Scuola Internazionale Superiore Studi Avanzati-SISSA
International School for Advanced Studies -ISAS
Trieste, Italy

Promotion of cortico-cerebral precursors expansion by artificial miRNAs against the *Emx2* locus

Thesis submitted for the degree of “Doctor Philosophiae”

Academic Year 2009/2010

Candidate
Assunta Diodato

Supervisor
Prof. Antonello Mallamaci

ABSTRACT **4**

INTRODUCTION **5**

1. Non coding RNA (ncRNA): an historical perspective 5

2. classical ncRNAs (rRNA, tRNAs, snRNAs, scRNA, snoRNAs) 7

3. Non-classical ncRNAs: basic classification 11

4. Short RNAs 11

4.1. miRNAs 11

4.2. siRNAs 14

4.3. piRNAs 14

4.4. PASRs and TASRs 15

4.5. tiRNAs 17

5. Long ncRNAs (lncRNAs) 17

5.1 Generalities 17

5.2 long ncRNAs regulating transcription 19

5.3 lncRNAs regulating gene expression levels post-transcriptionally 24

6. RNAa 25

7. Emx2: generalities and evolutionary aspects 28

8. Emx2 functions in CNS 29

8.1 Emx2 in cortical specification 30

8.2 Emx2 in cortical regionalization and arealization 32

8.3 Emx2 in cortical lamination 36

8.4 Emx2 in regulation of cortical precursors kinetics 37

9. Regulation of Emx2 expression 39

10. Emx2 as a regulator of gene expression 42

AIMS OF THE WORK **45**

MATERIALS AND METHODS **46**

Animal handling 46

Cell cultures 46
Primary cells. 46
Cell lines. 46
Lentiviral transfer-vector construction 47
Lentiviral vectors packaging, titration and usage 49
Cell counting 49
RNA extraction 50
Gene expression assays 50
cDNA preparation. 50
Quantitative PCR. 51
RACE Assay 52
ChIP-qPCR 53
RESULTS 55
DISCUSSION 63
FUTURE PERSPECTIVES 66
REFERENCES 67
ACKNOWLEDGMENTS 82

ABSTRACT

Emx2 encodes for a transcription factor implicated in several aspects of cerebral cortex development, such as morphogenetic field specification, arealization, and lamination. Its overexpression in cortico-cerebral precursors may promote stem cells self-renewal, inhibits their gliogenic commitment, protects neuronal progenitors and stimulates their differentiation, all activities of paramount relevance for purposes of gene-promoted brain repair.

In this study, a set of artificial miRNAs targeted against non-coding cis-active modules and/or phylogenetically conserved sequences of the *Emx2* locus were delivered to embryonic cortico-cerebral precursors, by lentiviral vectors, and consequences of that were evaluated. Three parameters were scored: (1) *Emx2* expression levels, by quantitative retrotranscription-polymerase chain reaction (RT-PCR), (2) binding of RNA polymerase II (RNApolII) to the *Emx2* locus, by chromatin immunoprecipitation-quantitative PCR, (3) kinetic behaviour of transduced precursors, by systematic cell counting. A subset of these miRNAs upregulated *Emx2*, possibly by recruiting RNApolII to the *Emx2* locus and stimulating transcription. Remarkably, this resulted in appreciable expansion of the proliferating precursors pool. By RT-PCR and rapid amplification of cDNA ends (RACE), two genomic regions targeted by these miRNAs were discovered to be naturally transcribed. Delivery of RNA fragments encoded by such regions to cortico-cerebral precursors, by lentiviral vectors, led to downregulation of *Emx2*-mRNA.

These results suggest that it is possible to stimulate transcription of *endogenous Emx2* by artificial miRNAs targeting its enhancers. This method, avoiding drawbacks of exogenous gene copies introduction, makes *Emx2* overexpression for purposes of brain repair a more feasible goal.

Moreover, this study shows that the *Emx2* enhancers are naturally transcribed and non coding RNAs stemming from them are apparently implicated in fine tuning of *Emx2*-mRNA levels.

INTRODUCTION

1. Non coding RNA (ncRNA): an historical perspective

Recent years have seen an increasing appreciation of the pervasive (genome-wide) transcription that occurs in most genomes and the multiple functional roles of RNAs within cells. Most of multicellular eukaryotes genomes is expressed in a cell and tissue-specific manner, and there is mounting evidence that much of this transcription is involved in the regulation of differentiation and development. In eukaryotic cells, the storage units of heritable information are compartmentalized in the nucleus as DNA and as epigenetic marks etched on the DNA itself and the associated chromatin proteins. RNA, by contrast, has been seen primarily as an ancillary molecule that accurately transfers information from the genome. Its functional roles, such as assisting in the synthesis of proteins by acting as a messenger (mRNA) or as a scaffold for protein synthesis (rRNAs), and helping to gather amino acids for protein synthesis (tRNA), have encouraged the view that the information stored in the genome is transferred to RNA in a co-linear fashion; in other words, the nucleotide sequences found in RNA transcripts are ordered in the same linear fashion as those found in the DNA genome. The discovery of splicing provided a more modular and non-contiguous view of this co-linear relationship, but even so the order of sequences in both DNA and RNA has been maintained. This co-linear organization seems logical and efficient given the perceived primacy of DNA in the genetic hierarchy. Additionally, underlying this organization of information in the genome and its co-linear transfer to RNA is the premise that the sequences that will be joined together in the mature RNAs reside on the same precursor RNA molecule. This seems to be the primary path of RNA processing from primary to mature transcripts. However, structural studies of RNAs in several species have revealed that the sequences that are ultimately joined together on the same mature transcript can be encoded in separately transcribed RNAs with multiple distinct genomic origins. Individual RNAs can be transcribed on separate chromosomes, on the same chromosome but with a different genomic order from that found in the mature RNA, on the same chromosome but transcribed from different strands, or on the same chromosome but from different alleles (Gingeras, 2009). In last years, much new information has been collected about RNA's world, especially after the non coding RNAs discoveries. RNAs can be divided into two distinct classes: protein-coding RNAs (mRNAs) and non-protein coding RNAs (ncRNAs).

It was traditionally believed that protein-coding RNAs controlled cellular functions, and the repertoire of protein-coding genes determined the complexity of an organism. Consequently, for much of the last twenty years up to recently completed genome projects, interest in the transcriptional activity of the genome has been focused almost exclusively on the discovery of these genes. For ncRNAs, it was previously believed that there were only a few species, such as tRNAs, rRNAs and spliceosomal RNAs, and they were considered accessory molecules involved in mediating transcription and translation. These traditional points of view have recently been challenged by the recent discovery of an huge repertoire of ncRNA types, characterized by an impressive variety of biogenetic, metabolic and functional properties. Large scale cDNA sequencing and comparative analyses showed that there is only a weak correlation between biological complexity and the number of protein-coding RNAs. If one uses the numbers of protein-coding RNAs as a basis for evaluating biological complexity, one would reach the conclusions that insects are less complex than nematodes, and rice is more complex than humans. It is clear that protein-coding RNAs are not the only source for determining biological functions. Large scale gene expression profiling, molecular cloning and tiling array analyses showed that around 63% of human genomic DNA is transcribed, but only 2% of it encodes proteins (Qi et al., 2006;Carninci, 2008). Non-protein-coding sequences increasingly dominate the genomes of multi-cellular organisms as their complexity increases, in contrast to protein-coding genes, which remain relatively static. Different classes of small and large noncoding RNAs (ncRNAs) have been shown to regulate almost every level of gene expression, including the activation and repression of homeotic genes and the targeting of chromatin-remodeling complexes. ncRNAs are involved in developmental processes in both simple and complex eukaryotes. While most of them have yet to be systematically studied, the emerging evidence suggests that there is a vast hidden layer of regulatory ncRNAs, that constitutes the majority of the genomic programming of multi-cellular organisms and plays a major role in controlling the epigenetic trajectories that underlie their ontogeny. During development, high organisms require coordinated actions of hundreds of genes to synthesize a specific cell type or tissue at a given time and place. In the past 30 years, functions of proteins as regulatory factors involved in such developmental control have been investigated and decoded to a large extent; now regulatory roles of RNAs is a new in growing issue. From an evolutionary point of view, RNAs as regulatory molecules have several advantages over proteins: (1) "simple and local" rules governing

interactions among nucleic acids as compared to protein-protein and protein-nucleic acid interactions make creation and recombination of new regulatory interfaces a much easier evolutionary task; (2) expression of RNA without subsequent translation possibly makes gene regulation more resistant to mutational noise (3) compared to polypeptides, ncRNAs offer high regulatory promptness and sensitivity, in front of dramatically reduced energetic costs. (Qi et al., 2006).

2. classical ncRNAs (rRNA, tRNAs, snRNAs, scRNA, snoRNAs)

As early as 1958, Crick (CRICK, 1958) predicted the existence of a functional RNA adaptor, which mediates between the triplet genetic code and the encoded amino acid. Crick's adaptor hypothesis was later biochemically demonstrated by Hoagland and his colleagues (Hoagland M B, Biol Chem 1958). Now, we call these adaptors transfer RNAs (tRNAs). tRNAs represent the second class of ncRNAs after ribosome RNAs (rRNAs) (BRACHET and CHANTRENNE, 1956). Because both tRNAs and rRNAs assist in translation they are called housekeeping RNAs. It was subsequently found that some other RNA species are also involved in housekeeping roles, such as small nuclear RNAs (snRNA) functioning in pre-RNA splicing, small nucleolar RNAs (snoRNA) functioning in RNA modifications, and RNAs important for the transport and insertion of proteins into membranes and telomeric sequence addition. These early discovered housekeeping RNAs are usually small, constitutively expressed and necessary for cell viability. Details of their biosynthesis and roles they play in eukaryotic cells have been studied in depth. Established knowledge of their biology is subject of molecular biology textbooks. Because of that, we will refer to them as to "classical" ncRNAs. A brief summary of these ncRNAs and their key features is provided below.

rRNA

Ribosomal ribonucleic acid (rRNA) is the RNA component of the ribosome, the protein manufacturing machinery of all living cells. Ribosomal RNA provides a mechanism for decoding mRNA into amino acids and interacts with tRNAs during translation by providing peptidyl transferase activity. In most eukaryotes, the ribosome is composed of two subunits, each one containing specific rRNAs: the large subunit includes the 5S,

5.8S and 28S rRNAs; the small one includes the 18S rRNA. Mammalian cells have 2 mitochondrial (12S and 16S) rRNA molecules and 4 cytoplasmic rRNAs (28S, 5.8S, 5S and 18S). 28S, 5.8S, and 18S rRNAs are encoded by a single transcription unit (45S) separated by 2 internally transcribed spacers. The 45S rDNA organized into 5 clusters (each has 30-40 repeats) on chromosomes 13, 14, 15, 21, and 22 in human. These are transcribed by RNA polymerase I. Separately, 5S rRNA also occurring in tandem arrays, is transcribed by RNA polymerase III. (Lewin, 1999).

tRNAs

The existence of tRNA was first hypothesized by Francis Crick before the genetic codec was discovered exactly. His hypothesis was based on the assumption that there must be an adapter molecule capable of mediating the translation of the RNA alphabet into the protein alphabet. In the human genome, genes encoding for tRNAs are clustered on all chromosomes, except 22 and Y. There are 22 mitochondrial tRNA genes; 497 nuclear genes encoding cytoplasmic tRNA molecules and there are 324 tRNA-derived putative pseudogenes. Cytoplasmic tRNA genes can be grouped into 49 families according to their anti-codon features. tRNA is a small polinucleotidic molecule composed of 75-85 nt. tRNAs are transcribed by RNA polymerase III as pre-tRNAs in the nucleus. Pre-tRNAs undergo extensive modifications inside the nucleus. Some pre-tRNAs contain introns; in bacteria these self-splice, whereas in eukaryotes and archaea they are removed by tRNA splicing endonuclease. The 5' sequence is removed by RNase P, whereas the 3' end is removed by the tRNase Z enzyme. Before tRNAs are exported into the cytoplasm by Los1/Xpo-t, tRNAs are aminoacylated. tRNAs give rise to a secondary, “clover-shape” structure, as well as to a tertiary, "L-shaped" structure, that allows the anti-codon and the acceptor “arms” to link up with the mRNA molecule and the specific amminoacid, respectively (Lewin, 1999).

snRNAs

In nuclear pre-messenger RNA (pre-mRNA) splicing, two sequential transesterification reactions excise an intron and ligate flanking exons, yielding mRNA (Small et al., 2006). These reactions are catalyzed by the spliceosome, a conserved, dynamic machine

composed of five small nuclear RNAs (snRNAs) U1, U2, U4, U5, and U6 and over 100 proteins (Staley and Guthrie, 1998;Jurica and Moore, 2003). Small nuclear ribonucleic acid (snRNA) is a class of small RNA molecules that are found within the nucleus of eukaryotic cells. Genes encoding for them, as well pseudogenes for these RNA species, are dispersed in the human and rodent genome. They are transcribed by RNA polymerase II or RNA polymerase III and are also involved in other important processes such as regulation of transcription factors or RNA polymerase II, and maintaining the telomeres. The association with specific proteins gives rises to complexes known as small nuclear ribonucleoproteins (snRNP) or sometimes as "snurps". The catalytic phase of splicing consists of two consecutive transesterification reactions: in the first step, the 2' hydroxyl of the branch site adenosine (BS-A) attacks the phosphodiester bond at the 5' splice site (5'SS), yielding an intermediate and a free 5' exon, which attacks the 3' SS in the second step to produce an excised intron and spliced mRNA. Catalysis is preceded by an extended assembly phase: the 5' SS and branch site (BS) are bound by U1 and U2 snRNPs, respectively, and the 3' SS by protein factors; the [U4/U6_U5] tri-snRNP joins the complex, and a series of ATP-dependent conformational rearrangements results in the release of U1 and then U4 (Burge et al., 1999;Smith et al., 2007). The recruitment of the CDC5L complex completes the formation of the catalytically competent spliceosome (Makarov et al., 2002).

scRNA

Small cytoplasmic RNAs are a heterogeneous not well known class of non coding RNA. It includes 7SL RNA, i.e. an abundant cytoplasmic RNA which functions in protein secretion as a component of the signal recognition particle (SRP). This is a complex that mediates co-translational insertion of proteins fated to membrane localization or secretion into the lumen of the endoplasmic reticulum. The SRP consists of 6 polypeptides and a 7SL RNA molecule, such as RN7SL1, that is partially homologous to Alu DNA (Ullu and Tschudi, 1984). Alu sequences are the most abundant family of human and rodent middle repetitive DNA sequences (Warren 1982). The primary structure of human 7SL RNA consists of an Alu sequence interrupted by a 155-base pair (bp) sequence that is unique to 7SL RNA. A model of the crystal structure of the Alu-domain of the mammalian SRP has been presented and it has been shown that the Alu

RNA stem-loops of the 5’ domain complex make tertiary interactions (Weichenrieder et al., 2000). A structural motif at the 5 ’ end of human 7SL RNA, called propeller-motif has been identified as an effective activator of RNA polymerase III transcription in vivo and in vitro. Two regulatory elements were identified, one upstream of the 7SL transcribed region and the other one within it; they confirm that 7SL RNA is a RNA pol III transcription product as other small stable RNAs (Englert et al., 2004).

snoRNAs

Small nucleolar RNAs (snoRNAs) are a class of small RNA molecules that primarily guide chemical modifications of other RNAs, mainly ribosomal RNAs, but also other cellular RNAs, including snRNAs, tRNAs and possibly even mRNAs. There are two main classes of snoRNA, the C/D box snoRNAs which are associated with methylation, and the H/ACA box snoRNAs which are associated with pseudouridylation. The majority of vertebrate snoRNA genes is encoded in the introns of proteins involved in ribosome synthesis or translation, and is synthesized by RNA polymerase II, but snoRNAs can also be transcribed from their own promoters by RNA polymerase II or III. In the human genome C/D box snoRNAs are found in tandem repeats within imprinted loci. These two loci (14q32, on chromosome 14, and 15q11, on chromosome 15) have been extensively characterized and in both regions multiple snoRNAs have been found to be located within introns of polypeptide-encoding genes, as clusters of closely related copies. The C/D and H/ACA snoRNAs exert their function through the formation of a canonical guide RNA duplex at the modification site (Bachellerie, 2002). In order to carry out modification, each snoRNA associates with at least four protein molecules in an RNA/protein complex referred to as a small nucleolar ribonucleoprotein (snoRNP). The proteins-RNAs combination depends on the type of snoRNA molecule. The snoRNA molecule contains an antisense element, a stretch of 10-20 nucleotides, which are base complementary to the sequence surrounding the base targeted for modification in the pre-RNA molecule. This enables the snoRNP to recognise and bind to the target RNA. Once the snoRNP has bound to the target site, the associated proteins are in the correct physical location to catalyze the chemical modification of the target base.

3. Non-classical ncRNAs: basic classification

Besides "classical" ncRNAs, ubiquitously expressed and quite exaustively characterized, there is another very large group of ncRNAs, whose expression often are modulated in space and time and whose regulatory functions are fully in the process to be decoded. These RNAs are highly heterogeneous, as for biosynthesis, structure and functions. Even if in some cases (e.g. miRNAs) a substantial advancement of their knowledge has been achieved in a relatively short time, we are still far from having a clear and complete comprehension of their biology. For this reason, we will collectively refer to them as to "non classical" ncRNAs. Actually, the existence of ncRNAs, controlling gene expression through complementary base paring (regulatory ncRNAs), was originally hypothesized by Jacob and Monod quite a lot of time ago, in 1961 (JACOB and MONOD, 1961), and subsequently re-proposed by Britte and Davidson in 1969 (Britten and Davidson, 1969). However, only with the advent of high-troughoutput studies, based on hybridization of genomic tiling arrays and massively parallel sequencing of the transcriptome, it has been shown that "non classical" ncRNAs (1) represent the main output of transcriptional activities, (2) are commonly present in all eukaryotic kingdoms and (3) seem to be involved in various genetic regulations, through multiple mechanisms (Qi et al., 2006).

Based on their lenght, over or below 200 bases, ncRNAs may be primarily classified as short and long ncRNAs (Kapranov et al., 2007). Besides, they may be distinguished, with regard to their subcellular location, as cytoplasmic/nuclear, as well, on the basis of the structure of their 5' and 3' termini, as cap+/cap- and polyA+/polyA-, respectively. These classification criteria are obviously related to the methodologies adopted in grounding studies, for their harvesting and purification. Details about the main classes of "non classical" ncRNAs, primarily categorized with regard to their average lenght, are provided below.

4. Short RNAs

4.1. miRNAs

MicroRNAs (miRNAs) are a large family of ~21 nucleotides long regulators of eukaryotic gene expression, emerging from complex biosynthetic pathways and able to impact on developmental and homeostatic processes, by modulating gene expression at a variety of regulatory levels. miRNAs are processed from precursor molecules termed

primary microRNAs (pri-miRNAs), which are transcribed by RNA polymerase II or, in a few cases, by RNA pol III (Borchert et al., 2006), from independent genes or introns of protein-coding genes. Transcription of miRNA genes is regulated in a similar manner to that of protein-coding genes, and is a major level of control responsible for tissue-specific or development specific expression of miRNAs (Krol et al., 2010). As for *canonical* post-transcriptional maturation, primary RNA precursors of miRNAs, or pri-miRNAs fold into hairpins, which are sequentially processed by the two RNase III family enzymes, Drosha and Dicer. The product of nuclear Drosha cleavage, a ~70-nucleotide long pre-miRNA, is exported thanks to exportin-5 to the cytoplasm, where Dicer processes it to a ~20-bp miRNA/miRNA* duplex. One strand of this duplex, the *driver*, representing the mature miRNA, is then incorporated into the miRNA-induced silencing complex (miRISC), the other one, the *passenger*, is discarded. Two *non canonical* pri-miRNA processing pathways have also been described. In case of *mirtrons*, the substrate of Dicer processing is a short intron generated by the standard splicing machinery in a Drosha-independent way (Okamura et al., 2007;Ruby et al., 2007;Berezikov et al., 2007). In the case of *miR-451*, the product of Drosha-dependent processing of the pri-miRNA is converted into the mature miRNA independently of Dicer, thanks to the nuclease activity of Argonaute 2 (Yang et al., 2010). Finally, miRNAs may undergo sequence editing. Remarkably, during miRNA maturation, both Drosha and Dicer are assisted by a number of cofactors or accessory proteins, some of them playing important regulatory functions (Winter et al., 2009).

It has been recognized from a long time that miRNAs mainly modulate gene expression a post-transcriptional levels, by negatively affecting mRNA stability and/or translatability. In doing that, they act as members of special ribonucleoproteic complexes termed miRISCs (miRNA-instructed silencing complexes), also including Argonaute (AGO) proteins, that directly interact with miRNAs, and the 182 kDa glycine-tryptophan-rich protein (GW182), playing as downstream repression effector. miRNAs act as adaptors for miRISC, making them able to recognize and regulate specific mRNAs. Most animal miRNAs bind with mismatches and bulges, although a key feature of recognition involves Watson-Crick base pairing of miRNA nucleotides 2–8, representing the seed region, to the target mRNA. In contrast, most plant miRNAs bind to nearly perfect complementary sites. The degree of miRNA-mRNA complementarity has been considered a key determinant of the regulatory mechanism. Perfect complementary

allows Ago-catalyzed cleavage of the mRNA strand, whereas central mismatches exclude cleavage and promote repression of mRNA translation. It has been thought that perfect complementary excludes translational repression because it enables cleavage, and it has contributed to the notion that plant and animal miRNAs act in fundamentally different ways (Carthew and Sontheimer, 2009). In addition to translational inhibition and mRNA destabilization, miRNAs may cause sequestration of inhibited mRNAs into nuclear P-bodies (Bhattacharyya et al., 2006;Chu and Rana, 2006;Eulalio et al., 2007). Moreover, in special circumstances (starvation), miRNAs may paradoxically stimulate mRNA translation (Vasudevan et al., 2007). Finally it has been recently shown that regulatory activity of miRNAs is not tightly restricted to post-transcriptional processes. In fact, natural miRNAs complementary to specific gene promoter sequences may stimulate gene transcription ("RNA activation", RNAa (Li et al., 2006)), as in the case of miR-373, targeting E-cadherin promoter and able to induce E-cadherin pre-mRNA transcription (Place et al., 2008).

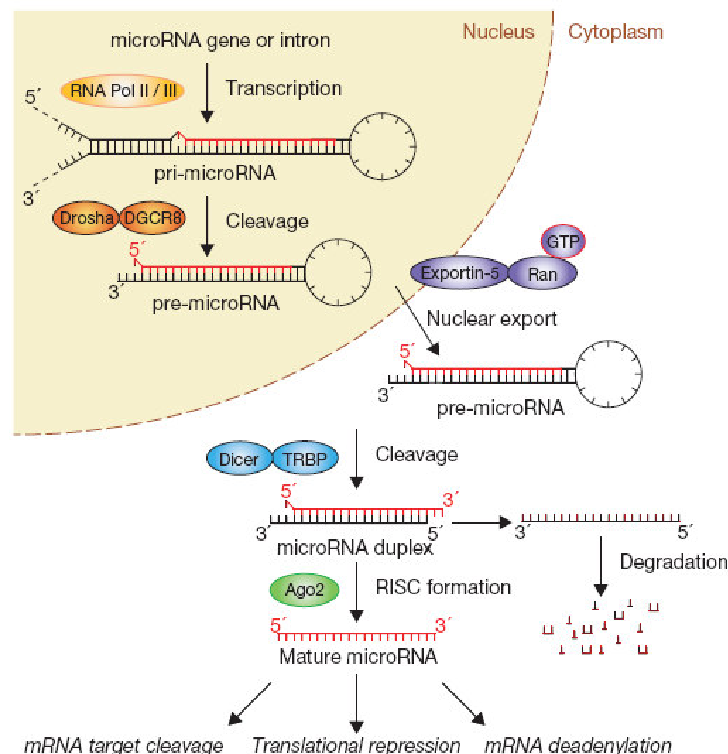


Figure 1. The 'linear' canonical pathway of microRNA processing. The canonical maturation includes the production ofthe primary miRNA transcript (pri-miRNA) by RNA polymerase II or III and cleavage of the pri-miRNA by the microprocessor complex Drosha–DGCR8(Pasha) in the nucleus. The resulting precursor hairpin, the pre-miRNA, is exported from the nucleus by Exportin-5–Ran-GTP. In the cytoplasm, the RNase Dicer in complex with the double-stranded RNA-binding protein TRBP cleaves the pre-miRNA hairpin to its

mature length. The functional strand ofthe mature miRNA is loaded together with Argonaute (Ago2) proteins into the RNA-induced silencing complex (RISC), where it guides RISC to silence target mRNAs through mRNA cleavage, translational repression or deadenylation, whereas the passenger strand (black) is degraded.(Winter et al 2009).

4.2. siRNAs

Small interfering RNAs (siRNA) are double-stranded ncRNAs that regulate gene expression by triggering the degradation of their target mRNAs. There are many common themes in the processing and biogenesis of miRNA and siRNA in the cytoplasm, such as cleavage of a long dsRNA to produce small dsRNA, unwinding of the RNA duplex, RISC loading of only one strand of 21-24 nt (guide strand), and finally the effector function on the target recognizing by Watson-Crick base pairing. About the effector step, miRNA and siRNA diverge in terms of their mechanism of action. While miRNAs mainly suppress translation of the target mRNA, siRNAs often promote the degradation of the target mRNA into pieces, inhibiting the translation as a consequence. However, siRNAs does not regulate gene expression only post-transcriptionally, but also at an upstream level, by modulating chromatin structure and transcription.

Endogenous siRNAs were originally known as defenders of genome integrity in response to foreign or invasive nucleic acids such as viruses, transposons, and transgenes (Carthew and Sontheimer, 2009). Then, chemically or biologically synthesized siRNAs, artificially introduced into cells, started to be largely employed for fast and cheap knock-down of endogenous genes (RNA interference, or RNAi (Farazi et al., 2008). More recently, it has been discovered that siRNAs may also activate gene transcription (RNA activation, or RNAa). All that makes siRNAs an extremely versatile molecular tool, of a paramount interest in the field of gene therapy (Turunen et al., 2009;Li et al., 2006).

4.3. piRNAs

Piwi interacting RNAs (piRNAs) are small non coding RNA of 24-30 nt of length. piRNAs ensure the genomic integrity of eggs and sperm, protecting the germ cell DNA from the double-stranded breaks and insertional mutagenesis caused by active transposons. They have been detected in both egg and sperm-producing cells in Drosophila and zebrafish and it was demonstrated that piRNA pathway is required for spermatogenesis in mouse (Theurkauf, 2008). piRNA genes exist in the genome in clusters; individual

clusters range from 1 to 100 kb in size and encode from 10 to 4500 piRNAs, demonstrating that thousands of piRNAs may be generated from one particular locus. The majority of piRNA clusters are mono-directional, i.e., within a given cluster all piRNAs are derived from one of the two strands of DNA. However, a few bidirectional clusters have also been identified, in which piRNAs derived from both strands (Choudhuri, 2010). It still does not understand how the piRNA long-single-stranded precursor transcripts are converted into small RNAs, but surely the piRNA's biogenesis is Dicer independent. A model about this small RNA class production is the following: piRNA precursor transcripts are fragmented and perhaps trimmed to yield primary piRNAs; primary piRNAs initiate an amplification loop (the 'ping-pong' cycle) that generates secondary piRNAs; and, finally, the resulting amplified piRNAs silence their regulatory targets, such as the mRNA transcripts of transposons, by guiding a specialized sub-class of Argonaute (Ago) proteins. These specialized Argonaute proteins are called PIWI proteins, after the founding member of the sub-family of Ago proteins in Drosophila, *P*-element-*I*nduced *Wi*mpy Testes or Piwi (Lin and Spradling, 1997;Zamore, 2010) Most details of current model for piRNA production and function remain to be confirmed.

4.4. PASRs and TASRs

Multiple transcripts at the 5' boundaries of genes were originally reported by Carninci et al. (Carninci et al., 2006), including unstable lRNAs, postulated to be involved in regulation of gene expression (Crowe et al., 2006;Davis and Ares, 2006;Martianov et al., 2007). PASRs (promoter-associated small RNAs) and TASRs (3' terminus-associated small RNAs) have been described by the group of Gingeras (Kapranov et al., 2007) who discovered them while investigating human nuclear and cytosolic polyadenylated RNAs longer than 200 nucleotides (nt) as well as whole-cell RNAs less than 200 nt, by hybridization to 5nt-dense tiling arrays. PASRs and TASRs are among the pletora of other small ncRNAs, intergenic, exonic as well intronic, reported by these authors and subsequently confirmed by other groups (Xu et al., 2009;Berretta and Morillon, 2009).

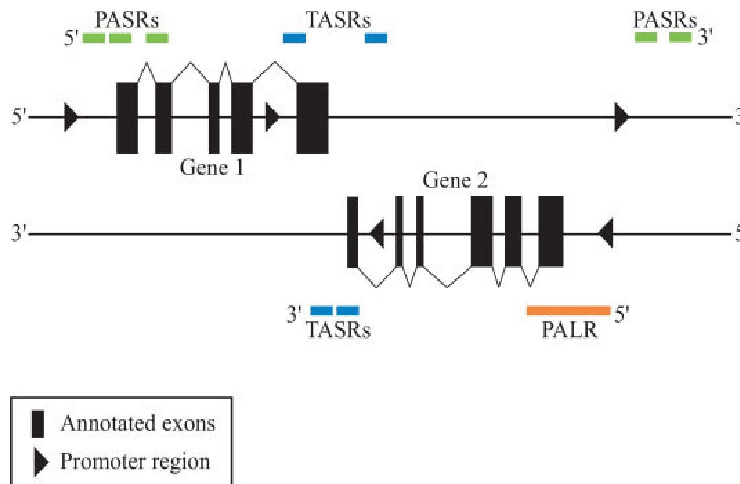


Figure 2. A schematic representation of three novel RNA classes in a genomic context of protein-coding genes. promoter-associated short RNAs (PASRs; green), 3' termini-associated short RNAs (TASRs; blue), and promoter-associated long RNAs (PALRs; orange)(Kapranov et al., 2007).

PASRs and TASRs are small ncRNAs, in the range of 22-200 bases, associated to the ends of polypeptide-encoding genes. They map to genomic regions <1kb, centered around the TSS and the terminus of the gene, respectively; sense PASRs and antisense TASRs prevalently map "inside" the associated gene, antisense PASRs and sense TASRs "outside" it. PASRs and TASRs are associated to almost half of human polypeptide-encoding genes and about 40% of them are conserved between man and mouse. Concerning their origin, it has been suggested (Kapranov et al., 2007) that a subset of them could originate as such, some others could derive from processing of longer non-coding transcripts. In particular PASRs might derive from longer ncRNAs, termed PALRs (promoter-associated long RNAs), which map to TSS, first exon and possibly first intron of the associated gene, and may share with PASRs their 5' end.

As for their function, this is presently obscure. Remarkably, density of all PASRs and antisense TASRs positively correlates with expression levels of the associated gene; the vast majority of silent genes have no associated PASR/TASR. This points to a possible their involvement in fine transcriptional regulation of the associated polypeptide-encoding genes. It was known that siRNAs directed to promoter regions can have a regulatory impact, sometimes silencing (Morris et al., 2004;Ting et al., 2005), sometimes activating (Janowski et al., 2007). That might happen via manipulation of PASR levels, consistent with accumulating evidence that destroying promoter-associated RNA

(PASRs) species can have both positive and negative impacts (Janowski et al., 2007). Finally, it has been proposed that PASRs and PALRs, which are transcribed in the same orientation as their associated protein-coding transcripts, could be that they represent upstream open reading frames (uORFs), encoding short regulatory polypeptides (Crowe et al., 2006).

4.5. tiRNAs

In metazoans, a new class of small RNAs, predominantly 18 nt in length, is derived from the regions adjacent to transcription start sites, *t*ranscription *i*nitiation RNAs (tiRNAs). As in human, also in chicken and fly tiRNAs upstream and antisense to TSSs are associated with bidirectional RNA polymerase II activity (Taft et al., 2009b). tiRNAs might be a signature of stalled or poised RNA Pol II, but the association is not strong. TFIIS cleavage activity has also been shown to be required to release stalled RNA to rapidly induce and resume transcription by Pol II (Adelman et al., 2005;Prather et al., 2005): so, one alternative hypothesis is that tiRNAs derive from the endonucleolytic cleavage of not-elongating nascent RNAs by the transcription elongation factor TFIIS, which must precede the restart of Pol II. This last biogenesis model correlates well with the association of tiRNAs with highly expressed genes, and makes specific predictions about fine mapping of tiRNAs. tiRNA density should peak slightly downstream to the average position of the +1 nucleosome, and relate to the average positioning of +1 nucleosomes in different species (Taft et al., 2009a). tiRNAs show some similarities with the previously described PASRs, including low abundance, a distribution skewed to the 3' of the TSS, the occurrence in more than one organism, the correlation with bidirectional transcription at particular promoters, the association with highly expressed genes and C+G-rich sequences. However, there is a significant difference in size between them, ~18 nt for tiRNAs, versus heterogeneous sizes ranging from ~22 to ~70 nt for PASRs. (Taft et al., 2009a).

5. Long ncRNAs (lncRNAs)

5.1 Generalities

Long ncRNAs (lncRNAs), generally longer than 200 nucleotides and sometimes over 100 kb, are mRNA-like, non-protein-coding RNAs that are pervasively transcribed throughout eukaryotic genomes, often at lower levels as compared to classical "coding" mRNAs.

Tiling array studies of the human genome revealed that the majority of transcribed genome, at least 80%, gives rise to long ncRNAS, often overlapping with, or interspersed among multiple protein-coding and non-coding transcripts (Kapranov et al., 2007). Besides, additional long ncRNAs were bioinformatically identified based on the chromatin signature peculiar to genes actively transcribed by RNA polymerase II (Pol II) (Guttman et al., 2009): trimethylation of lysine 4 of histone H3 (H3K4me3) at their promoter and trimethylation of lysine 36 of histone H3 (H3K36me3) along the length of the transcribed region (Mikkelsen et al., 2007). ncRNAs identified by this signature displayed clear conservation of nucleotide sequence and were wet-validated, showing to be often 5'-capped and poly-adenylated (Guttman et al., 2009).

In general many transcripts are classified as non-coding, simply when they do not harbor ORFs longer than 50–100 amino acids. However, this criterion may be misleading. For example, the 1.5-kb transcript originating from the Drosophila *tal* (tarsal-less) gene, harboring ORFs of <50 aminoacids and therefore originally classified as non coding, is actually translated into 11-amino-acid-long peptides, controlling key steps of tissue morphogenesis and pattern formation (Galindo et al., 2007; Kondo et al., 2007) (Pueyo and Couso, 2008). Moreover, the distinction between coding and non coding genes may not apply at all. That is the case of the *SRA* gene (encoding for Steroid receptor RNA Activator), yielding multiple RNA isoforms, some translated some not (Leygue, 2007), thus showing a gene may play functions carried out by both RNA and protein.

The biological meaning of lncRNAs is complex and still subject of intensive analysis. Long ncRNAs lack obvious features allowing - to date - a reliable a priori prediction of their functions; unlike protein-coding genes where sequence motifs are usually indicative of function, in fact, primary sequences of lncRNAs often contain insufficient information to predict their function. No doubt, some of them are simply precursors of shorter, non-coding regulatory molecules (e.g. the long transcsripts acting as "pri-miRNAs", or the promoter associated long RNAs (PALRs), possibly acting as

PASR precursors). However, a large fraction of them seems to be involved in regulation of gene expression *as such*, playing at different levels and according to distinct "styles". Therefore, a basic categorization of lncRNAs based on this criterion is proposed below. Before moving to its detais, it has to be emphasized here that such a clearcut categorization may not apply to some lncRNAs, apparently implicated in regulation of gene expression at multiple levels. This is the case of *Emx2OS*-ncRNA, the antisense ncRNA associated to the mRNA of the homeobox gene *Emx2*, co-expressed with the latter in the developing uro-genital and central nervous systems. Emx2OS-ncRNA seems - in fact - to stimulate transcription of its sense partner and down-regulates it post-transcriptionally, apparently promoting its Dicer-dependent destabilization (Spigoni et al., 2010).

5.2 long ncRNAs regulating transcription

The impact of ncRNAs on gene transcription may take place by a variety of mechanisms, experimentally demonstrated for a few prototype genes and now waiting to be generalized.

In some cases, poor sequence conservation but similar exon/intron organisation characterizing antisense transcripts from different species suggests that *transcription per se*, and not just the ncRNA molecule, may be crucial to the function. That has been demonstrated in the case of the silent human provirus HERV-K18, where antisense transcription promotes sense transcription (Leupin et al., 2005), as well as in a variety of other cases, where antisense transcription conversely inhibits sense transcription, because of competition between the two transcriptional machineries for shared cofactors, or due to collision between them (Mazo et al., 2007). In other cases, impressive conservation of primary RNA sequence occurring within specific modules of ncRNAs conversely suggests that the regulatory function may be exerted by the ncRNA molecule itself, rather than the *molecular labor* associated to its synthesis.

It has been theoretically proposed and experimentally verified that lncRNAs usually act as integral components of ribonucleoproteic complexes involved in different phases of transcriptional modulation, within which they may perform more distinct key functions (reviewed by Koziol and Rinn, 2010). lncRNAs may simply keep together different polypeptides (binding each of them thanks to specific polypeptide interaction domains), so allowing the formation of oligomeric proteins, which can perform highly sophisticated molecular computations (both epigenetic and co-transcriptional). lncRNAs

may further impact on the conformational state of their polypeptide partners, so allowing a fine allosteric modulation of their catalytic properties. Finally, lncRNAs may be implicated in proper targeting of polypeptide modulators of transcription to their appropriate genomic targets. They can exert this role: (1) as nascent transcripts, stemming from such target regions and still bound to them; (2) by specifically interacting with genomic regions distinct from their birthplace, straightly through a dsDNA/ssRNA triple helix interaction, or via pairing to other ssRNA molecules stemming from them. In case (1), lncRNAs play their role in *cis*, i.e., restricted to the genomic regions they originated from and their surroundings; in case (2), they act in *trans*, i.e., impinging on the expression state of genomic regions distinct from their birthplace neither closely associated to it.

Cis- and trans-active long ncRNAs may impact on transcription of their target genes, by regulating: (1) the accessibility of chromatin to transcription (X-chromosome inactivation, autosomal parental imprinting, DNA damage-triggered chromatin deacetylation, H3K4/H3K9/H3K27 methylation profile at specific loci), (2) the function of classical enhancers, (3) the activity of classical transcription factors. Examples for each of these mechanisms follow below.

Cis-active lncRNAs linked to X-chromosome inactivation. X-chromosome inactivation (XCI) is a classic epigenetic phenomenon associated with many large ncRNAs, by which one X chromosome is transcriptionally silenced in the female sex to ensure that XX and XY individuals have equivalent X-linked gene dosage (LYON, 1961;Wutz, 2003;Payer and Lee, 2008). During XCI, almost all of the ~1000 protein-coding genes on one of two chromosomes become transcriptionally inactivated *in cis* by a single control region known as the “X-inactivation centre” (Xic) (Cattanach and Isaacson, 1967;Rastan and Robertson, 1985). XCI includes at least 4 discrete steps: (1) "counting X cromosomes", (2) "choosing the X chromosome escaping XCI", (3) "silencing X chromosomes undergoing XCI", (4) "maintaing XCI". To date, at least seven distinct non-coding genes have been found within the Xic and surrounding regions and several have been ascribed specific and still not completely clarified functions during XCI.

First of all, there must be chromosome counting, i.e. the determination of whether the cell possesses one or two Xs (XY and XX, respectively) and whether it should therefore initiate XCI. An X: autosome (A) ratio of 1 trigger the XCI cascade (Kay et al., 1994;Boumil and Lee, 2001) while in male cells where X: A is 0.5 XCI is blocked. Evidence points to two ncRNA loci – *Xite* and *Tsix* – as X-linked dosage sensors (Morey

et al., 2004;Lee, 2005). Following the “counting”, a "choosing" mechanism randomly selects one X chromosome as the active X (Xa) and the other as inactive X (Xi) in a mutually exclusive manner (Lee, 2005). This mechanism requires a communication between the Xs *in trans* to ensure that no cell befalls the lethal outcome of creating two Xa or two Xi. In fact, prior the initiation of chromosome-wide silencing, the Xs briefly make contact at the Xic. Although *cis*-acting genes dominate the Xic, this evidence suggests that its function must also be extended *in trans* (Lee, 2005). The same two ncRNAs genes of the Xic, *Tsix* and *Xite* are involved in pairing. Finally, "silencing" factors must be recruited to the future Xi in a colinear fashion, spreading along the chromosome in a strictly *cis*-limited manner and without *trans* effect on homologous loci of the future Xa. The 17-kb *Xist* RNA is transcribed only from the Xi and since its transcription is required for XCI maintenance (Penny et al., 1996), it has been hypothesized that *Xist* recruits chromatin modeling complexes to silence Xi. Recently has been discovered a 1.6 kb ncRNA, *RepA*, which comprises sequences also contained in the 5’ region of *Xist* and which directly binds Polycomb proteins (PRC2) and recruits them to the Xic (Zhao et al., 2008). The actions of *RepA* and *Xist* RNAs are controlled by *Tsix*, a 40 kb ncRNA that is antisense to both RNAs (Lee, 2002). In pre-XCI cells, *RepA* initially recruits PCR2 to the future Xi, although the lncRNA *Tsix,* which is antisense to *Xist* and has an established role as a *Xist* antagonist, inhibits this interaction by binding PCR2, thus competing with *RepA* for this factor. At the onset of cell differentiation, *Tsix* persists only on the chromosome selected to become Xa. It is the persistence of *Tsix* RNA that prevents the up-regulation of *Xist* on Xa. On the future Xi, *Tsix* is downregulated, hence *RepA* can productively engage PCR2 and activate full-length *Xist* transcription. The upregulated *Xist* in turn preferentially binds to PCR2 through its *RepA* sequence, allowing the *RepA*-PCR2 complex to load onto the Xist chromatine and induce histone H3 lysine 27 trimethylation (H3K27,), an event that would then lead to activation of the Xist promoter, accumulation of Xist RNA, and its spread along the X (Zhao et al., 2008).

Cis-active lncRNAs linked to autosomal imprinting. Similar mechanisms have been observed during genomic imprinting of autosomal genes – a mono-allelic mechanism of gene silencing based on the parent-of-origin (Nagano et al., 2008). Examples are *Air* (Sleutels and Barlow, 2002) and *Kcnq1ot1* (Smilinich et al., 1999). Both the 108 kb *Air* and the 91 kb *Kcnq1ot1* ncRNAs are transcribed by RNAP II. These ncRNAs function to silence large domains of the genome epigenetically through their interaction with

chromatin. The Air ncRNA silences *in cis* the three paternally inherited genes Slc22a3, Slc22a2, and Igf2r and is transcribed in an antisense direction (Sleutels and Barlow, 2002). *Air* localizes to the silenced Slc22a3 promoter and recruits the KMT1C lysine methyltransferase, which leads to targeted H3K9 methylation and allele-specific gene silencing by chromatin remodeling (Seidl et al., 2006). The *Kcnq1ot1* mRNA is transcribed from intron 10 of the KCNQ1 gene in an antisense direction and silences several paternally inherited genes *in cis.* In addition, epigenetic silencing has been demonstrated to correlate with the interaction of *Kcnq1ot1* with both the PCR2 Polycomb complex and the KMT1C lysine methyltransferase, as well as with the enrichment of the repressive histone modifications H3K27me3 and H3K9me3 at the loci of silenced genes (Shin et al., 2008).

Cis-active lncRNAs involved in DNA damage sensing. TLS (for translocated in liposarcoma), serves as a key transcriptional regulatory sensor of DNA damage, acting in this context as a repressor of cyclin D1 gene (CCND1). Expression of CCND1 is downregulated in response to DNA damage signals, such as those arising from ionizing radiation (Agami and Bernards, 2000). In response to these signals, several ncRNAs are transcribed from multiple 5' regulatory regions of CCND1, where they remain tethered. TLS interacts with these ncRNAs and the resulting allosteric modification allows its N terminus to bind CREB-binding protein (CBP) and p300, so inhibiting CBP-p300-dependent stimulation of cyclin D1 (CCND1) transcription (Wang et al., 2008).

Trans-active lncRNAs as cofactors of H3K4/H3K9/H3K27 chromatin modifier enzymes. Several lincRNAs associate with chromatin-modifying complexes, so affecting gene expression. Maybe the best example of these lincRNAs is HOTAIR. Hundreds of HOX ncRNAs were identified along the human HOX loci (Rinn et al., 2007) among which, the 2.2 kb long HOTAIR (HOX antisense intergenic RNA) resides in a regulatory boundary in the HOXC locus. HOTAIR is spliced, polyadenylated and has very high nucleotide conservation in vertebrates. siRNA-mediated depletion of HOTAIR ncRNA showed that is required *in trans* to exert gene silencing of the HOXD locus. HOTAIR is transcribed from the HOXC locus and targets Polycomb Repressive Complex 2 (PRC2) to silence HOXD locus and select genes on other chromosomes (Rinn et al., 2007;Gupta et al., 2010). The genomic regions flanking HOXD are also bound by CoREST/REST repressor complexes (Lunyak et al., 2002), which contain LSD1, a demethylase that mediates enzymatic demethylation of H3K4me2 (Shi et al., 2004) required for proper repression of

Hox genes in *Drosophila* (Di Stefano et al., 2007). This suggested that HOTAIR may coordinately interact with both PRC2 and LSD1. In fact, it has been observed that HOTAIR is a modular bifunctional RNA that has distinct binding domains, a 5'domain for PRC2 and a 3'domain for LSD1 complexes. The presence of independent binding sites for PRC2 and LSD1 on HOTAIR suggests that HOTAIR may bridge PRC2 and LSD1 complexes. The ability to tether two distinct complexes enables RNA-mediated assembly of PRC2 and LSD1, and coordinates targeting of these complexes to chromatin for coupled histone H3 lysine 27 methylation and lysine 4 demethylation (Tsai et al., 2010). To date, the majority of trans-acting lincRNAs have been found to associate with the H3K27 methyltransferase PRC2 or the H3K9 methyltransferase G9a. However it is reasonable to hypothesize that other lincRNAs may function as partners of different chromatin modifier enzymes. The full range of biological diversity of these transcripts and their mechanism of action has to be still fully explored.

(Cis-active) enhancer RNAs (eRNA). Recruitment of RNA polII at classical enhancers and *transcription* of these enhancers was originally demonstrated in a number of individual cases, including those of beta-globin and MHC II genes (reveiwed by (Szutorisz et al., 2005;Koch et al., 2008). More recently, two teams showed that enhancer transcription actually is a quite general and genome-wide phenomenon. Kim et al (Kim et al., 2010) identified 12,000 neuronal activity-regulated enhancers, bound by the general transcriptional co-activator CBP. These enhancers are decorated by H3K4me(1) and, in 25% of cases, are bound by RNApolII, in an activity-dependent manner. They give rise to pairs of prevalently polyA-less divergent transcripts, called enhancer RNAs (eRNAs), which span about 2-4 kbs. eRNA levels positively correlate with mRNA transcription levels at nearby genes and may drop to zero if promoters of these genes are ablated. Ørom et al. (Ørom et al., 2010b) analyzed another set of about 3,000 non coding *transcribed* elements, conserved among vertebrates and provided with key features of classical enhancers (when associated to a minimal promoter, even if heterologous, they cis-stimulate transcription, in an orientation independent way). These elements are quite far from their cis-targets (about 100 kb), and are decorated by H3K4me(3) at their TSS and by H3K36 in their body. They are bound by CBP and RNA polII and give rise to unidirectional lncRNAsin the range of 0.9 to 9.0 kb, provided in about half of cases of polyA. Experimental depletion of a number of these transcripts leads to decreased expression of their neighboring protein-coding genes.

Trans-active lncRNAs as cofactors of transcription factors. Initial evidence that ncRNAs could function as transcriptional coregulators was provided by the identification of a conserved ncRNA, termed SRA, in a screen for nuclear receptor coactivators (Lanz et al., 1999). This trans-active lncRNA works as multivalent docking site for transcription factors and coregulators; extensive studies have demonstrated that SRA-mediated nuclear receptor (NR) coactivation does not require the expression of a SRA protein (Lanz et al., 1999). SRA ncRNA seems to act as a scaffold, bringing together NRs, coregulators and elements of the cell transcriptional machinery at NR target genes (Colley et al., 2008). Stem-loops were predicted within SRA, some of which crucial to SRA's activity (Lanz et al., 2002); RNA-interacting domains (RNA recognition motif, RRM), detected within NR coregulators, reasonably allow for their binding to SRA ncRNA (Colley et al., 2008). Remarkably, several other proteins identified in screens for coactivators of nuclear receptors (as well as other sequence-specific transcription factors) contain RNA-binding domains (Auboeuf et al., 2005), originally believed to play roles in co-transcriptional mRNA processing (Puigserver and Spiegelman, 2003). Now the possibility must be considered that these domains function as interactors of SRA-like ncRNAs, working either in *cis* or in *trans*.

Another case of lncRNAs serving as a "ligand" for transcription factors and acting as transcription co-activator in *trans* is that of *Evf2* ncRNA. Vertebrate *Dlx* genes are part of a homeodomain protein family related to the *Drosophila* Distalless gene (*dll*) (for review (Panganiban and Rubenstein, 2002) and play crucial roles in neuronal development and patterning (Feng et al., 2006). The Dlx genes are expressed in bigene clusters, and are regulated by two ultraconserved intergenic enhancers located in the *Dlx*5/6 and *Dlx*1/2 loci, one of them transcribed to a 3.8 kb ncRNA, *Evf*2. *Evf*2 specifically cooperates with the homeodomain protein Dlx2 to increase the activity of the Dlx5/6 enhancer, in a target and homeodomain-specific manner. Whether Dlx-2 binds both DNA and RNA during the cooperative interaction with *Evf2* ncRNA, or whether *Evf*2 sequesters a transcriptional inhibitor independently of binding to Dlx2 remains to be determined. However, the presence of *Evf*2/Dlx2 complexes within the nucleus supports a direct role of the *Evf*2 ncRNA on Dlx2 transcriptional activity (Feng et al., 2004).

5.3 lncRNAs regulating gene expression levels post-transcriptionally

Long ncRNAs may be implicated in post-transcriptional maturation of pre-mRNAs. In particular, antisense transcripts may modulate the splicing of its partner

sense pre-mRNA, as reported for the *TRa2* and *Zeb2* loci (Hastings et al., 2000;Beltran et al., 2008a). In case of *Zeb2*, for example, an antisense ncRNA can mask the 5′ splice site of the zinc finger homeobox pre-mRNA from the spliceosome, so allowing for intron retention. The translation machinery recognizes an internal ribosome entry site (IRE) in the retained intron and binds to it, resulting in efficient *Zeb2* translation and expression (Beltran et al., 2008b)

Long ncRNAs may regulate the half-life of coding mRNAs. Specifically within germ line, pairing of coding retrotransposon sense transcripts and their cognate non-coding antisense partners paves the way to Dicer-dependent cutting of resulting dsRNAs, followed by siRNA-instructed silencing of sense transcripts, from the same or paralogous loci (Tam et al., 2008;Okamura et al., 2008;Watanabe et al., 2008). On the other side, transcripts stemming from psudogenes may "buffer" miRNAs which target mRNAs originating from their cognate coding genes, so increasing mRNAs' half-life. This has been empyrically documented in the cases of the tumour suppressor gene PTEN and its pseudogene PTENP1, as well as of the oncogenic gene KRAS and its pseudogene KRAS1P. Moreover, this has been suggested to be a quite general mechanism, tuning the half-life of a huge number of mRNAs within the soma of mammalian organisms (Poliseno et al., 2010).

6. RNAa

Since 2003, it has shown that the inhibition of mammalian gene expression triggered by small interfering RNAs or RNA interference (RNAi), classically achieved by siRNA-driven endonucleolysis or translational repression of mRNA, may alternatively occur by transcriptional suppression of gene expression (TGS). More recently, new evidence has emerged that small dsRNAs targeted towards gene promoters may also stimulate gene transcription, in a phenomenon referred to as "dsRNA-induced gene expression" or "RNA activation" (RNAa) (reviewed by Morris, 2009). RNA interference has been rapidly developing into a promising new approach for battling cancer and other diseases. By RNAi, it is possible to block the production of mutant genes including dominant-negative proteins, aberrant splicing isoforms, or over expressed genes that have gain of function effects (Aigner, 2006). However, RNAi can only offer antagonism of specific molecular targets for disease treatment. Strategies that can provide agonism of specific gene targets, such as tumor suppressor genes, are equally crucial for the reversal of the disease. Although traditional gene therapy methods have the capacity of

correcting an abnormal copy of an endogenous gene or augmenting the expression of a normal gene, it has its inherent drawbacks including a tedious construction process and detrimental, sometimes even fatal, effects on the host genome. Similar to RNA interference, RNAa offers a fast, simple, and cost-effective approach to alter gene expression. Using dsRNAs, RNAa can activate silenced genes or augment the expression of less active genes.

In the majority of cases, RNAa has been elicited by synthetic dsRNAs, similar to those used for conventional RNAi. However, it has been experimentally proven that an endogenous miRNA complementary to a target sequence within the E-cadherin promoter, miR-373, may also trigger this phenomenon (Place et al., 2008). Moreover, other endogenous miRNAs, such as miR-17-5p and miR-20a, complementary to target sequences within gene promoters, may impact on transcription rates as well (Gonzalez et al., 2008). Remarkably, the effects elicited by the administration of small dsRNAs directed against gene promoters, RNA-triggered TGS or RNAa, are highly sensitive to the location of their target sequences within such promoters. Moreover, the same sRNAs may give rise to different outcomes, depending on the identity of the cell line subject of investigation or its functional state (Li et al., 2006;Schwartz et al., 2008;Turunen et al., 2009).

As for mechanisms of RNA-induced modulation of transcription, current models for interfering RNA-induced heterochromatin assembly propose that siRNAs guide the RNA-induced transcriptional silencing (RITS) complex to nascent centromeric RNAs, working as platforms for the recruitment of siRNA-mediated RITS complex. Then the Ago1-containing RITS complex recruits the RNA dependent RNA polymerase complex (RDRC) and Dicer, which mediates siRNA amplification. This results in further recruitment of histone3-lysin9-methyltransferase (H3K9-MT), trimethylation of H3K9 and creation of a binding site for HP1 proteins, required for heterochromatin formation (reviewed by Suzuki and Kelleher, 2009). Even if mammalian cells apparently lack several components of the RDRC (so that the same full mechanism hardly applies to them), nevertheless a similar recruitment of H3K9/27-MTs and histone deacetylases (HDACs), promoted by dsRNAs and Ago1 toward target chromatine regions, is commonly believed to underly TGS in our cells

As for RNAa, no straight interaction between small dsRNAs and genomic DNA has been reported. Conversely, as shown in cases of the *PR* and the *p21* genes, small RNAs activating transcription seem to interact with endogenous ncRNAs originating from the surroundings of the targeted promoter (Schwartz et al., 2008;Morris et al., 2008), or from genomic regions lying downstream of the 3' end of the gene (Yue et al., 2010). Remarkably, this may result in nuclear/post-transcriptional cutting of endogenous non coding transcripts by Ago2, as shown for the p21 gene (Morris et al., 2008). Thus, at least in some cases, dsRNAs might simply act by relieving the suppressive effect of antisense ncRNAs on sense mRNA synthesis.

However, in other cases (such as the *PR* gene) (Chu et al., 2010), the dsRNA does not destabilize its endogenous ncRNA target. In such cases the dsRNA could rather use its nascent ncRNA target as a molecular dock for directing to the appropriate chromatin region specific multiproteic complexes. These complexes might impact on the functional state of the chromatin in three main ways: (1) making it prone to transcription (2) stimulating transcription; (3) sterically interfering with the recruitment of complexes which antagonize transcription. Concerning polypeptide partners of dsRNAs contributing to RNAa, the Argonautes, also involved in in RNA-triggered TGS, are among them. During RNAa, Ago2 is recruited to the targeted promoter (Turunen et al., 2009) and/or to ncRNAs originating from its surroundings (Chu et al., 2010). Moreover, the same protein is strictly needed, for RNAa (like for RNAdep-TGS), as documented for p21, E-cadherin (Li et al., 2006), VEGF (Turunen et al., 2009) and PR (Schwartz et al., 2008;Chu et al., 2010). Differently from RNA-triggered TGS where Ago1 plays a major role (Janowski et al., 2006;Kim et al., 2006), however, Ago2 paralogs seem to be involved in RNAa only to a marginal extent. Following delivery of sRNAs against the PR promoter, in fact, binding to PR-ncRNAs has been only demonstrated for Ago1 (Chu et al., 2010). Moreover, knock-down of Ago1, Ago3 and Ago4 does not impair RNAa, as documented for *PR* (Chu et al., 2010), or affects it to a very modest extent, as reported for *p21* and *E-cadherin* (Li et al., 2006). Finally, little is still known about the precise chain of molecular events, triggered by dsRNAs and Ago2 during RNAa. In this respect, an increase of H3K4me2/3 and a decrease of H3K9me3 have been remarkably reported to be elicited by activating dsRNAs at the VEGF promoter (Turunen et al., 2009). Such report points to the recruitment of histone3-lysine4-methyltransferases (H3K4-MTs) and histone3-lysine9-demethylases (H3K9-DMs) as possible key steps of some forms of RNAa.

However, at present, how is this recuitment achieved and whether is it crucial to RNAa are still two unsolved issues.

7. ***Emx2*: generalities and evolutionary aspects**

The Empty spiracles homolog gene-2 (*Emx2*) is localized on murine chromosome 19, includes three exons and encodes for a homeodomain transcription factor, EMX2, playing multiple roles in the developing mouse embryo.

It was discovered together its paralog *Emx1* by homology-based screening of a murine brain cDNA library (Simeone et al., 1992a), by a probe originating from *empty spiracles* (*ems*), a gap gene controlling the formation of pre-antennal, antennal and intercalary segments of the *Drosophila* embryo (Dalton et al., 1989; Cohen and Jürgens, 1990).

Emx1 and *Emx2* are expressed in cephalic regions of mouse embryos including the developing cerebral cortex (Simeone et al., 1992a). *Emx2* is expressed also in primordia of the urogenital system and limb buds (Pellegrini et al., 1997; Simeone et al., 1992a). Similar expression patterns are displayed by *Emx* genes in other vertebrate models, including amphibians (Pannese et al., 1998).

The lack of *Emx2* is fatal. *Emx2* homozygous mutant mice (*Emx2*-/-) die soon after birth due to the absence of kidneys (Miyamoto et al., 1997; Pellegrini et al., 1996; Yoshida et al., 1997). They also suffer specific skeletal defects, including the absence of scapulae and ilia (Pellegrini et al., 2001). Knock-out of *Emx2* also deeply affects central nervous system (CNS) development. The cerebral hemispheres, olfactory bulbs and hippocampus are reduced and the dentate gyrus is absent, largely due to decreased cell proliferation (Pellegrini et al., 1996; Yoshida et al., 1997). There are also defects of migration, differentiation and innervations in specific neuronal populations (Boncinelli, 1997; Cecchi, 2002; Cecchi and Boncinelli, 2000; Gulisano et al., 1996). The role of Emx2 in CNS development was subject of further experimental investigations, whose results will be illustrated in details below. Finally, the functional *Emx2* domain is not restricted to the CNS only, but includes the peripheral nervous system (PNS) as well. Heterozygous null mice (*Emx2-/+*) suffer a minor morphological defect between the

incus and malleus within the middle ear and they have a small but significantly greater number of hair cells in the apical region of the cochlea (Rhodes et al., 2003).

8. *Emx2* functions in CNS

The development of the central nervous system is a complex process that begins during early embryogenesis with the formation of the neural plate and continues with its conversion into the neural tube (Fig 1).

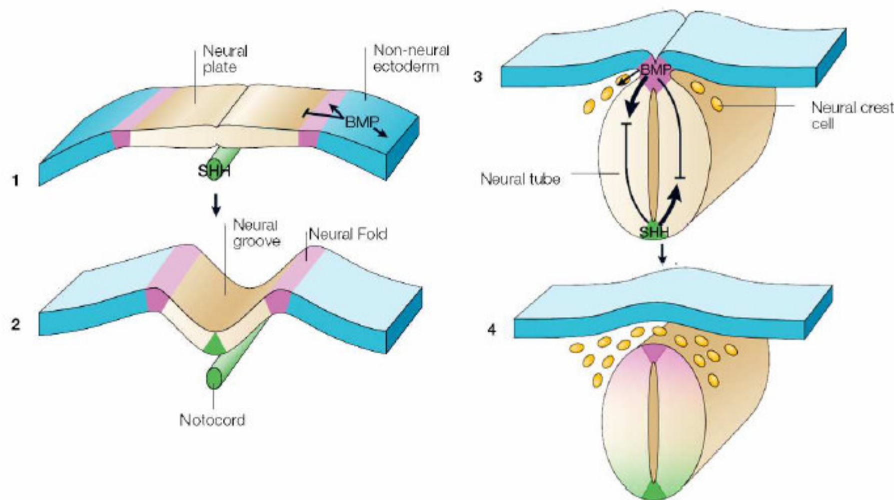


Figure 1. Schematic view of the neural tube formation and the signaling sources involved.1) opened neural plate; 2) neural grove formation; 3) closed the neural tube; 4) delaminating neural crests.

Before the closure of the neural tube, the neural plate becomes subdivided along the antero-posterior axis, into distinct domains corresponding to the three primary brain vesicles, plus the spinal cord anlage. These vesicles are: the prosencephalon (the forebrain), the mesencephalon (midbrain), and the rhombencephalon (the hindbrain). The prosencephalon will give rise to diencephalon and telencephalon, the rhombencephalon to metencephalon and mylencephalon. Finally, the telencephalon will further result into cerebral cortex and basal ganglia (striatum, pallidum and amygdala). These processes are schematically shown in Figure 2.

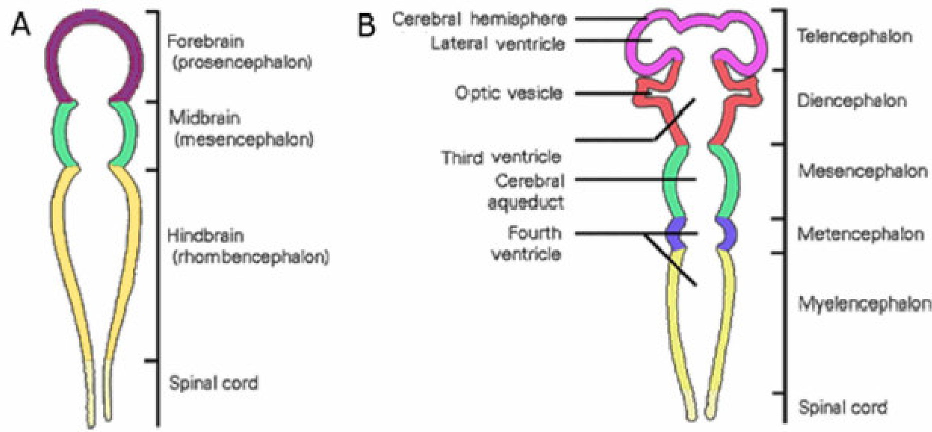


Figure 2 : Regional specification of the developing brain. (A) Early in gestation the neural tube is split in prosencephalon, mesencephalon and rhombencephalon. (B) Further development distinguishes the telencephalon and diencephalon from the prosencephalon. These subregions give rise to the rudiments of the major functional subdivisions of the brain, while the spaces they enclose will form the ventricles of the mature brain.

A pletora of genes are responsible for the molecular control of cerebral cortex morphogenesis. A subset of them harmonically co-regulates different morphogenetic aspects of cortical development, including cell death, self renewal, lineage fate choices, cell cycle control, differentiation, migration, etc. Among these genes there is *Emx2*, whose impact on the above mentioned processes will be subject of the following paragraphs.

8.1 Emx2 in cortical specification

The dorsal telencephalon (the pallium) gives rise to the archicortex (subiculum, hippocampus and dentate gyrus), the paleocortex (olfactory piriform cortex and enthorinal cortex) and the neocortex. The ventral telencephalon (the subpallium) mainly generates the basal ganglia and a part of the amygdala. The process leading the anterior neural plate to the specification of pallial and subpallial territories implies the determination and the specification of telencephalic identities along the R-C and D-V

axes. At mouse embryonal age E7.0, the early R-C patterning of the anterior neural tissue (anterior neural induction) is mediated by antagonistic signals coming from the primitive node (Hensen's node in the chicken) and the anterior visceral endoderm (AVE), required for neural induction and maintenance (Thomas and Beddington, 1996). The AVE is an extra-embryonic tissue that underlies the neural plate and secrete molecules, like cerberus and dickkopf, which antagonize the effects of posteriorizing molecules expressed by the neural plate at this stage, including Wnt and fibroblast growth factor (Fgfs) family members as well as retinoic acid (RA) (Altmann and Brivanlou, 2001;Sasai and De Robertis, 1997). After the anterior neural induction, cells at the junction between neural and non neural tissue of the most rostral part of the brain form the Anterior Neural Ridge (ANR), a secondary organizer necessary for forebrain induction and maintenance. Ablation of the ANR prevents in mice the expression of the telencephalic markers Foxg1 and Emx1 (Shimamura and Rubenstein, 1997). ANR activity in zebrafish is mediated by Tlc, a frizzled related protein acting as a Wnt antagonist. This suggests a general model, in which the default forebrain identity is posterior (diencephalic), and anterior telencephalic identity is achieved via Wnt signaling inhibition by the AVE and ANR patterning centers. In mammals Fgfs, secreted by the ANR, seem also to actively estabilish the telencephalic identity: in fact, when all three Fgf receptors are deleted, the telencephalon does not form (Paek et al., 2009). At E8.0, the mouse telencephalic primordium lies within the anterior third of the paired, downward-folded leaves of the neural plate, and the two sides of the primordium meet at the anterior midline. At this stage a dramatic set of morphogenetic movements and extensive cell proliferation transform the telencephalon into a set of paired vesicles. As a result of early R-C and D-V patterning events, the prosencephalon is subdivided in pallial and subpallial territories, each characterized by the expression of a specific set of TFs. Medial and lateral sectors of the subpallium, giving rise to the medial and the lateral ganglionic eminences (MGE and LGE), express the TFs *Nkx2*.1 and *Gsh2*, respectively, each promoting proper specification of the corresponding sector. The pallium, i.e. the forerunner of the cortex, conversely expresses the TFs Gli3, Pax6 and Emx2 (Simeone et al., 1992a;Walther and Gruss, 1991). Emx2 and Pax6 are both sufficient, even if alone, for dorsal telencephalon specification: *Emx2-/-Pax6-/-* mice exhibit an expansion of the choroidal roof and the subpallium at the expense of the cortex (Muzio et al., 2002) (Fig.3). Remarkably, *Emx2* is down-regulated in *Gli3*-/-mice (Theil et al., 1999) and *Gli3*-/-*Pax6*-/- mice have a

phenotype similar to *Emx2-/-Pax6-/-* mutants (Fuccillo et al., 2006), suggesting that *Emx2* is downstream to *Gli3*.

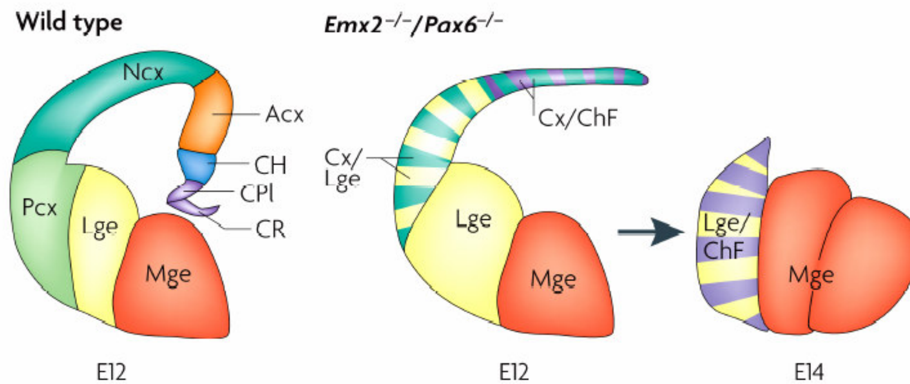


Figure 3: Emx2 involvement in dorsal forebrain specification.Loss of both Pax6 and Emx2 results in ventralization of cortical progenitors and the loss of the neocortical domain (Ncx), archicortex (Acx), cortical hem (CH) and choroid plexus (CPl), choroid field (choroid plexus and choroidal roof) (ChF) by embryonic day 14. Adapted from Muzio & Mallamaci, 2003 and Molyneaux et al. 2007.

8.2 Emx2 in cortical regionalization and arealization

Along its tangential dimension the neocortex is subdivided in distinct areas, showing different structural and functional properties. Four primary areas can be distinguished: visual (V1), somato-sensory (S1), auditory (A1) and motor (M1). An even slight affection of size, position and connectivity of these areas heavily affects brain functioning. The process leading to the specification of these areas is called cortical arealization. Cortical regionalization starts at early stages (E9.5 in mouse), with the specification of a primitive molecular protomap, set up according to specific positional and temporal cues. The codification of these signals initiates intrinsically to the cortical field, resulting from the interplay between soluble factors, secreted at the borders of this field, and transcription factors expressed along tangential gradients within it. Subsequently, (at E13.5 in mice) thalamo-cortical axons (TCA), relaying sensory information from distinct nuclei of the dorsal thalamus to different cortical regions, promote further inter-regional diversification, so leading to the properly called cortical arealization. There are five main patterning centers around the cortical field: rostrally, the anterior neural ridge and commissural plate (ANR/CoP), secreting fibroblastic growth factors (Fgf 8,17,18); ventrally, the precordal plate and the basal telencephalon, secreting Nodal and Sonic hedgehog (Shh); dorsocaudally, the roofplate and the cortical hem, secreting bone morphogenetic

proteins (Bmp2,4,6,7) and vertebrate orthologs of Drosophila wingless (Wnt 3a,5a,2b); laterally, outside the neural tube, the olfactory placode, secreting Retinoic acid (RA); lateroventrally, at the pallial subpallial border, the antihem, secreting epidermal growth factor family members (Tgf8, Nrg1 and Nrg3), Fgf7, as well as the Wnt signaling inhibitor Sfrp2 (Fig.4). Three of these centers commissural plate, cortical hem and cortical antihem are directly implicated in areal patterning of the cortical primordium.

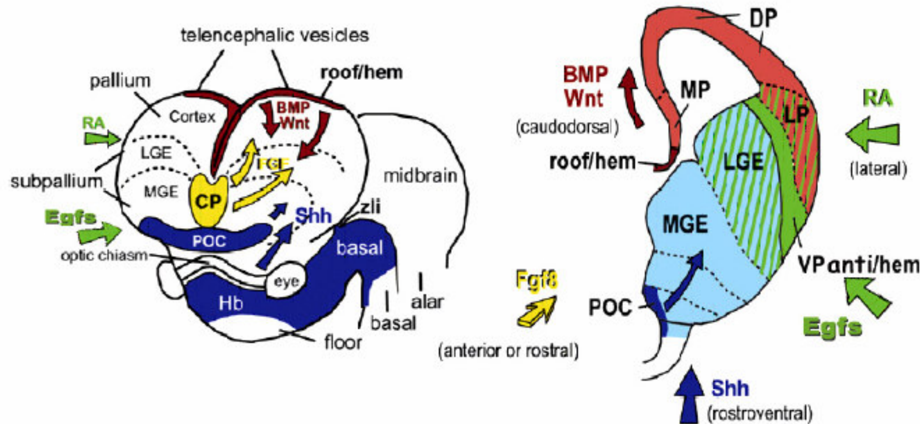


Figure 4: Signaling proteins and patterning centers involved in pallial patterning. Specification and area formation. Schematic diagrams of a mouse brain (left) or a frontal telencephalic hemisection (right) showing the signaling centers and proteins involved in pallial patterning. Modified from (Brox et al., 2004).

Secretion of soluble cues starts at very early stages and alteration of dosages of these growth factors have dramatic consequences: as an example, Fgf8 secretion by the ANR begins very early (E8.5 in mice) (Crossley and Martin, 1995) and its overexpression leads to a posterior shift of cortical areas, on the other hand, overexpression of the soluble form for one of its receptors (Fgfr3) shifts areas anteriorly, by inhibiting Fgf8 activity (Fukuchi-Shimogori and Grove, 2001).

In response to secreted factors coming from the borders of the cortical field, a small set of transcription factors expressed by this field along tangential gradients (Lhx2, Foxg1, Emx2, Emx1, Pax6, Coup-tf1, see Fig.5).) is also implicated in its regionalization, as proven by experimental perturbation of their gradients.

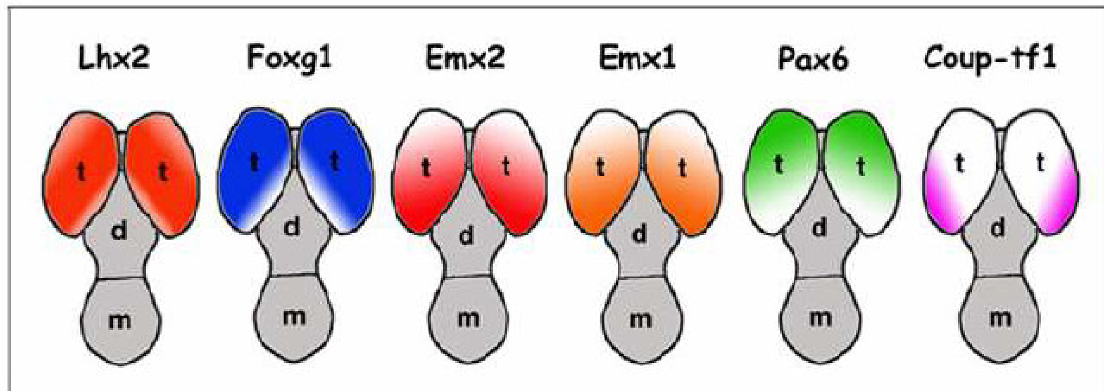


Figure 5: Graded transcription factors genes expression in the early cortical primordium. Schematic representation of E10 mouse brains. Abbreviations: t, telencephalon; d, diencephalon; m, mesencephalon. Image taken from (Mallamaci and Stoykova, 2006a).

Within the neocortical field, *Emx2*, *Pax6*, *COUP-TFI* and *Sp8* have a direct role in arealization determining size and position of cortical areas (O'Leary et al., 2007). In general, TFs confer different area identities to cortical cells within distinct parts of the field, allowing for proper expression of axon guidance molecules that control the area specific targeting of thalamo-cortical afferents. In particular, *Emx2* is expressed in the primary proliferative layer of the cortex along rostral/lateral low-to-caudal/medial high gradients (Gulisano et al., 1996), being more expressed in V1 and less in frontal/motor areas (Fig.6A). In *Emx2* knockout mice, occipital cortex and hippocampus are shrunken and frontal cortex is enlarged. Moreover, the areal distribution of the thalamo-cortical radiation is perturbed, coherently with such areal disproportions.

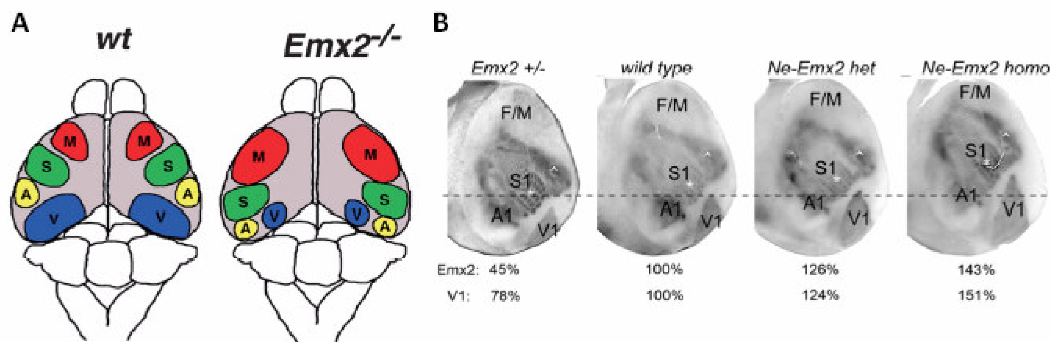


Figure 6 : Area identity shifts in cerebral cortices of perinatal and adult Emx2-/-. Pioneering work on Emx2 null mice showing cortical area shift with enlargement of anterior areas and shrunkage of posterior and medial areas (A). Direct involvement of Emx2 in cortical arealization has been later confirmed by

conditional gain of function and loss of function studies (B).(A) Adaptedfrom Muzio and Mallamaci 2003; (B) Adapted from Hamasaki et al. 2004.

This phenotype strongly suggests the *Emx2* capability to impart a posterior medial identity to the neocortical progenitors. Possible drawbacks, for this interpretation, is the activity exerted by *Emx2* in repressing Fgf8 signaling from the ANR (Fig.7 B) (Fukuchi-Shimogori and Grove, 2003) and in TCA path finding (López-Bendito et al., 2002). However pioneer works, claiming a direct implication of *Emx2* in cortical arealization, have been later confirmed by conditional loss of function (LOF) and gain of function (GOF) experiments (Hamasaki et al., 2004). These authors show that overexpression of *Emx2* by the nestin promoter is sufficient to cause anterior area shift in a dosage dependent manner, when compared to wild type and heterozygous mice. Since in pNes-Emx2 and in *Emx2+/-* mice no effects on Fgf8 expression and no aberrant TCAs are detectable, the direct implication of *Emx2* in cortical area specification has been confirmed (Fig.7).

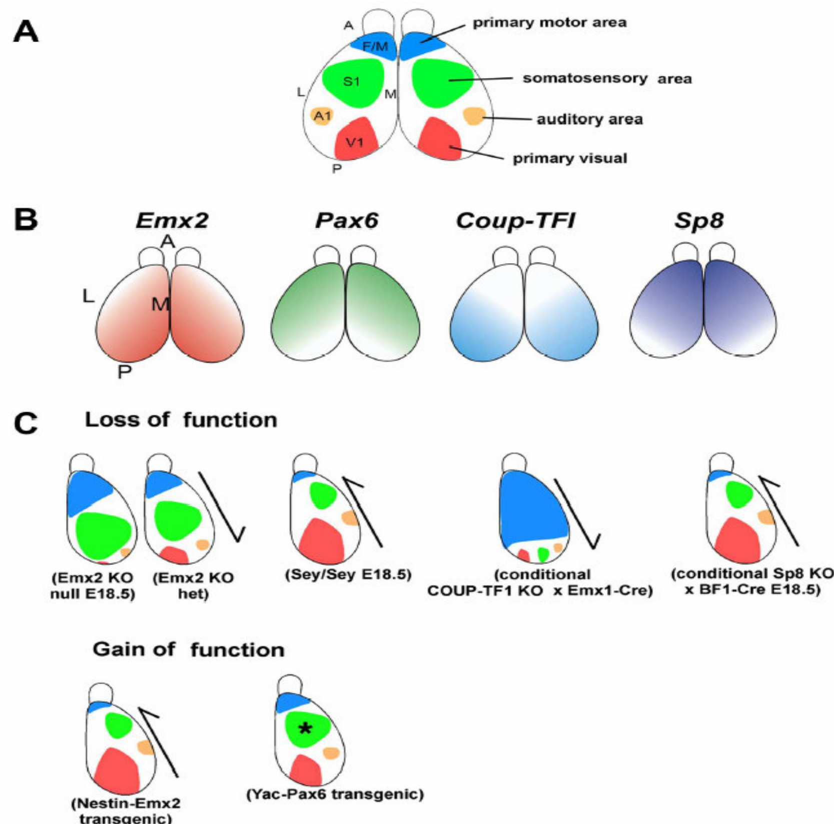


Figure 7.Summary of area patterning and mutant phenotypes. (A) Schematic diagram of anatomically and functionally distinct areas in the mouse. (B) Graded expression of transcription factors along the

anterior-posterior and lateral-medial axes (C) Summary of all reports of loss-of function or gain-of-function mice mutant for TFs that regulate area patterning. Image modified (O'Leary et al., 2007).

8.3 Emx2 in cortical lamination

The three main regions forming the cortex, archicortex, neocortex and paleocortex (the last two ones referrred to as allocortex), can be distinguished on the basis of their lamination: neurons in archicortex an paleocortex are organized in three horizontal layers, whereas those of the neocortex form six layers. Cortical layers located between the neocortex and the allocortex display three to six layers, reflecting their transitional nature. Each cortical layer contains two distinct neuronal types: projection (pyramidal or granule) and interneurons (non pyramidal). The former are glutamatergic and excitatory whereas the latter are GABAergic and inhibitory. Cortical projection neurons originate from progenitors located in the cortical ventricular zone. In contrast, if not all, the most cortical interneurons originate from progenitors located outside the cortex and primarily in the ventral telencephalon, at least in rodents (Marín and Rubenstein, 2001;Gorski et al., 2002). Following their birth in the ventral telencephalon, interneurons use multiple and complex routes of tangential migration to reach their final position in the developing cortex (Marín and Rubenstein, 2001). *Emx2* is expressed in the VZ, but not in the vast majority of neocortical neurons (Mallamaci et al., 1998;Simeone et al., 1992a), except Cajal-Retzius cells within the marginal zone. Cajal-Retzius cell formation is progressively impaired in *Emx2*-/- mice, resulting in aberrant cortical plate layering and pronounced radial glia abnormalities (Mallamaci et al., 2000a).

Abnormal radial migration displayed by of cortical plate neurons in *Emx2*-/- mice is similar to that observed in *reeler* mutant mice. In *reeler* mice, early cortical plate neurons do not penetrate the preplate, which is not split in marginal zone and subplate and give rise to the so called super-plate. Moreover, late born cortical plate neurons do not overcome earlier ones, so that the classical inside-out rule is not followed. In *reeler* mutants, these migratory defects originate from constitutive functional ablation of the *Reln* gene, whose expression in the cortical marginal zone is necessary and sufficient to properly orchestrate neocortical neuron layering (D'Arcangelo et al., 1995). In the *Emx2*-/- marginal zone, *Reln* mRNA expression is apparently normal at E11.5, it is reduced at E13.5 and completely absent since E15.5. In the same mutants, early phases of cortical plate radial migration are poorly affected whereas late phases are impaired in a *reeler*-

like way (Mallamaci et al., 2000a). It is reasonable to hypothesize that the same Cajal-Retzius neurons do not require the products of *Emx2* at the very beginning of their life and only subsequently become dependent on them for surviving and/or retaining their proper differentiation state. However, the increase of the absolute total number of *Reln*-expressing cells taking place in wild-type animals between E11.5 and E15 (Alcántara et al., 1998), as well as E10.5-E19 and E12-E19 birth dating survival data (Mallamaci et al., 2000), suggest that at least two different populations of *Reln* expressing cells do exist, which can be operationally distinguished on the basis of their dependence on *Emx2* function. There is an early transient population, prevalently generated before E11.0 and not dependent on the *Emx2* function, and a later one, still detectable at approximately birth, prevalently generated after E11.5 and dependent on the *Emx2* function for crucial steps of its development.

Remarkably, the caudomedial telencephalic wall and the cortical hem are one of the main sources of neocortical *Reln*-positive Cajal-Retzius (CR) cells (Takiguchi-Hayashi et al., 2004), being the other two the ventral pallium and the septum (Bielle et al., 2005). Cortical hem-born CR cells tangentially migrate beneath the pia mater, in an overall posterior-anterior direction, and finally distribute throughout the entire neocortex, along a caudomedial-high to rostrolateral-low gradient. Embryos lacking *Emx2* display an impaired development of their cortical caudomedial region. Thus, absence of CR cells in these mutants may be a subset of this regional phenotype, possibly arising from dramatic size-reduction of the caudal-medial proliferating pool that generates them.

8.4 Emx2 in regulation of cortical precursors kinetics

Initial studies on cortical precursors expressing abnormal Emx2 levels indicated that the impact of this TF on their histogenetic properties is quite complex. Acute inspection of embryonic brains knock-out for *Emx2* and short term analysis of embryonic neural precursors harboring abnormal dosages of this gene showed that Emx2 promotes cell cycle progression and inhibits premature neuronal differentiation (Heins et al., 2001;Muzio et al., 2005). Conversely, *Emx2* overexpression in more advanced neural stem cells (long-term passaged perinatal ones or adult ones) turned out to force neural precursors to stop proliferating and undergo neuronal differentiation (Galli et al., 2002;Gangemi et al., 2006).

Molecular and cellular mechanisms mediating pro-proliferative activity of Emx2 have been addressed in vivo to some extent. It has been shown that in *Emx2*-/- mice there is an elongation in neuroblast cycling time (TC) due to lengthening of TS (DNA synthesis phase), particularly pronounced in the early caudal-medial cortical primordium. Due to TC elongation, the proliferating pool of the mutant caudal medial cortex “loses” one cell cycle out of four/five, with respect to its wild type counter part. Moreover, cortical progenitors leave cell cycle more frequently. So, because of exaggerated neuronal differentiation, the caudal-medial proliferating pool is deprived of its components at even doubled rates. Those kinetic changes are associated to increased pro-neural/anti-neural gene expression ratio, down-regulation of lateral inhibition machinery and depression of canonical Wnt signalling (Muzio et al., 2005). Remarkably, by pharmacologically reactivating Wnt signaling in *Emx2*-/- mutants, the neurogenic rates are rescued. Wnt and Bmp signalling synergically promote *Emx2* transcription, through a beta-catenin/Smad1,4 binding module located within the *Emx2* telencephalic enhancer (Theil et al., 2002). In turn, *Emx2* up-regulates the final output of the canonical Wnt-signaling machinery, thanks to concerted modulation of ligands (*Wnt3a*, *Wnt8b*, *Wnt5a*, *and Wnt2b*), surface receptors (*Fzd9*, *Fzd10*), intracellular beta-catenin agonists (*Lef1*) and intracellular beta-catenin antagonists (*Groucho*) (Muzio et al., 2005). In this way, near the cortical hem, a positive regulatory loop establishes between *Emx2* and Wnt signaling, crucial for proper sizing of occipital cortex and hippocampus. All these phenomena are much more pronounced in caudal-medial than in rostral pallium, substantially contributing to selective hypoplasia of occipital cortex and hippocampus in the late gestation *Emx2* null embryos.

A more recent study (Brancaccio et al., 2010), confirmed the bimodal impact of Emx2 on neural stem cells (NSC). In fact, lentivirally-mediated overexpression of *Emx2* in dissociated E11.5 cortico-cerebral stem cells initially promotes their self-renewal and then addresses them to neuronogenesis, directly or via neuronal progenitors (NPs). This study also disclosed three new previously unkonw biological activities played by Emx2 in cortical precursors. *Emx2* commits NSCs to glial fates, while inhibiting further maturation of early bipotent glial progenitors; it protects NPs fom cell death; it strongly accelerates neuronal differentiation of NPs (Brancaccio et al., 2010). Molecular mechanisms underlying bimodal activity of Emx2 in NSCs are presently unknown. It is possible that, depending on the stage, embryonic or post-natal, Emx2 may undergo different post-

translational covalent modifications or also may be bound by different age-specific co-factors. That might modify its ability to be recruited to distinct chromatin loci and/or its transcription trans-modulating properties, so making it alternatively able to promote neural precursors self-renewal or neuronal differentiation. Remarkably, it has been shown that temporally structured overexpression of Emx2 (alone or in combination with the TF Foxg1) may promote the expansion of the neural stem pool and ameliorate the neuronogenic output of engineered cultures, while reducing their gliogenic one (Brancaccio et al., 2010). That might allow to circumvent two of the most basic problems arising in cell-mediated therapy of brain disease, the "collapse" of the proliferating stem pool and its exaggerated glial differentiation, so making the histogenetic activities of Emx2 of paramount interest for purposes of gene-assisted therapy of brain diseases.

9. Regulation of *Emx2* expression

Molecular mechanisms controlling *Emx2* expression have been studied in detail, in the developing CNS as well as in the urogenital system.

Emx2 expression in the developing CNS starts as early as at the three-somite stage, in an anterior stripe of neural plate, roughly corresponding to caudal prosencephalon. Slightly later, the gene is downregulated in the diencephalon (where its expression gets confined to a few specific subdomains) and is upregulated in the dorsal telencephalon (Suda et al., 2010). Within the developing cerebral cortex, *Emx2* expression is confined to the ventricular proliferating compartment, forming - from E11.5 onward - an expression gradient with the highest expression levels in the caudal/medial domain (Simeone et al., 1992b). Its expression remains strong in pioneer Cajal-Retzius cells of the marginal zone and fades out in other cortical abventricular layers, where - however - immunoreactivity for its protein can be still detected (Mallamaci et al., 1998). This does not apply to the human cortical plate, where, conversely, *Emx2*-mRNA is strongly upregulated (Bayatti et al., 2008). Finally, in rodents as well as in primates, *Emx2* remains active in proliferating neural precursors populating the anterior subventricular zone and the hippocampal subgranular layer, as well as in postmitotic granule cells of the dentate gyrus (Mallamaci et al., 1998;Galli et al., 2002;Tonchev et al., 2006). The *Emx2* cortical expression pattern apparently emerges as the result of multiple regulatory afferences, inferred on the basis of genetic dissection, however

investigated at molecular level only to a limited extent (reviewed by Mallamaci and Stoykova, 2006b).

To elucidate the transcriptional regulation underlying the cortical expression pattern, Theil et al. (Theil et al., 2002) tested genomic fragments from the mouse *Emx2* locus for enhancer activity. A 4.6 kb fragment was reported, immediately upstream of the *Emx2* translational start site, capable to direct *lacZ* reporter expression in the embryonic forebrain. This enhancer is active in the only diencephalon up to the 3-to-6-somite stage (circa E8.0) and subsequently extends its firing domain to the dorsal telencephalon, where a graded *lacZ* expression pattern, similar to the *Emx2* one, can be firmly detected starting from E11.5. It comprises two elements, 450 bp DT1 and 180 bp DT2, both of which are essential and, in combination, sufficient to direct the expression in dorsal telencephalon. The DT1 element contains binding sites for Tcf and Smad proteins, transcriptional mediators of the Wnt and Bmp signaling pathway, respectively. Transcriptional regulation of Wnt target genes occurs through nuclear translocation of a β-catenin/Tcf complex activating gene expression. Similarly, transmitting the Bmp signal involves phosphorylation, cytoplasm-to-nucleus translocation and binding to chromatin of the Bmp transducers Smad1, Smad5 or Smad8. Mutations of Tcf and Smad binding sites abolished DT1 telencephalic enhancer activity, while ectopic expression of these signaling pathways led to ectopic and synergistic activation of the enhancer. Consistently, null mutants for the *Gli3* gene, lacking *Bmp* and *Wnt* genes expression in the dorsal telencephalon, displayed a severe reduction of *Emx2* expression (Theil et al., 1999;Tole et al., 2000). Recently, another group (Suda et al., 2010) performed a systematic survey, scanning a number of non-coding domains conserved among mouse, human and chick *Emx2* loci for enhancer activity. They re-mapped DT1 (referred to as the θ or FB enhancer), which was found to lie not upstream of the *Emx2* ATG, as previously described, but about 1 kb downstream of the *Emx2* polyA site. Moreover, they found that this enhancer, well conserved among tetrapods, directed almost all the *Emx2* expression in forebrain: within caudal forebrain primordium at E8.5, dorsal telencephalon at E9.5- E10.5 and cortical ventricular zone after E12.5. However, it did not fire in cortical hem and its Cajal-Retzius cells derivatives; moreover, its cortical activity was not graded. Otx, Tcf, Smad and two unknown transcription factor binding sites were essential to all these activities. *Emx2* expression was greatly reduced, but persisted in the telencephalon of *θ enhancer-null* mutant. Such *Emx2* residual expression in *θ*

enhancer-null mutants and the not-graded *lacZ* expression sustained by θ in the cortex strongly suggest that another enhancer for *Emx2* expression, peculiar to the mammalian genome and firing in parallel with the θ element, should substantially contribute to the *Emx2* areal gradient.

Last year, Spigoni et al. (Spigoni et al., 2010) demonstrated that an antisense ncRNAs associated to *Emx2*, *Emx2OS*-ncRNA, is also apparently involved in the genesis of the cortical *Emx2* expression pattern in a complex, bimodal way. Such transcript colocalizes with its sense partner in periventricular neural precursors, is upregulated in their postmitotic progenies within the cortical plate and is absent in Cajal-Retzius cells within the marginal zone. *Emx2OS*-ncRNA promotes the activation of *Emx2* sense transcription, but subsequently destabilizes the mature mRNA product of it, possibly chelating its 5' and so allowing its Dicer-dependent degradation. In this way, it possibly contributes to shutting Emx2-mRNA down specifically within the cortical plate.

Emx2 is also expressed in the intermediate mesoderm and in its urogenital system derivative, from the beginning of its morphogenesis until the adult stage. In both rodents and primates, it is expressed within the uterine endometrium in a pulsating way: it is progressively upregulated during the proliferative half of the reproductive cycle, peaks in the first third of the secretory phase and is later rapidly downregulated (Troy et al., 2003).

Its expression negatively correlates with that of *HOXA10* (Taylor et al., 1998), with inhibits *Emx2* transcription, as shown by gain- and loss-of function approaches in Ishikawa cells. To do that, HOXA10 binds to a 30bp-element, located in the middle of the 300bp HOXA10-dependent *EMXB/C* silencer, lying in turn just downstream of the *Emx2*-mRNA transcriptional start site. Remarkably, this binding does not involve Pbx1 or Meis1 proteins (Troy et al., 2003), which usually modulate target recognition abilities of Hox proteins (Chang et al., 1996;Eklund et al., 2000; Chang et al., 1995; Knoepfler and Kamps, 1995; Lu and Kamps, 1996; Shen et al., 1997). This direct regulation of EMX2 by HOXA10 is very ancient. The *Emx2* 30bp-element binding HOXA10 is, in fact, strongly conserved through vertebrates (see: http://genome.ucsc.edu). Moreover, the interaction between *HOXA10* and *EMX2* is reminiscent of that occurring between *Abdominal-B* and *ems* in *Drosophila m*, where however the former stimulates the transcription of the latter. In this respect, it is likely that differences in the enhancer

sequence accumulating during evolution resulted in altered cofactor selection or cofactor-Hox interaction, thereby giving rise to oppositely oriented gene regulation between species.

10. *Emx2* as a regulator of gene expression

Emx2 has been first shown to modulate gene expression, promoting or inhibiting transcription, depending on contexts.

Emx2 was shown to act as a transcriptional promoter in the proximal limb bud mesenchime (Capellini et al., 2010), where it is implicated in proximo-distal and rostro-caudal specification (Bi et al., 2001; Pröls et al., 2004), promotes cell proliferation (Capellini et al., 2010) and is necessary for scapula and pelvis morphogenesis (Pellegrini et al., 2001). In this structure, *Emx2* is genetically downstream of the homeobox genes *Pbx1-3* and cooperates with *Pbx1* in activating transcription of *Alx1* within the rostral scapular blade. Emx2 binds to Pbx1, possibly thanks to its YPWL tetrapeptide, and synergistically with Pbx1, physically interacts with an evolutionary conserved module, located 5kb upstream of the Alx1 ATG. This was demonstrated by chromatin immunoprecipitation on limb bud tissue and confirmed in vitro, by electrophoretic mobility shift and supershift assays. Remarkably, such module, when linked to a minimal AdML promoter, cis-activates transcription of a luciferase reporter gene, in a combined Emx2/Pbx1-dependent way (Capellini et al., 2010).

Emx2 was also suggested to play complex transcriptional regulation in the developing rostral brain, paying special attention to the process of cortical arealization (reviewed by Mallamaci and Stoykova, 2006a). However, such regulatory function was investigated in detail only in a couple of cases, where the homeoprotein was shown to act as a transcriptional repressor. First, Emx2 regulates *Wnt1* transcription, by direct interaction with its locus. As demonstrated in vitro, by DNAse I footprinting assay, Emx2 strongly and specifically binds to a conserved 110 bp module encompassed in the 5.5kb CNS "3' enhancer", which lies downstream of the *Wnt1* polyA site (Iler et al., 1995). Remarkably, such interaction seems to be crucial to restrict *Wnt1* expression to diencephalon, as suggested by the expression domain of an associated transgenic reporter following mutagenesis of this module (Iler et al., 1995), as well as by spreading of endogenous *Wnt1* into dorsal telencephalon of *Emx2 null* mutants (Ligon et al., 2003). Second, Emx2 antagonizes the spreading of *Fgf8* expression toward intermediate and

caudal cortex. In this case, however, there is no straight interaction between Emx2 and its target. Emx2 simply chelates the zinc-finger transcriptional activator Sp8 and prevents it from binding *Fgf8*, so preventing its transcription (Zembrzycki et al., 2007).

It has to be stressed that a much more complex gene network implying feed-back and feed-forward mechanism among TFs and signaling centers, lies beyond the specification of cortical areas. As an example, rostral Fgf8 signaling from the ANR/CoP is maintained thanks to the Shh signaling from ventral telencephalon as well as by *Sp8* transcription factor, which is more expressed in anterior-medial areas. The Fgf8 signaling might be confined in more caudal areas by the concerted activity of Bmp secreted by the cortical hem, as well as by Sp8 protein sequestering by Emx2 (Sahara et al., 2007) (Fig 8).

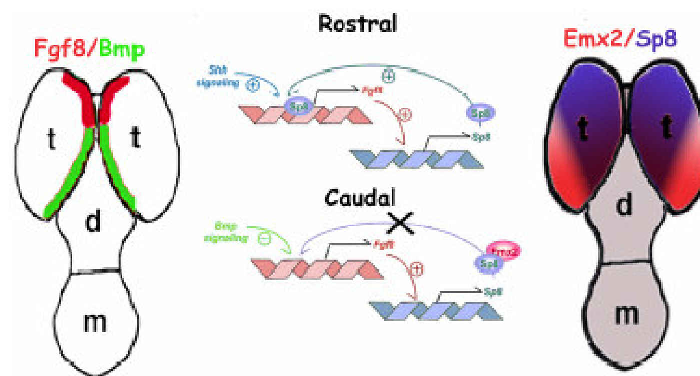


Figure 8: Secreted protein from patterning centres and graded transcription factors interactions. Morphogenes secreted from the patterning centres and transcription factors are involved in a complex web of positive and negative feedbacks. As an example Fgf8 expression could be restricted to the anterior forebrain by the concerted activity of Bmp proteins and Emx2 mediated sequestering of Sp8. Modified from (Mallamaci and Stoykova, 2006a)(O'Leary et al., 2007)

However, it has been shown that Emx2 immonoreactivity is not restricted to the nucleus, but can be detected in the cytoplasm of pyramidal cortical neurons (Mallamaci et al., 1998; Bayatti et al., 2008) and olfactory axons (Nédélec et al., 2004), so pointing to a possible involvment of this protein in post-transcriptional regulation of gene expression.

Consistent with this suggestion, it has been demonstrated that, within olfactory axons, Emx2 specifically binds the eukaryotic initiation factor 4E (eIF4E), via electrostatic interactions involving the conserved YXXXXLL/V module located within its aminoterminal region, not depending on integrity of cellular RNA. Actually, such interaction is not unique to Emx2, having been reported for a number of other "homeodomain transcription factors", among which Bicoid (Niessing et al., 2002), En2, Otx2 (Nédélec et al., 2004), HoxA9 (Topisirovic et al., 2005), PRH (Topisirovic et al., 2003), Prep1 (Villaescusa et al., 2009). It has been shown to allow accurate post-transcriptional regulation of gene expression, at level of nucleus-to-cytoplasm mRNA transport (Topisirovic et al., 2005) and mRNA translation (Niessing et al., 2002; Topisirovic et al., 2003 and 2005; Villaescusa et al., 2009). A similar involvement of Emx2 in gene regulatory circuits is therefore highly likely.

AIMS OF THE WORK

Emx2 overexpression in cortico-cerebral precursors alters their histogenetic behaviour, upregulating stem cells self-renewal and promoting neuronogenesis at expenses of gliogenesis, which is of paramount interest for purposes of gene-promoted brain repair. Transcription of endogenous genes may be stimulated by small RNAs targeted against non coding regions of these genes (RNA activation, or RNAa).

Primary aim of this work was to select artificial miRNAs targeted against "sensitive" regions of the *Emx2* locus, as valuable tools to upregulate its expression in cortico-cerebral precursors, so modulating their population kinetics in the absence of drawbacks arising from the introduction of exogenous transgenes into the cell genome.

Secondary aim was to cast light on molecular mechanisms mediating the impact of these miRNAs on *Emx2* transcription.

MATERIALS AND METHODS

Animal handling

Wild type (w.t) mice (strain CD1, purchased from Harlan-Italy) used in this study were maintained at the SISSA-CBM mouse facility. Embryos were staged by timed breeding and vaginal plug inspection. Animals handling and subsequent procedures were in accordance with European laws [European Communities Council Directive of November 24, 1986 (86/609/EEC)] and with National Institutes of Health guidelines. Embryos (E10.5-E18.5) were harvested from pregnant dames killed by cervical dislocation.

Cell cultures

<u>*Primary cells.*</u>

Cortical primordia and rhombo-spinal tracts were dissected from E10.5-12.5 mouse embryos and dissociated to single cells by gentle pipetting. Dissociated neural precursor cells were cultured in 24-multiwell plates, at 1000 cells/µl in DMEM-F12, 1X Glutamax (Gibco), 1X N2 (Invitrogen), 1 mg/ml BSA, 0.6% glucose, 2 µg/ml heparin (Stem Cell Technologies), 20 ng/ml bFGF (Invitrogen), 20 ng/ml EGF (Invitrogen), 1X Pen/Strept (Gibco), 10 pg/ml Fungizone (Gibco). When required, doxycyclin was added to the culture medium at 2µg/µl. (Clontech). Neural precursors were kept in culture up to 96 hours.

<u>*Cell lines.*</u>

HEK293T cells were cultured for lentivirus production and titration, following standard procedures in IMDM medium (Gibco) supplemented with 10% FBS.

NIH-3T3 cells line were cultured in 12-multiwell plates, at 500 cells/µl, in DMEM-Glutamax-ITM (Gibco) plus 10%FBS.

Lentiviral transfer-vector construction

Basic DNA manipulations (extraction, purification, ligation) as well as bacterial cultures and transformations were performed according to standard methods. Restriction and modification enzymes were obtained from New England Bioloabs and Promega; DNA fragments were purified from agarose gel by the QIAquick Gel Extraction Kit (Qiagen); small and large scale plasmid preparations were done by DNA plasmid purification kit (Qiagen); plasmids were cloned and amplified in TOP10 E.Coli cells (Invitrogen).

As for miRNAs overexpression, the plasmid encoding for the constitutive lentiviral miR expressor, pLVmiR.23, was assembled as follows. The "BfuAI-stuffer" (5' TGCTTCGTGCAGGTCTGCAGGAATTCACCTGCGGACCAGG 3') was cloned into pcDNA^{TM}6.2-GW/EmGFP-miR (Invitrogen) and the DraI/EcoRV 155 bp fragment of the resulting plasmid was transferred into the ΔXhoI-ΔBamHI derivative of pCCL-SIN-18PPT.Pgk.EGFP-Wpre (Follenzi and Naldini, 2002) (gift from L. Naldini), cut by SalI and filled in by Klenow enzyme, in sense orientation. Then, pri-miR-cDNAs, designed by “BLOCK-iT™ RNAi Designer" (Invitrogen) and listed below, were cloned into BfuAI-digested LVmiR.23.

miRNA	*pri-miRNA cDNA insert (underlined, the mature miR)*
Seq2OS.120	TGCTGTTTAGCAGCTTCTTACCGAAGGTTTTGGCCACTGACTGACCTTCGGTAAAGCTGCTAAACAGG
Seq2S.9	TGCTGATATCAGGCGGAAAGATGGCAGTTTTGGCCACTGACTGACTGCCATCTCCGCCTGATATCAGG
DT2S.81	TGCTGTATCACAGCCATTGACTTTGAGTTTTGGCCACTGACTGACTCAAAGTCTGGCTGTGATACAGG
DT2S.164	TGCTGATTAGGACCACTAAACTCTTCGTTTTGGCCACTGACTGACGAAGAGTTGTGGTCCTAATCAGG
Seq3S.239	TGCTGAGTCCAAGGACAATCCATGGAGTTTTGGCCACTGACTGACTCCATGGAGTCCTTGGACTCAGG
Seq3OS.251	TGCTGAAGAAGCGAACACTTCCATGGGTTTTGGCCACTGACTGACCCATGGAAGTTCGCTTCTTCAGG
DT1S.136	TGCTGGATTAATGCAGCCTATCGGGAGTTTTGGCCACTGACTGACTCCCGATACTGCATTAATCCAGG
DT1S.172	TGCTGAACAGACGAGGTTTCTCTATCGTTTTGGCCACTGACTGACGATAGAGACCTCGTCTGTTCAGG
DT1OS.232	TGCTGTGATAGAGAAACCTCGTCTGTGTTTTGGCCACTGACTGACACAGACGATTTCTCTATCACAGG
DT1OS.279	TGCTGATTTGCTTCCTATCCCGATAGGTTTTGGCCACTGACTGACCTATCGGGAGGAAGCAAATCAGG

Seq12S.8	TGCTGTATAGGTTCAGTCGCCACGATGTTTTGGCCACTGACTGACATCGTGGCCTGAACCTATACAGG
Seq12OS.360	TGCTGAAACATTCTAGCAGGCGGGATGTTTTGGCCACTGACTGACATCCCGCCCTAGAATGTTTCAGG

Finally, the negative control pri-miR expressor lentivector was obtained, by transferring the SalI-XhoI cDNA fragment from “pcDNATM 6.2-GW/EmGFP-miR_neg_control_ plasmid" (Invitrogen) into SalI-digested pCCL-SIN-18PPT.Pgk.EGFP-Wpre, in sense orientation.

As for DTs tags overexpression, the two DT1 (400bp) and DT2 (236bp) fragments were amplified from murine genomic DNA, by the primers listed below

DT1-S	*Fw-DT1/AgeI* CCCACCGGTAGCCTTATAAGGCAAGCATTCTGAGAGATCTTC *RevDT1/XhoI* CCGCTCGAGCAGTGTAAAGGGTGGGGGAGATTGTCCTAAATTATGTC
DT1-OS	*Fw-DT1/XhoI* CCGCTCGAGAGCCTTATAAGGCAAGCATTCTGAGAGATCTTC *Rev-DT1/AgeI* CCCACCGGTCAGTGTAAAGGGTGGGGGAGATTGTCCTAAATTATGTCC
DT2-S	*FW-DT2 AgeI* CCCACCGGTCTGGTACAGAGAGGTGAGGGATCAAGTAAGTC *Rev-DT2 BamHI* CGCGGATCCGTGAGTCTTGTCTTGAATGGGTCTGCAATGCTG
DT2-OS	*FW-DT2 BamHI* CGCGGATCCCTGGTACAGAGAGGTGAGGGATCAAGTAAGTC *Rev-DT2 AgeI* CCCACCGGTGTGAGTCTTGTCTTGAATGGGTCTGCAATGCTG

The resulting PCR products were digested by AgeI/XhoI (DT1) and AgeI/BamHI (DT2) and cloned, in sense and antisense orientations, into the "LV:LTR-TREt-(---)-IRES-eGFP-Wpre-LTR" expressor plasmid described in Fig. 1C of (Spigoni et al., 2010), digested by XmaI/XhoI (DT1) or XmaI/BamHI (DT2). The empty LV:LTR-TREt-(---)-IRES-eGFP-Wpre-LTR" was used as negative control. "LV:LTR-TREt-(---)-IRES-eGFP-

Wpre-LTR" expressor derivatives were used in combination with the "LV:LTR-pPgk1-rtTA2^{S}-M2-Wpre-LTR" driver (Spigoni et al., 2010), at the same moi.

Finally, for each construct, DT- or pri-miR inserts and their surroundings were checked by double strand sequencing.

Lentiviral vectors packaging, titration and usage

Third generation self-inactivating (SIN) lentiviral vectors were produced as previously described (Follenzi and Naldini, 2002) with some modifications. Briefly, 293T cells were colipofected (Lipofectamine 2000, Invitrogen) with the transfer vector plasmid plus three auxiliary plasmids (pMD2 VSV.G; pMDLg/pRRE; pRSV-REV). The conditioned medium was collected after 24 and 48hs, filtered and ultracentrifuged at 50000 RCF on a fixed angle rotor (JA 25.50 Beckmann Coulter) for 150 min at 4°C. Viral pellets were resuspended in PBS without BSA (Gibco).

A subset of EGFP-expressing lentiviral vectors was titrated on HEK293T cells, by end point fluorescence titration, as previously described (Follenzi and Naldini, 2002) and titer expressed as transducing units per ml (TU/ml). All viruses viruses were titrated by Real Time quantitative PCR after infection of HEK293T cells, as previously reported (Sastry et al., 2002). One end point fluorescence-titrated lentivirus was included in each PCR titration session and PCR-titers were converted into fluorescence-equivalent titers throughout the study.

Finally, recombinant lentivireuses were delivered to primary neural and NIH/3T3 cells as such and to HEK293T cells in the presence of 9 µg/ml polybrene

Cell counting

Per each sample, cell density was determined by an operator blind to sample identity, just before lentiviral infection (day0) and 96 hours later (day4), scoring about 1,000 cells by a Bürker Counting chamber. Then, day4/day0 density ratios from 3 independent biological replicates were averaged and statistically evaluated, by by one-way ANOVA.

RNA extraction

RNA was extracted from CNS explants and cell cultures by Trizol™ (Invitrogen), according to manufacturer instructions. It was quantified, on Agarose gel as well as by NanoDrop ND-1000, and qualitatively analyzed, by Agilent 2100 Bioanalyzer. As for detection of intronless transcript fragments, the RNA preparation was further treated by DNAseI™ (Promega).

Gene expression assays

cDNA preparation.

A standard amount of RNA from each sample (1,5 µg, when not otherwise stated; 2,5 µg, as for intronless transcript fragments, DT-ncRNAs and Emx2-pre-mRNA) was retro-transcribed by SuperScriptIII™ (Invitrogen), according to manufacturer instructions, with minor modifications. Retrotranscriptions were generally primed by 5µM random hexamers. To assay orientation of DT transcribed tags, these ones were replaced by 0.1µM transcript specific oligos plus 0.1µM Tbp-b-RT. Oligo sequences are listed below:

oriented transcribed tag	*oligo*
ncRNA-DT2 sense	*DT2/R-OUT2* TGAGAAGGTGAGTCTTGTC
ncRNA-DT2 antisense	*DT2/F-OUT2* CAGTACACCCTGGTACAG
ncRNA-DT1 sense	*DT1/R-OUT2* AGATTGTCCTAAATTATGTCCAG
ncRNA-DT1 antisense	*DT1/F-OUT2* CTGAGAGATCTTCCACTCTTA
mRNA-Tbp	*Tbp-b-RT* CTTGCACGAAGTGCAATG

Quantitative PCR.

For every PCR reaction, cDNA representing 30 ng of total RNA was analyzed by the SybrGreen™ qPCR platform (Biorad). In case of the secondary PCR, for detection of Emx2-pre-mRNA, 1/10th of primary PCR product was used as a substrate. Reactions were performed in 10-20 µl, according to the general thermal program "[95℃/3 min] (1X); [T(a)/t(a); 72℃/t(e) sec; 95℃/10 sec] (40X; only 15X, in case of primary PCR for Emx2-pre-mRNA)". Plates were generally read at 72℃, except DT2-ncRNA, read at 68℃. Primers, annealing temperatures (T(a)), annealing times (t(a)) an elongation times (t(e)) were as below:

amplicon (size, bps)	*oligo [working concentration]*	*T(a) °C*	*t(a) sec*	*t(e) sec*
Emx2-mRNA (145)	*E2S/N2F [500 nM]* GGAAAGGAAGCAGCTGGCTCACAGTCTCAGTCTTAC *E2S/N2R [500 nM]* GTGGTGTGTCCCTTTTTTCTTCTGTTGAGAATCTGAGCCTTC	65	20	20
Tbp-mRNA (174)	*Tbp-b/Fw [500 nM]* ATTCTCAAACTCTGACCACTGCACCGTTG *Tbp-b/Rev [500 nM]* TTAGGTCAAGTTTACAGCCAAGATTCACGGTAG	60	20	20
DT1-ncRNA (167)	*DT1/F [500 nM]* AGGTTGTTTTTTGCATGCTTCATTTGCTTCCTATCC *DT1/R [500 nM]* ATATTCCTGGTATGATAATTGCTTAAACTGATTTGCAC	55	20	20
DT2-ncRNA (160)	*DT2/F [250 nM]* TGGGAGTTTCAATCAAATCTTCCAAACAGGTCTGG *DT2/R [250 nM]* TGACTACAAATTAGGACCACTAAACTCTTCACTCAG	68	40	0
Emx2-pre-mRNA (primary PCR)	*I1 [500 nM]* GTCTCTGAAGCTCGTTTGGGTTACTG 3 *I4 [500 nM]*	55	10	40

	AGTGAGTGTAGAGCAGAGTTGAAGTCC			
Emx2-pre-mRNA (secondary PCR)	*I2[500 nM]* GCGAGGTCTTTGAATCCTGTTTC *I3[500 nM]* GCAGAGTTGAAGTCCAGTGAACC	55	10	40

Each PCR reaction was run at least in technical triplicate, results were averaged by MJ Opticon Monitor 3.1 software (Biorad) and averages were normalized against TATA binding protein (Tbp) mRNA. In case of intronless transcripts, normalized results were further diminished by the normalized background signal yielded by RT(-) negative controls. Normalized data from at least 3 independent biological replicates were averaged and statistically evaluated, by one-way ANOVA, run by Excel Mac 2008 software (Microsoft).

RACE Assay

3’ and 5’ rapid amplification of cDNA ends (RACE) was performed by SMARTer™ technology (Clontech), according to manufacturer’s instructions, with minor modifications. 1μg of RNA originating from a pool of 16 E11.5 cortices was used. Synthesis of tagged cDNA for 5' and 3' RACE was primed by random primers and "3'-RACE CDS Primer A", respectively. Primary and secondary PCRs were primed by Universal Primer A Mix (UPM) and Nested Universal Primer (NUP), respectively, appropriately combined with the gene-specific primers listed below:

oriented transcribed tag	*oligo*	*used in*	
		1st PCR	*2nd PCR*
DT1-S/5' & DT1-OS/3'	*DT1/RACE4* CTGATTTCACTTTCACAAAAGCTCACAGGACGC	+	
	DT1/RACE5 CACAGGACGCTTTGTAGCTCGAACAGAACAGACG	+	+
	DT1/RACE6 CAGAACAGACGAGGTTTCTCTATCAATGG		+
DT1-S/3' & DT1-OS/5'	*DT1/RACE1* CCATTGATAGAGAAACCTCGTCTGTTCTGTTCGAGCTAC	+	
	DT1/RACE2	+	+

	CAAAGCGTCCTGTGAGCTTTTGTGAAAG		
	DT1/RACE3 GTGAAAGTGCAAATCAGTTTAAGCAATTATCATACCAGG		+
DT2-S/5' & DT2-OS/3'	*DT2/R* TGACTACAAATTAGGACCACTAAACTCTTCACTCAG	+	
	DT2/R-INT CAGCATTAGCATCATTAGCATTAGTACGCTTG		+
DT2-S/3' & DT2-OS/5'	*DT2/F* TGGGAGTTTCAATCAAATCTTCCAAACAGGTCTGG	+	
	DT2/F-INT TGCCGCTCAAAGTCAATGGCTGTGATAC		+

Primary and secondary PCRs were performed by Advantage 2 (Clontech) and Taq (Promega) polymerases, respectively. Negative control reactions were run on SMARTase(-), DNAseI-treated RNA samples. RACE PCR fragments were purified using Nucleo Trap Gel (Clontech) and QIAquick Gel (Qiagen) Extraction Kits and sequenced, straightly or upon their cloning into pGEM®-T Easy cloning vectors (Promega). Finally, sequences were aligned with the murine genome (NCBI37/mm9), by the USCS blat browser.

ChIP-qPCR

Chromatin immunoprecipitation quantitative polimerase chain reaction assays (ChIP-qPCRs) were performed on chromatin extracted from neural cell cultures, infected with bio-active and control lentiviruses and kept in culture for 96 hours, according to the MAGnify™ Chromatin Immunoprecipitation System protocol (Invitrogen), with minor modifications. Each ChIP assay was run on chromatin from 10^6 cells. Chromatin was fixed in cell for 10 min and, after cellular lysis, sonicated by a Soniprep 150 apparatus (30 sec ON on ice, 30 sec OFF on ice, oscillation amplitude 10 microns, 7 cycles) into ~300bp fragments. Sonicated chromatin was immunoprecipitated by the following antibodies: α-RNApol II (Abcam, 5408), 2,5 µg; α-IgG Rabbit (Invitrogen), 1 µg. Cross-linking was reversed overnight at 65°C. Immunoprecipitated DNAs were purified according to manufacturer's instructions.

1/30 of each immunoprecipitated DNA sample was amplified by qPCR. Reactions were performed in 30 µl, according to the general thermal program "[95°C/3

min] (1X); [T(a)/t(a); 72℃/20 sec; 95℃/10 sec] (40X)". Plates were generally read at 72℃. Primers, annealing temperatures (T(a)) and annealing times (t(a)) were as below:

amplicon *(size, bps)*	*primer name [working concentration]* *5'-3' sequence*	*T(a)* *°C*	*t(a)* *sec*
α (80)	*E2S/N1F [250 nM]* GCGAGTAATAGCGACCAATCATCAAGCCATTTACCAGGCTTCG *Emx2-A0/Rev [500 nM]* GCCTAATTAGTGCGGGGATCACATAAACAGCTTC	68	40
β (133)	*Emx2-A1Fw [500 nM]* CTCATGCTAGGGGTAAAAACAACCCACGATAAAAGAC *Emx2-A1/Rev [500 nM]* TCAATCTCTCCAACCACTAAAAGGAAAAGTTGACTTG	60	20

Limited to Emx2-pA, the 72℃ elongation step was omitted and the plate was read at 68℃. Per each sample, reactions were run at least in technical triplicates and the average number of amplicons was calculated.

For each experimental session, the increase of RNApolII binding to chromatin, or "enrichment", induced by any bioactive miRNA was calculated according to the following formula:

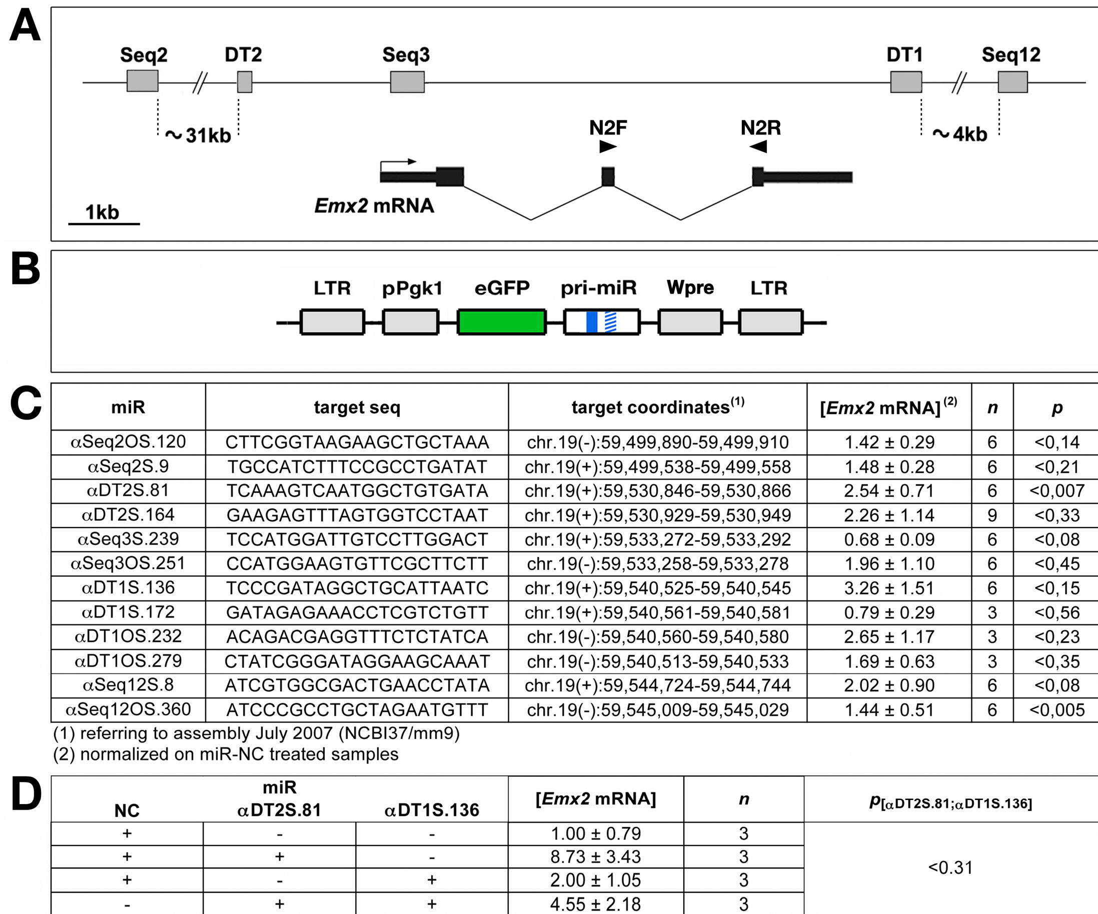
$$\log_2(Ab_X/IgG_X)-\log_2(Ab_{NC}/IgG_{NC}),$$

where: (1) Ab and IgG are the number of amplicons immuno-precipitated by α-RNApolII and IgG, respectively; (2) X is the miRNA under examination; (3) NC is the negative control miRNA.

RESULTS

As proposed by Britten and Davidson (Britten and Davidson, 1969) and experimentally verified in plants, yeast and insects (Mette et al., 2000; Sijen et al., 2001; Volpe et al., 2002; Pal-Bhadra et al., 2004), siRNAs and miRNAs targeted against non coding conserved sequences surroundings vertebrate polypeptide-encoding genes may perturb transcription of these genes, in a positive (RNAa, RNA activation) (Li et al., 2006; Kuwabara et al., 2004; Janowski et al., 2007; Chen et al., 2008; Place et al., 2008; Schwartz et al., 2008; Morris et al., 2008; Turunen et al., 2009) or negative (RNA-directed TGS, transcriptional gene silencing) (Kawasaki and Taira, 2004; Morris et al., 2004; Park et al., 2004;Ting et al., 2005; Janowski et al., 2006; Morris et al., 2008; Gonzalez et al., 2008) way.

We wondered if siRNAs/miRNAs against conserved non coding sequences (NCSs) and/or transcriptionally cis-active elements of the *Emx2* locus were able to modulate expression levels of its main transcript. For this purpose, we selected the 3 most conserved *Emx2* NCSs, located >30kb upstream of the *Emx2* TSS (Seq2), just downstream this TSS (Seq3) and >5kb downstream of the *Emx2* polyA site (Seq12), by USCS online software. Moreover, we took into account the two telencephalic enhancers flanking the *Emx2* transcription unit, DT2 and DT1, previously described by Theil et al (Theil et al., 2002) (Fig. 1A). We designed miR-155-based pri-miRs, harboring artificial miRNAs against both strands of these modules, by Invitrogen Block-iT$^{(TM)}$ software. Such pri-miRs were transferred into the lentiviral vector pLVMiR.23 (Fig. 1B) and the resulting recombinant lentiviruses were delivered to floating cultures of mouse E12.5 cortico-cerebral precursors, at m.o.i. = 20. Four days later, these cultures were profiled for *Emx2* mRNA levels, by quantitative RT-PCR (qRT-PCR). Except miR-aSeq3S.239, which downregulated *Emx2* to 0.69±0.09, with $p<0.08$, the majority of tested miRNAs elicited an opposite effect. In particular, miR-aDT2S.81 and miR-aDT1S.136 upregulated *Emx2* by 2.54±0.71 and 3.26±1.51, with $p<0.007$ and $p<0.15$, respectively (Fig. 1C). Coinfection of cortical precursors by both miR-aDT2S.81 and miR-aDT1S.136, did not show any statistically significant functional interaction between these miRNAs (Fig. 1D).



miR	target seq	target coordinates(1)	[*Emx2* mRNA](2)	*n*	*p*
αSeq2OS.120	CTTCGGTAAGAAGCTGCTAAA	chr.19(-):59,499,890-59,499,910	1.42 ± 0.29	6	<0,14
αSeq2S.9	TGCCATCTTTCCGCCTGATAT	chr.19(+):59,499,538-59,499,558	1.48 ± 0.28	6	<0,21
αDT2S.81	TCAAAGTCAATGGCTGTGATA	chr.19(+):59,530,846-59,530,866	2.54 ± 0.71	6	<0,007
αDT2S.164	GAAGAGTTTAGTGGTCCTAAT	chr.19(+):59,530,929-59,530,949	2.26 ± 1.14	9	<0,33
αSeq3S.239	TCCATGGATTGTCCTTGGACT	chr.19(+):59,533,272-59,533,292	0.68 ± 0.09	6	<0,08
αSeq3OS.251	CCATGGAAGTGTTCGCTTCTT	chr.19(-):59,533,258-59,533,278	1.96 ± 1.10	6	<0,45
αDT1S.136	TCCCGATAGGCTGCATTAATC	chr.19(+):59,540,525-59,540,545	3.26 ± 1.51	6	<0,15
αDT1S.172	GATAGAGAAACCTCGTCTGTT	chr.19(+):59,540,561-59,540,581	0.79 ± 0.29	3	<0,56
αDT1OS.232	ACAGACGAGGTTTCTCTATCA	chr.19(-):59,540,560-59,540,580	2.65 ± 1.17	3	<0,23
αDT1OS.279	CTATCGGGATAGGAAGCAAAT	chr.19(-):59,540,513-59,540,533	1.69 ± 0.63	3	<0,35
αSeq12S.8	ATCGTGGCGACTGAACCTATA	chr.19(+):59,544,724-59,544,744	2.02 ± 0.90	6	<0,08
αSeq12OS.360	ATCCCGCCTGCTAGAATGTTT	chr.19(-):59,545,009-59,545,029	1.44 ± 0.51	6	<0,005

(1) referring to assembly July 2007 (NCBI37/mm9)
(2) normalized on miR-NC treated samples

miR			[*Emx2* mRNA]	*n*	$p_{[\alpha DT2S.81;\alpha DT1S.136]}$
NC	αDT2S.81	αDT1S.136			
+	-	-	1.00 ± 0.79	3	<0.31
+	+	-	8.73 ± 3.43	3	
+	-	+	2.00 ± 1.05	3	
-	+	+	4.55 ± 2.18	3	

Figure 1 Artificial miRs against the *Emx2* locus and their impact on *Emx2* transcription. (A) Schematic representation of the mouse *Emx2* locus, with previously characterized cis-active regulatory modules (DT2, DT1), selected non coding, evolutionary conserved, sequences (Seq2, Seq3, Seq12), and primers used for qRT-PCR evaluation of *Emx2* mRNA levels. (B) Lentiviral vector driving constitutive expression of primary transcripts encoding for artificial miRNAs. LTR, long terminal repeat; pPgk1, (human) phospho-glycero-kinase 1 promoter; eGFP, enhanced green fluorescent protein; WPRE, Woodchuck hepatitis virus Posttranscriptional Regulatory Element. (C) Emx2 mRNA levels in E12.5 embryonic cerebral cortex precursor cells, infected by lentiviral expressors of miRNAs against listed sequences, at m.o.i. = 20, cultured as floating neurospheres over 4 days and profiled by qRT-PCR (data normalized against TATA-binding protein mRNA (*Tbp*) and further normalized against negative control miRNA-treated samples). (D) Assaying for functional interaction between miR-αDT2S.81, miR-αDT1S.136 in E12.5 cortico-cerebral precursors. miR-expressing lentiviruses were used in different combinations, at total m.o.i. = 30. Infected cells, cultured as floating neurospheres over 3 days, were profiled by qRT-PCR. miR-NC lentivector encodes for an artificial miR targeting no mammalian transcripts. Both miR-αDT2S.81 and miR-αDT1S.136 up-regulated *Emx2*-mRNA; however no functional interaction between them was detectable, by two-ways ANOVA assay.

Besides, when miR-aDT2S.81 and miR-aDT1S.136 were delivered, combined or alone, to rhombospinal precursors or non-neural NIH/3T3 cells (Fig. 2), no *Emx2* upregulation took place. This means that the effects of these miRNAs on *Emx2* transcription are highly context-dependent, as already shown for RNAa at the PR, E-cad and VEGF loci (Li et al., 2006; Schwartz et al., 2008).

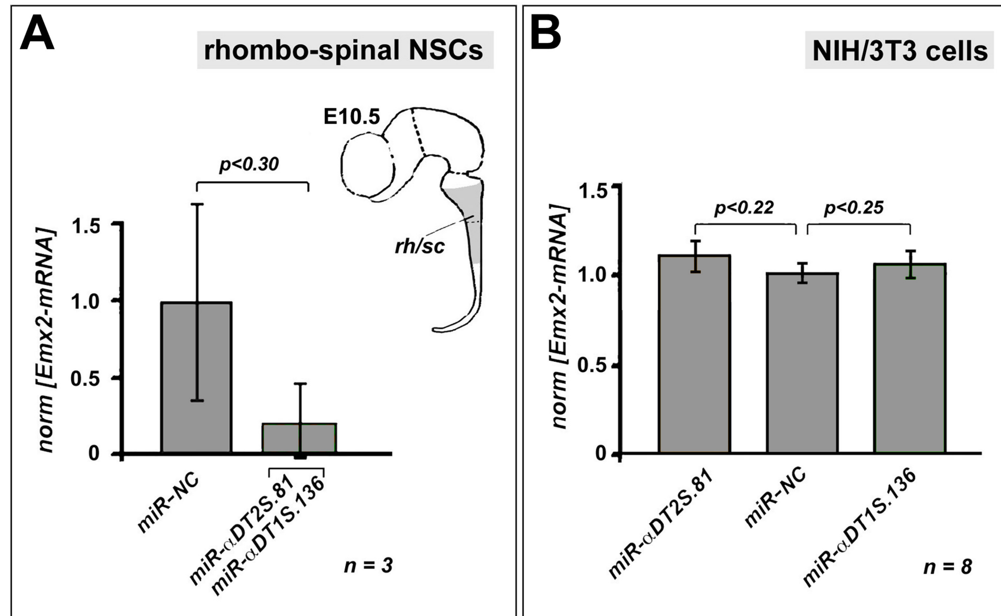


Figure 2. Tissue specificity of RNAa by artificial miRs against the *Emx2* locus. (A) E10.5 rhombencephalic neuroblasts, expressing very low levels of *Emx2* mRNA, were co-infected by lentivectors driving constitutive expression of miR-αDT2S.81 and miR-αDT1S.136, at m.o.i.= 15+15, and cultured as floating neurospheres over 3 days. Compared to control, this treatment did not elicit any statistically relevant up-regulation of *Emx2* mRNA. (B) NIH-3T3 cells were infected by lentivectors driving constitutive expression of miR-αDT2S.81 or miR-αDT1S.136, at m.o.i.= 20, and cultured over 3 days. Compared with control, these treatments did not elicit any statistically relevant modulation of *Emx2* mRNA.

It was previously shown that *Emx2* stimulates embryonic neural stem cells self-renewal, protects neuronal progenitors from cell death, antagonizes their early exit from cell cycle and inhibits gliogenesis (Heins et al.,, 2001; Muzio et al., 2005; Brancaccio et al., 2010), so resulting in an overt expansion of the cortical proliferating pool. To assay if the moderate *Emx2* upregulation elicited by RNAa is sufficient to trigger these phenomena (all of obvious potential interest for purposes of neural repair), we monitored the size of cortical cultures infected by miR-aDT2S.81 and miR-aDT1S.136 (Fig 3) and, in both cases, found it increased by >3-fold over 4 days ($p<0.05$).

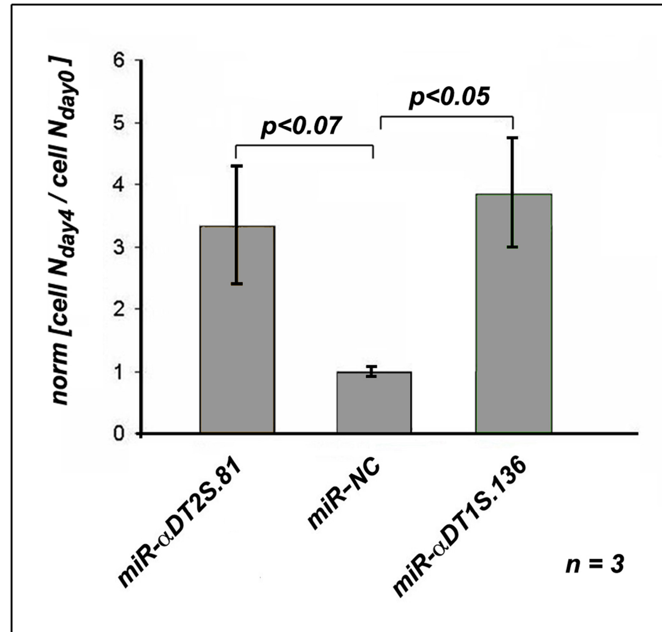


Figure 3. Modulation of cortico-cerebral precursors population kinetics by miR-αDT1S.136 and miR-αDT2S.81 expression. E12.5 cortico-cerebral precursors, infected by lentiviruses driving miR-αDT2S.81 or miR-αDT1S.136 expression, at m.o.i.= 20, and cultured over 4 days as floating neurospheres gave rise to at least 3.5 more progenies, compared with controls.

The fact that miR-αDT2S.81 and miR-αDT1S.136 targets fall outside the main *Emx2* transcription unit suggests that *Emx2* upregulation exerted by these miRNAs could be due to pre- and/or co-transcriptional mechanisms. We tested this inference, by measuring levels of primary, unspliced *Emx2* transcripts, via qRT-PCR on intronic sequences. As expected, we found that miR-αDT2S.81 and miR-αDT1S.136 upregulated these levels by about 2 and 4 times, respectively (Fig.4C). Consistently, these miRNAs also increased binding of DNA-dependent RNA polymerase II (RNApolII) to the *Emx2* locus. The $\log_2$enrichment of both TSS and 3' terminus for this enzyme fell - in fact - in the range of 1 to 5 (Fig. 4B, sites α and β), as evaluated by chromatin immunoprecipitation (ChIP).

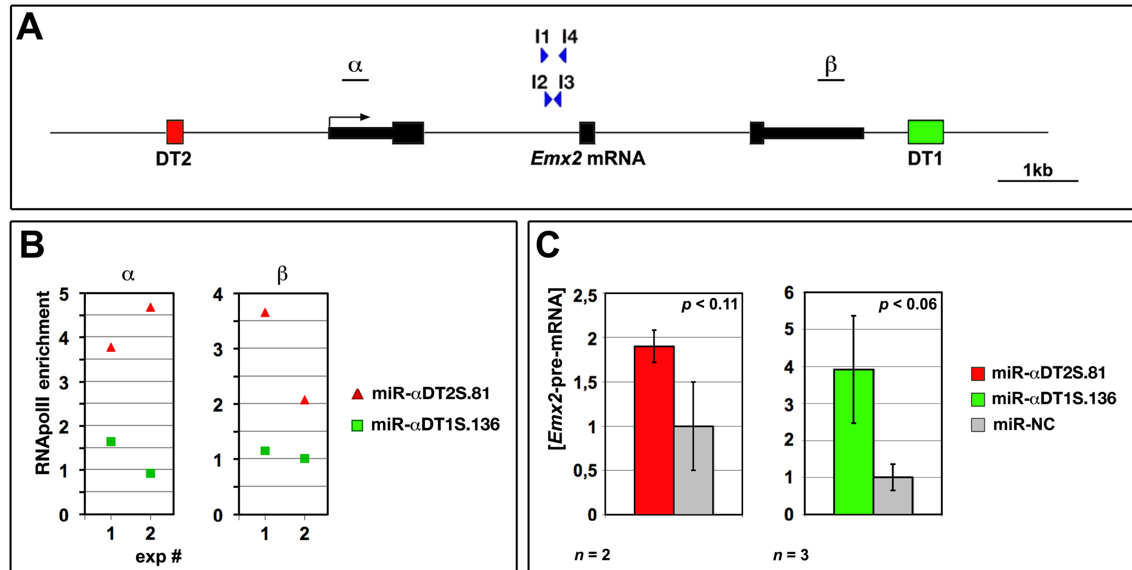


Figure 4. Stimulation of *Emx2* transcription by miR-αDT2S.81 and miR-αDT1S.136. (A) The *Emx2* locus with the a and b amplicons interrogated in chromatin immunoprecipitation-quantitative PCR (ChIP-qPCR) assays (not at scale), as well as the I1-4 primers employed for nested qRT-PCR detection of *Emx2*-pre-mRNA. (B) Enrichment of α and β amplicons for RNA polymerase II (RNApolII) induced by the two artificial miRNAs. The enrichment was calculated as in Materials and Methods; exp# is the experimental session number. (C) *Emx2*-pre-mRNA levels detected upon overexpression of the same miRNAs. All these assays were run on E12.5 cortico-cerebral precursors, infected by miRNA-expressing lentiviruses at m.o.i.= 20, and cultured over 4 days as floating neurospheres. Briefly, both miR-αDT2S.81 and miR-αDT1S.136 promoted RNApolII binding to the *Emx2* locus, apparently resulting in promotion of its transcription.

Such impact of miR-αDT2S.81 and miR-αDT1S.136 on *Emx2* expression levels hardly relies on straight pairing of these miRNAs to genomic targets within the *Emx2* locus (Morgan and Wells, 1968); conversely, it should be mediated by non coding transcripts originating from these targets. To cast light on this issue, we scored the developing neural tube for cumulative DT2 and DT1 transcript levels, by random-primed qRT-PCR. Both elements resulted to be specifically transcribed by structures which also express *Emx2* mRNA, such as pallium and its derivatives, but not by neural regions devoided of it, such as rhombencephalon. Moreover, DT2- and DT1-ncRNAs progressively declined from E11.5 to E18.5, like *Emx2* mRNA (Fig. 5A). Strand-specific qRT-PCR showed coexistence, at E11.5, of sense and antisense transcripts originating from both DT2 and DT1 regions. Besides, 5' and 3' rapid amplification of cDNA ends (RACE) yielded multiple putative TSSs and 3' termini associated to DT tags; a subset of them was mapped (Fig. 5B).

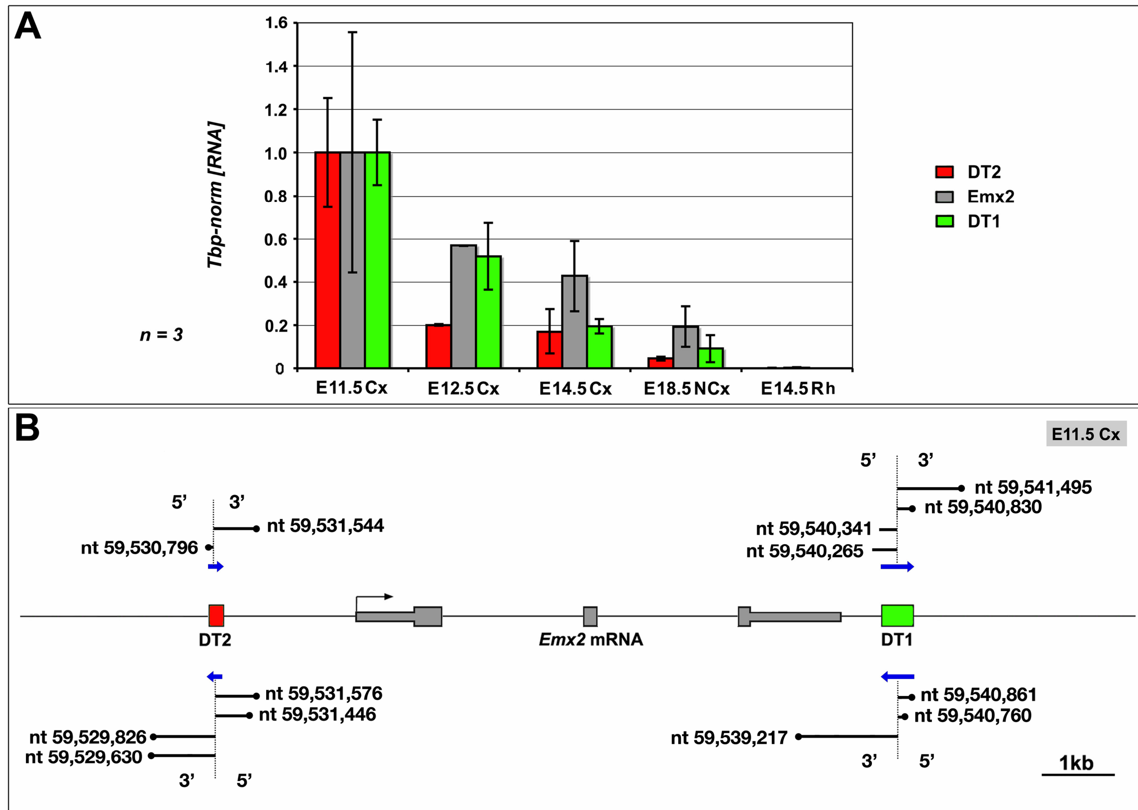
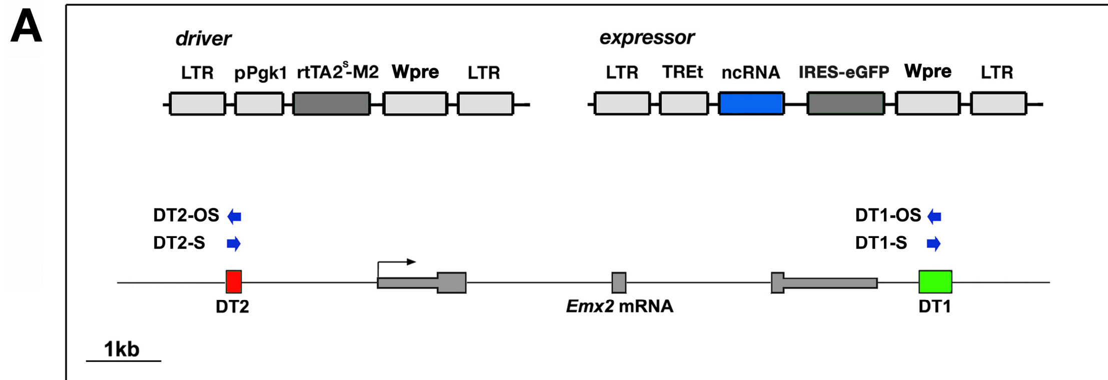


Figure 5. DTi-ncRNA expression in the developing CNS: cumulative time course profiles, quantitation of transcripts orientation, RACE analysis. (A) Expression profiles of DT2, *Emx2* and DT1 transcripts in the developing CNS, as assayed by random-primed qRT-PCR of total RNA (data normalized against *Tbp* mRNA and further normalized against E11.5 Cx). Cx, cortex; NCx, neocortex; Rh, rhombencephalon. (B) DT1 and DT2 transcript tags (oriented blue arrows), as assayed by specific strand-primed qRT-PCR of E11.5 cortico-cerebral total RNA. Associated to each transcript tag, shown are its mapped 5' and 3' RACE extensions, with genomic coordinates of their ends (bulleted are fully mapped extensions).

Similarities among DT2-, DT1-ncRNAs and *Emx2*-mRNA spatio-temporal expression profiles suggest that the former two transcripts may be implicated in natural fine tuning of *Emx2*-mRNA. Consistently with this prediction, overexpression of antisense DT1 and DT2-ncRNA fragments specifically downregulated *Emx2*-mRNA (Fig. 6B).

A



B

ncRNA	ncRNA coordinates	[*Emx2* mRNA]	*p*	*n*
DT1-S	chr.19(+):59,540,370-59,540,769	1.18 ± 0.82	0.20	7
DT1-OS	chr.19(-):59,540,370-59,540,769	0.74 ± 0.61	0.06	7
DT2-S	chr.19(+):59,530,756-59,530,991	0.60 ± 0.28	0.10	4
DT2-OS	chr.19(-):59,530,756-59,530,991	0.49 ± 0.35	0.03	4

(1) referring to assembly July 2007 (NCBI37/mm9)
(2) normalized on control-treated samples

Figure 6. Overexpression of DT1 and DT2 tags in cortico-cerebral precursors. (A) sense and antisense DT1 and DT2-ncRNA fragments were delivered to E12.5 cortico-cerebral precursors, by "LV:LTR-TREt-(---)-IRES-eGFP-Wpre-LTR" derivatives and "LV:LTR-pPgk1-rtTA2^S-M2-Wpre-LTR", each at m.o.i.=12. (B) Infected cells, cultured as floating neurospheres over 3 days, were profiled by qRT-PCR; data were normalized against *Tbp* mRNA and further normalized against negative control samples. Both antisense ncRNA fragments specifically downregulated *Emx2*-mRNA.

Endogenous ncRNAs associated to polypeptide-encoding genes may allow modulation of their transcription by small RNAs via two classes of mechanisms: ncRNAs are *themselves* involved in transcriptional control and small RNAs simply destabilize them; alternatively, ncRNAs act as *landing pads* for small RNAs, un turn conveying transcription factors to chromatine. Remarkably, no changes of DT2 and DT1-ncRNA cumulative levels were found upon lentiviral delivery of miR-αDT2S.81 and miR-αDT1S.136 to E12.5 cortico-cerebral precursors (Table 1). As about 90% of total DT1-ncRNA is sense-oriented (data not shown), this suggests that at least miR-αDT1S.136 might act according to the second mechanism.

miR	ncRNA	[ncRNA]	*n*	*p*
αDT1S.136	DT1	0.92 ± 0.10	2	<0.84
αDT1S.136	DT2	0.93 ± 0.11	2	<0.87
αDT2S.81	DT2	1.17 ± 0.08	2	<0.70
αDT2S.81	DT1	1.13 ± 0.10	2	<0.76

Table 1. Quantification of DT1 and DT2 transcripts upon miR-αDT1S.136 and miR-αDT2S.81 expression. The assays were run on E12.5 embryonic cerebral cortex precursor cells, infected by miR-lentiviral expressors at m.o.i. = 20, cultured as floating neurospheres over 4 days and profiled for DT1 and

DT2 non coding RNAs (ncRNAs) levels, by random primed RT- quantitative PCR. Data were normalized against *Tbp* mRNA and further normalized against negative control miRNA-treated samples. No statistically significant changes of ncRNA levels were detected.

DISCUSSION

Here we report the results of a functional screening aimed at finding out artificial miRNAs modulating mRNA levels of *Emx2*, a TF gene implicated in cortical specification of the dorsal telencephalon and subsequent arealization, lamination and histogenetic progression of the early cortico-cerebral primordium (Bishop et al., 2000; Mallamaci et al., 2000b; Mallamaci et al., 2000c; Heins et al., 2001; Muzio et al., 2005; Galli et al., 2002; Brancaccio et al., 2010). Roughly one half of the interrogated miRNAs, targeted against non coding conserved sequences of the murine *Emx2* locus, were found to be active, the majority of them promoting *Emx2* expression. Properties of two of these miRNAs, miR-αDT2S.81 and miR-αDT1S.136, directed against DT2 and DT1, were investigated more in depth, in primary cultures of cortico-cerebral precursors. They upregulated *Emx2* mRNA, in a context-dependent way. This phenomenon was sufficiently strong to elicit an appreciable expansion of the neural precursor pool. Their activity was apparently mediated by recruitment of RNApolII to the *Emx2* locus and consequent promotion of gene transcription. It also resulted from this analysis that DT2 and DT1 are endogenously transcribed and their RNA products may contribute to natural regulation of *Emx2* mRNA expression.

Promotion of transcription by siRNAs/miRNAs directed against non coding regions of mammalian genes, or "RNA activation" (RNAa), has been already reported in a variety of cases, in cell lines (as for genes Cdh1 and PR) as well as in vivo (as for VEGF) (Li et al., 2006; Janowski et al., 2007; Place et al., 2008; Schwartz et al., 2008; Turunen et al., 2009; Yue et al., 2010). RNAa for specific genes resulted to be restricted to specific tissues and cell lines subject of investigation (Janowski et al., 2005; Janowski et al., 2007; Li et al., 2006; Turunen et al., 2009); moreover, it was deeply sensitive to their functional state (Schwartz et al., 2008). Therefore, looking for miRNAs modulating *cortical Emx2* expression, we carried our assays on dorsal telencephalic precursors, kept as high density floating cultures, under standard Fgf2/Egf, for up to 96 hours. These culture conditions, in fact, preserve precursors' positional identity (Kelly et al., 2009; Onorati et al., 2010), promote their proliferation and recreate the richness of cell-cell interactions which characterizes embryonic periventricular proliferative layers (Brancaccio et al., 2010).

Actually we already reported, in a previous study, one artificial miRNA, miR-aEmx2OS-774, able to upregulate *Emx2* in cortico-cerebral precursors (Spigoni et al., 2010). However this miRNA acts post-transcriptionally, by destabilizing *Emx2OS*-ncRNA, i.e. the endogenous antisense transcript (Spigoni et al., 2010) associated to *Emx2*-mRNA, and so "protecting" to some extent the latter from the Dicer-dependent degradation triggered by the former (Spigoni et al., 2010). On the contrary, bio-active miRNAs reported in the present study neither can target known antisense transcripts associated to *Emx2* mRNA, nor fall within the *Emx2* mRNA sequence (so shielding this mRNA from possible degradation triggered by unknown overlapping antisense transcripts). This suggests that they should act pre- or co-transcriptionally. We confirmed this prediction, by showing that at least two of them, miR-αDT2S.81 and miR-αDT1S.136, ameliorate the recruitment of RNApolII along the *Emx2* locus and increase the concentration of *Emx2*-pre-mRNA. These miRNAs do not apparently destabilize endogenous transcripts stemming from non coding regions. This further suggests that they do not act by suppressing a hypothetic inhibitory role exerted by DT-RNAs, but possibly recruit pro-active factors to chromatin, using these nascent transcripts as docking sites or interacting straightly with DNA.

Transcription of classical enhancers was originally reported in cases of beta-globin and MHC II genes (reveiwed by Szutorisz et al., 2005; Koch et al., 2008) and more recently shown to be a pervasive genome-wide phenomenon (Kim et al., 2010; Ørom et al., 2010a). Endogenous transcripts we mapped at *Emx2*-DT elements are structurally reminiscent of pairs of divergent transcripts, called enhancer RNAs (eRNAs), described by (Kim et al., 2010) and differ from unidirectionally oriented lncRNAs reported by (Ørom et al., 2010a). Similarly to eRNAs and lncRNAs, *Emx2*-DT-RNAs also display a positive correlation with the main mRNA product of their own locus, so suggesting that they may promote its transcription. *Emx2*-mRNA downregulation observed upon overexpression of DT-antisense tags might be accounted for by distinct mechanisms. Such tags might compete with full lenght endogenous *antisense* transcripts, for interaction with factors promoting *Emx2*-mRNA expression (dominant negative effect). Alternatively, they might chelate endogenous *sense* transcripts, so destabilizing them or antagonizing their natural transcription-promoting activity. Unfortunately, miRNAs against sense and antisense DT1 transcripts evaluated in this study do not replicate the effects of the exogenous DT1-antisense fragment, so inhibiting us from distinguishing

between these two hypotheses. Screening additional miRNAs able to destabilize endogenous DT1, sense and antisense, transcripts will help solving this issue.

Finally, it has been reported that structured *Emx2* overexpression in cortico-cerebral precursors allows for promotion of stem cells self-renewal, inhibition of their gliogenic commitment, protection of neuronal progenitors and stimulation of their differentiation, all activities of paramount interest for purposes of gene-promoted brain repair (Brancaccio et al., 2010). The possibility to stimulate overexpression of *endogenous Emx2* by small RNAs, so preventing drawbacks arising from the introduction of *exogenous* copies of the whole gene (Nienhuis et al., 2006), makes therapeutic exploitation of this gene manipulation a more feasible goal.

FUTURE PERSPECTIVES

In the future, we intend to:

(1) clarify molecular mechanisms by which artificial miRNAs against Emx2 enhancers stimulate transcription

We will assay binding of artificial biotinylated miRNAs to endogenous ncRNAs originating from these enhancers as well as their interaction with genomic DNA, by biotin/streptavidin-mediated RNA or DNA precipitation, followed by (RT)PCR quantification

Moreover, we will study *Emx2* chromatin dynamics upon delivery of artificial miRNAs, by scoring methylation levels of H3K4, H3K9, H3K27, H4K20 as well as acetylation levels of H3 and H4, via standard ChIP-qPCR

(2) cast light on regulatory functions exerted by natural non coding transcripts stemming from Emx2 enhancers

This will be done by a loss-of-function approach, by selecting new artificial miRNAs, able to specifically destabilize endogenous, sense and antisense, transcripts.

Additional gain-of-function investigations will be performed too, via overexpression of full-lenght copies of these transcripts

(3) assess the possibility to exploit artificial miRNAs to modulate neural precursors behaviour in vivo

The best two performing miRNAs described in this study will be delivered to endogenous cortical precursor cells in vivo, by adeno-associated viral vectors, and their ability to modulate proliferation/differentiation kinetics will be assayed, by appropriate BrdU-immunoprofiling

REFERENCES

Adelman, K., Marr, M. T., Werner, J., Saunders, A., Ni, Z., Andrulis, E. D., and Lis, J. T. (2005). Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS. Mol. Cell *17*, 103-112.

Agami, R., and Bernards, R. (2000). Distinct initiation and maintenance mechanisms cooperate to induce G1 cell cycle arrest in response to DNA damage. Cell *102*, 55-66.

Aigner, A. (2006). Gene silencing through RNA interference (RNAi) in vivo: strategies based on the direct application of siRNAs. J. Biotechnol *124*, 12-25.

Alcántara, S., Ruiz, M., D'Arcangelo, G., Ezan, F., de Lecea, L., Curran, T., Sotelo, C., and Soriano, E. (1998). Regional and cellular patterns of reelin mRNA expression in the forebrain of the developing and adult mouse. J. Neurosci *18*, 7779-7799.

Altmann, C. R., and Brivanlou, A. H. (2001). Neural patterning in the vertebrate embryo. Int. Rev. Cytol *203*, 447-482.

Auboeuf, D., Dowhan, D. H., Dutertre, M., Martin, N., Berget, S. M., and O'Malley, B. W. (2005). A subset of nuclear receptor coregulators act as coupling proteins during synthesis and maturation of RNA transcripts. Mol. Cell. Biol *25*, 5307-5316.

Bachellerie, J. (2002). The expanding snoRNA world. Biochimie *84*, 775-790.

Bayatti, N., Sarma, S., Shaw, C., Eyre, J. A., Vouyiouklis, D. A., Lindsay, S., and Clowry, G. J. (2008). Progressive loss of PAX6, TBR2, NEUROD and TBR1 mRNA gradients correlates with translocation of EMX2 to the cortical plate during human cortical development. Eur. J. Neurosci *28*, 1449-1456.

Beltran, M., Puig, I., Peña, C., García, J. M., Alvarez, A. B., Peña, R., Bonilla, F., and de Herreros, A. G. (2008a). A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. Genes Dev *22*, 756-769.

Beltran, M., Puig, I., Peña, C., García, J. M., Alvarez, A. B., Peña, R., Bonilla, F., and de Herreros, A. G. (2008b). A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. Genes Dev *22*, 756-769.

Berezikov, E., Chung, W., Willis, J., Cuppen, E., and Lai, E. C. (2007). Mammalian mirtron genes. Mol. Cell *28*, 328-336.

Berretta, J., and Morillon, A. (2009). Pervasive transcription constitutes a new level of eukaryotic genome regulation. EMBO Rep *10*, 973-982.

Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I., and Filipowicz, W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. Cell *125*, 1111-1124.

Bi, W., Huang, W., Whitworth, D. J., Deng, J. M., Zhang, Z., Behringer, R. R., and de Crombrugghe, B. (2001). Haploinsufficiency of Sox9 results in defective cartilage primordia and premature skeletal mineralization. Proc. Natl. Acad. Sci. U.S.A *98*, 6698-6703.

Bielle, F., Griveau, A., Narboux-Nême, N., Vigneau, S., Sigrist, M., Arber, S., Wassef, M., and Pierani, A. (2005). Multiple origins of Cajal-Retzius cells at the borders of the developing pallium. Nat. Neurosci *8*, 1002-1012.

Bishop, K. M., Goudreau, G., and O'Leary, D. D. (2000). Regulation of area identity in the mammalian neocortex by Emx2 and Pax6. Science *288*, 344-349.

Boncinelli, E. (1997). Homeobox genes and disease. Curr. Opin. Genet. Dev *7*, 331-337.

Borchert, G. M., Lanier, W., and Davidson, B. L. (2006). RNA polymerase III transcribes human microRNAs. Nat Struct Mol Biol *13*, 1097-1101.

Boumil, R. M., and Lee, J. T. (2001). Forty years of decoding the silence in X-chromosome inactivation. Hum. Mol. Genet *10*, 2225-2232.

BRACHET, J., and CHANTRENNE, H. (1956). The function of the nucleus in the synthesis of cytoplasmic proteins. Cold Spring Harb. Symp. Quant. Biol *21*, 329-337.

Brancaccio, M., Pivetta, C., Granzotto, M., Filippis, C., and Mallamaci, A. (2010). Emx2 and Foxg1 inhibit gliogenesis and promote neuronogenesis. Stem Cells *28*, 1206-1218.

Britten, R. J., and Davidson, E. H. (1969). Gene regulation for higher cells: a theory. Science *165*, 349-357.

Brox, A., Puelles, L., Ferreiro, B., and Medina, L. (2004). Expression of the genes Emx1, Tbr1, and Eomes (Tbr2) in the telencephalon of Xenopus laevis confirms the existence of a ventral pallial division in all tetrapods. J. Comp. Neurol *474*, 562-577.

Capellini, T. D., Vaccari, G., Ferretti, E., Fantini, S., He, M., Pellegrini, M., Quintana, L., Di Giacomo, G., Sharpe, J., Selleri, L., et al. (2010). Scapula development is governed by genetic interactions of Pbx1 with its family members and with Emx2 via their cooperative control of Alx1. Development *137*, 2559-2569.

Carninci, P. (2008). Non-coding RNA transcription: turning on neighbours. Nat. Cell Biol *10*, 1023-1024.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C. A. M., Taylor, M. S., Engström, P. G., Frith, M. C., et al. (2006). Genome-wide analysis of mammalian promoter architecture and evolution. Nat. Genet *38*, 626-635.

Carthew, R. W., and Sontheimer, E. J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642-655.

Cattanach, B. M., and Isaacson, J. H. (1967). Controlling elements in the mouse X chromosome. Genetics *57*, 331-346.

Cecchi, C., and Boncinelli, E. (2000). Emx homeogenes and mouse brain development. Trends Neurosci *23*, 347-352.

Cecchi, C. (2002). Emx2: a gene responsible for cortical development, regionalization and area specification. Gene *291*, 1-9.

Chang, C. P., Brocchieri, L., Shen, W. F., Largman, C., and Cleary, M. L. (1996). Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. Mol. Cell. Biol *16*, 1734-1745.

Chang, C. P., Shen, W. F., Rozenfeld, S., Lawrence, H. J., Largman, C., and Cleary, M. L. (1995). Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. Genes Dev *9*, 663-674.

Chen, Z., Place, R. F., Jia, Z., Pookot, D., Dahiya, R., and Li, L. (2008). Antitumor effect of dsRNA-induced p21WAF1/CIP1 gene activation in human bladder cancer cells. Molecular Cancer Therapeutics *7*, 698-703.

Choudhuri, S. (2010). Small noncoding RNAs: biogenesis, function, and emerging significance in toxicology. J. Biochem. Mol. Toxicol *24*, 195-216.

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

Chu, Y., Yue, X., Younger, S. T., Janowski, B. A., and Corey, D. R. (2010). Involvement of argonaute proteins in gene silencing and activation by RNAs complementary to a non-coding transcript at the progesterone receptor promoter. Nucleic Acids Res *38*, 7736-7748.

Cohen, S. M., and Jürgens, G. (1990). Mediation of Drosophila head development by gap-like segmentation genes. Nature *346*, 482-485.

Colley, S. M., Iyer, K. R., and Leedman, P. J. (2008). The RNA coregulator SRA, its binding proteins and nuclear receptor signaling activity. IUBMB Life *60*, 159-164.

CRICK, F. H. (1958). On protein synthesis. Symp. Soc. Exp. Biol *12*, 138-163.

Crossley, P. H., and Martin, G. R. (1995). The mouse Fgf8 gene encodes a family of polypeptides and is expressed in regions that direct outgrowth and patterning in the developing embryo. Development *121*, 439-451.

Crowe, M. L., Wang, X., and Rothnagel, J. A. (2006). Evidence for conservation and selection of upstream open reading frames suggests probable encoding of bioactive peptides. BMC Genomics *7*, 16.

Dalton, D., Chadwick, R., and McGinnis, W. (1989). Expression and embryonic function of empty spiracles: a Drosophila homeo box gene with two patterning functions on the anterior-posterior axis of the embryo. Genes Dev *3*, 1940-1956.

D'Arcangelo, G., Miao, G. G., Chen, S. C., Soares, H. D., Morgan, J. I., and Curran, T. (1995). A protein related to extracellular matrix proteins deleted in the mouse mutant reeler. Nature *374*, 719-723.

Davis, C. A., and Ares, M. (2006). Accumulation of unstable promoter-associated

transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U.S.A *103*, 3262-3267.

Di Stefano, L., Ji, J., Moon, N., Herr, A., and Dyson, N. (2007). Mutation of Drosophila Lsd1 disrupts H3-K4 methylation, resulting in tissue-specific defects during development. Curr. Biol *17*, 808-812.

Eklund, E. A., Jalava, A., and Kakar, R. (2000). Tyrosine phosphorylation of HoxA10 decreases DNA binding and transcriptional repression during interferon gamma - induced differentiation of myeloid leukemia cell lines. J. Biol. Chem *275*, 20117-20126.

Englert, M., Felis, M., Junker, V., and Beier, H. (2004). Novel upstream and intragenic control elements for the RNA polymerase III-dependent transcription of human 7SL RNA genes. Biochimie *86*, 867-874.

Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007). P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Mol. Cell. Biol *27*, 3970-3981.

Farazi, T. A., Juranek, S. A., and Tuschl, T. (2008). The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. Development *135*, 1201-1214.

Feng, J., Bi, C., Clark, B. S., Mady, R., Shah, P., and Kohtz, J. D. (2006). The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. Genes Dev *20*, 1470-1484.

Feng, J., White, B., Tyurina, O. V., Guner, B., Larson, T., Lee, H. Y., Karlstrom, R. O., and Kohtz, J. D. (2004). Synergistic and antagonistic roles of the Sonic hedgehog N- and C-terminal lipids. Development *131*, 4357-4370.

Fuccillo, M., Rutlin, M., and Fishell, G. (2006). Removal of Pax6 partially rescues the loss of ventral structures in Shh null mice. Cereb. Cortex *16 Suppl 1*, i96-102.

Fukuchi-Shimogori, T., and Grove, E. A. (2001). Neocortex patterning by the secreted signaling molecule FGF8. Science *294*, 1071-1074.

Fukuchi-Shimogori, T., and Grove, E. A. (2003). Emx2 patterns the neocortex by regulating FGF positional signaling. Nat. Neurosci *6*, 825-831.

Galindo, M. I., Pueyo, J. I., Fouix, S., Bishop, S. A., and Couso, J. P. (2007). Peptides encoded by short ORFs control development and define a new eukaryotic gene family. PLoS Biol *5*, e106.

Galli, R., Fiocco, R., De Filippis, L., Muzio, L., Gritti, A., Mercurio, S., Broccoli, V., Pellegrini, M., Mallamaci, A., and Vescovi, A. L. (2002). Emx2 regulates the proliferation of stem cells of the adult mammalian central nervous system. Development *129*, 1633-1644.

Gangemi, R. M. R., Daga, A., Muzio, L., Marubbi, D., Cocozza, S., Perera, M., Verardo, S., Bordo, D., Griffero, F., Capra, M. C., et al. (2006). Effects of Emx2 inactivation on the gene expression profile of neural precursors. Eur. J. Neurosci

23, 325-334.

Gingeras, T. R. (2009). Implications of chimaeric non-co-linear transcripts. Nature *461*, 206-211.

Gonzalez, S., Pisano, D. G., and Serrano, M. (2008). Mechanistic principles of chromatin remodeling guided by siRNAs and miRNAs. Cell Cycle *7*, 2601-2608.

Gorski, J. A., Talley, T., Qiu, M., Puelles, L., Rubenstein, J. L. R., and Jones, K. R. (2002). Cortical excitatory neurons and glia, but not GABAergic neurons, are produced in the Emx1-expressing lineage. J. Neurosci *22*, 6309-6314.

Gulisano, M., Broccoli, V., Pardini, C., and Boncinelli, E. (1996). Emx1 and Emx2 show different patterns of expression during proliferation and differentiation of the developing cerebral cortex in the mouse. Eur. J. Neurosci *8*, 1037-1050.

Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., Tsai, M., Hung, T., Argani, P., Rinn, J. L., et al. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature *464*, 1071-1076.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M. F., Feldser, D., Huarte, M., Zuk, O., Carey, B. W., Cassady, J. P., et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223-227.

Hamasaki, T., Leingärtner, A., Ringstedt, T., and O'Leary, D. D. M. (2004). EMX2 regulates sizes and positioning of the primary sensory and motor areas in neocortex by direct specification of cortical progenitors. Neuron *43*, 359-372.

Hastings, M. L., Ingle, H. A., Lazar, M. A., and Munroe, S. H. (2000). Post-transcriptional regulation of thyroid hormone receptor expression by cis-acting sequences and a naturally occurring antisense RNA. J. Biol. Chem *275*, 11507-11513.

Heins, N., Cremisi, F., Malatesta, P., Gangemi, R. M., Corte, G., Price, J., Goudreau, G., Gruss, P., and Götz, M. (2001). Emx2 promotes symmetric cell divisions and a multipotential fate in precursors from the cerebral cortex. Mol. Cell. Neurosci *18*, 485-502.

Iler, N., Rowitch, D. H., Echelard, Y., McMahon, A. P., and Abate-Shen, C. (1995). A single homeodomain binding site restricts spatial expression of Wnt-1 in the developing brain. Mech. Dev *53*, 87-96.

JACOB, F., and MONOD, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol *3*, 318-356.

Janowski, B. A., Huffman, K. E., Schwartz, J. C., Ram, R., Hardy, D., Shames, D. S., Minna, J. D., and Corey, D. R. (2005). Inhibiting gene expression at transcription start sites in chromosomal DNA with antigene RNAs. Nat. Chem. Biol *1*, 216-222.

Janowski, B. A., Huffman, K. E., Schwartz, J. C., Ram, R., Nordsell, R., Shames, D. S., Minna, J. D., and Corey, D. R. (2006). Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. Nat. Struct. Mol. Biol *13*, 787-792.

Janowski, B. A., Younger, S. T., Hardy, D. B., Ram, R., Huffman, K. E., and Corey, D. R. (2007). Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. Nat. Chem. Biol *3*, 166-173.

Jurica, M. S., and Moore, M. J. (2003). Pre-mRNA splicing: awash in a sea of proteins. Mol. Cell *12*, 5-14.

Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., Stadler, P. F., Hertel, J., Hackermuller, J., Hofacker, I. L., et al. (2007). RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. Science *316*, 1484-1488.

Kawasaki, H., and Taira, K. (2004). Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. Nature *431*, 211-217.

Kay, G. F., Barton, S. C., Surani, M. A., and Rastan, S. (1994). Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development. Cell *77*, 639-650.

Kelly, T. K., Karsten, S. L., Geschwind, D. H., and Kornblum, H. I. (2009). Cell Lineage and Regional Identity of Cultured Spinal Cord Neural Stem Cells and Comparison to Brain-Derived Neural Stem Cells. PLoS ONE *4*, e4213.

Kim, D. H., Villeneuve, L. M., Morris, K. V., and Rossi, J. J. (2006). Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. Nat. Struct. Mol. Biol *13*, 793-797.

Kim, T., Hemberg, M., Gray, J. M., Costa, A. M., Bear, D. M., Wu, J., Harmin, D. A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*, 182-187.

Knoepfler, P. S., and Kamps, M. P. (1995). The pentapeptide motif of Hox proteins is required for cooperative DNA binding with Pbx1, physically contacts Pbx1, and enhances DNA binding by Pbx1. Mol. Cell. Biol *15*, 5811-5819.

Koch, F., Jourquin, F., Ferrier, P., and Andrau, J. (2008). Genome-wide RNA polymerase II: not genes only! Trends Biochem. Sci *33*, 265-273.

Kondo, T., Hashimoto, Y., Kato, K., Inagaki, S., Hayashi, S., and Kageyama, Y. (2007). Small peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA. Nat. Cell Biol *9*, 660-665.

Koziol, M. J., and Rinn, J. L. (2010). RNA traffic control of chromatin complexes. Curr. Opin. Genet. Dev *20*, 142-148.

Krol, J., Loedige, I., and Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet. Available at: http://www.nature.com/doifinder/10.1038/nrg2843.

Kuwabara, T., Hsieh, J., Nakashima, K., Taira, K., and Gage, F. H. (2004). A small modulatory dsRNA specifies the fate of adult neural stem cells. Cell *116*, 779-793.

Lanz, R. B., McKenna, N. J., Onate, S. A., Albrecht, U., Wong, J., Tsai, S. Y., Tsai, M. J.,

and O'Malley, B. W. (1999). A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. Cell *97*, 17-27.

Lanz, R. B., Razani, B., Goldberg, A. D., and O'Malley, B. W. (2002). Distinct RNA motifs are important for coactivation of steroid hormone receptors by steroid receptor RNA activator (SRA). Proc. Natl. Acad. Sci. U.S.A *99*, 16081-16086.

Lee, J. T. (2002). Homozygous Tsix mutant mice reveal a sex-ratio distortion and revert to random X-inactivation. Nat. Genet *32*, 195-200.

Lee, J. T. (2005). Regulation of X-chromosome counting by Tsix and Xite sequences. Science *309*, 768-771.

Leupin, O., Attanasio, C., Marguerat, S., Tapernoux, M., Antonarakis, S. E., and Conrad, B. (2005). Transcriptional activation by bidirectional RNA polymerase II elongation over a silent promoter. EMBO Rep *6*, 956-960.

Lewin, B. (1999). Il gene VI (Bologna: Zanichelli).

Leygue, E. (2007). Steroid receptor RNA activator (SRA1): unusual bifaceted gene products with suspected relevance to breast cancer. Nucl Recept Signal *5*, e006.

Li, L., Okino, S. T., Zhao, H., Pookot, D., Place, R. F., Urakami, S., Enokida, H., and Dahiya, R. (2006). Small dsRNAs induce transcriptional activation in human cells. Proceedings of the National Academy of Sciences *103*, 17337-17342.

Ligon, K. L., Echelard, Y., Assimacopoulos, S., Danielian, P. S., Kaing, S., Grove, E. A., McMahon, A. P., and Rowitch, D. H. (2003). Loss of Emx2 function leads to ectopic expression of Wnt1 in the developing telencephalon and cortical dysplasia. Development *130*, 2275-2287.

Lin, H., and Spradling, A. C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development *124*, 2463-2476.

López-Bendito, G., Chan, C., Mallamaci, A., Parnavelas, J., and Molnár, Z. (2002). Role of Emx2 in the development of the reciprocal connectivity between cortex and thalamus. J. Comp. Neurol *451*, 153-169.

Lu, Q., and Kamps, M. P. (1996). Structural determinants within Pbx1 that mediate cooperative DNA binding with pentapeptide-containing Hox proteins: proposal for a model of a Pbx1-Hox-DNA complex. Mol. Cell. Biol *16*, 1632-1640.

Lunyak, V. V., Burgess, R., Prefontaine, G. G., Nelson, C., Sze, S., Chenoweth, J., Schwartz, P., Pevzner, P. A., Glass, C., Mandel, G., et al. (2002). Corepressor-dependent silencing of chromosomal regions encoding neuronal genes. Science *298*, 1747-1752.

LYON, M. F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature *190*, 372-373.

Makarov, E. M., Makarova, O. V., Urlaub, H., Gentzel, M., Will, C. L., Wilm, M., and Lührmann, R. (2002). Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. Science *298*, 2205-2208.

Mallamaci, A., Iannone, R., Briata, P., Pintonello, L., Mercurio, S., Boncinelli, E., and Corte, G. (1998). EMX2 protein in the developing mouse brain and olfactory area. Mech. Dev *77*, 165-172.

Mallamaci, A., Mercurio, S., Muzio, L., Cecchi, C., Pardini, C. L., Gruss, P., and Boncinelli, E. (2000a). The lack of Emx2 causes impairment of Reelin signaling and defects of neuronal migration in the developing cerebral cortex. J. Neurosci *20*, 1109-1118.

Mallamaci, A., Mercurio, S., Muzio, L., Cecchi, C., Pardini, C. L., Gruss, P., and Boncinelli, E. (2000b). The lack of Emx2 causes impairment of Reelin signaling and defects of neuronal migration in the developing cerebral cortex. J. Neurosci *20*, 1109-1118.

Mallamaci, A., Muzio, L., Chan, C. H., Parnavelas, J., and Boncinelli, E. (2000c). Area identity shifts in the early cerebral cortex of Emx2-/- mutant mice. Nat. Neurosci *3*, 679-686.

Mallamaci, A., and Stoykova, A. (2006a). Gene networks controlling early cerebral cortex arealization. Eur. J. Neurosci *23*, 847-856.

Mallamaci, A., and Stoykova, A. (2006b). Gene networks controlling early cerebral cortex arealization. Eur. J. Neurosci *23*, 847-856.

Marín, O., and Rubenstein, J. L. (2001). A long, remarkable journey: tangential migration in the telencephalon. Nat. Rev. Neurosci *2*, 780-790.

Martianov, I., Ramadass, A., Serra Barros, A., Chow, N., and Akoulitchev, A. (2007). Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. Nature *445*, 666-670.

Mazo, A., Hodgson, J. W., Petruk, S., Sedkov, Y., and Brock, H. W. (2007). Transcriptional interference: an unexpected layer of complexity in gene regulation. J. Cell. Sci *120*, 2755-2761.

Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A., and Matzke, A. J. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. EMBO J *19*, 5194-5201.

Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T., Koche, R. P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Miyamoto, N., Yoshida, M., Kuratani, S., Matsuo, I., and Aizawa, S. (1997). Defects of urogenital development in mice lacking Emx2. Development *124*, 1653-1664.

Morey, C., Navarro, P., Debrand, E., Avner, P., Rougeulle, C., and Clerc, P. (2004). The region 3' to Xist mediates X chromosome counting and H3 Lys-4 dimethylation within the Xist gene. EMBO J *23*, 594-604.

Morgan, A. R., and Wells, R. D. (1968). Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. J. Mol. Biol *37*, 63-80.

Morris, K. V. (2009). RNA-directed transcriptional gene silencing and activation in human cells. Oligonucleotides *19*, 299-306.

Morris, K. V., Chan, S. W., Jacobsen, S. E., and Looney, D. J. (2004). Small interfering RNA-induced transcriptional gene silencing in human cells. Science *305*, 1289-1292.

Morris, K. V., Santoso, S., Turner, A., Pastori, C., and Hawkins, P. G. (2008). Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. PLoS Genet *4*, e1000258.

Muzio, L., Soria, J. M., Pannese, M., Piccolo, S., and Mallamaci, A. (2005). A mutually stimulating loop involving emx2 and canonical wnt signalling specifically promotes expansion of occipital cortex and hippocampus. Cereb. Cortex *15*, 2021-2028.

Muzio, L., DiBenedetto, B., Stoykova, A., Boncinelli, E., Gruss, P., and Mallamaci, A. (2002). Conversion of cerebral cortex into basal ganglia in Emx2(-/-) Pax6(Sey/Sey) double-mutant mice. Nat. Neurosci *5*, 737-745.

Nagano, T., Mitchell, J. A., Sanz, L. A., Pauler, F. M., Ferguson-Smith, A. C., Feil, R., and Fraser, P. (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science *322*, 1717-1720.

Nédélec, S., Foucher, I., Brunet, I., Bouillot, C., Prochiantz, A., and Trembleau, A. (2004). Emx2 homeodomain transcription factor interacts with eukaryotic translation initiation factor 4E (eIF4E) in the axons of olfactory sensory neurons. Proc. Natl. Acad. Sci. U.S.A *101*, 10815-10820.

Nienhuis, A. W., Dunbar, C. E., and Sorrentino, B. P. (2006). Genotoxicity of retroviral integration in hematopoietic cells. Mol. Ther *13*, 1031-1049.

Niessing, D., Blanke, S., and Jäckle, H. (2002). Bicoid associates with the 5'-cap-bound complex of caudal mRNA and represses translation. Genes Dev *16*, 2576-2582.

Okamura, K., Balla, S., Martin, R., Liu, N., and Lai, E. C. (2008). Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. Nat. Struct. Mol. Biol *15*, 581-590.

Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M., and Lai, E. C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. Cell *130*, 89-100.

O'Leary, D. D. M., Chou, S., and Sahara, S. (2007). Area patterning of the mammalian cortex. Neuron *56*, 252-269.

Onorati, M., Binetti, M., Conti, L., Camnasio, S., Calabrese, G., Albieri, I., Febo, F., Toselli, M., Biella, G., Martynoga, B., et al. (2010). Preservation of positional identity in fetus-derived neural stem (NS) cells from different mouse central nervous system compartments. Cell. Mol. Life Sci. Available at: http://www.springerlink.com/index/10.1007/s00018-010-0548-7.

Ørom, U. A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F.,

Zytnicki, M., Notredame, C., and Huang, Q. (2010a). Long Noncoding RNAs with Enhancer-like Function in Human Cells. Cell *143*, 46-58.

Ørom, U. A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q., et al. (2010b). Long noncoding RNAs with enhancer-like function in human cells. Cell *143*, 46-58.

Paek, H., Gutin, G., and Hébert, J. M. (2009). FGF signaling is strictly required to maintain early telencephalic precursor cell survival. Development *136*, 2457-2465.

Pal-Bhadra, M., Leibovitch, B. A., Gandhi, S. G., Rao, M., Bhadra, U., Birchler, J. A., and Elgin, S. C. R. (2004). Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science *303*, 669-672.

Panganiban, G., and Rubenstein, J. L. R. (2002). Developmental functions of the Distal-less/Dlx homeobox genes. Development *129*, 4371-4386.

Pannese, M., Lupo, G., Kablar, B., Boncinelli, E., Barsacchi, G., and Vignali, R. (1998). The Xenopus Emx genes identify presumptive dorsal telencephalon and are induced by head organizer signals. Mech. Dev *73*, 73-83.

Park, C. W., Chen, Z., Kren, B. T., and Steer, C. J. (2004). Double-stranded siRNA targeted to the huntingtin gene does not induce DNA methylation. Biochem. Biophys. Res. Commun *323*, 275-280.

Payer, B., and Lee, J. T. (2008). X chromosome dosage compensation: how mammals keep the balance. Annu. Rev. Genet *42*, 733-772.

Pellegrini, M., Mansouri, A., Simeone, A., Boncinelli, E., and Gruss, P. (1996). Dentate gyrus formation requires Emx2. Development *122*, 3893-3898.

Pellegrini, M., Pantano, S., Fumi, M. P., Lucchini, F., and Forabosco, A. (2001). Agenesis of the scapula in Emx2 homozygous mutants. Dev. Biol *232*, 149-156.

Pellegrini, M., Pantano, S., Lucchini, F., Fumi, M., and Forabosco, A. (1997). Emx2 developmental expression in the primordia of the reproductive and excretory systems. Anat. Embryol *196*, 427-433.

Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature *379*, 131-137.

Place, R. F., Li, L., Pookot, D., Noonan, E. J., and Dahiya, R. (2008). MicroRNA-373 induces expression of genes with complementary promoter sequences. Proceedings of the National Academy of Sciences *105*, 1608-1613.

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J., and Pandolfi, P. P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature *465*, 1033-1038.

Prather, D. M., Larschan, E., and Winston, F. (2005). Evidence that the elongation factor TFIIS plays a role in transcription initiation at GAL1 in Saccharomyces cerevisiae. Mol. Cell. Biol *25*, 2650-2659.

Pröls, F., Ehehalt, F., Rodriguez-Niedenführ, M., He, L., Huang, R., and Christ, B. (2004). The role of Emx2 during scapula formation. Dev. Biol *275*, 315-324.

Pueyo, J. I., and Couso, J. P. (2008). The 11-aminoacid long Tarsal-less peptides trigger a cell signal in Drosophila leg development. Dev. Biol *324*, 192-201.

Puigserver, P., and Spiegelman, B. M. (2003). Peroxisome proliferator-activated receptor-gamma coactivator 1 alpha (PGC-1 alpha): transcriptional coactivator and metabolic regulator. Endocr. Rev *24*, 78-90.

Qi, L., Li, X., Zhang, S., and An, D. (2006). Genetic regulation by non-coding RNAs. SCI CHINA SER C *49*, 201-217.

Rastan, S., and Robertson, E. J. (1985). X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. J Embryol Exp Morphol *90*, 379-388.

Rhodes, C. R., Parkinson, N., Tsai, H., Brooker, D., Mansell, S., Spurr, N., Hunter, A. J., Steel, K. P., and Brown, S. D. M. (2003). The homeobox gene Emx2 underlies middle ear and inner ear defects in the deaf mouse mutant pardon. J. Neurocytol *32*, 1143-1154.

Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E., et al. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.

Ruby, J. G., Jan, C. H., and Bartel, D. P. (2007). Intronic microRNA precursors that bypass Drosha processing. Nature *448*, 83-86.

Sahara, S., Kawakami, Y., Izpisua Belmonte, J. C., and O'Leary, D. D. M. (2007). Sp8 exhibits reciprocal induction with Fgf8 but has an opposing effect on anterior-posterior cortical area patterning. Neural Dev *2*, 10.

Sasai, Y., and De Robertis, E. M. (1997). Ectodermal patterning in vertebrate embryos. Dev. Biol *182*, 5-20.

Schwartz, J. C., Younger, S. T., Nguyen, N., Hardy, D. B., Monia, B. P., Corey, D. R., and Janowski, B. A. (2008). Antisense transcripts are targets for activating small RNAs. Nat Struct Mol Biol *15*, 842-848.

Seidl, C. I. M., Stricker, S. H., and Barlow, D. P. (2006). The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. EMBO J *25*, 3565-3575.

Shen, W. F., Rozenfeld, S., Lawrence, H. J., and Largman, C. (1997). The Abd-B-like Hox homeodomain proteins can be subdivided by the ability to form complexes with Pbx1a on a novel DNA target. J. Biol. Chem *272*, 8198-8206.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-953.

Shimamura, K., and Rubenstein, J. L. (1997). Inductive interactions direct early

regionalization of the mouse forebrain. Development *124*, 2709-2718.

Shin, J., Fitzpatrick, G. V., and Higgins, M. J. (2008). Two distinct mechanisms of silencing by the KvDMR1 imprinting control region. EMBO J *27*, 168-178.

Sijen, T., Vijn, I., Rebocho, A., van Blokland, R., Roelofs, D., Mol, J. N., and Kooter, J. M. (2001). Transcriptional and posttranscriptional gene silencing are mechanistically related. Curr. Biol *11*, 436-440.

Simeone, A., Gulisano, M., Acampora, D., Stornaiuolo, A., Rambaldi, M., and Boncinelli, E. (1992a). Two vertebrate homeobox genes related to the Drosophila empty spiracles gene are expressed in the embryonic cerebral cortex. EMBO J *11*, 2541-2550.

Simeone, A., Gulisano, M., Acampora, D., Stornaiuolo, A., Rambaldi, M., and Boncinelli, E. (1992b). Two vertebrate homeobox genes related to the Drosophila empty spiracles gene are expressed in the embryonic cerebral cortex. EMBO J *11*, 2541-2550.

Sleutels, F., and Barlow, D. P. (2002). The origins of genomic imprinting in mammals. Adv. Genet *46*, 119-163.

Small, E. C., Leggett, S. R., Winans, A. A., and Staley, J. P. (2006). The EF-G-like GTPase Snu114p Regulates Spliceosome Dynamics Mediated by Brr2p, a DExD/H Box ATPase. Molecular Cell *23*, 389-399.

Smilinich, N. J., Day, C. D., Fitzpatrick, G. V., Caldwell, G. M., Lossie, A. C., Cooper, P. R., Smallwood, A. C., Joyce, J. A., Schofield, P. N., Reik, W., et al. (1999). A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. Proc. Natl. Acad. Sci. U.S.A *96*, 8064-8069.

Smith, D. J., Query, C. C., and Konarska, M. M. (2007). trans-Splicing to Spliceosomal U2 snRNA Suggests Disruption of Branch Site-U2 Pairing during Pre-mRNA Splicing. Molecular Cell *26*, 883-890.

Spigoni, G., Gedressi, C., and Mallamaci, A. (2010). Regulation of Emx2 expression by antisense transcripts in murine cortico-cerebral precursors. PLoS ONE *5*, e8658.

Staley, J. P., and Guthrie, C. (1998). Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell *92*, 315-326.

Suda, Y., Kokura, K., Kimura, J., Kajikawa, E., Inoue, F., and Aizawa, S. (2010). The same enhancer regulates the earliest Emx2 expression in caudal forebrain primordium, subsequent expression in dorsal telencephalon and later expression in the cortical ventricular zone. Development *137*, 2939-2949.

Suzuki, K., and Kelleher, A. D. (2009). Transcriptional regulation by promoter targeted RNAs. Curr Top Med Chem *9*, 1079-1087.

Szutorisz, H., Dillon, N., and Tora, L. (2005). The role of enhancers as centres for general transcription factor recruitment. Trends Biochem. Sci *30*, 593-599.

Taft, R. J., Glazov, E. A., Cloonan, N., Simons, C., Stephen, S., Faulkner, G. J.,

Lassmann, T., Forrest, A. R. R., Grimmond, S. M., Schroder, K., et al. (2009a). Tiny RNAs associated with transcription start sites in animals. Nat Genet *41*, 572-578.

Taft, R. J., Kaplan, C. D., Simons, C., and Mattick, J. S. (2009b). Evolution, biogenesis and function of promoter-associated RNAs. Cell Cycle *8*, 2332-2338.

Takiguchi-Hayashi, K., Sekiguchi, M., Ashigaki, S., Takamatsu, M., Hasegawa, H., Suzuki-Migishima, R., Yokoyama, M., Nakanishi, S., and Tanabe, Y. (2004). Generation of reelin-positive marginal zone cells from the caudomedial wall of telencephalic vesicles. J. Neurosci *24*, 2286-2295.

Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R. M., et al. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature *453*, 534-538.

Taylor, H. S., Arici, A., Olive, D., and Igarashi, P. (1998). HOXA10 is expressed in response to sex steroids at the time of implantation in the human endometrium. J. Clin. Invest *101*, 1379-1384.

Theil, T., Alvarez-Bolado, G., Walter, A., and Rüther, U. (1999). Gli3 is required for Emx gene expression during dorsal telencephalon development. Development *126*, 3561-3571.

Theil, T., Aydin, S., Koch, S., Grotewold, L., and Rüther, U. (2002). Wnt and Bmp signalling cooperatively regulate graded Emx2 expression in the dorsal telencephalon. Development *129*, 3045-3054.

Theurkauf (2008). piRNA function in germline development. StemBook. Available at: http://www.stembook.org/node/479.

Thomas, P., and Beddington, R. (1996). Anterior primitive endoderm may be responsible for patterning the anterior neural plate in the mouse embryo. Curr. Biol *6*, 1487-1496.

Ting, A. H., Schuebel, K. E., Herman, J. G., and Baylin, S. B. (2005). Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation. Nat. Genet *37*, 906-910.

Tole, S., Goudreau, G., Assimacopoulos, S., and Grove, E. A. (2000). Emx2 is required for growth of the hippocampus but not for hippocampal field specification. J. Neurosci *20*, 2618-2625.

Tonchev, A. B., Yamashima, T., Sawamoto, K., and Okano, H. (2006). Transcription factor protein expression patterns by neural or neuronal progenitor cells of adult monkey subventricular zone. Neuroscience *139*, 1355-1367.

Topisirovic, I., Culjkovic, B., Cohen, N., Perez, J. M., Skrabanek, L., and Borden, K. L. B. (2003). The proline-rich homeodomain protein, PRH, is a tissue-specific inhibitor of eIF4E-dependent cyclin D1 mRNA transport and growth. EMBO J *22*, 689-703.

Topisirovic, I., Kentsis, A., Perez, J. M., Guzman, M. L., Jordan, C. T., and Borden, K. L. B. (2005). Eukaryotic translation initiation factor 4E activity is modulated by HOXA9 at multiple levels. Mol. Cell. Biol *25*, 1100-1112.

Troy, P. J., Daftary, G. S., Bagot, C. N., and Taylor, H. S. (2003). Transcriptional repression of peri-implantation EMX2 expression in mammalian reproduction by HOXA10. Mol. Cell. Biol *23*, 1-13.

Tsai, M., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., Shi, Y., Segal, E., and Chang, H. Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science *329*, 689-693.

Turunen, M. P., Lehtola, T., Heinonen, S. E., Assefa, G. S., Korpisalo, P., Girnary, R., Glass, C. K., Vaisanen, S., and Yla-Herttuala, S. (2009). Efficient Regulation of VEGF Expression by Promoter-Targeted Lentiviral shRNAs Based on Epigenetic Mechanism: A Novel Example of Epigenetherapy. Circulation Research *105*, 604-609.

Ullu, E., and Tschudi, C. (1984). Alu sequences are processed 7SL RNA genes. Nature *312*, 171-172.

Vasudevan, S., Tong, Y., and Steitz, J. A. (2007). Switching from repression to activation: microRNAs can up-regulate translation. Science *318*, 1931-1934.

Villaescusa, J. C., Buratti, C., Penkov, D., Mathiasen, L., Planagumà, J., Ferretti, E., and Blasi, F. (2009). Cytoplasmic Prep1 interacts with 4EHP inhibiting Hoxb4 translation. PLoS ONE *4*, e5213.

Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I. S., and Martienssen, R. A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science *297*, 1833-1837.

Walther, C., and Gruss, P. (1991). Pax-6, a murine paired box gene, is expressed in the developing CNS. Development *113*, 1435-1449.

Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M. G., Glass, C. K., and Kurokawa, R. (2008). Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature *454*, 126-130.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., et al. (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature *453*, 539-543.

Weichenrieder, O., Wild, K., Strub, K., and Cusack, S. (2000). Structure and assembly of the Alu domain of the mammalian signal recognition particle. Nature *408*, 167-173.

Winter, J., Jung, S., Keller, S., Gregory, R. I., and Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat. Cell Biol *11*, 228-234.

Wutz, A. (2003). RNAs templating chromatin structure for dosage compensation in

animals. Bioessays *25*, 434-442.

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Münster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., and Steinmetz, L. M. (2009). Bidirectional promoters generate pervasive transcription in yeast. Nature *457*, 1033-1037.

Yang, J., Phillips, M. D., Betel, D., Mu, P., Ventura, A., Siepel, A. C., Chen, K. C., and Lai, E. C. (2010). Widespread regulatory activity of vertebrate microRNA* species. RNA. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21177881.

Yoshida, M., Suda, Y., Matsuo, I., Miyamoto, N., Takeda, N., Kuratani, S., and Aizawa, S. (1997). Emx1 and Emx2 functions in development of dorsal telencephalon. Development *124*, 101-111.

Yue, X., Schwartz, J. C., Chu, Y., Younger, S. T., Gagnon, K. T., Elbashir, S., Janowski, B. A., and Corey, D. R. (2010). Transcriptional regulation by small RNAs at sequences downstream from 3' gene termini. Nat. Chem. Biol *6*, 621-629.

Zamore, P. D. (2010). Somatic piRNA biogenesis. EMBO J *29*, 3219-3221.

Zembrzycki, A., Griesel, G., Stoykova, A., and Mansouri, A. (2007). Genetic interplay between the transcription factors Sp8 and Emx2 in the patterning of the forebrain. Neural Dev *2*, 8.

Zhao, J., Sun, B. K., Erwin, J. A., Song, J., and Lee, J. T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science *322*, 750-756.

ACKNOWLEDGMENTS

I want to thank my supervisor, Prof. A. Mallamaci for giving me the possibility to join in his lab, allowing me to grow up the scientific abilities.

I want to thank some people who partially contribute to this work:

Marco Brancaccio for technical help in lentiviruses production and scientific discussions; Moira Pinzan for her technical contribution in RACE experiments.

I want to express my regards to all the people of the Mallamaci's lab, the oldest and the youngest team that helped me and supported me during these years of PhD, creating a really very nice atmosphere in the lab.

Great thanks go to my best friends Claudia and Marilena with whom I condivided difficult and good time of PhD's adventure in Trieste, they have been essential for me.

I want to thank my sisters an my mother for believing in me always.

Thank you!