Second Part: The microRNA in the mammary gland

I-Introduction

I-I The miRNA

I-I-a RNA silencing and miRNA

From the discovery of the structure of DNA (1953, Watson and Crick) to our days, big steps have been done in discovering the biological mechanisms with which DNA can carry the genetic information, can transmit it from one cell to one other and transfer it in a molecule of RNA and later in the structure of a protein, the final functional actor of the biology of a cell.
Even if the recent completion of the human, mouse and other eukaryotic genomes were important scientific milestones towards the understanding of eukaryotic biology, it's not easy to assess which regions of DNA have simply structural functions, which are really transcribed and code from a protein, how many genes are presents in a genome. In the human the last reported genome annotation has identified only 20000-25000 protein-coding genes (International Human Genome Sequencing Consortium, 2004), in contrast with previous higher estimate (Fields et al., 1994) and this raise some questions about the real definition of an eukaryotic gene. A possible answer could be found in the presence of alternative-splicing and of many non-coding RNA gene that do not have any clear "Open Reading Frame"(ORF) and are very difficult to predict from genomic sequences (Costa, 2005).
For many years RNA was considered to be just accessory molecules involved in mediating transcription and translation. RNA molecules are very versatile and their chemical properties allow them to form complex tertiary structures capable of performing several roles that were thought to be under the exclusive domain of proteins (Szymanski et al., 2003). They can interact with different proteins forming ribocomplexes, they can associate with specific DNA and/or other RNA sequences, controlling several aspects of gene regulation and different molecular connections in cells that are only partially discovered (Mattick, 2004).
In 1969, Britten and Davidson for the first time proposed that RNAs can solve the problem of eukaryotic gene regulation determine which genes are turned off and on by base-pairing against DNA (Britten and Davidson, 1969). The idea was abandoned with the discovery of a large class of protein transcription factors.
It was only in the 1990 that two different groups discovered for the first time the mechanism of RNA silencing like an internal mechanism of defense in petunia observing a 'cosuppression' of a transgenic and an endogenous gene after the introduction of the first one in the plant (Napoli et al, 1990; van der Krol et al., 1990). In some years this phenomenon was discovered in a broad spectrum of eukaryotes, from fungi to flies (Zamore and Haley, 2005) and it was shown to be involved in a plethora of mechanism like the regulation of

transcription, of chromatin structure, of genome integrity and, most commonly, of mRNA stability.

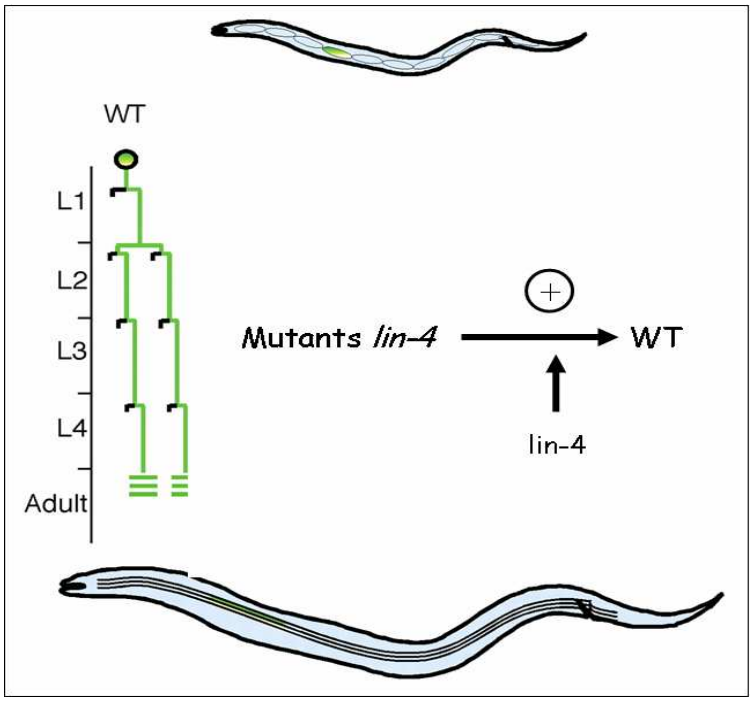
More precisely double stranded RNA-mediated gene silencing is a general term that refer to several pathways by which double stranded RNA can orchestrate epigenetic changes, repress translation, and direct mRNA degradation in a sequence-specific manner. These diverse effects of non-coding RNA on gene expression have been termed RNA interference (RNAi) (Rao and Sockanathan, 2005). It is carried out by tree different class of small non-coding RNA: small interfering RNAs (siRNAs), repeated-associated small interfering RNAs (rasiRNAs) and microRNAs (miRNAs) that are distinguished by their origin, but that share a common set of proteins in the mechanism of production and action.
In the main lines RNA interference is triggered by double strand RNA (dsRNA) precursor that vary in length and origin and that is processed in the cytosol by a specific ribonuclease called Dicer into a short RNA duplex of 21 to 28 nucleotides in length which determines in a sequence-specific way which mRNA should be degradated. This short double stranded RNA guides a protein complex to the recognized mRNA target that is silenced by cleavage or translational repression.
In nature double stranded RNA can be produced by RNA polymerization starting from a viral RNA, for example, or by hybridization of overlapping transcripts, for example from repetitive sequence such as transgene arrays or transposons. Such dsRNA give rise to siRNAs or rasiRNAs which generally guide mRNA degradation or chromatine modification.
In the contrary endogenous transcripts that contain complementary or near-complementary 20-to 50-base-pair inverted repeats fold back on themselves to form dsRNA hairpins. These are processed in miRNA, that in the most part of cases repress translation, but that can also guide the degradation of mRNA (see review: Meister and Tuschl, 2004). This class is predicted to regulate alone one third of all human genes.

I-I-b The discovery of miRNAs

MiRNA are a class of evolutionary conserved, small (19-25 nt) non-coding RNAs that negatively regulate gene expression at the post-transcriptional level.
The finding member of miRNA family, *lin-4*, was identified in *C.elegans* through a genetic screen of mutants for defects in the temporal control of post-embryonic development (Chalfie et al., 1981; Ambros, 1989).
In *C.elegans* the post-embryonic development pass through four different larval stages (L1-L4) in which cell lineages have distinct characteristics. *Lin-4* encodes a 22-nt non coding RNA that is partially complementary to a short (7 nt) conserved site in the 3'-untranslated region (3'UTR) of the *lin-14* gene, its target (Lee et al., 1993; Wightman et al., 1991). *Lin-14* encodes a nuclear protein that is normally downregulated at the end of the first larval stage to

allow the developmental progression into the second larval stage (Ruvuk and Giusto, 1989) (figure 1).

Figure 1. The stages of development of *C.elegans* and the mechanism of action of lin-4.



Mutants for *lin-4* do not progress in the second larval stage showing reiteration of specific cell-division pattern of first larval stage even late in the development. Opposite phenotypes were observed in mutants deficient for *lin-14* and even before the molecular identification of *lin-4* and *lin-14* these genes were placed in the same regulatory pathway on the basis of opposite phenotypes and antagonistic genetic interaction. After a series of molecular and biochemical studies was demonstrated that the direct and imprecise binding of *lin-4* to the 3'UTR of *lin-14* was able to reduce the amount of the LIN-14 protein without changing in the level of *lin-14* mRNA (Olsen and Ambros, 1999). These evidences supported a model in which the *lin-4* RNA pair to the 3'UTR of *lin-14* to specify translational repression of it as part of the regulatory pathway that control the timing of development in the worm. Also another target of *lin-4* was discovered, *lin-28*, a cold-shock-domain protein that initiates the developmental transition between L2 and L3 stages (Moss et al., 1997).

For seven years any others miRNA was identified in nematodes and there was no evidence of any similar non coding RNAs beyond nematodes.

In 2000, the second miRNA was discovered, *let-7*, also using forward genetics in *C.Elegans*. *let-7* encodes a temporally regulated 21- nt RNA that binds to the 3' UTR of *lin-41* and *lin-57*, inhibiting their translation (Lin et al., 2003; Abrahante et al., 2003; Slack et al., 2000; Vella et al., 2004). *let-7* controls the transition from L4 stage to the adult stage (Reinhart et al., 2000). The identification of *let-7* not only suggested the existence of a new class of molecular regulator of the timing of developmental transition, but also it opened the way to

the discovering of many *let-7* homologs in other species. Pasquinelli et al. (2000) found, first through BLASTN searches, the existence of homologs of *let-7*, later experimentally by Northern blot, their expression in all stages of development of *D.Melanogaster* and in all tissues of human. They went on to find homologs in all vertebrates studied and in the same time siRNA were discovered and it was shown that components of the siRNA apparatus processing RNA are also involved in *lin-4* and *let-7* expression. This suggested that these small RNA could be more common than just *lin-4* and *let-7*. In less than one year, thanks to the work of three different labs, approximately one hundred of miRNA were cloned from flies (20 in *Drosophila*), worms (60) and human cells (30) (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). This first group of miRNA identified showed the same length, the same way of production from an endogenous precursor, and they were generally evolutionary conserved, some quite broadly, others only in more closely related species such as *C.elegans* and *C.briggsae*. Even in this first group some showed a tissue or cellular-specific expression, differently from *lin-4* and *let-7*, whose expression is temporally-regulated.

Intensified cloning efforts have revealed numerous additional miRNA genes in plants, mammals, fish, worms, flies and even virus (Lagos-Quintana et al., 2002, 2003; Mourelatos et al., 2002; Ambros et al., 2003; Aravin et al., 2003; Dostie et al., 2003; Houbaviy et al., 2003; Kim et al., 2003; Lim et al., 2003) giving origin to the first microRNA registry, a public list to catalog miRNA and to facilitate the naming of newly identified genes (Griffiths-Jones, 2004). More than 330 miRNA have been cloned in humans (Griffiths-Jones et al., 2006; Hsu et al., 2006) and bioinformatic tools predict that they could be 1.000 and can represent up to the 3-5% of all the genes in the human genome.

Their high number, their spatiotemporal, tissue- and cell-type expression and the extensive conservation strongly indicated an important role in development, like it was supported later (Bartel 2004; He and Hannon, 2004).

I-I-c Biogenesis and mechanism of action

MiRNAs genes are widespread in the genome and their genomic localization and organization vary together with their mode of transcription (Bartel, 2004).

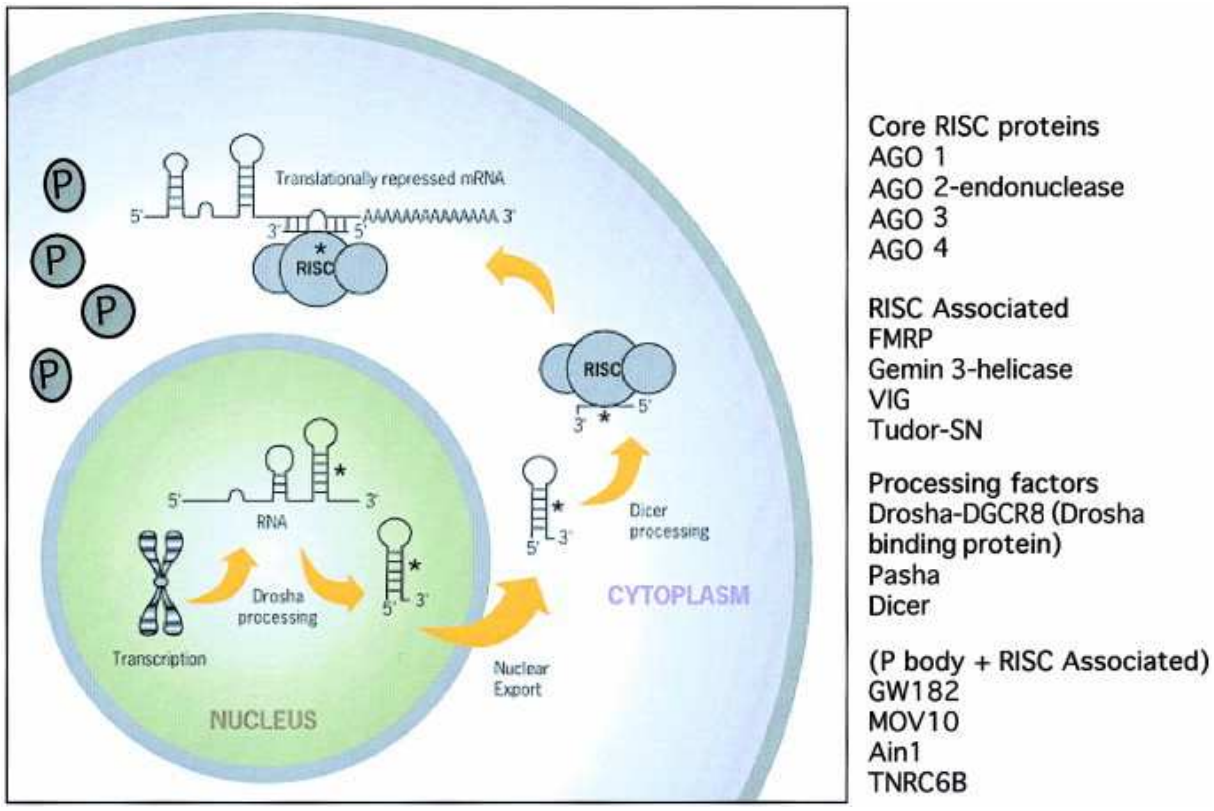
Most mammalian miRNA genes come from regions of the genome quite distant from previously annotated genes and are considered as independent transcription units with specific promoter core elements and polyadenylation signals (Pasquinelli, 2002; Cullen, 2004; Kim and Nam, 2006).

The remaining are located part in long non coding RNA transcript or, in the majority of cases, in introns of protein encoding genes. These are not transcribed from their own promoter, but they are transcribed together with their host genes, for example processed from introns by

alternative splicing (Aravin et al., 2003; Lagos-Quintana et al., 2003; Lau et al., 2001; Lee and Ambros, 2001).
They could be present singularly or in cluster. The most part of human and *Drosophila* miRNA are clustered. These clusters are single transcription units and produce poly-cistronic transcripts. Often the miRNAs within the same genomic cluster are related to each other, like it happens, for example, for the orthologs of *C.Elegans lin-4* and *let-7*, that are coexpressed from the same cluster in fly and human genomes (Aravin et al., 2003; Bashirullah et al., 2003; Sempere et al., 2003). This suggests that in the same group miRNAs with no apparent sequence homology could have functional relationship.

Not all the molecular steps that elucidate the biogenesis of a miRNA starting from its transcription to its maturation are well established and the general model often refers to the biogenesis of the first miRNA identified, *lin-4* (Conrad et al., 2006).
The generation of miRNAs is a complex multistep process that occurs in two separate cellular compartement, the nucleus and the cytoplasm, and during which miRNA pass four different stages and structure: primary miRNA, precursor miRNA, duplexed miRNA and active miRNA (figure 2).
A primary miRNA (pri-miRNA) of 100 to more then 1000 nucleotides in length, is transcribed from the genome by RNA polymerase II in the nucleus (Song et Tuan, 2006).

Figure 2. Biosynthesis and mechanism of actions of miRNAs and the main molecular components involved



Initially RNA polymerase III was the candidate for the transcription of miRNAs, like it happens for some of the shorter noncoding RNAs, including tRNAs, 5S ribosomal RNA and the U6 snRNA, but numerous evidences supported the activity of the RNA polymerase II (see review: Di Leva et al., 2006). In the beginning several long transcripts comprising miRNAs were identified in expressed sequence tags and their complex expression control was typical of those transcribed by RNA polymerase II, later the association with this enzyme was demonstrated clearly. Until now only a few different pri-miRNA have been isolated and characterized, three from human, one from *C.Elegans*, one from plants and they all are capped, polyadenylated and apparently noncoding (Cullen, 2004).

In the nucleus the pri-miRNA is converted to precursor miRNA or pre-miRNA , a 60-70 nucleotides stem loop intermediate, through the cleavage activity of the Drosha enzyme, a nuclear Ribonuclease III endonuclease which cleaves the flank regions of pri-miRNA (Lee et al., 2002, 2003; Zeng and Cullen, 2003).
Drosha can cut only pri-miRNAs that have a large terminal loop (>10 nt) in the hairpin and a stem region one turn bigger than the precursor, 5' and 3' single-stranded RNA extensions at the base of the future miRNAs (Filipowicz et al., 2005, Tomari and Zamore 2005). The hypothesis is that Drosha recognizes the primary precursor through the stem-loop structure and then cleaves the stem at a fixed distance from the loop, liberating the pre-miRNA and determining one end of the mature miRNA. It's not clear in which way Drosha recognizes the pri-miRNA stem-loop from stem-loop of other RNAs. The pre-miRNA presents a 5' phosphate and 3'hydroxy termini and two or three nucleotides with single-stranded overhanging ends, classic characteristics of Ribonuclease III cleavage of dsRNAs (Di Leva et al., 2006). It is 60-70 nucleotides long with an imperfect stem-loop structure. In the stem part of one of the two arm is present the sequence of the mature miRNA and in the other arm the near complementary miRNA* that will be later eliminated.
The pre-miRNA is actively transported from the nucleus to the cytoplasm. Pre-miRNA interacts with the export receptor Exportin-5 (Exp5) (Kim, 2004) and RanGTP forming a nuclear heterotrimer that promote the stabilization of pre-miRNA and is exported to the cytoplasm. Once the heterotrimer reach the cytoplasm through the nuclear pore, the RanGTP is hydrolyzed to RanGDP and the pre-miRNA is released (Di Leva et al., 2006).

In the cytoplasm the pre-miRNA is processed into 18-22 nucleotide imperfect double strand RNA duplex (miRNA: miRNA*) by the cytoplasmic Ribonuclease III, Dicer, that acts, in humans, with the trans activator RNA (tar)-binding protein , TRBP (Chendrimada et al. 2005).
Dicer contains a putative helicase domain, a DUF283 domain, a PAZ (Piwi-Argonaute-Zwille) domain, two tandem RNase-III domains and a dsRNA-binding domain 8 (dsRBD). The PAZ

domain is responsible of the interaction of Dicer with the 2-nucleotide 3' overhangs of dsRNA such as the pre-miRNA. The efficient Dicer cleavage also requires the presence of the overhangs and a minimal stem length. The model assumes that the PAZ domain of Dicer recognizes the end of the pre-miRNA and can position the site of the second cleavage on the stem of the precursor. The variable size of the product, from 18 to 22 nt, results from the presence of bulges and mismatches on the pre-miRNA. Efficient cleavage requires dimerized RNase III domains because the functional catalytic site resides in the interface of the dimmer (see review: He and Hannon, 2004).
Like Dicer also Drosha has two tandem RNase-II domains. The exact biochemical mechanism that guides the cleavage has not been elucidate, but it's probable that it shares closely related mechanism for processing miRNA.
In plant no Drosha homologues have been found and it suggests that the maturation of miRNA from long primary transcript should occur differently comparing to the animal model. However there are four Dicer homologues in *Arabidopsis thaliana*, DCL1, DCL2, DCL3, DCL4, two of which contain nuclear localization signals. It seems possible that in plant the Drosha function is carried out by one or more specialized Dicer. In plants deficient for DCL1 not only the production of some miRNA is reduced, but also is not detected the accumulation of the corresponding pre-miRNA. A model in which Dicer specialised enzymes catalyse both Drosha and Dicer cleavage for the maturation of miRNAs inside the nucleus has been built. The functional specificity of different Dicer enzymes in organisms with multiple Dicer homologues has recently been indicated also in *Drosophila* and the function of Dicer seems not simply restricted to the cleavage, but also correlated to the initiation of RNA-silencing in the effector complex (see review: He and Hannon, 2004).

Only one strand of the dsRNA contains the miRNA that preferentially enters the RNA-induced silencing complex (RISC), the effector protein complex in which the miRNA pairs the mRNA target and produces its degradation or the inhibition of its translation into a protein (see review: He and Hannon, 2004).
This effector complex shares so much core components with that of siRNA that it's generally called RISC for siRNA and miRNA even if in humans it's called miRNP, miRNA ribonucleoprotein, after the identification of the proteins that constitute it (Hutvagner and Zamore, 2002; Mourelatos et al., 2002). Several proteins have been purified and identified as essential components of RISC, but only a few have been functionally characterized (see review: Di Leva et all., 2006). RISC has been purified in many organisms and it always contains a member of the Argonaute protein family, which is thought to be a core component. Many Argonaute proteins were already identified in worms, fungi and plants and shown to be implicated in the mechanism of RNA silencing. These evolutionary conserved proteins of approximately 100 kDa are called also PPD proteins because they all share the PAZ and PIWI domains. The first one domain has the function to bind weakly to

single-stranded RNA and also to double stranded RNA; this suggest that this protein can have the ability to bind miRNA before and after its association with the mRNA target. Structural and biochemical studies have proved that the Argonaute proteins are the target-cleaving endonuclease of RISC, and that the complex is coordinated also by other proteins whose functions are not really understood, like the RNA-binding protein VIG, the Fragile-X related protein in *Drosophila*, the exonuclease Tudor-SN, and many other putative helicases (Nelson et al., 2003). In humans miRNP is constituted by the Argonaute protein called eIF2C2 (Martinez et al., 2002), and other two helicases, Gemin3 and Gemin4.

When the miRNA strand of the miRNA: miRNA* duplex is loaded into the RISC the miRNA* is unwind and rapidly degraded. The target specificity and probably also the functional efficiency of a miRNA requires that the mature miRNA strand from the duplex be selectively incorporated into the RISC for target recognition (see review: Bartel, 2004). What is the mechanism for choosing which of the two strand enters the RISC? Some evidences show that the strand that enter is nearly always the one whose 5' end is less tightly paired (Khvorova et al., 2003; Schwarz et al., 2003). After the cleavage of Dicer the stability of the 5' ends of the two strand of the duplex is usually different. It seems that helicases present in the RISC take with the same frequencies both the end of the two strand before beginning to unwind the duplex and that finally the relative ease of unwind the less stable facilitate its preferential incorporation into the RISC, determining the asymmetrical assembly of the complex.

The precise mechanism that underlies post-transcriptional repression by miRNAs still remain unknown. We know that two processes exist for miRNA-mediated gene regulation: degradation of the target mRNA and translational repression, depending largely on the degree of the complementarity between the miRNA and the target. The miRNA will specify cleavage if the mRNA has sufficient complementarity to the miRNA or it will repress productive translation if the mRNA is only partially complementary, but has a sufficient number of miRNA complementary sites (Hutvagner and Zamore, 2002; Zeng et al., 2002, 2003; Doench et al., 2003). This model is supported from many evidences, but it can not be considered a general rule because there was at least one case of a plant miRNA, miR-172 in *A.thaliana*, that regulates APETALA2 via translational repression despite the near-perfect complementarity between the miRNA and the target (Aukerman and Sakai, 2003; Chen, 2003).
When miRNA guides the cleavage the cut happens in a precise position, between the nucleotide pairing to residues 10 and 11 of the miRNA, like it happens for siRNA (Kasschau et al., 2003; Hutvagner and Zamore, 2002; Llave et al., 2002) and it appears to be determined relative to miRNA residues, not to miRNA: target base pairs. After cleavage of the mRNA the miRNA remains intact and can guide the recognition and destruction of additional mRNA (Tang et al., 2003).

The mechanism of translational inhibition was first observed and studied looking at the RNA-silencing of *lin-4* over *lin-14* in *C.Elegans*. It was observed that in the animal kingdom miRNA typically mediate translational repression rather than mRNA cleavage, that is more common in plants, even if recently one miRNA in *Drosophila*, miR-196, was found to direct mRNA cleavage of its target, Hoxb8 (Yekta et al., 2004). It seems that in animal generally the miRNAs have a lower degree of complementary to mRNA targets comparing to the nearly perfect base pairing of plant miRNA to the corresponding target, that in plants generally guides to its destruction (Hake, 2003). It was observed that the cooperative action of multiple RISC provide the most efficient translational inhibition (Doench et al., 2003). This correlates with the presence of multiple miRNA complementary sites in most genetically and computationally identified targets of metazoan miRNA. It has been proposed that different miRNA can regulate the same target and that exists a combinatorial control (Reinhart et al., 2000; Abrahante et al., 2003; Lin et al., 2003).
We know that the complementary sites for the known metazoan targets reside in the 3' UTR of mRNA, in contrast with the target complementary sites in plants, that are located throughout the transcribed regions of the target gene (see review: Bartel, 2004). It was demonstrated that in metazoan the most important site of complementarity to the target on the miRNA sequence is a short portion at the 5' end of seven nucleotides, from residues 2 to 8. Actually this short sequence is the most conserved among homologous metazoan miRNA (Lewis et al., 2003; Lim et al., 2003), it was observed to be perfectly complementarity to multiple 3' UTR sites involved in post-transcriptional repression also in invertebrate (Lai, 2000), moreover this heptamer seems to be the most useful to productive prediction of target mRNA and is the most important complementary site also in plant miRNA. We don't know why the complementarity to the 5'end is so universally important, but understanding the mechanism of pairing of miRNA to the mRNA in the RISC will also help to reveal the process of translational repression.

I-I-d Approaches to microRNA discovery (see review : Berezikov et al., 2006)

The first step to discover and understand the biology of miRNA is to isolate and identify the miRNAs expressed in the cells and organisms of interest. Since the discovery of the first miRNA, *lin-4*, in *C. Elegans*, many miRNAs have been identified or predicted in a wide array of organism. In 2003, the rapid growth of the number of miRNA genes led to the creation of a comprehensive an searchable database of published miRNA sequences via a web interface : miRNA Registry (http://microrna.sanger.ac.uk/sequences) (Griffiths-Jones, 2004). This was created with two objectives: to avoid to assign unique names to distinct miRNAs and to provide a database for all miRNAs sequences, including the stem-loop sequences, the genomic location, homologous sequences and possible target predictions. The miRNAs are annotated with numerical identifiers based on sequence similarity. For homolog miRNAs in

different organisms, itm is usual to assign the same nameon the similarity of the 22-nt mature sequence. Identical mature forms are assigned the same name and, if identical miRNAs are produced from different genomic loci they are differentiated by suffixes, such as “miR-16-1” and “miR-16-2”; if there are differences in one or two bases they are denominated with a final different letter such as “miR-181a” and “miR-181b”. If two miRNA derive from the two arms of the same precursor it is added to the miRNA name the suffix 5p and 3p to identify the two different arms, until the data will confirm which form is predominantly expressed, while the species less expressed is normally denoted by an asterix (Di Leva et al., 2006).
In October 2006 the miRNA Registry contained 4.361 entries from 40 organism including viruses and mammals, counting 332 human miRNA (Croce, 2006).

It was necessary a uniform definition of miRNA to annotate new “candidate miRNA” like true miRNA to prevent misclassification of other types of small non coding RNA like miRNAs (see review: Berezikov et al., 2006).
MiRNAs were defined as non coding RNAs that fulfill the following combination of expression and biogenesis criteria:

1) mature miRNA should be expressed as a distinct transcript of approximately 22 nt that is detectable experimentally (by Northern blot analyses, cloning, real-time quantitative PCR. in situ hybridisation, primer extension..);
2) mature miRNA should originate from a precursor with a characteristic secondary structure, such as hairpin or fold-back, without any large internal loops or bulges, and miRNAs should occupy the stem part of the hairpin;
3) mature miRNA should be processed by Dicer (Ambros et al., 2003).

The definition implies that miRNA should have a demonstrated function, however biological function has been elucidated only for a few miRNA and the criteria established for miRNA classification (Ambros et al., 2003) do not include the requirement of a biological role. Instead, an optional but commonly used criterion is the phylogenetic conservation of the mature miRNA, an indirect indication of a possible function. Strictly speaking the term “candidate miRNA” should be used as long as the function of miRNA is unknown, but practically evidences of expression of a 22-nt transcript and of the presence of an hairpin precursor are sufficient to classify a sequence as a miRNA.
All approaches to discovering miRNAs are based on these definition and can be split in two groups:

1) experiment-driven methods, in which the expression of small miRNA is first experimentally established and structural requirements for the precursor are searched later by bioinformatic tools;
2) computation-driven approaches, in which candidate miRNA are first predicted in genome sequence using structural features and expression is demonstrated later experimentally (Berezikov et al., 2006).

In the beginning forward genetics methods were able to identify the first miRNAs genes, *lin-4* and *let-7*, but since then only four additional miRNAs, *bantam*, *miR-14* and *miR-278* in *Drosophila melanogaster* (Brennecke et al., 2003; Xu et al., 2003; Teleman and Cohen, 2006) and *lys-6* in *C.elegans* (Johnston and Hobert, 2003), have been discovered by forward genetics approaches. The inefficiency of this methods can be explained by the difficulties in targeting by spontaneous or induced mutagenesis the miRNA sequence and specially the "seed sequence" of 7 nucleotide determinant for their functionality, considering their redundancy (Abbott et al., 2005), which often tolerate mutations that does affect the 'seed sequence' and does not result in strong variation in the phenotype of the mutants.

The preferred approach to the identification of miRNAs is to sequence size-fractionated cDNA libraries. An initial protocol useful for cloning small interfering RNA (Elbashir et al., 2001) was shown to be adapt also for identifying many miRNA (Lagos-Quintana, 2001). Later little variations of it were developed independently (Lau et al., 2001; Lee and Ambros, 2001) but all these successful protocols follow the same principles (Aravin and Tuschl, 2005). In the main lines an RNA sample is separated in a denaturing polyacrylamide gel and the size fraction corresponding to the miRNA is recovered. Then 5' and 3' adapters were attached to the RNAs and a RT-PCR is carried out. The fragments are cloned into vectors to create a cDNA library. The clones are sequenced and analyzed to search the genomic origin of the small RNAs. Bioinformatic tools are required to check if the hairpin precursor is encoded in the genomic regions where the small RNAs have been localized and if this precursor is conserved in other species. This analyses is complicated because hairpin structure are common in eukaryotes and are not a unique features of miRNAs (Lin et al., 2006; Shen et al., 1995) and moreover miRNAs should be distinguished from others types of endogenous small RNAs (Aravin and Tuschl, 2005; Kim and Nam, 2006).
The limit of cloning protocol is the difficulty to discover miRNAs that are expressed at a low level, in specific stages or specific cell types. Moreover some miRNAs could be hard to clone due to physical properties correlated to sequence composition or post-transcriptional modifications, such as editing or methylation (Luciano et al., 2004; Yang et al., 2006).
Surveying genomic sequences to predict miRNAs became popular after initial cloning efforts generated sufficient information about miRNA properties to recognize a set of distinctive miRNA features (Berezikov and Plasterk, 2005; Bentwich, 2005). On the basis of the particular features of miRNAs different approaches have been developed to predict miRNA, but all of them use secondary structure information, many rely on phylogenetic conservation of both miRNA sequence and structure, other methods asses the thermodynamic stability of hairpins and refers to sequence and structure similarity of known miRNAs, or search miRNAs on the genome near known miRNAs already localized.
Many software, like MiRScan, snarloop, miRSeeker were developed to search miRNA on the basis of conservation criteria referring to hairpins structure. The genome of *C. Elegans* (Lim et al., 2003; Grad et al., 2003), of *D. melanogaster* (Lai et al., 2003) and the human (Lim et

al., 2003) were analysed and the number of predicted miRNA strongly extended on the basis of haipin sequence similarity to experimentally confirmed miRNAs. The potential target sequences of miRNAs have been analyzed in the 3'UTR of genes in search of complementary sequence to 'seed sequences' of known miRNAs. Using conserved motifs that did not match to any known miRNAs it was possible to predict other miRNAs candidates in human (Xie et al., 2005). Similar approaches have been recently applied to the prediction of miRNAs in *A.Thaliana* (Adai et al., 2005), flies and worms (Chan et al., 2005).

On the basis of the lower free energies of folding of miRNAs comparing to tRNAs and rRNAs (Bonnet et al., 2004) was set a new software that combines thermodynamic stability criteria with conservation criteria, Rnas (Washietl et al., 2005a, 2005b), that was successful to predict additional miRNAs in various organisms (Hsu et al., 2006; Missal et al., 2006).

Recently other alignment-type methods have been develop to identify homologs of known miRNAs "aligning" potential miRNA with known one at both sequence and structural level (Legendre et al., 2005; Nam et al., 2005; Wang et al., 2005).

All methods that rely on phylogenetic conservation can not predict non conserved miRNAs. For this reason some new *ab initio* approaches were developed that use only intrinsic structural features of miRNAs and not external informations (Bentwich et al., 2005; Sewer et al., 2005; Xue et al., 2005; Pfeffer et al., 2005). Each of these methods builds classifiers that can measure how a candidate miRNA is similar to known miRNAs on the basis of several features, such as free energy of folding, length of the perfect longest stem, average size of symmetrical loops, proportion of different nucleotides in the stem, etc..(Sewer et al., 2005), to which a model assigns different weights and an overall score result is measured for each candidate miRNA.

All these computationally predicted miRNAs need to be validate experimentally. Validation approaches can be split in two categories: those that demonstrate only miRNAs expression, like Northern blot analyses (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001) and *in situ* hybridization (Wienholds et al., 2005; Kloosterman et al., 2006; Nelson et al., 2006), and others determine the exact sequence of the mature miRNA in the precursor sequence, like cloning strategy and primer extension assay (Seitz et al., 2004) or RNA-primed array-based Klenow extension (RAKE) (Nelson et al., 2004).

I-I-e Strategy to determine biological functions (see review : Krutzfeldt et al., 2006)

The elucidation of the general mechanism of miRNA function in the regulation of gene expression suggests a gene regulatory model in which nuclearly encoded genetic information is not only transcribed and translated into proteins, but at the same time regulates these processes through non coding miRNA. This paradigm adds a new level of regulation and fine control of gene expression that is likely to be important for the maintenance of many, if not all, cellular functions.

In spite of our ability to identify miRNA and elucidate their biogenesis and the basic mechanism of actions, very little is known regarding miRNA function. With the near completion of the miRNA inventory the focus is shifting to elucidation of their biological role. For this purpose other scientific aspects have been studied, and the corresponding technology have been developed, including the analyses of the possible miRNA target by bioinformatics prediction algorithms, preliminary inspections of the localization and the effects of their expression by *in vitro* or *in vivo* expression studies, by reporter assays, *in situ* hybridizations, over expression and silencing technologies.

Different complementary strategies can be useful to begin to study a miRNA of interest: it is possible to examine the profile of its expression in different cellular contexts, and indirectly make a first hypothesis on its possible role, predict or identify its molecular target, strictly connected to its function, before over expressing or silencing a miRNA *in vivo* and *in vitro*, that is a task that requires a defined model and precise technical competences, even if the technology is now available.

From the first study in *C. Elegans* and in *D. Melanogaster* was evident that miRNA has not only-spatio-temporal, but also tissue and cell-type specific expression. For this reason many commercially miRNA microarrays including the content found in Sanger miRBase 7.0 (http://microorna.sanger.ac.uk) were developed making it possible to monitor tissue-specific miRNA expression and regulatory changes in developmental, physiological and disease states. First results using oligonucleotide microarray confirmed the existence of several tissue-specific miRNAs (BaskerVille and Bartel, 2005; Barad et al., 2004; Nelson et al., 2004; Thomson et al., 2004) that may suggest that some of them have an organ- or cell type-specific functions. Microarrays have also been used to study miRNA expression profile during differentiation of cells, such as myoblasts (Chen et al., 2006) and preadipocytes (Esau et al., 2004) or in disease state, most notably human cancer, such as in B cell chronic lymphocytic leukemia (Calin et al., 2002), in colon carcinomas (Michael et al., 2003), in small lung carcinomas (Johnson et al., 2005; Takamizawa et al., 2004). It was shown for few specific miRNA that their level of expression can decrease or increase typically in a particular differentiation cellular stage or neoplastic tissue raising the idea that miRNA profiling could contribute to more precise tumor classifications and predict therapeutic outcomes in the future.

Several techniques have been developed to visualize miRNA expression *in vivo*. In *C. elegans*, it was possible to study the activity of the miRNA promoter of miRNAs using reporter cassette that contains the miRNA promoter regions fused with the sequence of the green fluorescence protein (GFP) or β-galactosidase, the reporter gene (Johnson et al., 2003). A method to detect the presence of a specific miRNA in tissues uses the “sensor” transgene, which constitutively express a reporter gene that contains sequences complementary to the miRNA of interest in its 3’ UTR region (Mansfield et al., 2004). In this case in the tissues or cells in which the miRNA is expressed the activity of the reporter gene will be blocked. This

method has potentially excellent spatial temporal resolution, but it is not known if it can be used for the detection of low expressed miRNAs. The most frequently applied method to visualize miRNA expression to date is *in situ* hybridization and in particular a variation of it that uses special probes Locked Nucleic Acid- (LNA) modified probes able to detect short sequences like miRNAs. This technique has already been applied successfully to identify miRNA expression in mouse embryos (Kloostermam et al., 2006).

Even if several independent groups have established computational algorithms designed to predict target genes of miRNA sequences (John et al., 2004; Kiriakidou et al., 2004; Krek et al., 2005; Rajewsky, 2006; Lewis et al., 2003) there is a big lack of experimental evidence that validate this sequences like the corresponding target of miRNAs.

Moreover, also due to the functional mechanism of miRNA, that do not require necessary a strong degree of sequence complementarity to the target and that do not exclude binding of multiple miRNA to the same mRNA, the computational prediction of a target is difficult and on average 200 genes have been predicted to be regulated by a single miRNA (Krek et al., 2005; Lewis et al., 2003).

To date several methods have been established to show experimentally the miRNA regulation of a putative target. One of the most used is the luciferase reporter construct, containing the target 3'UTR with the putative binding site downstream of the reporter coding region. These constructs are used to transfect cells expressing the relevant miRNA, along with vectors carrying mutant versions of binding sites. Evidence for miRNA activity can be demonstrated if wild-type reporters have less activity than their respective mutants. One other approach uses antisense 2'-O-methyl modified oligoribonucleotides to inhibit miRNA expression and provokes some loss-of-function effect (Chen et al., 2006; Poy et al., 2004; Schratt et al., 2006). Another approach tries to determine miRNA target increasing the intracellular concentration of a miRNA by transfection of homologous synthetic short interfering RNAs or recombinant adenoviral infection and measuring differential gene expression by microarray (Krutzfeldt et al., 2005; Lim et al., 2005).

Induced expression of miRNAs was the initial step that identified miRNA function in many model organisms or mammalian cell systems.

Transient overexpression in cell-based assays is easily achieved by transfection of a double-stranded RNA similar to Dicer cleavage product, but long-term overexpression in cultured cells or mouse models depends on the transfection of a plasmidic vector that carry a specific construct for the expression of a miRNA. This construct is relatively simple and is the same used for protein-coding mRNA; introducing the sequence of the precursor together with a strong constitutive promoter is sufficient to overexpress a miRNA (He et al., 2005; Hayashita et al., 2005). It is possible to introduce these vectors into adenovirus or retrovirus (Chen et al., 2004) system and then transfect cultured cells or inject them in mouse tissues *in vivo*. Tissue-

specific overexpression of a miRNA *in vivo* can be obtained also generating transgenic mice, even if the technology requires time and higher competences.
Studies that are based only on overexpression must be taken with caution because misexpression of miRNA could target genes otherwise not affected in physiological context and the results should be confirmed by loss-of-function experiments.

The technologies to silence miRNA and generate loss-of-function mutants can be divided in genetic and nongenetic approaches.
The first class developed with the parallel development of DNA recombinant technology and the recent generation of mutant mice has been invaluable in the elucidation of miRNA gene function.
In particular three class of experiment can be design to disrupt miRNA-mediated gene regulation and infer their possible role.

1) The generation of mice with mutated alleles of Dicer1 leads to the deficiency of all mature miRNAs (Kanellopoulou et al., 2005; Harfe et al., 2005; Harris et al., 2006); the phenotypes analysed show the important collective functional role of miRNAs in many developing tissues, but can not provide information on the exact role of individual miRNA. For this reason the injection of a singular miRNA in Dicer-null mutant can restore the expression of a specific miRNA and discover the contribution of individual miRNA. This approach has been successful in zebrafish (Giraldez et al., 2005).
2) The generation of a knockout mutant for a miRNA that exist in cluster can interfere with the

proper folding and processing of the polycistronic transcript, affecting the expression of neighboring miRNAs and provoking a visible changing in the phenotype (Ying et al., 2005).
To date there is no evidence for any miRNA knockout in any animal models.

3) The generation of mutants with mutating binding sites in the 3'UTR of the target gene, but also this approach is not actuated till now.

Nongenetic approaches to silence miRNA function use the transfection or injection of synthetic oligonucleotides that act like chemical inhibitors of miRNAs.
1) 2'-O-methyl-modified oligoribonucleotides complementary to the miRNA act irreversibly like stoichiometric inhibitors and have been used in cell lines (Meister et al., 2004), *C. elegans* (Hutvagner et al., 2004) and *D. melanogaster* (Leaman et al., 2005).
2) 'antagomiRs', cholesterol-conjugated single-stranded RNAs complementary to miRNAs have been injected in mouse tissues *in vivo* (Krutzfeldt et al., 2005). This silencing is dose-dependent, can be observed within 24 hours and last 3 weeks. Even if it is not clear till now in which way the cells take up antagomiRs this approach can allow the study of individual miRNA. This technology was successful in the mouse for the functional study of miR-122.

4) Recently new antisense oligonucleotides, (ASOs) (Esau et al., 2006), unconjugated single-stranded RNAs that carry complete phosphorothioate backbones and 2'-O-methoxyethyl modifications, have been developed to target miRNAs *in vivo* and for the

moment they have been used only in liver tissues, confirming the results about miR-122 obtained with the 'antagomiRs'.

I-I-f MiRNAs and Cell differentiation in mammalian development (see review : Song and Tuan, 2006)

Animal development requires the establishment of a highly regulated spatiotemporal gene expression network in order to convert the totipotent zygote into an animal containing various specialized tissues functioning in a concerted manner (Lee et al., 2006). One important feature of this regulatory scheme is the specific expression of factors that are required for each developmental window. However, it is also crucial to inhibit the expressions of genes that are not required for particular developmental stages or which may promote alternative differentiation pathways (O'Rourke et al., 2006).
At a cellular level the tissue development is produced by cell differentiation. The ability of a precursor cell to differentiate into different cell types actually depends on upregulation of factors required for one lineage combined with the downregulation of others specific for a different fate.
It has long been thought that this process is regulated primarily at the level of transcription, but it is also possible that posttranscriptional mechanism are required to drive cell commitment. miRNAs may serve as a switch to determine the developmental program of precursor cells; alternatively they may function to maintain the identity of differentiated cell types (O'Rourke et al., 2006).
Here I will summarize the recent data on miRNA involved in cellular differentiation during mammalian development.

Embryonic stem (ES) cells are totipotent cells that resides in the inner cell mass of the blastocysts that have the capacity to generate precursor cells of endoderm, ectoderm and mesoderm. In vitro, ES murine and human cells lines can be induced to generate germ and somatic cells of both the three layers, reproducing part of the in vivo embryonic development. ES cells present great potential in clinical and biological applications, even if the molecular mechanism governing their differentiation is poorly understood. A possible role for miRNA came from the identification of ES cell-specific miRNAs in mouse screening of two miRNA libraries from undifferentiated and differentiated mouse ES cells (Houbaviy et al., 2003). 6 miRNAs (miR-290, miR-291-as, miR-292-as, miR-293, miR-294, miR-295) were found expressed in pluripotent cells and silenced or downregulated after differentiation. Later miRNA libraries from human ES cells were analysed and 17 novel miRNA were

identified (Suh et al., 2004). 12 of them were found localized in two genomic cluster that are transcribed in two polycistronic primary transcript whose level decreases when ES cells develop into embryonic bodies, suggesting the specificity of these miRNAs for undifferentiated cells.
miRNAs play a key role also in maintaining the differentiation state of pluripotent ES cells as demonstrated by the loss of stem cells and early lethality of Dicer-null mice embryos (Murchison et al., 2005). In vitro Dicer-null ES cells did not differentiate indicating that lack of Dicer and endogenous miRNAs compromise the differentiation potential of ES cells.

The first indication of the involvement of miRNAs in limb development in mammals came from an *in situ* hybridization study on mouse embryos (Schulman et al., 2005). An ortholog of the worm *lin-41* was detected and localized in the embryo in the posterior region of the limb bud, while in embryo at the same stage *let-7*, the corresponding negative regulator, is expressed in the anterior region of the limbs. Their reciprocal expression patterns implied that they play a role in limb development.
Using microarray analyses Hornstein et al. (2005) identified a miRNA, miR-196, which is preferentially expressed in the hindlimb. It was shown that miR-196 can direct the cleavage of a known transcription factor that mediate anteroposterior polarity in fore- and hindlimb buds, Hoxb8, only in the hindlimb, but not in the forelimb, regulating the limb development.

Adipocyte differentiation can be reproduced *in vitro* using primary subcutaneous preadypocytes
that exposed to hormonal stimuli develop in mature adipocytes. Esau et al. (2004) use this systems to assess the effects of several miRNAs on adypocyte differentiation transfecting preadipocytes with 2'-O-methoxy-ethyl phosphorothioate-modified antisense RNA oligonucleotides targeting specific miRNAs. They found that miR-143 is involved in the maturation of preadipocytes, like it was confirmed by its upregulation in differentiated adipocytes. ERK5 is probable the gene target and its protein level is elevated in cells with decreased miR-143 expression. This was the first time that antisense oligonucleotides were used successfully to determine the function of miRNAs.

Cardiomyocyte differentiation and cardiogenesis require sequential activation and repression of transcriptional factors such as serum response factor (SFR), MyoD, Mef2. Zhao et al. (2005) identified, by a combination of *in silico* and experimental approaches, miR-1-1 and miR-1-2, whose expression was specific of in the heart and skeletal muscle of adult mice. Overexpression of miR-1 demonstrated is essential role in cardiogenesis and its molecular target, Hand2. It was also analysed the regulation of its expression and it was found that potential binding sites exist in the enhancer region of miR-1 for the transcription factors Mef2, SFR and MyoD. A model was proposed in which miR-1 functions in the SFR-myocardin-dependent pathways in cardiac progenitor cells and is responsive to MyoD/Mef2

in skeletal muscle precursor. Recently another miRNA, miR-181, was discovered to be upregulated during terminal differentiation of myoblasts (Naguibneva et al., 2006). Loss-of-function assays *in vitro* using antisense oligonucleotide against miR-181 completely abolished the myoblast differentiation of cells. Since miR-181 expression was not detected in resting muscle cells *in vivo* it is probable that it plays a role only in the establishment of the differentiated phenotype, but not in its maintenance.

Regulation of hematopoietic differentiation is a complex process that involves the commitment, proliferation, apoptosis and maturation of hematopoietic stem/progenitor cells and a variety of regulatory molecules including miRNA. Several miRNA were detected preferentially in specific hematopoietic cell lineages: miR-181 in differentiated B lymphocytes, miR-142s in B-lymphoid and myeloid, miR-223 in myeloid (Chen et al., 2004). In particular miR-181 seems to be a positive regulator for B-cell differentiation, as its ectopic expression led to a doubling of cells in the B-lymphoid lineage without changing of T-lymphoid lineage both in vitro and in vivo.

Mice with defective Dicer function show abnormal epithelium morphogenesis, both in the skin and in the lung. Yi et al. (2006) saw that mice with defective Dicer activity in their skin progenitor cells exhibited abnormal epidermis and hair follicles. In a other study Harris et al. (2006) observed that Dicer inactivation led to dramatic branching effects in the lung. An increased and prolonged cell death was observed in both skin and lung epithelia in the mutant mice, but it is not known how this
contributes to the abnormal morphogenesis and which are the miRNA responsible.

Several miRNAs were found exclusively (miR-9, miR-142a, miR-124b, miR-135, miR-153, miR-183, miR-219) or highly expressed (miR-9*, miR-125a, miR-125b, miR-128, miR-132, miR-137, miR-139) in mouse and human brain tissues and some are also upregulated in embryonal carcinoma cells (Sempere et al., 2004). Overexpression and inactivation of three of them, miR-124a, miR-9, miR-125b, in neuronal progenitor cells dramatically change the relative fraction of astroglial-like cells and neuronal cells, confirming their critical role in neuronal differentiation.
To explain their mechanism of function it was proposed that they promote neuronal differentiation by suppressing the expression of non-neuronal transcripts. Experiment of overexpression and inactivation of the same miRNAs led to trace a model in which miR-125a and miR-125b are responsible of neuronal differentiation targeting the 3'UTR of lin-28 and altering both its translational efficiency and its mRNA levels. It was also demonstrated (Krichevsky et al., 2006) that the phosphorylation state of STAT3, a transcription factor that when phosphorylated inhibits neuronal terminal differentiation and promotes glial-like cells differentiation, is controlled by miR-9. To balance the formation of neuronal and glial cells in mammalian brain, expression of miRNA is tightly controlled. It was shown that the RE1

silencing transcription factor, REST, can switch the differentiation lineages cells between neuronal and glial cells and that it is correlated to the expression of the miRNA studied. Moreover miR-134 was found to be a negative regulator of dendritic spine development in hippocampal neurons and the protein kinase Limk1 was proposed as its target (Schratt et al., 2006).

I-I-g MiRNAs and cancer (see review : Esquela-Kerscher and Slack, 2006)

Cancer is caused by uncontrolled proliferation and inappropriate survival of damaged cells, which results in tumour formation. Many regulatory factors switch on or off genes that direct cellular proliferation and differentiation. Damage to these genes, which are referred to as tumor-suppressor genes and oncogenes, is elected for in cancer. Recent evidence indicates that miRNAs might also function as tumor suppressor and oncogenes. They have been shown to control cell growth, differentiation and apoptosis, consequently impaired miRNA expression has been implicated in tumorigenesis (Esquela-Kerscher and Slack, 2006).

Components of the miRNA-machinery have been found involved in tumorigenesis. For example, expression of Dicer has been shown to be downregulated in lung cancer (Karube et al., 2005), suggesting a potential indirect role of Dicer in tumor formation that result from the collective reduction of miRNAs. The Argonaute proteins have also been associated with various cancer. Three
human Argonaute genes are frequently deleted in Wilms tumors of the kidney and have been also associated with neuroectodermal tumors. In particular it is supposed that Argonaute 1 (AGO1) is involved in developing lung, kidney and renal tumors. An additional human argonaute gene, HIWI, maps to a genomic region associated with testicular germ-cell cancer and might normally control the proliferation and maintenance of germ cells.

The biological role of only a small fraction of identified miRNAs have been elucidated to date. These miRNAs regulate cancer-related processes such as cell-growth and tissue differentiation and therefore might themselves function as oncogenes.
Interestingly the mammalian homologues of *lin-4* and *let-7* have been shown to control cell proliferation in human cell lines (Lee et al., 2005; Takamizawa et al., 2004) and are also associated with various cancer (Johnson et al., 2005; Calin et al., 2004; Iorio et al., 2005; Sonoki et al., 2005). In *D. Melanogaster, bantam* induces tissue growth by both stimulating cell proliferation and inhibiting apoptosis (Brennecke et al., 2003; Hipfner et al., 2002), *miR-14* suppress strongly apoptosis (Xu et al., 2003), and these are features of oncogenes. Other characterized miRNAs have essential functions during development and differentiation of cells into various tissues.

A recent study showed that about 50% of annotated human miRNAs are located in areas of the genome, known as fragile sites, that are associated with cancer. For example miR-125b-1 is located in a region that is deleted in a subset of patients with breast, lung, ovarian and cervical cancers (Calin et al., 2004) and recently it has also been associated with leukaemia.
The first indication that miRNAs could function as tumor suppressors came from a report that showed that patients diagnosed with the B-cell chronic lymphocytic leukaemia, (CLL), often have deletions or downregulation of two clustered miRNA genes, miR-15a and miR-16-1 (Calin et al., 2002). Deletions within this locus occur also in 50% of mantle cell lymphomas cases, 16-40% of multiple myelomas and 60% of prostate cancers cases. It was predicted that a tumor-suppressor gene reside in this region. Later Cimmino et al. (2005) showed that miR-15 and miR-16-1 negatively regulate BCL2, an anti-apoptotic gene that is often overexpressed in many types of human cancers, including leukaemias and lymphomas, supporting a tumor-suppressor role for these two miRNAs.
Additional studies have shown a strong correlation between abrogated expression of miRNAs and oncogenesis. For example, miR-143 and miR-145 are significantly reduced in colorectal tumours (Michael et al., 2003).
The let-7 miRNAs family were the first group of miRNAs shown to regulate the expression of an oncogene, the Ras gene. Ras protein are membrane associated GTPase signalling proteins that regulate cellular growth and differentiation. About 15-30% of human tumors possess mutations in Ras genes. The 12 human homologous miRNAs of let-7 family map to fragile sites associated with
lung, breast, urothelial and cervical cancers (Calin et al., 2004). In particular the transcripts of certain let-7 were downregulated in human lung cancer (Takamizawa et al., 2004). Later studies in *C. Elegans* found that the 3'UTR of Ras genes contains multiple complementary sites for the let-7 family and that let-7 and Ras expression is inversely correlated in tumours (Grosshans et al., 2005; Johnson et al., 2005).
The MYC oncogene, which encodes a basic helix-loop-helix transcription factor, is often mutated or amplified in human cancers and has been shown to function as an important regulator of cell growth owing to its ability to induce both cell proliferation and apoptosis (Pelengaris et al., 2002). It seems that there is a correlation between miRNAs and the increased expression of MYC in the development of B-cell malignancies. MiR-142 and miR-155 are associated to MYC overexpression in the development of B-cell cancers, in Burkitt and Hodgkin lymphoma. MiR-155 is also involved in breast carcinomas, indicating other roles for this miRNA outside of the hematopoietic system. Recently He et al. (2005) and O'Donnell et al. (2005) describe a more direct relationship between miRNAs, MYC and cancer identifying a transcript that was preferentially upregulated in cancers and that encode the miR-17-92 clusters. By overexpression experiment it was shown that miRNAs within the miR-17-19b-1 cluster function cooperatively as oncogenes, possibly by targeting apoptotic factors activated in response to MYC overexpression and thus indirectly provoking uncontrollable cell proliferation. Surprisingly two miRNA gene in this cluster were shown to

block indirectly the cell proliferation acting on the transcription factor E2F1. The double nature of the miR-17-92 cluster, the tumor-suppressing and the oncogenic one, emphasizes the complexity of cancer progression as well as the intricacies of miRNA-mediated gene regulation. These results might also reflect the fact that a single miRNA can control many unrelated gene targets, resulting in the control of opposing activities such as cellular proliferation and differentiation. A recent report (Felli et al., 2005) describes the ability of miR-221 and miR-222 to downregulate the KIT oncogene and future studies will reveal that miRNA function as key regulators of many cancer-related genes like BCL2, Ras, E2F1, MYC and KIT. Therefore miRNAs might be powerful drug target that could be used in a broad range of cancer therapies.

As Northern blot and microarrays analyses have already been used to determine tissue specific 'signatures' of miRNA genes in humans (Pasquinelli et al., 2000; Lagos-Quintana et al., 2003; Lim et al., 2003; Liu ret al., 2004; Nelson et al., 2004; Thomson et al., 2004; Krichevsky et al., 2003; Miska et al., 2004; Sempere et al., 2004; SmiRnova et al., 2005; Sun et al., 2004; Monticelli et al., 2005; Babak et al., 2004), researchers are now using miRNA-expression signatures to classify cancers and to define miRNA markers that might predict favourable prognosis (Takamizawa et al., 2004; Iorio et al., 2005; Calin et al., 2002; Lu et al., 2005; Ciafre et al., 2005; Chan et al., 2005; He et al., 2005; O'Donnel et al., 2005; Calin et al., 2005).
A recent report from Lu et al. (2005) found that the expression profile of a relatively few miRNAs (200) can be sufficient to accurately classify human cancers.
Following comparison of the expression level of miRNAs in normal and tumorous tissues it was shown that in general miRNA are downregulated in tumorous tissues, supporting a model in which miRNAs drive cells in more differentiated state and can be marker of the degree of cell differentiation. These studies define miRNA more like oncomiRs and imply that abnormalities in miRNA expression might directly result in de-differentiation of cells, allowing tumour formation (Esquela-Kerscher and Slack, 2006).

The emergence of miRNAs as important cancer-prevention genes is likely to have a large effect on gene therapies designed to block tumour progression. Large-scale expression screen to compare miRNA levels in tumours versus normal tissues will be useful in identifying novel miRNAs involved in cancer. In the future the administration of synthetic anti-sense oligonucleotides that encodes sequences complementary to oncogenic miRNAs, the anti-miRNA-oligonucleotides (AMOs), could inactivate miRNAs in tumours or slow their growth. The antagomiRs, that are AMOs conjugated with cholesterol, have already used to inhibit miRNA activity in various organs after injection into mice (Krutzfeldt et al., 2005), and might be a promising therapeutic agents. At contrary, techniques to overexpressed tumor-suppressor could be used to treat specific tumours. More development of these methods is needed before miRNAs treatment can move from the laboratory bench to the bedside. Even if we do not

know if miRNAs will become a ‘magic bullet’ in the future, research in this area will undoubtedly provide insight into the underlying mechanism of oncogenesis.

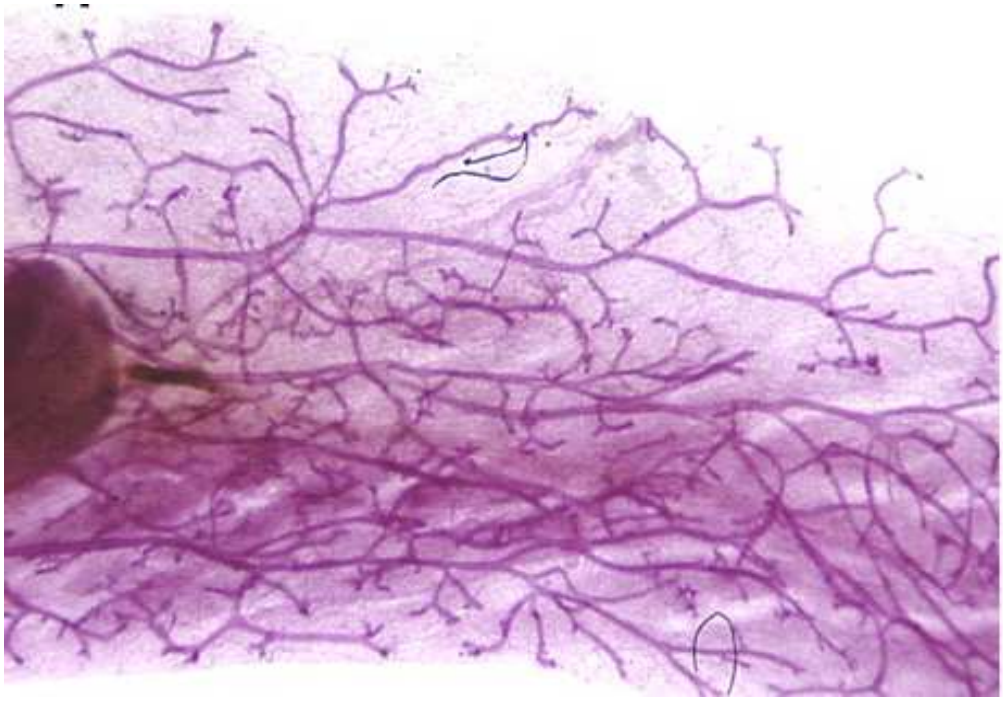
I-II The mammary gland

I-II-a The mammary gland: structure and cellular composition

Mammalian evolution has been accompanied by the formation of a unique organ: the mammary gland. In fact, on a phylogenetic scale, this organ is a recent acquisition: it appeared 200 million years ago with the appearance of mammals to provide nourishment to the newborn in the form of milk (Hennighausen and Robinson, 1998). Unlike other branched organs, the most part of its development takes place post-natal rather than in embryonic life to accomplish the unique capacity of producing and secreting milk during the lactation (Sternlicht et al., 2006).

The number and location of mammary glands vary strongly between different species, but the structure and cellular composition is very similar. This organ is constituted from two tissue compartments: the epithelial one, that will give origin to ducts and to milk-producing alveolar cells, and the connective one (stroma or mammary fat pad) composed of adipocytes, fibroblast, cells of the haematopoietic systems, blood vessels and also neurons (see review: Hennighausen and Robinson, 2005) (figure 3).

Figure 3. Carmine-stained whole mounts of a section of mammary gland : in violet the epithelial tissue and its ducts, in white the stroma.



In general the epithelial tissue, at the parturition, is differentiated in cells constituting ducts, elongated canals transporting milk, and luminal secretory and myoepithelial cells that together constitute the central lumen and the outer layer of the alveoli, the functional secretory structural unit of mammary gland. Many grouped alveoli constitute a lobule and many lobules are grouped in many bigger lobuloalveolar units. This branched structure is similar to the lung structure.
Each alveoli has a spherical structure inside which a monolayer of epithelial cells secrete milk in the central lumen. The milk is transported into the ducts by the contractile actions of myoepithelial cells, and is delivered to the body surface through the nipple. The extensive system of ducts and alveoli is embedded in the stroma, that supports the epithelial tissue and provides nourishment to epithelial cells (see review: Hennighausen and Robinson, 2005).

In the mouse there are five pairs of mammary gland located just below the skin, which extend from the thoracic (three pairs) to the inguinal (two pairs) regions of the animal along what is termed the milk or mammary line (Richert et al., 2000). Apart a nipple and a ductal-alveolar system, each gland has a limph node that is often used as a landmark when examining histological sections or whole mounts (Russo IH and Russo J, 1996). There is a gradient of differentiation among the glands, with the first thoracic gland being the least differentiated and the fifth inguinal gland the most (Bolander, 1990).

I-II-c The development of mammary gland

The mammary gland is a dynamic organ the structure of which changes throughout the female reproductive cycle. The development of the gland occurs in distinct stages, defined fundamentally by hormones, that are connected to the sexual development and reproduction: embryonic, prepubertal and pubertal stages, pregnancy, lactation and involution (see review: Hennighausen and Robinson, 1998).
In the main lines at birth the anlage consists of a few rudimentary ducts in the vicinity of nipple, pronounced ductal outgrowth and branching commences at puberty, in pregnancy and expanded lobulo-alveolar compartment develops. Functional differentiation of the secretory epithelium coincides with parturition and large amounts of milk are produced and secreting during lactation.
After weaning of the young the entire alveolar epithelium compartment is remodelled to resemble a virgin-like state. With each pregnancy a new round of lobulo-alveolar development occurs.

The epithelium and the surrounding stroma are derived from ectoderm and mesoderm, respectively (Parmar and Cunha, 2004).

In mice the mammary gland first appears embryonically as an epithelium bud that penetrates the underlying mesenchyme. The first morphological signs of mammary rudiments are lens-like placodes that form around embryonic day 11 and protrude slightly from the body wall (Robinson, 2004). This rudiment becomes bulb-shaped, they elongates and invades the mesenchym to form a simple ductal tree with several branching ducts. This first phase of development is indipendent of hormonal signals (Richert et al., 2000).

In mice at birth the mammary gland consists of the epithelial cords and the stroma. While the first one is rudimentary (Topper and Freeman, 1980; Russo IH and Russo J, 1996) the stroma is thick and dense around epithelial structures and consists of eosinophilic fibrous connective tissue and fibroblast and in the early stage of development is filled with large adipocytes. Also present are lymphatics and blood vessels, the last will increase in number during pregnancy and lactation (Matsumoto et al., 1992).
The period of most rapid growth occurs during puberty from approximately 3-6 weeks of age in the mouse. The ducts lengthen and branch to form secondary and tertiary ducts that ultimately extend to fill the mammary fat pad by approximately 3 months of age. The terminal end buds (TEBs) appear at 3 weeks at the tips of growing ducts and are the sites of highest epithelial proliferation in the gland (Richert et al., 2000). From this bulbous structure cells are capable to migrate, to proliferate and differentiate in luminal and ductal epithelial cells (Daniel and Silberstein, 1987). This migration and proliferation result both in elongation of ducts and invasion of the fat pad; the differentiation in the TEBs is also responsible of branching (Gordon and Bernfield, 1980; Silberstein and Daniel, 1982) and formation of lateral and alveolar buds, that eventually subdivide to form rudimentary alveolar structure in the post-pubertal glands, after 10-12 weeks of age, in response to cyclic secretion of ovarian hormones at each estrous cycle (Andres and Strange, 1999).
The peak of mammary differentiation occurs during the 19-21 days of pregnancy and culminates with formation of alveoli and a fully lactating gland at parturition (Nandi, 1958).
In the beginning of pregnancy, a massive proliferation of ductal branches and the formation of alveolar buds, like in the postpubertal stage, could be observed. The epithelial to adipocyte ratio increase.
During the second half of pregnancy the alveolar buds progressively cleave and differentiate into individual alveoli that in the late pregnancy fill the majority of the fat pad. By the day 18 of pregnancy the alveolar epithelial cells are producing milk proteins and lipid, in preparation for lactation. The amount of stroma is greatly decreased, allowing more contact of the epithelium with adipocytes (Neville et al., 1998; Elias et al., 1973).

As lactation begins the milk in the lumen of alveoli is forced into the ducts (Asch HL and Asch BB, 1985; Richardson, 1949; Dulbecco et al., 1986), the fat in the adipocytes is metabolized and the alveoli expand to completely fill the gland (Neville, 1999). In normal condition the process of lactation continues for approximately 3 weeks, until the pups are

weaned. At this moment the gland goes through a process of death and remodelling, the involution. This process is initiated by milk stasis once milk removal has ceased (Quarrie et al., 1996).
Forced weaning is often chosen as a model for involution because it is more controlled than natural weaning and allows for more precise timing of structural changes.

In the first day of involution big morphological changes is not observed, except for the flattening of the epithelium due to engorgement of the alveoli with milk. After 2 days the gland begins the irreversible sequence of cell death and remodelling: the secretory epithelial cells of the alveoli go in apoptosis and can be cleared by neighbouring epithelial cells or invading macrophages (Burwen and Pitelka, 1980; Richards and Benson, 1971; Fadok, 1999). At the day 4 the alveoli collapse into clusters of epithelial cells, while the adipocytes appear to be refilling. The epithelium progressively disorganize and decrease while adipocytes and stroma increase (Richert et al., 2000). At the day 6 of involution all the alveoli have collapsed and both epithelium and stroma are rearranged (Strange et al., 1992) as the majority of cell death has already occurred; the involution of alveoli continue till the day 21 of involution, when the gland resemble the prepregnant mature gland.
With each pregnancy a new round of lobulo-alveolar development occurs, together with the cycle of proliferation-secretion and involution of the epithelial tissue.

I-II-d Endocrine control on mammary development

While in embryo the initial stages of mammary development are independent of systemic endocrine signals and rather depends on reciprocal signalling between the epithelium and the mesenchym, the most part of development, that occurs after birth during pregnancy, is under control of steroid and peptide hormones.
Both the role of systemic hormones and the influence of the stroma on mammary epithelium have been recognized for some time (Mackie et al., 1987), actually the study of endocrine control of mammopoiesis and lactogenesis began more than 100 years ago.
The first demonstration that ovarian steroids and pituitary hormones can determine breast development and lactation came from an experiment of ovariectomy and transplantation of ovaries in mouse in 1900 (Halban and Knauer, 1900). The responsible bioactive compounds extracted were the progesterone and estrogen (Allen, 1924).
Later it became clear that other factors than ovarian hormones were required for mammopoiesis and in 1928, Stricker and Grueter induced milk secretion artificially in castrated virgin rabbits by injection of pituitary extract (Stricker and Grueter, 1928). Five years later Riddle and collegues purified the prolactin from this extract (Riddle et al., 1933).
From the 1906 it was known that also the placenta can secrete mammotrophic substances (Lane-Claypton and Starling, 1906), like placental lactogens, estrogens, progesterone and

gonadotrophins. It was shown for the first time in 1980, by the introduction of *in vitro* mammary organ cultures, that it is a synergy of insulin, hydrocortisone and prolactin that controls the differentiation of secretory mammary epithelium (Topper and Freeman, 1980). In the same year steroid and peptide hormone receptors were cloned and in 1990 downstream signalling components were identified, providing a basis for the understanding of signal transduction pathways.

Ductal elongation in the first days after birth originates from a few small TEBs and is probably the result of residual effects of maternal and fetal hormones (Hennighausen and Robinson, 1998). The acceleration of ductal growth during puberty and the strong lobulo-alveolar proliferation during pregnancy are controlled mainly by ovarian steroid hormones (Daniel and Silberstein, 1987), respectively by the oestrogen and progesterone, that act regulating cell proliferation and cellular turnover.

Progesterone is secreted in the beginning of pregnancy from the yellow body and its level is low in the beginning, increases during this phase and decreases brutally near the parturition, when the placenta and the yellow body involute (Martinez and Houdebine, 1994, chap.1).

The level of estradiol is high during puberty, in pregnancy the concentrations of estrogens secreted from the placenta is lower, but sufficient to cooperate with the progesterone in inducing the growth of lobulo-alveolar systems till the parturition, when estrogen level decreases rapidly (Martinez and Houdebine, 1994, chap.1).

Both estrogen and progesterone have pleiotropic actions in the uterus, ovaries and the hypothalamic-pituitary axis in regulation of sexual development. Since the need for a functioning mammary gland is dependent on a successful pregnancy, the evolutionary process use the same set of hormones for both developmental process (Hennighausen and Robinson, 2001).

The primary mechanism of steroid hormone action is through their specific nuclear receptors, which function as transcription factors when bound to their ligands (Hennighausen and Robinson, 2005).In post-natal mammary tissue not only most epithelial cell express receptors for estrogen (ER) and progesterone (PR), but also cells of stroma.

Both ER and PR have two isoform, ERα and β, PR-A and -B, that have different functions during the development of mammary gland.

Studies from knockout mice for ERα demonstrated that both stromal and epithelial ERα are required for normal ductal elongation and outgrowth during puberty (Bocchinfuso et al., 2000), even if ERα is not necessary for pregnancy alveolar expansion (Mueller et al., 2002). Recombinant tissue experiment showed that estradiol elicits epithelial mitogenesis indirectly through ER stromal cells (Cunha et al., 1997).

Knockout mice showed that is the PR-B form responsible of proliferative effects on mammary epithelium, in particular to expansion of the alveolar compartment, and only in minor part to ductal elongation and branching (Mulac-Jericevic et al., 2003).

In early pregnancy PR cells are found in closely proximity to proliferating cells, suggesting a paracrine effects for progesterone. Progesterone seems to induce the production of a signal that guides the proliferation of neighbouring cells. One possible candidate is the receptor activator of nuclear factor kB (NF-kB)-ligand or RANK-L (Mulac-Jericevic et al., 2003), belonging to the tumor necrosis factor (TNF) family.
It is now clear that estrogen induce the receptor for progesterone in epithelial cells, increasing the sensibility of cells to this hormone.

Prolactin (PRL) signalling is essential for the proliferation and functional differentiation of lobulo-alveolar structures during pregnancy (Topper and Freeman, 1980).
PRL is produced mainly by the lactotrophs in the anterior pituitary gland, even if also local production of PRL by mammary epithelium has been reported (Vonderhaar, 1999).
Its level is relatively low during the most part of the pregnancy, but in the last part increases and reaches high level at the parturition (Martinez and Houdebine, 1994, chap.1).
It has two roles in reproduction : the maintenance of corpus luteum, through which the secretion of estrogen and progesterone is ensured, and the induction of mammary development. After birth PRL is essential for maintaining lactation.
By the use of knockout mice the four independent components of prolactin pathway have been identified : the ligand itself (Horseman et al., 1997), the receptor (PRLR) (Ormandy et al., 1997), a transmembrane protein of the class I cytokine receptor family, the transcription factors Stat5a (Liu et al., 1997) and Stat5b (Udy et al., 1997).
Binding of PRL to its receptor leads to receptor dimerization and the activation of the Janus kinase 2 (JAK2), Fyn, a specific tyrosine kinase associated to the PRLR. JAK2 phosphorylates the two Stat5 isoforms that dimerize and migrate in the nucleus to induce transcription of target genes, such as genes for the caseines and genes containing γ-interferon activation sites (GAS). As well as Stat5, PRLR can signal through the mitogen-activated protein kinase (MAPK) pathways and others that are dependent of JAK2 (Hennighausen and Robinson, 2005).
Current evidence indicates that PRL present a generic signal that activates transcriptional programmes that are shared between several cytokine receptors, and even if these pathways have some cell-specific components they mediate general responses like proliferation and cell survival. Moreover not only PRL activate STAT5, leading to a developmental program that ends with the production of milk-secreting cells, but also other placental lactogens and members of the EGF family, whose effecte is mediated by EGF receptor such as ERBB1 and ERBB4, both necessary for mammalian development during pregnancy. In particular ERBB4 was shown to have a more prominent role in the functional luminal cell during lactation than PRL has (Long et al., 2003).
The signalling pathway activated by hormones is quite understood, but the mechanism by which it is negatively modulated is not well known. Recently evidences suggest that member

of the SOCS family are involved in the inhibition of PRL signalling (Linderman et al., 2001; tonko-Geymayer et al., 2002).

Mammary development is not only controlled by systemics hormones, like estrogen, progesterone and PRL, but also by peptide that are produced either in the stromal or epithelial compartment, such as the osteoclast differentiation factor RANKL (Fata et al., 2000), inhibinβB (Robinson et al., 1997) and member of the TGFβ family (Nguyen and Pollard et al., 2000).
Several evidence from knockout mice suggest that RANKL, compared to PRL, induces identical or related developmental programs during pregnancy (Humphreys et al., 1999; Fata et al., 2000).
The growth factors, like transforming growth factor α and β, TGFα and TGFβ, mammary derived growth factor 1, MDGF1, and epidermal growth factor, EGF, are present in the mammary epithelium, secreted by the epithelial cells. The MDGF1, TGFα and the TGFβ are autocrine and mitogenic factors secreted by epithelial cells in order to stimulate the production of the collagen IV, an essential component of the basal membrane, where epithelial cells lie and proliferate in a polarized way during the alveolar development (Martinez and Houdebine, 1994, chap.1). Estrogen control indirectly the synthesis of collagen IV and the activity of the growth factors throughout the degradation of the basal membrane which supports epithelial cells.

At the moment of parturition strong changing in the concentrations of hormones occurs :
The progesterone, that negatively controls the PRL secretion and the local synthesis of caseins and other milk components, disappears, while PRL reaches high concentration; the level of estrogen increase progressively and stimulate the secretion of PRL; glucocorticoids are produced to amplify the PRL action; other hormones not specific of the lactation are involved, like the growth factor, (GH), and thyroid hormones (Martinez and Houdebine, 1994, chap.1).

I-II-e Role of extracellular matrix on mammary development

The multihormonal control on the mammary epithelium development and on the secretion of milk proteins was observed and studied relatively early due to the fact that the glands could be analysed easily in vivo (Dembinski and Shiu, 1987; Houdebine et al., 1985; Neville and Neifert, 1983; Topper and Freeman, 1980). However it must be recognized that a substantial proportion of epithelial hormonal responses reflects the modulation imparted by a complex extracellular compartment, that can exert its influence on mammary epithelium through several mechanisms : the mediation of hormonal signals via stromal hormone receptors; the local elaboration of soluble agonist/antagonist factors; the provision of a supporting vascular network; the contribution to a bed of basement membrane proteins on which epithelial cells

are positioned (Russell and Vonderhaar, 2002). Moreover the importance of cell to cell interactions and cell to extracellular compartment interactions is gaining importance since functional cell cultures in vitro were developed (Martinez and Houdebine, 1994, chap.4).

At every stage of mammary development the duct or the alveoli lie on a basal membrane. It is possible that the interactions stroma-epithelium in vivo are mediated through the structure and composition of this extracellular matrix (MEC), that is the surface and the region of contact between the two tissues. The study of the biochemical composition and the structure of the MEC (Hassell et al., 1985; Kleinman et al., 1986; Miller and Gay, 1987) show that this basal membrane is not a passive layer, but in the contrary is an active membrane that receive structural and functional message to direct the behaviour of stromal and epithelial cells (Bissel and Aggeler, 1981; Bissel and Hall, 1987; Bissel aet al., 1982; Hay, 1981; IngBer and Jamieson, 1985; Wicha, 1984) .

This basement membrane underlying epithelial cells in vivo consists of 3 separate layers (Sakakura, 1991) : in contact with epithelium there is the lamina lucida, a thin space under which is located the lamina densa. Together they constitute the basal lamina. Adjacent to the basal lamina is the stroma-associated layer of variable thickness, the reticula lamina.

Based on in vitro study and immunolocalization experiments it was long assumed that components of the basal lamina, such as laminin, heparin sulphate proteoglycans and type IV collagen, were all derived from epithelial cells and that components of reticula lamina, such as collagen types I and III, fibronectin and tenascin, were derived from the stroma (Russell and Vonderhaar, 2002). Recent studies have assessed that the stroma is the primary source of extracellular matrix proteins and that also collagen I, IV and laminin derived from stroma (Keely et al., 1995). These finding define even the time of production of these macromolecule: collagen I is expressed in early puberty and early pregnancy, collagen IV during pregnancy and laminin during lactation. Moreover the expression of fibronectin from the stroma seems to be regulated by ovarian steroid hormones in association with epithelial-stromal interactions (Woodward et al., 2001). It is not clear if this dynamic construction of basal membrane during the mammary gland development is the result or the cause of the epithelial morphogenesis.

It is clear that various components of this basal membrane regulates the formation and function of epithelial cells and their response to external signals, such as ovarian steroid hormones or growth factors (Woodward et al., 2000). Even if we do not know the exact contribution of these extracellular proteins at a cellular level a general model establishes that the extracellular matrix exert its influence interacting with transmembrane proteins, able to communicate with the cytoscheleter and the nucleus of epithelial cells (Martinez and Houdebine, 1994, chap.4).

I-II-f The miRNAs in the mammary gland

An implication of miRNA in mammary gland biology is suggested from the data of some few recent reports, most of them focusing more on pathological situations, such as the appearance of breast cancer, than on the normal mammary development.

Liu et al. (2004) analyzed the gene expression profile of 18 adult and 2 fetal normal human tissues using a microchip containing the oligonucleotides for 248 miRNA (161 derived from human, 84 from mouse, 3 from Arabidopsis). They showed that each tissue has a specific pattern of miRNome expression (defined like the totality of miRNA present in a cell) that can be quantified. The mammary gland was one of the tissue analyzed and it was revealed that its specific signatures is characterized by the expression profile of only 23 miRNAs, the lowest number of miRNA detected in any tissue.

Other indirect evidences of miRNA involvement in the biology of mammary gland come from studies about breast tumors.

It was analyzed the genomic localization of 186 human miRNA (Calin et al., 2004), 52,5 % of them are present in cancer-associated genomic regions or in fragile sites and between them 15 miRNA are located in regions involved in human breast cancers. It was quantified (Jiang et al., 2005) by real-time PCR the expression of 222 pre-miRNA in 32 human cancerous cell lines, 5 derived from breast cancer, and it was observed that let-7f-1 expression was 7-fold higher in epithelial-derived breast, lung and colorectal cancer cells comparing to the mean of the remaining cell lines. Moreover another study used microarray technology to measure the differential expression of miRNA in normal and neoplastic human breast tissue and 29 miRNAs were found to be differentially regulated, 15 of which could be used with 100% accuracy to predict the tumor (Iorio et al., 2005). In particular miR-125b have a decreased expression level in samples derived from breast cancer primary tumors comparing to normal breast tissue (Lee et al., 2005).

To date any reports deals about the expression of miRNAs in normal mammary gland during the stages of its development.

II-Objective

To establish the genetic and functional network of a more comprehensive developmental model of the mammary gland the genomics approaches should identify new putative control genes and gene manipulation, in combination with tissue transplants, should evaluate their physiological role. It should be important to evaluate also the time windows during which a particular gene product is needed.

Taking in mind that many genetics pathways that control the development of mammary tissue are used in organ systems that appeared earlier in evolution and considering the big evolutionary conservation of miRNAs throughout every kingdoms and their involvement in various mechanism of organogenesis, it was chosen to address the attention to miRNAs, in order to discover putative regulatory molecule of the mammary gland development.

The study of miRNA in the mammary gland began analysing the expression of a first group of conserved miRNAs, during different stages of mammary development in mouse; then the expression profile during all the gland development has been studied in search for their potential regulatory role in determining the passage from one phase to one other. Later it was examined the cellular origin of their production.

The second objective of this work was the identification of mammary gland specific miRNA, the idea was supported from the finding of organ- and tissue- specific miRNA (Lagos-Quintana et al., 2002; Liu et al., 2004; Sempere et al., 2004; Pay et al., 2004; Frederikde et al., 2006; Ryan et al., 2006; Chen et al., 2006, Ramkisson et al., 2006; Coutinho et al., 2006; Xu et al., 2006; Gu et al., 2006) and also from the recent discovery of new specific primate miRNAs (Devor, 2006). After having constructed a cDNA library of small RNA extracted at different stages of mouse mammary gland, the expression of ‘candidate miRNAs’ was characterized and a composite analyses, in part using bioinformatics and experimentally tools, has been realized in order to validate them like miRNAs.