follow the effect of 5-AZA-CdR on these particular cells. Therefore, for the purposes of the current work, the KIR over-expression observed on day 26 of culture on the 5-AZA-CdR treated cells sample was considered as a good indicator of a global DNA demethylation effect of the 5-AZA-CdR in NK cells (dotplots n.3 on Fig. 14). Besides the desired DNA demethylating effect, 5-AZA-CdR is reported to be cytotoxic leading to increased cell mortality. This was also observed during the experiments performed. It is interpreted as a consequence of 5-AZA-CdR to be an analogue of deoxicytidine. On the cells it can intercalate with DNA forming adducts with a role on DNA synthesis arrest and, consequently, activating programmed cell death. The increased cell mortality is more evident when working with primary cells than with cell lines which show higher tolerance to the drug. However, in the present work, the cells were sorted prior to further use and only living cells were used in subsequent steps, so the increased mortality was not considered as limiting aspect of the drug usage.

The usage of DNA demethylating drugs such as 5-AZA-CdR, leads to global DNA demethylation. Despite this, early works on the effect of epigenetic factors on miRNA expression focused on the epigenetic alterations only of defined DNA regions that surround miRNAs. In spite of the crucial role that some enzymes have in the miRNA maturation pathway, there are no studies that consider the possibility that the observed changes in miRNA expression after drug treatment, could be a consequence of a regulatory

role of DNA methylation on the expression of these key enzymes. The currently increasing number of publications on the miRNA maturation pathway suggests that the enzymes act in a sequence-independent but structure-dependent way[233,234]. At the same time, the data available do not suggest that these enzymes have a higher affinity for a particular class or group of miRNAs. For this reason, if 5-AZA-CdR treatment would influence the expression of any of those transcripts, it would be expected to cause a generalized change and a drastic effect on the global miRNAs expression. Due to its crucial importance, a putative impact of 5-AZA-CdR on the expression of the enzymes of the miRNA processing machinery was covered in the present study especially after the analysis of the genomic organization of these transcripts, which support indeed an eventual regulatory role of DNA methylation in their expression (Fig. 15). However, the results of transcript detection by real-time PCR showed only a mild induction of one transcript (Dicer) by 5-AZA-CdR (Fig. 16). This is the first time that such observations are reported and, in globally, they indicate that, in spite of the visible correlation between the colocalization of the first exon of the transcripts involved in miRNA processing, and classical CpG islands, no induction of the majority of these transcripts is observed under the action of the DNA demethylating agent 5-AZA-CdR in NK cells. Therefore, it might be assumed that these transcripts are not under regulation of DNA methylation at least in healthy NK cells and that the alterations detected in miRNA expression are consequence of demethylation of other important genomic areas.

The analysis of miRNAs global expression on control samples and samples treated with 5-AZA-CdR was performed using the recently developed multiplex TaqMan® MicroRNA assay. In order to allow direct comparison of TaqMan® MicroRNA assays between samples, a novel normalization method was used. Until the time this analysis was made, several normalization methods had been proposed for miRNA expression data. Those were based on the expression of some miRNAs like miRNA-16, let-7a or miRNA-24 thought to be present across all cell types[235-237]. The main problem concerning the usage of these "house-keeping miRNAs" is the lack of consensus concerning their expression levels. In 2002 Vandesompele *et al.* published a method to determine the best reference genes in a given set[238]. This method was designed to work with a small number of possible control genes and it was so far not applied to miRNA expression datasets. Due to the lack of data about miRNA expression in NK cells and the absence of a consensual miRNA normalization factor, the referred method was seen as a valuable alternative providing consistency to the data. Additionally, every study works with different experimental models and conditions and the big advantage of this method is that it allows the definition of the best control genes for every set of samples and for every group of genes on which the study focuses.
None of the 5 miRNAs for which the expression was the most stable across all the tested samples, have been so far used for normalization purposes (Fig. 17). This fact does not compromise the obtained results neither the applicability of this method to miRNAs

expression analysis and it is seen as a consequence of the homogeneity of the samples tested (only NK cells) that contrast with more heterogenic sample groups used by others. To see if the method would point the same miRNAs when a more heterogeneous group of samples is considered, the same method was applied to a group of real-time PCR measurements from CD34+ cells and NK cells (data not shown). This resulted in a different set of miRNAs as been the most stable ones (miRNA-130b, -150, -16, let7d and let-7f) and that group included some of the "house-keeping miRNAs" proposed by other publications. This reinforces the idea that the miRNAs chosen for normalization of the NK cell experiments were exceptionally stably expressed in the particular experimental model used, and so can be considered as good for normalization purposes on this system.
The different expression levels that the 5 chosen miRNAs (Ct values Fig. 17) show among the samples, is justified by different amount of RNA obtained from the initial 1000-cell samples or by different accuracy on handling the two steps protocol of the TaqMan® MicroRNA assay. However, relevant to the normalization method is the fact that, although different in absolute values, all the Ct values maintain a very similar correlation between each other in all samples, which can be observed in the common shape of the graphs. The normalization factor constituted for each sample, and based on the geometric mean expression of these 5 miRNAs in that sample, allowed the normalization of the data from all miRNAs, the subsequent comparison of the expression between samples by the ΔΔCt method, and the determination of the normalized standard deviation. This is the first time that this method is applied to a miRNA data set and, according to the obtained results the method proves to be suitable for normalizing big real-time miRNA expression datasets.

After normalization of the data within all samples, the comparison of miRNA expression between them show that the vast majority of miRNAs (74,8%) was not influenced by 5-AZA-CdR in NK cells and only 11,1% (n=26) were up-regulated. The significant fraction of unaffected miRNAs might be reasoned by a complete independence on DNA methylation, but other facts may be playing a role, too. miRNAs are thought to have a high stability due to the secondary structures of the precursor forms as well as to the short size of the mature form. That might lead to a long biologic life time and the detection of a certain amount of miRNA in a certain moment might therefore be the consequence of a residual expression and consequent accumulation of such long-lived molecule. In such cases, the effect of a stimulus normally able to interfere with transcription level, like DNA demethylation, might be masked to some extent by that natural abundance of that RNA molecule.
Concentrating on the 26 up-regulated miRNAs, when the same results where compared with another biological replicate of the control sample, the number of miRNA considered to be up-regulated after 5-AZA-CdR treatment of NK cells decreased to 16 (Fig. 19). Although at the time this study was initiated, there was no evidence on miRNAs regulation by epigenetic factors like DNA methylation, some studies in cell lines were published in the mean while. However, cell lines are usually derived from carcinomas with extensive

proliferative history, and it is reasonable to think that their response to demethylating agents might be different from the response of a primary cell to the same agents. This is supported by existing evidence for an altered methylation patterns in cancer cells (global genome hypomethylation is normally observed[239]) and also for changes in DNA methylation upon continuous cell culturing[240].

In spite of the differences in the cell model used in the present study and the above referred ones, interestingly, the percentage of miRNAs found to be up-regulated in the present study (16/234 = 6,8%), lies in the same range of values as presented in the cell line studies. The group of P.A Jones reported an up-regulation of 17/313 = 5,4% miRNAs after treatment of T24 bladder cancer cells with 5-AZA-CdR together with PBA[241]. Knocking out DNMT1 and DNMT3b in HCT-116 cells, a colon cancer cell line, lead to an up-regulation of 18/320 = 5,6% miRNAs[242]. One other recent work reported the up-regulation of 21/197 = 10,7% miRNAs after 5-AZA-CdR treatment of malignant cholangiocytes MzChA-1[243].

Together these results suggest that, independently from the cell model used, the number of miRNAs up-regulated after treatment with the DNA demethylating agent 5-AZA-CdR is similar and modest. However, that small number can not be taken as an indication for minor cellular consequences. As a single miRNA can interact with several different mRNAs, a change on a small number of miRNAs might actually lead to a significant effect on multiple proteins and, subsequently, on the cell phenotype.

Comparing the group of 16 miRNAs up-regulated in the present work with the group of miRNAs up-regulated on the other works, only two miRNAs are found in common (miRNA-200c and miRNA-99a)[244,245]. These observations also suggest that different cells react differently to the same stimulus having a cell specific sensitivity to 5-AZA-CdR treatment.

The developed work allowed the identification of 16 up-regulated miRNAs which were the basis for further investigation on the role of DNA methylation on their expression.

5.1 - Low impact of 5-AZA-CdR on the expression of the miRNA's surrounding transcripts

As already mentioned, the DNA demethylating effect of 5-AZA-CdR is not confined to specific genomic regions but it affects the whole genome. So, an up-regulation of a certain miRNA by this drug can be a consequence of alterations on methylation possibly occurring at different genomic regions. One of the candidate regions is the one where it might be located the promoter of a transcription factor necessary for the transcription of that miRNA. Activating this transcription factor, would constitute an indirect effect of 5-AZA-

CdR on the activation of that miRNA. Alternatively, the alteration on DNA methylation can occur directly on the promoter responsible for the miRNA expression, and therefore it is necessary to locate it. On that respect, some important information can be extracted from previous publications. In 2005 the standard of knowledge was that miRNA would be encoded mostly in intergenic regions of their own transcriptional unit[246]. However, later it was recognized that the previous interpretation had been biased mainly by partial lack of knowledge of the genome. More recent and detailed genomic analysis demonstrated that the vast majority (around 70%) of the human miRNA loci are located within the intronic regions (intragenic miRNAs) of predicted or annotated transcript (protein-coding and non-coding ones) [247-249]. The exact percentage may be subject to debate since that most of these transcription units (TUs) are still not validated for their transcription capacity. Additionally, the authors of some of these studies demonstrate also that intragenic miRNAs show the same orientation of the surrounding TUs and that transcription of both largely coincides, indicating that the intronic miRNAs and their hosting genes may be co-regulated and generated from a common precursor transcript[250,251]. In this respect, the most comprehensive study is the one form Baskerville and Bartel[252], where the expression of 175 human miRNAs and their surrounding transcripts was analysed in samples from 24 different human organs. Pointing to a coordinated expression of the certain miRNAs and their surrounding transcripts, these studies suggest also that both RNA species share the same promoter which is located close to the first exon of the surrounding transcript and not necessarily close to the miRNA itself. Although suggestive, the conclusions of these publications had not been considered in later studies on the DNA methylation effect on the expression of specific miRNAs[253-255]. These studies concentrate only on the region surrounding the miRNA looking for the miRNA private promoter and ignoring its genomic context in terms of the presence of any surrounding transcripts and, additionally, their putative correlation with CpG islands.

On the present study, the analysis done to the genomic location of the 16 up-regulated miRNAs allowed to divide them in two groups: intragenic (associated with a longer surrounding transcript); and intergenic (not associated with any know transcript). Among the up-regulated intragenic miRNAs it is possible to observe that, for the big majority, the first exon of the surrounding transcript is located within a CpG island (Fig. 20). Here also this strong correlation is suggestive of a potential role of DNA methylation on expression of these transcripts, however, the analysis of the transcript expression done in the present study by real-time PCR does not support that role (Fig. 21). Only the transcript surrounding miRNA-128a showed slight induction by 5-AZA-CdR concordantly in part with the observations of Baskerville and Bartel[256] that refer this miRNA as one presenting the highest correlation with its surrounding transcript. Nevertheless, contrarily to the result of this cited study, the high inducibility verified for the miRNA in the present work is not comparable with the much lower one verified for the surrounding transcript. This can be a cell specific effect that the authors of the cited study did not report probably because they

worked with RNA from bulk cell populations from different organs and not with defined cell populations.

In addition to the differences observed in the inducibility of the miRNAs and the surrounding transcripts, the results show also differences between the expression profile of the two RNA species (Fig. 19A and Fig 21). In two cases while the miRNAs were not detected in the control NK samples, their surrounding transcripts were detected with a significant abundance (miRNA-128a and 33). Contrarily, for one other case the miRNA was detected in NK cells treated with 5-AZA-CdR while the corresponding surrounding transcript was not detected in any experimental condition including that one (miRNA-449). Summing up, the transcripts differ in most instances from the corresponding miRNAs in respect of expression and inducibility by 5-AZA-CdR in NK cells. The results also provide evidence that at least some intragenic miRNAs are regulated independently of the TU in which they are situated in. Once that the information on the regulation of longer transcripts is already abundant, the idea is very attractive that miRNAs would be transcribed together with these as part of their introns. However, by the means of the experimental evidence discussed above, and also collected by others[257], it seems that there are no general rules common to all miRNAs, and the possibility should not be excluded that of the existence of multiple overlapping transcription units and/or that a miRNA-producing precursor may not be the same molecule as the mRNA-producing one. Alternative splicing can provide a mechanism for generation of multiple transcription units.

In order to gain more insight on the putative role of DNA methylation on the expression of the surrounding transcripts of miRNAs, was performed an analysis of the methylation status of the CpG island in which the first exon of some of those transcripts is located in. The results show that the loci analysed, are almost completely demethylated in control NK cells not supporting any direct role of DNA methylation on the regulation of transcript expression in NK cells (Fig. 22). Since only methylated CpG can be affected directly by DNA demethylating drugs as 5-AZA-CdR whereby being able to affect the expression of the corresponding transcript, some methylation on the CpG island associated with a transcript that showed some inducibility by 5-AZA-CdR would be expected (transcript R3HD1). Despite the initial expectations, the analysis revealed that that locus is also already almost completely demethylated in control NK cell samples suggesting that the observed inducibility of this transcript might be a consequence of the action of other factors (eg. R3HD1 activators) induced themselves by 5-AZA-CdR. Interestingly, the methylation status found in the locus of the transcript Q36Y33 (transcript that surrounds miRNA-449) indicates that the locus can be actively transcribed, although no expression was detected. This fact suggests that, besides an epigenetically favourable status of the locus, other elements necessary for the transcription may be absent in NK cells or the locus might be under effect of a transcription repressor. A demethylated locus does not necessarily imply that the transcript is transcribed. The possibility that other epigenetic factors controlling

chromatin structure (like histone modifications) may also be playing a role in regulating the transcriptional activity of this locus, should also not be excluded.

Excluding the role of DNA methylation on the regulation of the longer transcripts surrounding miRNAs, further investigations were carried trying to locate the private promoter of the 16 up-regulated candidate miRNAs that was assumed to be in a close proximity of the miRNA.

5.2 - Some of the up-regulated miRNAs are located on CG-rich regions and DNA methylation status of those areas correlate with its expression in different cell types

At this time point, the present study focused on the role of the genomic region around the miRNAs themselves where the miRNA private promoter can be located. This was done in a way similar to some published work[258-262]. Investigating the genomic area localized around the up-regulated miRNAs, again an evident colocalization between the miRNAs and CG-rich regions (fitting into the classical CpG island definition) is observed (Fig. 23). Interestingly, that is visible for all the intergenic miRNAs and is visible only for one intragenic (miRNA-140). The extensive demethylation found for the CpG islands co-localizing with miRNA-132 and 148a loci contrast with the lack of or low expression of these miRNAs in the control NK cell samples. As discussed previously, the methylation status of the locus should be interpreted only as an indication that the locus can be actively transcribed although expression might not be observed if other necessary factors are not present. By contrast, the high methylation detected on the other 4 loci (containing 6 miRNAs) in the control NK cells sample, suggests that these miRNAs may be, in this way, repressed in these cells, which can provide explanation for the observed inducibility via demethylation. In support of this, a strong hint about a correlation between the methylation status of the locus and the expression levels of miRNAs 200c/141 and 194-2/192 was obtained after analyzing the bisulfite-converted DNA from MCF7 and HepG2 cells respectively (Fig. 26). These cells were chosen once that they show a high relative abundance of those two miRNAs, as reported by the data from the smiRNAdb (Fig. 25) and also by two previous publications[263,264]. The obvious differences in the methylation status in NK, MCF7 and HepG2 cells provide support for the differences in miRNAs levels between the corresponding cells. Demethylation of the miRNA locus seems to be associated with high miRNA expression.
Based on the high level of DNA methylation of the loci containing miRNAs 200c, 194, 34a and 140 in NK cells, one would expect very little or even no expression of those miRNAs. Nonetheless, these miRNAs were detected in NK cells (Fig. 19A). This can probably

happen because the extensive methylation alone is not able to shut off the expression entirely, and only reduces it to residual levels. The expression levels detected in the NK control sample could then result from accumulation of miRNA over time due to their high stability, and not as the outcome of fully active loci. Another possible explanation is that, for each of these miRNAs, several transcriptional units (TUs) might exist as supported by later experimental evidence. From these, some might not be under the influence of DNA methylation and would be able to drive the detected basal expression, while others, being controlled by DNA methylation, could be responsible for the observed up-regulation.

5.3 - The transcription starting site of the pri-miRNA-200c/141 is located in close proximity to the pre-miRNAs loci

In order to investigate the functional importance of the CpG islands that show a methylation status that correlate with the observed miRNA expression in different samples, an *in silico* analysis was made to the 4 loci containing the 6 candidate miRNAs (200c/141, 194-2/192, 140 and 34a) (Fig. 27). Not surprisingly, the results of the analysis clearly show that all the 6 miRNAs are widely conserved among species, showing consistency with all published studies that lead to the current definition of this small RNA molecules[265]. Additionally to the miRNA loci, the conservation analysis of the region surrounding the miRNAs was already used by other authors who successfully identified functionally important regions in the regulation of miRNA expression[266]. The concept behind it is that if miRNAs are so well evolutionary conserved, it is reasonable to suggest the same for the promoter fragments governing their expression. On the basis of this, the observation of conserved motifs that partly colocalize with some predicted Pol II promoters and with the CpG island, in 2 of the analysed loci (containing miRNA-200c/141 and 194-2/192), supports the idea that the promoter responsible for the expression of the pri-miRNAs might be located on the regions directly surrounding the miRNAs sequences. This hypothesis found also support in a recent study that estimated that miRNA promoter would preferentially lie in the region up to 500bps upstream of the miRNA[267]. The results of the 5' RACE experiment confirmed that hypothesis supporting also the idea of a putative role of that CG-rich region in transcription regulation since both major transcription starting sites are located in the CpG island or very close to it. Additionally, these sites colocalize also with conserved motifs supporting the idea that the key elements of the transcriptional units might show conservation along evolution. Interestingly those motifs do not contain any conventional Pol II promoter element in a position that could point to a crucial functional relevance for the main transcripts detected. That allow us to

say that we might be analysing a non-conventional eukaryotic promoter region similarly to other reported cases[268-270].

In addition to the discussed, noticeably the 5' RACE results show clearly the existence of several transcriptional units. Analysing the number of clones obtained in the experiment, one can expect that the expression of miRNA-141 in MCF7 cells is almost always accompanied by the expression of miRNA-200c, but it seems possible to exist expression of 200c without expression of 141. The detected transcription units might be under different regulatory mechanisms and this can be very well the explanation for the differences verified in expression levels of the two miRNAs on the initial miRNA profiling on NK cells, as well as the differences observed in the data extracted from the smiRNAdb database (Fig. 19A and Fig. 25). Interestingly, similar differences in the expression of these same miRNAs were also observed by others on different cell types, but were not discussed[271]. On the one hand, the close proximity of the miRNAs in the pairs suggests that both miRNAs share a common TU, but, on the other hand, the existence of alternative TUs, possibly generated by alternative splicing seems able to generate differences in the miRNAs expression levels (Fig. 35). This would give miRNA regulation another level of complexity.

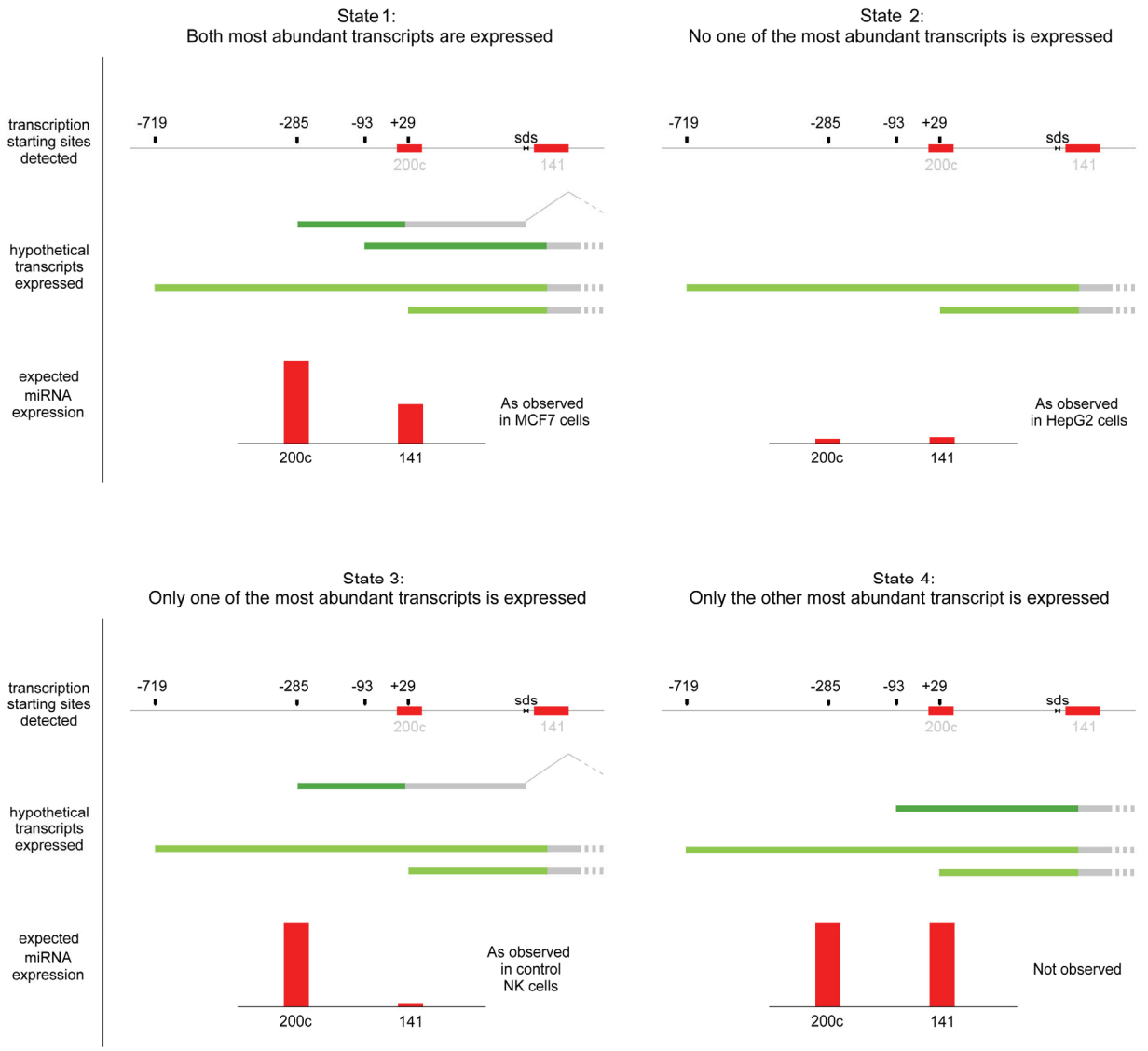


Fig. 35 - Four possible situations resulting from the possible expression of the transcripts. On the upper part of each scheme are represented the position of the miRNAs (red blocks) on the genomic DNA, the position of the transcription starting sites detected on the 5' RACE experiment (black marks) and the position of the splicing donor site (sds). On the middle part of each scheme are represented the hypothetical transcripts expressed. In green is represented the part of the transcript detected on the 5' RACE experiment while in grey is represented the hypothetical part. The two upper transcripts (darker green) are the most abundant one according to the same RACE experiment, while the two down transcripts (lighter green) are the marginal ones. On the lower part of each scheme is represented the expected miRNA expression that would result from the transcripts expressed.

The splicing donor site (boundary exon/intron) found between miRNA-200c and miRNA141 strongly support the hypothesis of existence of alternative splicing events. Possibly, the high abundant TU starting at -285 contains an intron that includes miRNA-141 and that is spliced out during mRNA maturation. As discussed previously for the case of the surrounding transcripts, seems to exist a difference between the expression levels of a transcript and a miRNA located in its intronic portion. In this case, a differential processing of the exonic and intronic part of the TU could be responsible for differential expression of both miRNAs. Therefore it is predictable that, if the TU starting at position -285 alone would be transcribed, only miRNA-200c could be expressed without the

simultaneous expression of miRNA-141 (situation 3). In turn, if the two main transcriptional units would be transcribed, since miRNA-200c can be generated from both, one would expect more copies of this miRNA (situation 1).
The idea of coexpression of closely clustered miRNAs has been used by other authors as evidence that they derive from a common primary transcript[272]. The data presented here for the case of miRNAs-200c and 141, is compatible with this idea but additionally suggest that closely clustered miRNAs can also derive from different transcripts.

Alternatively the distinct expression observed for the closely located miRNAs could also lie in the differential processing of a common TU by Drosha. The proximity of the two miRNA of a pair could be an obstacle for the right processing of both miRNAs simultaneously, and could possibly lead to the preferable processing of one. However for the right cropping of pre-miRNA from primary long transcripts only a small number of nucleotides flanking the pre-miRNA are required[273] and the space between miRNAs in both pairs is much larger. Moreover, the ability of Drosha enzyme to act also on already partially sliced RNAs[274] suggests that even after the cropping of one miRNA the remaining RNA molecule can be used in further cropping reactions.

Summing up, the data collected so far allowed the identification of several transcription units likely to be responsible for the expression of the miRNAs 200c and 141. The transcription starting sites detected support an important role of a CpG island to their expression helping to explain the observed correlation between methylation and transcription level.

5.4 - Identification of the miRNA-200c/141 promoter and its susceptibility to DNA methylation

Gene reporter assays were performed in order to gain experimental evidence to prove the existence of a promoter located in the region nearby the miRNAs and under the influence of the CpG island. The results obtained from the experiments clearly show that the fragment containing miRNA-200c and 141 and the up-stream sequence, has an extremely strong Pol II promoter activity. Moreover, the capacity of this promoter to lead also to an over-expression of the mature forms of the miRNAs was later confirmed by northern blot (Fig. 31). The detected promoter was particularly active in MCF7 cells and that is consistent with the strong natural abundance of the two miRNAs in these cells. Interestingly, the same fragment showed also strong promoter activity in HepG2 and Hela cells. The promoter activity detected on these two cell types would make expectable a natural expression of these two miRNAs as it is indicative of the presence of the necessary

machinery for promoter transcription. However, expression of these two miRNAs is not reported on the smiRNAdb for HepG2 or Hela cells (Fig. 25 and data not shown). This might be a consequence of epigenetic silencing of that locus on those cells or to the action of more distal repressors not present on the cloned fragment.
In its turn, the lack of promoter activity of the same fragment in primary NK cells, contrasts with the initial miRNA profiling results showing that miRNA-200c is expressed in control NK cell sample at low amount. This suggests that, for any reason, the transcriptional machinery present on the NK cells is insufficient to drive alone an over-expression in NK cells using this artificial system, although it can keep a very low basal expression. A possible explanation is that, the epigenetic repression relevant for the expression of this miRNAs on NK cells is not confined to the miRNA locus but also to loci of transcription factors required for their expression. If these loci would also be epigenetically repressed by DNA methylation, the expression levels of the transcription factors would increase after 5-AZA-CdR, contributing to the miRNA up-regulation verified in NK cells. Possibly due to the initial low abundance of the transcription factors, a period longer than 4h after transfection, as used on our experiments, might be needed in order to see any significant over-expression in NK cells. Of notice is the fact that in the smiRNAdb, no expression of these two miRNAs is reported in NK cells (data not shown).

The experiments performed with *in vitro* methylated promoter provide clear experimental evidence for the direct role of DNA methylation on the activity of the identified promoter and, consequently, on the expression of the miRNAs 200c and 141 (Fig. 34). The *in vitro* methylation of the construct prior to its use in transfection experiments was sufficient to silence the promoter. Once that the effect of methylation is dependent on the number and density of CpG dinucleotides[275-277], and as this region show a significant number of those, the observations can be interpreted as a direct consequence of the promoter methylation. The observed effect of methylation on the activity of the promoter can be due to an interference with the binding of transcription factors recognizing CG-containing motifs (cis-effect) or, alternatively, to a recruitment of m^5CpG-binding proteins known to have repressor activity.
Noticeable, the low expression of the luciferase gene in the control vector was also reduced after *in vitro* methylation. As the basal expression verified for this vector was very low, any slight reduction of it had a significant impact in terms of percentage. This observation suggests that the basal expression can be driven by a basic promoter containing CpG dinucleotides located in the backbone of the vector. Indeed, analysis of the vector backbone revealed that several CpG dinucleotides are present in the 2 Kb region upstream of the multiple cloning site 1 used (region containing the pMB1ori and the Ampicillin resistance gene (Fig. 12 and data not shown). But, the reduction observed on the pMOD vector does not reduce the importance of the results obtained for the pMOD_200c/141 construct since the promoter activity is unquestionably higher and that is not a consequence of any putative marginal promoter but rather of the cloned one. So, if the

silencing effect would be uniquely based on methylation of the vector backbone and methylation would not interfere with the promoter activity of the cloned fragment, still a strong luciferase activity should be detected due to this last because it is located in a proximal position. Therefore, the silencing obtained in the pMOD_200c/141 construct can be seen as a consequence of a direct effect on the methylation of the cloned fragment.
Summing up, after the identification of the transcription starting sites colocalizing with a CG rich region, these last results prove the location of a strong promoter in that region and point to a crucial importance of the methylation status for the activity of that promoter and, consequently for expression of miRNAs-200c and 141.

6 – Conclusions and Perspectives

As stated in the introductory chapters, the initial aim of this work was to contribute to the rapidly increasing knowledge about the regulation of miRNA expression with the focusing on the influence of DNA methylation on that. In this respect, the work presented here provides several new experimental pieces of evidence characterizing the effect of the DNA demethylating agent 5-AZA-CdR on the expression of miRNAs in a primary, healthy and highly differentiated cell type as NK cells.
The comparative analysis of miRNA expression in cells kept in standard conditions and cells cultured in the presence of 5-AZA-CdR allowed to conclude that only a small number of miRNAs are up-regulated by this DNA demethylating agent in NK cells. As the usage of DNA demethylating agent 5-AZA-CdR can cause demethylation of the whole genome and not of a defined region, the present work started by exploring the possibility that DNA demethylation could affect the expression of the key enzymes involved in the miRNA maturation pathway. The results allowed to exclude this hypothesis and to conclude that the effect responsible for alteration on the expression was occurring in other genomic regions. Other regions of putative importance considered in this study were the regions regulating the expression of the transcripts that surround some up-regulated miRNAs. The results show that DNA methylation was also not responsible for the regulation of these transcripts. Indeed the data suggest that at least some intragenic miRNAs can be expressed and regulated independently of the transcriptional unit in which they are located in.

Analyses of the genomic area where 5-AZA-CdR-upregulated miRNAs are located, revealed that some were found in CG-rich regions. The fact that the methylation status of those regions correlates with the expression of the respective miRNAs, was a first finding arguing for a direct influence of DNA methylation on their transcription. Later, reinforcing the importance of the CG-rich region, for candidate locus containing the two clustered miRNA-200c and 141, experimental evidence allowed to identify transcription starting sites in close proximity to the region differentially methylated. Conclusive support for the importance of the CG-rich region was the ability of DNA methylation to repress the activity of a strong promoter present in that region, employing an independent cell model.

In general the results support the conclusion that, at least for some specific cases, DNA methylation is able to interfere in a decisive and direct way with the regulation of miRNA expression. The results demonstrate that the criteria used, including analysis of the genomic context of the miRNA-200c/141 region, led to the identification of the functionally relevant promoter region. However, the same criteria failed to identify miRNA promoters in other cases. These observations together with previous published works support the idea that miRNA promoter regions, similar to conventional RNA Polymerase II-regulated mRNA promoters, do not share specific features that are common to all miRNA promoter elements. In contrast, the data presented here support the view that each miRNA is regulated by a specific combination of factors that can involve already characterized or even novel transcription factors.

The observation of a highly divergent DNA methylation status between healthy cells vs. malignant breast cancer cells opens the possibility that epigenetic regulation of miRNA-200c and 141 expression could play a role in the pathological processes involved in malignant transformation.

Future perspectives:

Different elements of the miRNA-200 family, including the two miRNA-200c and 141, have been implicated in cellular mechanisms involved in epithelial to mesenchymal transition (EMT). This transition is seen as a facilitator of tissue remodelling during embryonic development but also as an essential early step in tumour metastasis involved in breast cancer as well as several forms of ovarian cancer. The observation suggests that a down-regulation of these miRNAs leads to an up-regulation of their experimentally verified targets such as zinc finger E-box binding homeobox 1 (ZEB1) and Smad interacting protein 1 (SIP1) with subsequent down-expression of E-cadherin (normally repressed by ZEB1)[278-282]. This cascade of events finally conduct to a commitment of the epithelial cell into mesenchymal transition[283]. It was hypothesized that the major function of miRNA-200 family is to prevent the expression of ZEB1, which in turn would prevent EMT and the switch into an invasive phenotype.

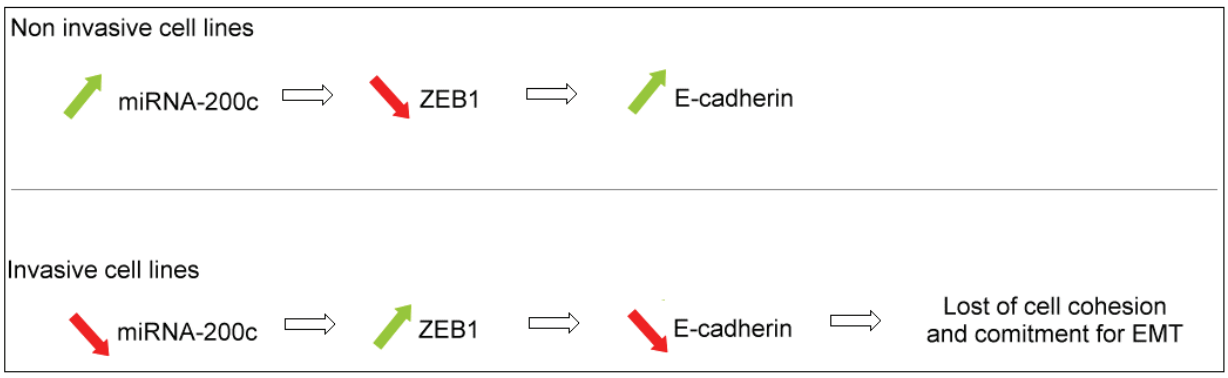


Fig. 36 – Cascade of events initiated by a different miRNA-200c and taking place in cells with different phenotypes.

A recent publication from 2008 gave the first insight in the regulation of miRNA-200c/141 expression[284]. The authors show that the target protein of miRNA-200c, ZEB1, is able to repress miRNA-200c transcription by binding to several sites upstream of this miRNA. The reciprocal inhibitory effect of miRNA-200c and ZEB1 creates a feed-forward loop that, once started, leads to the above mentioned EMT and to an invasive phenotype.

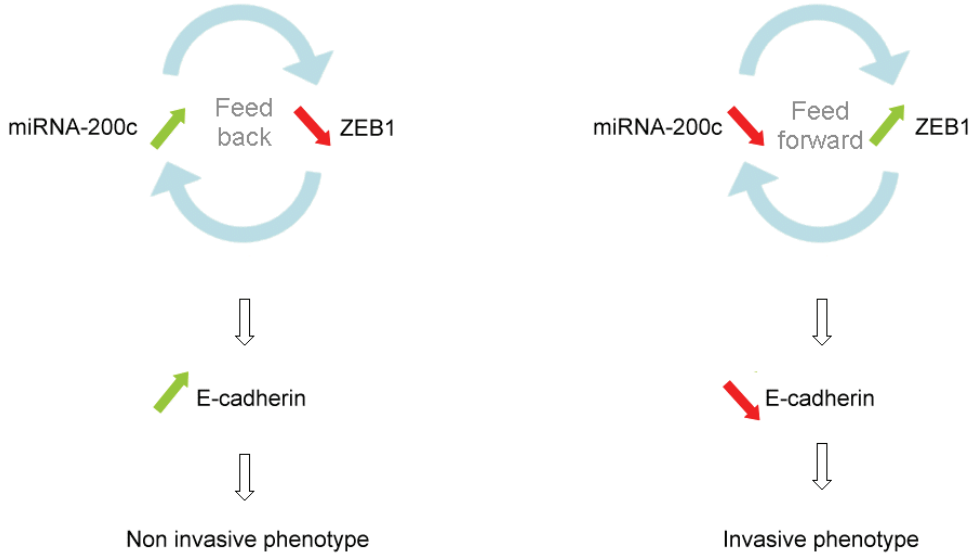


Fig. 37 – The inversion of a feedback loop constituted by miRNA-200c and ZEB1 and resulting in a low expression of this last, into a feedforward loop leads to an invasive phenotype of the cells.

Interestingly, ZEB1 binding sites are in close proximity to the most active transcription starting sites that have been identified in the present work and four of these are located in the same CG-rich region analysed, confirming the importance of our findings.
The idea of a feedforward loop able to turn a non-invasive cell into an invasive one is interesting and raises new questions. Among these questions is definitely: What triggers this loop to start? The present work provides clear evidence for the importance of DNA methylation on the expression regulation of miRNAs-200c and 141. As in cancer cells the homeostasis of the genome methylation is altered, on the basis of the work here presented, it is likely that an aberrant DNA methylation of miRNA-200c locus can be the decisive event leading to initiation of the feedforward loop. The aberrant methylation of the private promoter of miRNA-200c and 141 could lead to reduction of miRNAs levels by the recruitment of m^5CpG-binding proteins known to have repressor activity or to a reduced capacity of activator transcription factors to bind. Once the expression of miRNA-200c starts to diminish, the expression of ZEB1 starts to increase and this higher level of ZEB1 protein contributes together with the methylation to ensure the silencing of the miRNA-200c locus. It is known that ZEB1 binding sites do not contain any CG dinucleotides and so the methylation of the locus is unlikely to have a repulsive effect on the binding of ZEB1.

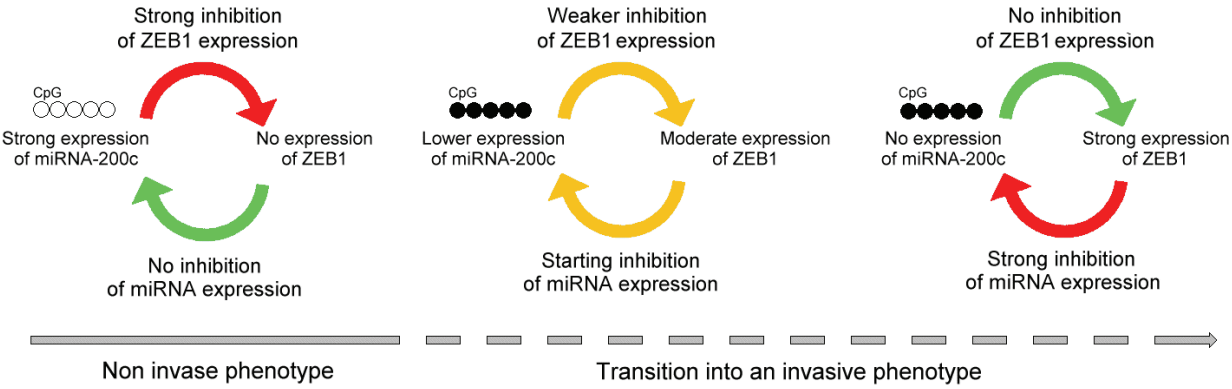


Fig. 38 – Proposed model showing how methylation of the miRNA-200c locus can invert the initial feedback into a feedforward loop with subsequent transition into an invasive phenotype. Methylation of the locus (represented by the closed circles) can interrupt the feedback loop lowering the expression of miRNA-200c and causing an increased expression of ZEB1. Once moderate levels of ZEB1 are reached, it starts to contribute also to miRNA silencing, resulting in the feedforward loop. Red arrows represent strong inhibition; orange arrow represents moderate levels of inhibition; and green arrow represents no inhibition.

The present work already presents a first experimental hint supporting this hypothesis. MCF7 cells are derived from a breast cancer tumour with a non-invasive phenotype and show a completely demethylated miRNA-200c locus (Fig. 26). It seems of crucial interest to analyze other cell lines with different phenotypes in order to more closely explore this idea. Naturally, it would also be interesting to extend the analysis to several patient samples (primary cells) with different clinical prognosis and correlate the results of methylation with tumour stage and capacity for metastasis formation. In case of a correlation, this could open very promising perspectives to help the grading of the respective tumour and consequently the choice of the most appropriate therapeutical approach. The analysis of the methylation status by genomic sequencing of bisulfite converted DNA (as it was done here) is time consuming, but the acquired knowledge enables usage of faster and higher-throughput techniques such as methylation specific PCR (MS-PCR) or denaturing high performance liquid chromatography (dHPLC)[285]. Based on this work, the usage of demethylating agents, as the clinical approved 5-AZA-CdR, deserve to be considered for future therapy of breast cancer as it has the potential to keep the levels of miRNA-200c high, preventing the appearance of cells with invasive phenotype and metastasis formation.

Although this work was based on haematopoietic and breast cancer cell models, the DNA methylation mechanisms are universal to all human cell populations and changes in DNA methylation has been widely described for many loci involved in many cancer types. Therefore it is very likely that the conclusions of the present study are also applicable to cells involved in other forms of cancer.

7 – List of Abbreviations

5-AZA-CdR - *5-aza-2'-deoxycytidine*
AGO - *argonaute protein family*
AML - *acute myeloid leukaemia*
ARE - *AU-rich element*
bps - *base pairs*
BRE - *TFIIB recognition element*
C17orf91 - *chromosome 17 open reading frame 91*
C21orf34 - *chromosome 21 open reading frame 34*
ChIP-on-chip - *chromatin immunoprecipitation-on-chip*
CML - *chronic myeloid leukaemia*
Ct - *cycle threshold*
DCE - *downstream core element*
DGCR8 - *DiGeorge critical region 8 protein*
dHPLC - *denaturing high performance liquid chromatography*
DNA - *desoxyribonucleic acid*
DNMTs - *DNA methyltransferases*
DPE - *downstream promoter element*
dsRNA - *double stranded RNA*
EDC - *1-ethyl-3-(3-dimethylaminopropyl) carbodiimide*
EMT - *epithelial to mesenchymal transition*
Exp5 - *Exportin-5*
FMRP - *fragile X mental retardation protein*
HAT - *histone acetyltransferases*
HDAC - *histone deacetylases*
IPTG - *isopropyl β-D-1-thiogalactopyranoside*
KIR - *killer cell immunoglobulin-like receptors*
LB - *Lennox L Broth*
m^4C - *N4-methylcytosine*
m^5C - *5-methylcytosine*
m^6A - *N6-methyladenine*
MACS - *magnetic activated cell sorting*
MDS - *myeloid dysplastic syndrome*
MED-1 - *multiple start site element downstream 1*
miRISC - *miRNA induced silencing complex*
miRNA - *micro RNA*
M-MLV RT - *moloney-murine leukemia virus reverse transcriptase*
MOPS - *3-morpholinopropane-1-sulfonic acid*
mRNA - *messenger RNA*
MS-PCR - *methylation specific PCR*
nt - *nucleotide*
P-bodies - *processing-bodies*

PABP1 - *polyadenylate-binding protein 1*
PBA - *4-phenylbutyric acid*
PBMCs - *peripheral blood mononuclear cells*
PCR - *polimerase chain reaction*
pre-miRNA - *precursor micro RNA*
pri-miRNA - *primary micro RNA*
PPTases - *phosphatases*
PSE - *proximal sequence element*
R3HD1 - *R3H domain containing 1*
RACE - *rapid amplification of cDNA ends*
REST - *RE1 silencing transcript factor*
RNA - *ribonucleic acid*
SAM - *S-adenosylmethionine*
SDS - *sodium dodecyl sulfate*
sds - *splicing donor site*
SIP1 - *Smad interacting protein 1*
SP - *specific primer*
SREBF2 - *sterol regulatory element binding protein-2*
SRF - *serum response factor*
TFBS - *transcription factors binding sites*
TRE - *temporal regulatory element*
TRBP - *TAR RNA binding protein*
TSA - *trichostatin A*
TU - *transcription unit*
X-Gal - *5-bromo-4-chloro-3-indolyl-β-D-galactoside*
ZEB1 - *zinc finger E-box binding homeobox 1*

7 - Bibliography

1. Waddington,C.H. The epigenomotype. *Endeavour* 18-20 (1942).

2. Russo,V.E.A., Martienssen,R.A. & Riggs,A.D. Epigenetic mechanisms of gene regulation. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1996).

3. Holliday,R. Epigenetics: an overview. *Dev. Genet* **15**, 453-457 (1994).

4. Allis,C.D., Jenuwein,T. & Reinberg,D. Epigenetics., pp. 23-50 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY,2007).

5. Jaenisch,R. & Bird,A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* **33 Suppl**, 245-254 (2003).

6. Kornberg,R.D. Chromatin structure: a repeating unit of histones and DNA. *Science* **184**, 868-871 (1974).

7. Kouzarides,T. Chromatin modifications and their function. *Cell* **128**, 693-705 (2007).

8. Sterner,D.E. & Berger,S.L. Acetylation of histones and transcription-related factors. *Microbiol. Mol. Biol. Rev.* **64**, 435-459 (2000).

9. Finnin,M.S., Donigian,J.R. & Pavletich,N.P. Structure of the histone deacetylase SIRT2. *Nat Struct. Biol.* **8**, 621-625 (2001).

10. Zhang,Y. & Reinberg,D. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* **15**, 2343-2360 (2001).

11. Nowak,S.J. & Corces,V.G. Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet* **20**, 214-220 (2004).

12. WYATT,G.R. Occurrence of 5-methylcytosine in nucleic acids. *Nature* **166**, 237-238 (1950).

13. Riggs,A.D. X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet* **14**, 9-25 (1975).

14. Holliday,R. & Pugh,J.E. DNA modification mechanisms and gene activity during development. *Science* **187**, 226-232 (1975).

15. Egger,G., Liang,G., Aparicio,A. & Jones,P.A. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **429**, 457-463 (2004).

16. Morgan,H.D., Santos,F., Green,K., Dean,W. & Reik,W. Epigenetic reprogramming in mammals. *Hum. Mol. Genet* **14 Spec No 1**, R47-R58 (2005).

17. Bird,A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6-21 (2002).

18. Bird,A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6-21 (2002).

19. Lyko,F., Ramsahoye,B.H. & Jaenisch,R. DNA methylation in Drosophila melanogaster. *Nature* **408**, 538-540 (2000).

20. Gowher,H., Leismann,O. & Jeltsch,A. DNA of Drosophila melanogaster contains 5-methylcytosine. *EMBO J.* **19**, 6918-6923 (2000).

21. Gruenbaum,Y., Naveh-Many,T., Cedar,H. & Razin,A. Sequence specificity of methylation in higher plant DNA. *Nature* **292**, 860-862 (1981).

22. Vanyushin,B.F. DNA methylation in plants. *Curr. Top. Microbiol. Immunol.* **301**, 67-122 (2006).

23. DOSKOCIL,J. & SORMO'VA,Z. The occurrence of 5-methylcytosine in bacterial deoxyribonucleic acids. *Biochim. Biophys. Acta* **95**, 513-515 (1965).

24. DUNN,D.B. & SMITH,J.D. The occurrence of 6-methylaminopurine in deoxyribonucleic acids. *Biochem. J.* **68**, 627-636 (1958).

25. Ehrlich,M. *et al.* DNA methylation in thermophilic bacteria: N4-methylcytosine, 5-methylcytosine, and N6-methyladenine. *Nucleic Acids Res* **13**, 1399-1412 (1985).

26. Ehrlich,M. *et al.* Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res* **10**, 2709-2721 (1982).

27. Lander,E.S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).

28. Venter,J.C. *et al.* The sequence of the human genome. *Science* **291**, 1304-1351 (2001).

29. Bird,A.P. CpG islands as gene markers in the vertebrate nucleus. *Trends Genet* **13**, 469-470 (1987).

30. Antequera,F. & Bird,A. Number of CpG islands and genes in human and mouse. *Proc. Natl Acad. Sci. U. S. A* **90**, 11995-11999 (1993).

31. Marino-Ramirez,L., Spouge,J.L., Kanga,G.C. & Landsman,D. Statistical analysis of over-represented words in human promoter sequences. *Nucleic Acids Res* **32**, 949-958 (2004).

32. Davuluri,R.V., Grosse,I. & Zhang,M.Q. Computational identification of promoters and first exons in the human genome. *Nat Genet* **29**, 412-417 (2001).

33. Antequera,F. & Bird,A. Number of CpG islands and genes in human and mouse. *Proc. Natl Acad. Sci. U. S. A* **90**, 11995-11999 (1993).

34. Ratel,D., Ravanat,J.L., Berger,F. & Wion,D. N6-methyladenine: the other methylated base of DNA. *Bioessays* **28**, 309-315 (2006).

35. Jones,P.A. & Takai,D. The role of DNA methylation in mammalian epigenetics. *Science* **293**, 1068-1070 (2001).

36. Futscher,B.W. *et al.* Role for DNA methylation in the control of cell type specific maspin expression. *Nat Genet* **31**, 175-179 (2002).

37. Kass,S.U., Landsberger,N. & Wolffe,A.P. DNA methylation directs a time-dependent repression of transcription initiation. *Curr. Biol.* **7**, 157-165 (1997).

38. Boyes,J. & Bird,A. Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. *EMBO J.* **11**, 327-333 (1992).

39. Hsieh,C.L. Dependence of transcriptional repression on CpG methylation density. *Mol. Cell Biol.* **14**, 5487-5494 (1994).

40. Kass,S.U., Landsberger,N. & Wolffe,A.P. DNA methylation directs a time-dependent repression of transcription initiation. *Curr. Biol.* **7**, 157-165 (1997).

41. Tate,P.H. & Bird,A.P. Effects of DNA methylation on DNA-binding proteins and gene expression. *Curr. Opin. Genet Dev.* **3**, 226-231 (1993).

42. Bell,A.C. & Felsenfeld,G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **405**, 482-485 (2000).

43. Lewis,J.D. *et al.* Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69**, 905-914 (1992).

44. Meehan,R.R., Lewis,J.D. & Bird,A.P. Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Res* **20**, 5085-5092 (1992).

45. Meehan,R.R., Lewis,J.D., McKay,S., Kleiner,E.L. & Bird,A.P. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* **58**, 499-507 (1989).

46. Bird,A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6-21 (2002).

47. Lock,L.F., Takagi,N. & Martin,G.R. Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. *Cell* **48**, 39-46 (1987).

48. Futscher,B.W. *et al.* Role for DNA methylation in the control of cell type specific maspin expression. *Nat Genet* **31**, 175-179 (2002).

49. Vilches,C. & Parham,P. KIR: diverse, rapidly evolving receptors of innate and adaptive immunity. *Annu. Rev. Immunol.* **20**, 217-251 (2002).

50. Valiante,N.M. *et al.* Functionally and structurally distinct NK cell receptor repertoires in the peripheral blood of two human donors. *Immunity.* **7**, 739-751 (1997).

51. Santourlidis,S. *et al.* Crucial role of DNA methylation in determination of clonally distributed killer cell Ig-like receptor expression patterns in NK cells. *J. Immunol.* **169**, 4253-4261 (2002).

52. Trompeter,H.I. *et al.* Three structurally and functionally divergent kinds of promoters regulate expression of clonally distributed killer cell Ig-like receptors (KIR), of KIR2DL4, and of KIR3DL3. *J. Immunol.* **174**, 4135-4143 (2005).

53. De Smet,C. *et al.* The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation. *Proc. Natl Acad. Sci. U. S. A* **93**, 7149-7153 (1996).

54. De Smet,C., Lurquin,C., Lethe,B., Martelange,V. & Boon,T. DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. *Mol. Cell Biol.* **19**, 7327-7335 (1999).

55. Panning,B. & Jaenisch,R. DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev.* **10**, 1991-2002 (1996).

56. Reik,W., Dean,W. & Walter,J. Epigenetic reprogramming in mammalian development. *Science* **293**, 1089-1093 (2001).

57. Monk,M., Boubelik,M. & Lehnert,S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99**, 371-382 (1987).

58. Sanford,J.P., Clark,H.J., Chapman,V.M. & Rossant,J. Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. *Genes Dev.* **1**, 1039-1046 (1987).

59. Duncan,B.K. & Miller,J.H. Mutagenic deamination of cytosine residues in DNA. *Nature* **287**, 560-561 (1980).

60. Xu,G.L. *et al.* Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402**, 187-191 (1999).

61. Chen,R.Z., Pettersson,U., Beard,C., Jackson-Grusby,L. & Jaenisch,R. DNA hypomethylation leads to elevated mutation rates. *Nature* **395**, 89-93 (1998).

62. Sapienza,C., Paquette,J., Tran,T.H. & Peterson,A. Epigenetic and genetic factors affect transgene methylation imprinting. *Development* **107**, 165-168 (1989).

63. Allen,N.D., Norris,M.L. & Surani,M.A. Epigenetic control of transgene expression and imprinting by genotype-specific modifiers. *Cell* **61**, 853-861 (1990).

64. Engler,P. *et al.* A strain-specific modifier on mouse chromosome 4 controls the methylation of independent transgene loci. *Cell* **65**, 939-947 (1991).

65. Reik,W. *et al.* Adult phenotype in the mouse can be affected by epigenetic events in the early embryo. *Development* **119**, 933-942 (1993).

66. Zhu,B. *et al.* Overexpression of 5-methylcytosine DNA glycosylase in human embryonic kidney cells EcR293 demethylates the promoter of a hormone-regulated reporter gene. *Proc. Natl Acad. Sci. U. S. A* **98**, 5031-5036 (2001).

67. Bhattacharya,S.K., Ramchandani,S., Cervoni,N. & Szyf,M. A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* **397**, 579-583 (1999).

68. Weiss,A., Keshet,I., Razin,A. & Cedar,H. DNA demethylation in vitro: involvement of RNA. *Cell* **86**, 709-718 (1996).

69. Okano,M., Bell,D.W., Haber,D.A. & Li,E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-257 (1999).

70. Monk,M., Boubelik,M. & Lehnert,S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99**, 371-382 (1987).

71. Sanford,J.P., Clark,H.J., Chapman,V.M. & Rossant,J. Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. *Genes Dev.* **1**, 1039-1046 (1987).

72. Howlett,S.K. & Reik,W. Methylation levels of maternal and paternal genomes during preimplantation development. *Development* **113**, 119-127 (1991).

73. Kafri,T. *et al.* Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev.* **6**, 705-714 (1992).

74. Okano,M., Xie,S. & Li,E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet* **19**, 219-220 (1998).

75. Okano,M., Bell,D.W., Haber,D.A. & Li,E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-257 (1999).

76. Okano,M., Bell,D.W., Haber,D.A. & Li,E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-257 (1999).

77. Lei,H. *et al.* De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122**, 3195-3205 (1996).

78. Li,E., Bestor,T.H. & Jaenisch,R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915-926 (1992).

79. Stancheva,I. & Meehan,R.R. Transient depletion of xDnmt1 leads to premature gene activation in Xenopus embryos. *Genes Dev.* **14**, 313-327 (2000).

80. Riggs,A.D., Xiong,Z., Wang,L. & LeBon,J.M. Methylation dynamics, epigenetic fidelity and X chromosome structure. *Novartis. Found. Symp.* **214**, 214-225 (1998).

81. Jones,P.A. & Baylin,S.B. The fundamental role of epigenetic events in cancer. *Nat Rev. Genet* **3**, 415-428 (2002).

82. Esteller,M. & Herman,J.G. Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. *J. Pathol.* **196**, 1-7 (2002).

83. Jones,P.A. & Laird,P.W. Cancer epigenetics comes of age. *Nat Genet* **21**, 163-167 (1999).

84. Baylin,S.B., Herman,J.G., Graff,J.R., Vertino,P.M. & Issa,J.P. Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv. Cancer Res* **72**, 141-196 (1998).

85. Issa,J.P., Baylin,S.B. & Herman,J.G. DNA methylation changes in hematologic malignancies: biologic and clinical implications. *Leukemia* **11 Suppl 1**, S7-11 (1997).

86. Melki,J.R. & Clark,S.J. DNA methylation changes in leukaemia. *Semin. Cancer Biol.* **12**, 347-357 (2002).

87. Jones,P.A. & Takai,D. The role of DNA methylation in mammalian epigenetics. *Science* **293**, 1068-1070 (2001).

88. Jones,P.A. & Takai,D. The role of DNA methylation in mammalian epigenetics. *Science* **293**, 1068-1070 (2001).

89. Santini,V., Kantarjian,H.M. & Issa,J.P. Changes in DNA methylation in neoplasia: pathophysiology and therapeutic implications. *Ann. Intern. Med.* **134**, 573-586 (2001).

90. Melki,J.R. & Clark,S.J. DNA methylation changes in leukaemia. *Semin. Cancer Biol.* **12**, 347-357 (2002).

91. Gonzalgo,M.L. & Jones,P.A. Mutagenic and epigenetic effects of DNA methylation. *Mutat. Res* **386**, 107-118 (1997).

92. Jones,P.A. & Laird,P.W. Cancer epigenetics comes of age. *Nat Genet* **21**, 163-167 (1999).

93. Xu,G.L. *et al.* Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402**, 187-191 (1999).

94. Ehrlich,M. *et al.* High frequencies of ICF syndrome-like pericentromeric heterochromatin decondensation and breakage in chromosome 1 in a chorionic villus sample. *J. Med. Genet* **38**, 882-884 (2001).

95. Christman,J.K. 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene* **21**, 5483-5495 (2002).

96. Richel,D.J., Colly,L.P., Lurvink,E. & Willemze,R. Comparison of the antileukaemic activity of 5 aza-2-deoxycytidine and arabinofuranosyl-cytosine in rats with myelocytic leukaemia. *Br. J. Cancer* **58**, 730-733 (1988).

97. Willemze,R. *et al.* A randomized phase II study on the effects of 5-Aza-2'-deoxycytidine combined with either amsacrine or idarubicin in patients with relapsed acute leukemia: an EORTC Leukemia Cooperative Group phase II study (06893). *Leukemia* **11 Suppl 1**, S24-S27 (1997).

98. Kantarjian,H.M. *et al.* Decitabine studies in chronic and acute myelogenous leukemia. *Leukemia* **11 Suppl 1**, S35-S36 (1997).

99. Issa,J.P. *et al.* Phase 1 study of low-dose prolonged exposure schedules of the hypomethylating agent 5-aza-2'-deoxycytidine (decitabine) in hematopoietic malignancies. *Blood* **103**, 1635-1640 (2004).

100. Zaharko,D.S., Covey,J.M. & Muneses,C.C. Experimental chemotherapy (L1210) with 5-aza-2'-deoxycytidine in combination with pyran copolymer (MVE-4), an immune adjuvant. *J. Natl Cancer Inst.* **74**, 1319-1324 (1985).

101. Quddus,J. *et al.* Treating activated CD4+ T cells with either of two distinct DNA methyltransferase inhibitors, 5-azacytidine or procainamide, is sufficient to cause a lupus-like disease in syngeneic mice. *J. Clin. Invest* **92**, 38-53 (1993).

102. Lee,R.C., Feinbaum,R.L. & Ambros,V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-854 (1993).

103. Chalfie,M., Horvitz,H.R. & Sulston,J.E. Mutations that lead to reiterations in the cell lineages of C. elegans. *Cell* **24**, 59-69 (1981).

104. Ambros,V. & Horvitz,H.R. The lin-14 locus of Caenorhabditis elegans controls the time of expression of specific postembryonic developmental events. *Genes Dev.* **1**, 398-414 (1987).

105. Reinhart,B.J. *et al.* The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **403**, 901-906 (2000).

106. Slack,F.J. *et al.* The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Mol. Cell* **5**, 659-669 (2000).

107. Pasquinelli,A.E. *et al.* Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **408**, 86-89 (2000).

108. Zhao,T. *et al.* A complex system of small RNAs in the unicellular green alga Chlamydomonas reinhardtii. *Genes Dev.* **21**, 1190-1203 (2007).

109. Molnar,A., Schwach,F., Studholme,D.J., Thuenemann,E.C. & Baulcombe,D.C. miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. *Nature* **447**, 1126-1129 (2007).

110. Bartel,D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-297 (2004).

111. Ambros,V. *et al.* A uniform system for microRNA annotation. *RNA.* **9**, 277-279 (2003).

112. Griffiths-Jones,S., Grocock,R.J., van Dongen,S., Bateman,A. & Enright,A.J. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* **34**, D140-D144 (2006).

113. Bentwich,I. *et al.* Identification of hundreds of conserved and nonconserved human microRNAs. *Nat Genet* **37**, 766-770 (2005).

114. Lewis,B.P., Burge,C.B. & Bartel,D.P. Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. *Cell* **120**, 15-20 (2005).

115. Rajewsky,N. microRNA target predictions in animals. *Nat Genet* **38 Suppl**, S8-13 (2006).

116. Lee,Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051-4060 (2004).

117. Lee,Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051-4060 (2004).

118. Cai,X., Hagedorn,C.H. & Cullen,B.R. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA.* **10**, 1957-1966 (2004).

119. Tam,W. Identification and characterization of human BIC, a gene on chromosome 21 that encodes a noncoding RNA. *Gene* **274**, 157-167 (2001).

120. Lagos-Quintana,M., Rauhut,R., Lendeckel,W. & Tuschl,T. Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853-858 (2001).

121. Lee,Y., Jeon,K., Lee,J.T., Kim,S. & Kim,V.N. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.* **21**, 4663-4670 (2002).

122. Lau,N.C., Lim,L.P., Weinstein,E.G. & Bartel,D.P. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science* **294**, 858-862 (2001).

123. Lee,R.C. & Ambros,V. An extensive class of small RNAs in Caenorhabditis elegans. *Science* **294**, 862-864 (2001).

124. Lagos-Quintana,M., Rauhut,R., Lendeckel,W. & Tuschl,T. Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853-858 (2001).

125. Rodriguez,A., Griffiths-Jones,S., Ashurst,J.L. & Bradley,A. Identification of mammalian microRNA host genes and transcription units. *Genome Res* **14**, 1902-1910 (2004).

126. Lee,Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-419 (2003).

127. Denli,A.M., Tops,B.B., Plasterk,R.H., Ketting,R.F. & Hannon,G.J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**, 231-235 (2004).

128. Gregory,R.I. *et al.* The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**, 235-240 (2004).

129. Han,J. *et al.* The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.* **18**, 3016-3027 (2004).

130. Lee,Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-419 (2003).

131. Zeng,Y. & Cullen,B.R. Sequence requirements for micro RNA processing and function in human cells. *RNA.* **9**, 112-123 (2003).

132. Yi,R., Qin,Y., Macara,I.G. & Cullen,B.R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **17**, 3011-3016 (2003).

133. Bohnsack,M.T., Czaplinski,K. & Gorlich,D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* **10**, 185-191 (2004).

134. Gwizdek,C. *et al.* Terminal minihelix, a novel RNA motif that directs polymerase III transcripts to the cell cytoplasm. Terminal minihelix and RNA export. *J. Biol. Chem.* **276**, 25910-25918 (2001).

135. Gwizdek,C. *et al.* Exportin-5 mediates nuclear export of minihelix-containing RNAs. *J. Biol. Chem.* **278**, 5505-5508 (2003).

136. Lund,E., Guttinger,S., Calado,A., Dahlberg,J.E. & Kutay,U. Nuclear export of microRNA precursors. *Science* **303**, 95-98 (2004).

137. Yi,R., Qin,Y., Macara,I.G. & Cullen,B.R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **17**, 3011-3016 (2003).

138. Bernstein,E., Caudy,A.A., Hammond,S.M. & Hannon,G.J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363-366 (2001).

139. Hutvagner,G. *et al.* A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**, 834-838 (2001).

140. Ketting,R.F. *et al.* Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes Dev.* **15**, 2654-2659 (2001).

141. Knight,S.W. & Bass,B.L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. *Science* **293**, 2269-2271 (2001).

142. Wu,H., Xu,H., Miraglia,L.J. & Crooke,S.T. Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. *J. Biol. Chem.* **275**, 36957-36965 (2000).

143. Ma,J.B., Ye,K. & Patel,D.J. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**, 318-322 (2004).

144. Lingel,A., Simon,B., Izaurralde,E. & Sattler,M. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature* **426**, 465-469 (2003).

145. Zhang,H., Kolb,F.A., Jaskiewicz,L., Westhof,E. & Filipowicz,W. Single processing center models for human Dicer and bacterial RNase III. *Cell* **118**, 57-68 (2004).

146. Khvorova,A., Reynolds,A. & Jayasena,S.D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209-216 (2003).

147. Schwarz,D.S. *et al.* Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199-208 (2003).

148. Landgraf,P. *et al.* A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**, 1401-1414 (2007).

149. Peters,L. & Meister,G. Argonaute proteins: mediators of RNA silencing. *Mol. Cell* **26**, 611-623 (2007).

150. Chu,C.Y. & Rana,T.M. Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol.* **4**, e210 (2006).

151. Eulalio,A. *et al.* Target-specific requirements for enhancers of decapping in miRNA-mediated gene silencing. *Genes Dev.* **21**, 2558-2570 (2007).

152. Pillai,R.S. *et al.* Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* **309**, 1573-1576 (2005).

153. Chu,C.Y. & Rana,T.M. Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol.* **4**, e210 (2006).

154. Leung,A.K., Calabrese,J.M. & Sharp,P.A. Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. *Proc. Natl Acad. Sci. U. S. A* **103**, 18125-18130 (2006).

155. Sen,G.L. & Blau,H.M. Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. *Nat Cell Biol.* **7**, 633-636 (2005).

156. Liu,J., Valencia-Sanchez,M.A., Hannon,G.J. & Parker,R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol.* **7**, 719-723 (2005).

157. Liu,J. *et al.* A role for the P-body component GW182 in microRNA function. *Nat Cell Biol.* **7**, 1261-1266 (2005).

158. Leung,A.K., Calabrese,J.M. & Sharp,P.A. Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. *Proc. Natl Acad. Sci. U. S. A* **103**, 18125-18130 (2006).

159. Lee,R.C., Feinbaum,R.L. & Ambros,V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-854 (1993).

160. Arasu,P., Wightman,B. & Ruvkun,G. Temporal regulation of lin-14 by the antagonistic action of two other heterochronic genes, lin-4 and lin-28. *Genes Dev.* **5**, 1825-1833 (1991).

161. Wightman,B., Ha,I. & Ruvkun,G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell* **75**, 855-862 (1993).

162. Ha,I., Wightman,B. & Ruvkun,G. A bulged lin-4/lin-14 RNA duplex is sufficient for Caenorhabditis elegans lin-14 temporal gradient formation. *Genes Dev.* **10**, 3041-3050 (1996).

163. Olsen,P.H. & Ambros,V. The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. *Dev. Biol.* **216**, 671-680 (1999).

164. Lee,R.C., Feinbaum,R.L. & Ambros,V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-854 (1993).

165. Petersen,C.P., Bordeleau,M.E., Pelletier,J. & Sharp,P.A. Short RNAs repress translation after initiation in mammalian cells. *Mol. Cell* **21**, 533-542 (2006).

166. Richter,J.D. & Sonenberg,N. Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* **433**, 477-480 (2005).

167. Cho,P.F. *et al.* A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP. *Cell* **121**, 411-423 (2005).

168. Wells,S.E., Hillner,P.E., Vale,R.D. & Sachs,A.B. Circularization of mRNA by eukaryotic translation initiation factors. *Mol. Cell* **2**, 135-140 (1998).

169. Derry,M.C., Yanagiya,A., Martineau,Y. & Sonenberg,N. Regulation of poly(A)-binding protein through PABP-interacting proteins. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 537-543 (2006).

170. Kiriakidou,M. *et al.* An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell* **129**, 1141-1151 (2007).

171. Wu,L., Fan,J. & Belasco,J.G. MicroRNAs direct rapid deadenylation of mRNA. *Proc. Natl Acad. Sci. U. S. A* **103**, 4034-4039 (2006).

172. Bagga,S. *et al.* Regulation by let-7 and lin-4 miRNAs Results in Target mRNA Degradation. *Cell* **122**, 553-563 (2005).

173. Giraldez,A.J. *et al.* Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312**, 75-79 (2006).

174. Yekta,S., Shih,I.H. & Bartel,D.P. MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304**, 594-596 (2004).

175. Jing,Q. *et al.* Involvement of microRNA in AU-rich element-mediated mRNA instability. *Cell* **120**, 623-634 (2005).

176. Mishima,Y. *et al.* Differential regulation of germline mRNAs in soma and germ cells by zebrafish miR-430. *Curr. Biol.* **16**, 2135-2142 (2006).

177. Leung,A.K. & Sharp,P.A. Function and Localization of MicroRNAs in Mammalian Cells. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 29-38 (2006).

178. Yekta,S., Shih,I.H. & Bartel,D.P. MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304**, 594-596 (2004).

179. Hornstein,E. *et al.* The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development. *Nature* **438**, 671-674 (2005).

180. Lim,L.P. *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769-773 (2005).

181. Jing,Q. *et al.* Involvement of microRNA in AU-rich element-mediated mRNA instability. *Cell* **120**, 623-634 (2005).

182. Cullen,B.R. Transcription and processing of human microRNA precursors. *Mol. Cell* **16**, 861-865 (2004).

183. Pasquinelli,A.E. *et al.* Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **408**, 86-89 (2000).

184. Lund,E., Guttinger,S., Calado,A., Dahlberg,J.E. & Kutay,U. Nuclear export of microRNA precursors. *Science* **303**, 95-98 (2004).

185. Lee,Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-419 (2003).

186. Cai,X., Hagedorn,C.H. & Cullen,B.R. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA.* **10**, 1957-1966 (2004).

187. Lee,Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051-4060 (2004).

188. Johnson,S.M., Lin,S.Y. & Slack,F.J. The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. *Dev. Biol.* **259**, 364-379 (2003).

189. Johnston,R.J. & Hobert,O. A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans. *Nature* **426**, 845-849 (2003).

190. Johnson,S.M., Lin,S.Y. & Slack,F.J. The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. *Dev. Biol.* **259**, 364-379 (2003).

191. Johnston,R.J. & Hobert,O. A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans. *Nature* **426**, 845-849 (2003).

192. Johnson,S.M., Lin,S.Y. & Slack,F.J. The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. *Dev. Biol.* **259**, 364-379 (2003).

193. Lee,Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051-4060 (2004).

194. Smale,S.T. & Kadonaga,J.T. The RNA polymerase II core promoter. *Annu. Rev. Biochem.* **72**, 449-479 (2003).

195. Lee,Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051-4060 (2004).

196. Zhao,Y., Samal,E. & Srivastava,D. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* (2005).

197. O'Donnell,K.A., Wentzel,E.A., Zeller,K.I., Dang,C.V. & Mendell,J.T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**, 839-843 (2005).

198. Fazi,F. *et al.* A minicircuitry comprised of microRNA-223 and transcription factors NFI-A and C/EBPalpha regulates human granulopoiesis. *Cell* **123**, 819-831 (2005).

199. Conaco,C., Otto,S., Han,J.J. & Mandel,G. Reciprocal actions of REST and a microRNA promote neuronal identity. *Proc. Natl Acad. Sci. U. S. A* **103**, 2422-2427 (2006).

200. Zhou,X., Ruan,J., Wang,G. & Zhang,W. Characterization and Identification of MicroRNA Core Promoters in Four Model Species. *PLoS Comput. Biol.* **3**, e37 (2007).

201. Scott,G.K., Mattie,M.D., Berger,C.E., Benz,S.C. & Benz,C.C. Rapid alteration of microRNA levels by histone deacetylase inhibition. *Cancer Res* **66**, 1277-1281 (2006).

202. Saito,Y. *et al.* Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell* **9**, 435-443 (2006).

203. Yanaihara,N. *et al.* Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* **9**, 189-198 (2006).

204. Diederichs,S. & Haber,D.A. Sequence variations of microRNAs in human cancer: alterations in predicted secondary structure do not affect processing. *Cancer Res* **66**, 6097-6104 (2006).

205. Meng,F., Wehbe-Janek,H., Henson,R., Smith,H. & Patel,T. Epigenetic regulation of microRNA-370 by interleukin-6 in malignant human cholangiocytes. *Oncogene* (2007).

206. Lujambio,A. *et al.* Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res* **67**, 1424-1429 (2007).

207. Lehmann,U. *et al.* Epigenetic inactivation of microRNA gene hsa-mir-9-1 in human breast cancer. *J. Pathol.* (2007).

208. Kozaki,K., Imoto,I., Mogi,S., Omura,K. & Inazawa,J. Exploration of tumor-suppressive microRNAs silenced by DNA hypermethylation in oral cancer. *Cancer Res* **68**, 2094-2105 (2008).

209. Fazi,F. *et al.* Epigenetic silencing of the myelopoiesis regulator microRNA-223 by the AML1/ETO oncoprotein. *Cancer Cell* **12**, 457-466 (2007).

210. Antequera,F., Boyes,J. & Bird,A. High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. *Cell* **62**, 503-514 (1990).

211. Kahle,P., Wernet,P., Rehbein,A., Kumbier,I. & Pawelec,G. Cloning of functional human T lymphocytes by limiting dilution: impact of filler cells and interleukin 2 sources on cloning efficiencies. *Scand. J. Immunol.* **14**, 493-502 (1981).

212. Kahle,P., Wernet,P., Rehbein,A., Kumbier,I. & Pawelec,G. Cloning of functional human T lymphocytes by limiting dilution: impact of filler cells and interleukin 2 sources on cloning efficiencies. *Scand. J. Immunol.* **14**, 493-502 (1981).

213. Vandesompele,J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, RESEARCH0034 (2002).

214. Gregory,R.I. *et al.* The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**, 235-240 (2004).

215. Pall,G.S., Codony-Servat,C., Byrne,J., Ritchie,L. & Hamilton,A. Carbodiimide-mediated cross-linking of RNA to nylon membranes improves the detection of siRNA, miRNA and piRNA by northern blot. *Nucl. Acids Res.* **35**, e60 (2007).

216. Griffiths-Jones,S. The microRNA Registry. *Nucleic Acids Res.* **32**, D109-D111 (2004).

217. Griffiths-Jones,S., Grocock,R.J., van Dongen,S., Bateman,A. & Enright,A.J. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* **34**, D140-D144 (2006).

218. Takai,D. & Jones,P.A. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl Acad. Sci. U. S. A* **99**, 3740-3745 (2002).

219. Takai,D. & Jones,P.A. The CpG island searcher: a new WWW resource. *In Silico. Biol.* **3**, 235-240 (2003).

220. Kent,W.J. *et al.* The human genome browser at UCSC. *Genome Res* **12**, 996-1006 (2002).

221. Barski,A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837 (2007).

222. Santourlidis,S. *et al.* Crucial role of DNA methylation in determination of clonally distributed killer cell Ig-like receptor expression patterns in NK cells. *J. Immunol.* **169**, 4253-4261 (2002).

223. Chen,C. *et al.* Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* **33**, e179 (2005).

224. Vandesompele,J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, RESEARCH0034 (2002).

225. Ota,T. *et al.* Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nat Genet* **36**, 40-45 (2004).

226. Shimomura,I., Shimano,H., Horton,J.D., Goldstein,J.L. & Brown,M.S. Differential expression of exons 1a and 1c in mRNAs for sterol regulatory element binding protein-1 in human and mouse organs and cultured cells. *J. Clin. Invest* **99**, 838-845 (1997).

227. McDonald,F.J. *et al.* Ubiquitin-protein ligase WWP2 binds to and downregulates the epithelial Na(+) channel. *Am. J. Physiol Renal Physiol* **283**, F431-F436 (2002).

228. Gardiner,K., Slavov,D., Bechtel,L. & Davisson,M. Annotation of human chromosome 21 for relevance to Down syndrome: gene structure and expression analysis. *Genomics* **79**, 833-843 (2002).

229. The MGC Project Team. The Status, Quality, and Expansion of the NIH Full-Length cDNA Project: The Mammalian Gene Collection (MGC). *Genome Res.* **14**, 2121-2127 (2004).

230. Landgraf,P. *et al.* A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**, 1401-1414 (2007).

231. Kass,S.U., Pruss,D. & Wolffe,A.P. How does DNA methylation repress transcription? *Trends Genet* **13**, 444-449 (1997).

232. Santourlidis,S. *et al.* Crucial role of DNA methylation in determination of clonally distributed killer cell Ig-like receptor expression patterns in NK cells. *J. Immunol.* **169**, 4253-4261 (2002).

233. Lee,Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-419 (2003).

234. Zeng,Y. & Cullen,B.R. Sequence requirements for micro RNA processing and function in human cells. *RNA.* **9**, 112-123 (2003).

235. Mattie,M.D. *et al.* Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies. *Mol. Cancer* **5**, 24 (2006).

236. Bandres,E. *et al.* Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues. *Mol. Cancer* **5**, 29 (2006).

237. Lawrie,C.H. MicroRNAs and haematology: small molecules, big function. *Br. J. Haematol.* **137**, 503-512 (2007).

238. Vandesompele,J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, RESEARCH0034 (2002).

239. Esteller,M. & Herman,J.G. Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. *J. Pathol.* **196**, 1-7 (2002).

240. Antequera,F., Boyes,J. & Bird,A. High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. *Cell* **62**, 503-514 (1990).

241. Saito,Y. *et al.* Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell* **9**, 435-443 (2006).

242. Lujambio,A. *et al.* Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res* **67**, 1424-1429 (2007).

243. Meng,F., Wehbe-Janek,H., Henson,R., Smith,H. & Patel,T. Epigenetic regulation of microRNA-370 by interleukin-6 in malignant human cholangiocytes. *Oncogene* (2007).

244. Lujambio,A. *et al.* Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res* **67**, 1424-1429 (2007).

245. Meng,F., Wehbe-Janek,H., Henson,R., Smith,H. & Patel,T. Epigenetic regulation of microRNA-370 by interleukin-6 in malignant human cholangiocytes. *Oncogene* (2007).

246. Kim,V.N. MicroRNA biogenesis: coordinated cropping and dicing. *Nat Rev. Mol. Cell Biol.* **6**, 376-385 (2005).

247. Rodriguez,A., Griffiths-Jones,S., Ashurst,J.L. & Bradley,A. Identification of mammalian microRNA host genes and transcription units. *Genome Res* **14**, 1902-1910 (2004).

248. Kim,Y.K. & Kim,V.N. Processing of intronic microRNAs. *EMBO J.* (2007).

249. Baskerville,S. & Bartel,D.P. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* **11**, 241-247 (2005).

250. Baskerville,S. & Bartel,D.P. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* **11**, 241-247 (2005).

251. Rodriguez,A., Griffiths-Jones,S., Ashurst,J.L. & Bradley,A. Identification of mammalian microRNA host genes and transcription units. *Genome Res* **14**, 1902-1910 (2004).

252. Baskerville,S. & Bartel,D.P. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* **11**, 241-247 (2005).

253. Saito,Y. *et al.* Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell* **9**, 435-443 (2006).

254. Lujambio,A. *et al.* Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res* **67**, 1424-1429 (2007).

255. Brueckner,B. *et al.* The human let-7a-3 locus contains an epigenetically regulated microRNA gene with oncogenic function. *Cancer Res* **67**, 1419-1423 (2007).

256. Baskerville,S. & Bartel,D.P. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* **11**, 241-247 (2005).

257. Kim,Y.K. & Kim,V.N. Processing of intronic microRNAs. *EMBO J.* (2007).

258. Saito,Y. *et al.* Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell* **9**, 435-443 (2006).

259. Brueckner,B. *et al.* The human let-7a-3 locus contains an epigenetically regulated microRNA gene with oncogenic function. *Cancer Res* **67**, 1419-1423 (2007).

260. Lujambio,A. *et al.* Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res* **67**, 1424-1429 (2007).

261. Scott,G.K., Mattie,M.D., Berger,C.E., Benz,S.C. & Benz,C.C. Rapid alteration of microRNA levels by histone deacetylase inhibition. *Cancer Res* **66**, 1277-1281 (2006).

262. Meng,F., Wehbe-Janek,H., Henson,R., Smith,H. & Patel,T. Epigenetic regulation of microRNA-370 by interleukin-6 in malignant human cholangiocytes. *Oncogene* (2007).

263. Hurteau,G.J., Carlson,J.A., Spivack,S.D. & Brock,G.J. Overexpression of the microRNA hsa-miR-200c leads to reduced expression of transcription factor 8 and increased expression of E-cadherin. *Cancer Res* **67**, 7972-7976 (2007).

264. Hurteau,G.J., Spivack,S.D. & Brock,G.J. Potential mRNA degradation targets of hsa-miR-200c, identified using informatics and qRT-PCR. *Cell Cycle* **5**, 1951-1956 (2006).

265. Pasquinelli,A.E. *et al.* Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **408**, 86-89 (2000).

266. Zhao,Y., Samal,E. & Srivastava,D. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* (2005).

267. Zhou,X., Ruan,J., Wang,G. & Zhang,W. Characterization and Identification of MicroRNA Core Promoters in Four Model Species. *PLoS Comput. Biol.* **3**, e37 (2007).

268. Lee,Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051-4060 (2004).

269. Johnson,S.M., Lin,S.Y. & Slack,F.J. The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. *Dev. Biol.* **259**, 364-379 (2003).

270. Johnston,R.J. & Hobert,O. A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans. *Nature* **426**, 845-849 (2003).

271. Gregory,P.A. *et al.* The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol.* **10**, 593-601 (2008).

272. Lau,N.C., Lim,L.P., Weinstein,E.G. & Bartel,D.P. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science* **294**, 858-862 (2001).

273. Lee,Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-419 (2003).

274. Kim,Y.K. & Kim,V.N. Processing of intronic microRNAs. *EMBO J.* (2007).

275. Kass,S.U., Landsberger,N. & Wolffe,A.P. DNA methylation directs a time-dependent repression of transcription initiation. *Curr. Biol.* **7**, 157-165 (1997).

276. Boyes,J. & Bird,A. Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. *EMBO J.* **11**, 327-333 (1992).

277. Hsieh,C.L. Dependence of transcriptional repression on CpG methylation density. *Mol. Cell Biol.* **14**, 5487-5494 (1994).

278. Nam,E.J. *et al.* MicroRNA Expression Profiles in Serous Ovarian Carcinoma. *Clin. Cancer Res* **14**, 2690-2695 (2008).

279. Taylor,D.D. & Gercel-Taylor,C. MicroRNA signatures of tumor-derived exosomes as diagnostic biomarkers of ovarian cancer. *Gynecol. Oncol.* **110**, 13-21 (2008).

280. Iorio,M.V. *et al.* MicroRNA signatures in human ovarian cancer. *Cancer Res* **67**, 8699-8707 (2007).

281. Gregory,P.A. *et al.* The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol.* **10**, 593-601 (2008).

282. Hurteau,G.J., Carlson,J.A., Spivack,S.D. & Brock,G.J. Overexpression of the microRNA hsa-miR-200c leads to reduced expression of transcription factor 8 and increased expression of E-cadherin. *Cancer Res* **67**, 7972-7976 (2007).

283. Gregory,P.A. *et al.* The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol.* **10**, 593-601 (2008).

284. Burk,U. *et al.* A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep.* **9**, 582-589 (2008).

285. Betz,B. *et al.* Denaturing high-performance liquid chromatography (DHPLC) as a reliable high-throughput prescreening method for aberrant promoter methylation in cancer. *Hum. Mutat.* **23**, 612-620 (2004).

8 - Acknowledgments

There were two groups of people that contributed at different levels to make this work possible. In the first group I include all those who directly contributed for the work. Among those, I would like to dedicate the first words to **Markus** (of course!). You gave me the opportunity to join your laboratory and to get contact with the fascinating world of miRNAs. The discussions maintained with you during these 3,5 years for sure provided me a good basis for learning and improving some skills. Joining your group I joint also the TRANS-NET Marie-Curie European network and that was definitely very important to me. Each meeting with the rest of the consortium represented an extra motivation to keep my work in progress and oriented to very specific goals. The contact with the consortium partners also represented a fresh air that helped me to keep my mind open. Markus, sorry for anything and thanks for everything.

On the hierarchy of people with direct responsibilities on my work, now should be mentioned **Simon** (the Greek). Well… what can I say to you? It is fare to say here that the idea of connecting miRNAs to epigenetics came from you. Your passionate epigenetics point of view worked as an inspiration and as a strong engine, forcing me to generate more and more data and to keep my mind open to new results and new ways of conducting research. Besides that, and although not recognized by many people in the group, you have great capacities for motivating and stimulating people. There are times when some words can easily make one person to overcome bad experimental results. It is important to ear them on the right moment from an older colleague. You, as no one else, were most of the times there, when necessary. You were aware of my capacities and fragilities and gave me a priceless help pushing up my strength and helping me to focus on the work. Thanks for all this and also for the help gave in reviewing my thesis. With all your help I almost forgot the final of Euro2004. – I said almost!

On the beginning of my work, I dedicated about one year trying to develop a new miRNA cloning strategy. During this time, I had countless discussions with **Ingo** (the Mac addicted). Amazing how different is the picture that I "painted" from you based on what people had previously said about you and the picture that later I "painted" from you after those discussions. That give me the feeling that, if people would concentrate more on what is really important on a working colleague, they could profit much more from having you around. Your experience-based knowledge was very important and I learned lots of technical aspects on molecular biology. Later on, already in the final of the work here presented, I counted again with your and **Kasia** (the horse rider) help to perform the Northern Blots. You both made my life much easier and a "big thanks" for that!

One other important contribution for this work came from **Sandra**. She made all the TaqMan miRNA measurements. I wouldn't be completely fare with the facts if I wouldn't say here that this contribution represented a big headache to me (and probably to you) because was not spontaneous and desired by my part but rather imposed. At the same time, to be fare I have also to say that the data collected by you was extremely important for all my work. Thanks for that.

I would like to address a special note to **Christophe** (the French chef), my TRANS-NET colleague and desk neighbour. Always so close but most of the times so far! Amazing how our projects developed and ended up so differently. The only thing I can say to you is to wish you the best success in this new period. After all, there is more life outside the laboratory. About the work… thanks for sharing knowledge and some of the molecular biology skills that you quickly developed here. And what about **Julia** (the Kaza!)?… I miss you. As a technician, you had lots of responsibilities here in this laboratory. Even so, you kept always interest on research and helped me mainly with your deep knowledge and care on cell culture. Thanks for that. Thanks also for your patience when the things were not so fine. **Teodora**, (the pessimist Bulgarian girl) could not be forgotten. You helped me a lot reviewing my thesis. I am quite sure that your critics led to an improvement of it. Also a short note should go to **Gaby**, **Gabi** and **Damian** (the most dedicated models of sequencers that ABI ever tried to make). How many sequencing reactions did you run for me during these 3,5 years? I counted: 1482!!. A small thanks for every single one of those make an enormous thanks after all.
Additionally, I would like to mention all the other people that participated in the group meetings and helped to discuss my work: **Martina**, **Nina**, **Christa**, **Kathrin**, **Judith**, **Jens**, **Foued**, **Christien**, **Natália** and **Margareta**.

As I wrote in the beginning of this section, there is still another group of people that contributed at a different level to make this work possible. In my opinion, a scientist should be, before all, a person. Therefore, not just the working networks are important but the social networks that we are able to build and maintain play a decisive role in everything we make. On that respect I would like to address some notes to very important people that have constituted a solid emotional basis during this time here in Germany.
In the very first place has to be mentioned **Joana**. I am quite sure that is nothing that I could tell here that I have not said you personally but, just for the others to know, you were decisive on the decision to come to Germany. After all it was not my decision but ours and that was what made this time so special. Thanks for every single moment. We will keep growing together. Next, I would like to mention my parents. **Luis** e **Ilda**. Thanks for supporting our decision to come to Germany and for your open mind to accept some other decisions not very concordant with your conceptions. Thanks also for all the patience during my short and always busy visits to Coimbra. I am not sure if I was able to find the right balance between the time spent with the friends and the time that the family wanted to spend with me. After my return to Portugal I will compensate all those moments.
To **Ana**, **Pedro** and **André**… what mixed feelings. The time will help to turn some previous attitudes in something else. Do not ask me "in what" nor "when". As I already wrote previously, there is more life outside the laboratory and, in the context of my personal projects, I just have to say you thanks! Your help has been priceless.
There is also a "big thanks" that should be given to Joana's parents, **Francisco** e **Clara** (the best future mother in law that one could have). You support us very much (as well as

the rest of your family) and make us feel closer from Portugal with every phone call and every package that you sent by post. The visit of old friends had also the capacity to bring us closer from Portugal and also closer from the times as student. Was very nice to have you here **Pardilhó, Dulce, Ricardo**, **Barros**, **Orlando** and **Gonçalo**. To all of you… the best success on your careers and please stay around.

On the group of people from the university, I would like to mention **Pedro**, it was nice to keep contact with you. All the moments smoking shisha and discussing, helped to keep my mind open. So many plans about the "...ases" were made by us, that someday one will become real.

I could not forget a group of people that I got to know here and that were wonderful teachers on those lovely disciplines about "friendship" and "world food": **Anupan** (from India), **Marina** (from France), **Luis** (from Brazil), **Margareta** (from Portugal), **Natália** (from Spain), **Teodora** (from Bulgaria), **Simon** and **Jenny** (from Greece), **Bärbel** and **Alexander** (Germans). Those times passed together around tables, as well as the tasty food, will not be forgotten. Please keep in contact. A special thanks to **Julia** and **Carsten**. Besides the nice dinners and evenings chatting, I do not forget that you introduced us some dancing steps. OK… probably we were not the most dedicated students but you can be sure that we had lots of fun!!

One other special note to all the people from the **TRANS-NET consortium**. How many interesting people are around! Good luck for everyone!

Obrigado!