UNIVERSITÀ DEGLI STUDI DI TRIESTE

FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI

Classe delle Lauree specialistiche in Biologia 6S

Corso di Laurea specialistica in Neuroscienze

TESI SPERIMENTALE

Regulation of Emx2 expression by non coding RNAs

Laureando/a:

Moira Pinzan

Relatore:

Prof. Antonello Mallamaci

ANNO ACCADEMICO 2009-10

INDEX

1. INTRODUCTION.....6

1.1. Emx2 in the embryonic central nervous system development 6

1.1.1. R-C (rostro-caudal) and D-V (dorso-ventral) identity specification: telencephalic induction. 9

1.1.2. Arealization 16

1.1.3. Cell cycle control, neural fate and cortical lamination.....19

1.1.4. Wiring.....26

1.2. EMX2 in the adult central nervous system 30

1.3. Emx2 transcriptional regulation in the cerebral cortex 33

1.4. Non-coding RNAs (ncRNAs):.....35

1.4.1. Classical non coding RNAs: 39

1.4.2. Non classical ncRNAs: 43

2. MATERIAL AND METHODS.....62

2.1. Embryonic tissue retrieval.....62

2.2. Cell cultures.....62

2.3. RNA extraction.....64

2.4. Reverse transcription 64

2.5. Quantitative RT-PCR 65

2.6. RACE.....67

2.7. Cloning techniques.....77

2.8. Transforming Competent cells 80

2.9. Purification of plasmidic DNA and gel extraction.....80

2.10. TET-ON Gene Expression System.....81

2.11. Lentiviral transfer vector construction.....82

2.12. Transient lipofection of Hek293T cells for Lentiviral Vectors Production 86

2.13. Titering lentiviral vectors 87

2.13.1. Fluorescence titration 87

2.13.2. Titration of lentiviral vectors by genomic Real Time PCR 89

2.14. Neurospheres infection 92

2.15. Statistical analysis 92

3. RESULTS 93

3.1. Time course analysis of DT2 and DT1 expression in vivo 93

3.2. Structural analysis of transcripts. 94

3.3. Overexpression of artificial DT transcripts in embryonic cortico-cerebral precursors 97

4. DISCUSSION 99

5. FUTURE AIMS 104

6. REFERENCES 105

ABSTRACT

Emx2 encodes for an evolutionary conserved transcription factor controlling several key aspects of embryonic cortico-cerebral morphogenesis. As emerging from ongoing investigations running in our lab, small artificial RNAs targeted against non coding conserved regions of the *Emx2* locus may impact on expression levels of its mRNA, by modulating its transcription rates. An interaction among such artificial RNAs and endogenous non coding transcripts emerging from cis-active elements of the *Emx2* locus are suspected to mediate such phenomena.

Aims of this study were: (1) to reconstruct transcriptional organization of the two main enhancers, DT2 and DT1, driving Emx2 expression in the developing rostral CNS; (2) to preliminarly assess any involvement of ncRNA stemming from these enhancers in transcriptional regulation of the gene.

It resulted from this analysis that: (1) both DT2 and DT1 enhancers are transcribed during CNS development, according to a spatio-temporal profile mimicking that of the main *Emx2* mRNA transcript; (2) both enhancers give rise to complex arrays of sense and antisense transcripts; (3) artificial overexpression of DT2 and DT1 antisense ncRNA tags downregulate *Emx2* mRNA levels.

These results are the basement of future, more in depth studies aimed at reconstructing molecular mechanisms by which the above non coding tags impact on *Emx2* transcription and clarifying how endogenous non coding transcripts centered on them really act during cortico-cerebral development.

RIASSUNTO

Emx2 codifica per un fattore trascrizionale conservato nel corso dell'evoluzione, che controlla numerosi aspetti chiave della morfogenesi cortico-cerebrale embrionale. Come risulta da studi in corso nel nostro laboratorio, piccoli segmenti di RNA, mirati contro regioni conservate di *Emx2* non codificanti possono avere un impatto sul livello di espressione del suo mRNA, modulandone la trascrizione. Si suppone che un'interazione tra questi frammenti di RNA artificiali e i trascritti non coding endogeni risultanti dagli elementi cis-attivi del locus di *Emx2*, possa mediare tali fenomeni.

Gli obiettivi di questo studio includevano: (1) la ricostruzione dell'organizzazione trascrizionale dei due principali enhancer, il DT2 e il DT1 che guidano l'espressione di *Emx2* nel SNC rostrale in via di sviluppo; (2) la valutazione preliminare di qualsiasi coinvolgimento dell'RNA non coding avente origine da questi enhancer nella regolazione trascrizionale del gene.

Da questa analisi è risultato che: (1) gli enhancer DT2 e DT1 vengono trascritti durante lo sviluppo del SNC, mimando il profilo spazio-temporale del trascritto principale di *Emx2*; (2) entrambi gli enhancer danno luogo a complesse serie di trascritti, ad orientamento senso ed antisenso; (3) la sovraespressione artificiale dei frammenti di RNA non coding corrispondenti a DT2 e DT1 antisenso sottoregola i livelli dell'mRNA di *Emx2*.

Questi risultati costituiscono la base per studi futuri, più approfonditi, con l'obiettivo di ricostruire i meccanismi molecolari tramite i quali i sucitati frammenti non coding influenzano la trascrizione di *Emx2* e spiegare in che modo i trascritti endogeni non coding centrati su di essi agiscano realmente durante lo sviluppo cortico-cerebrale.

1. INTRODUCTION

1.1. Emx2 in the embryonic central nervous system development

The vertebrate Central Nervous System (CNS) originates from the embryonic dorsal ectoderm. The differentiation of the neural epithelium from the ectoderm and the formation of the neural plate constitute the first phase of a complex process called neurulation which culminates in the formation of the neural tube, the anlage of the CNS (J L Smith & Schoenwolf 1997).

Ectodermal cells have the capacity to undergo neural differentiation as their default state. In fact, neural differentiation must be suppressed in the lateral ectoderm by signals transmitted between neighboring cells, in order to develop as epidermis. These molecular signals are members of the bone morphogenetic protein (BMP) a subclass of transforming growth factor β (TGF- β)-related proteins (Wittler & Kessel 2004). Next, diffusible signaling proteins that block the action of these growth factors are released from a particular "organizer" region of the gastrula. These proteins (known as follistatin, noggin, and chordin) allow a signaling cascade to proceed in cells near the organizer that promotes neural differentiation, forming the neural plate (Sasai 1998).

Recent studies using chick embryos have shown that neural induction really begins prior to the formation of the organizer region and thus must be initiated by signals derived from other cellular areas.

Members of other families of signaling molecules, notably the fibroblast growth factors (FGFs), have now been proposed as early-acting factors, which initiate neural induction by progressive sequence of molecular interactions.

First, the presumptive neural plate area is established by the Fgf8 activity coming from the primary endoderm.

Subsequently, the suppression of BMP signaling maintains rather than initiates the process of neural differentiation. These molecular interactions together with the participation of Hox genes (Woltering & Antony J Durston 2008)(Hooiveld et al. 1999) during the process of gastrulation regulate cellular inductive events leading to the definition of the antero-posterior and dorso-ventral axes of the embryo and to the generation of the three blastodermal layers: ectoderm, mesoderm and endoderm (Claudio D Stern et al. 2006)

Thus, in the central area of the embryo (at its prospective dorsal region), ectoderm cells are induced to develop into neural plate cells as a result of these progressive cellular and molecular interactions, acting via planar and vertical induction.

Subsequently, the process of neurulation involves cell shape changes and epithelial rearrangement which result in the bending of the neural plate and opposition of its latent edges to form the neural tube.

Cells at the interface between the dorsal neural tube and the overlying epithelium are called neural crest cells. Such cells acquire a migratory behavior and are bound to give rise to peripheral nervous system (PNS), melanocytes and other cell types (Dupin et al. 2007).

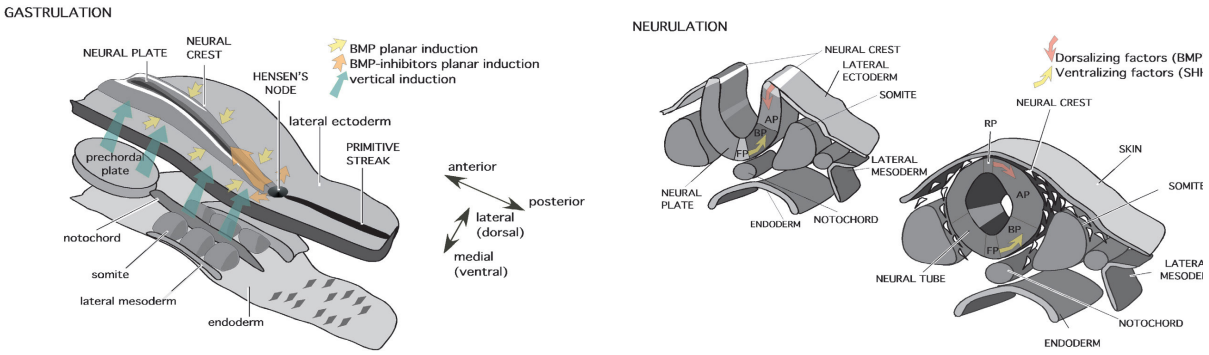


Fig. 1.1. **The neurulation process** (Vieira et al. 2010)

(A) At neural plate stage, vertical induction (green arrows) from the underlying axial mesendoderm (notochord and prechordal plate), together with planar induction from

Hensen's node (orange arrows) and ectoderm (yellow arrows) regulate dorsoventral polarity and the initial steps of antero-posterior regionalization in the neuroepithelium. **(B)** During neurulation, neural folds close at the dorsal midline. Neural crests cells delaminate and migrate from the neural folds before closure and the neural grove becomes the lumen of the neural tube. Planar information from the ventral midline (floor plate; FP; yellow arrow) and dorsal midline (roof plate; RP; red arrow) plays a fundamental role in the establishment of definitive dorso-ventral regionalization, using sonic hedgehog (SHH) and bone morphogenetic proteins (BMP) as signaling molecules. As a consequence of these inductive events, the lateral wall of the neural tube is subdivided into two columnar domains: the basal plate (close to the floor plate) and the alar plate (close to the roof plate). AP, alar plate; BP, basal plate.

As development proceeds, the anterior (cephalic or rostral) portion of the neural tube undergoes a series of swellings, constrictions and flexures that form anatomically defined regions of the brain. Initially there are three vesicles at the anterior end of the neural tube: the prosencephalon (or forebrain), the mesencephalon (or midbrain) and the rhombencephalon (or hindbrain).

Later, the prosencephalic vesicle gives rise to two laterally enlarged bulges, termed telencephalon and a medial secondary vesicle, the diencephalon; on the other hand, the rhombencephalon gives rise to the metencephalon and mielencephalon.

The caudal portion of the neural tube retains a relatively simple tubular structure and forms the spinal cord (Pombero & Salvador Martinez 2009).

The telencephalon is subsequently subdivided into the ventral and the dorsal telencephalon.

The dorsal telencephalon gives rise to a thin sheet, the pallium, from which the cerebral cortex develops.

The ventral telencephalon, or subpallium, develops two hill-like extensions, the Lateral and Medial Ganglionic Eminence (LGE/MGE), which are the forerunners of the basal ganglia (J L Rubenstein et al. 1998).

1.1.1. R-C (rostro-caudal) and D-V (dorso-ventral) identity specification: telencephalic induction.

What has been discovered so far is that transient signaling centers produce diffusible cues that create positional information. Recipient cells translate these signals through the induction of combinatorial codes of transcriptions factors. As a result they acquire specific cellular identities (T Edlund & T M Jessell 1999).

Rostro-caudal (RC) patterning is the process that leads to the generation of distinct transverse domain at different axial positions in the central neural system. It begins during early gastrulation.

The early patterning of both anterior and posterior neural tissue is mediated through signals that emanate from the primitive node or organizer. The node is a major organizing center in primitive – streak – stage embryos that regulates pattern formation. It is known as Hensen's node in chicks and the Spemann organizer in frogs.

Studies in mammals indicate that, in addition to the organizer, the anterior visceral endoderm (AVE) is required for head induction and maintenance (de Souza & Niehrs 2000)(P Thomas & Beddington 1996).
The AVE (Fig. 1.2) is the extra-embryonic tissue that underlies the future neural plate.
Removal of the AVE from mouse embryos of early stages of gastrulation leads to a loss or reduction of forebrain marker expression (P Thomas & Beddington 1996).
Also, several mutants that lack genes which are normally expressed in the AVE (for example, Hesx1, Lim1 and Otx2) fail to develop anterior structures, including the forebrain (Beddington & Robertson 1999).
On the other hand, transplantation of the mouse AVE into chick embryos results in the expression of forebrain markers (C D Stern 2001).

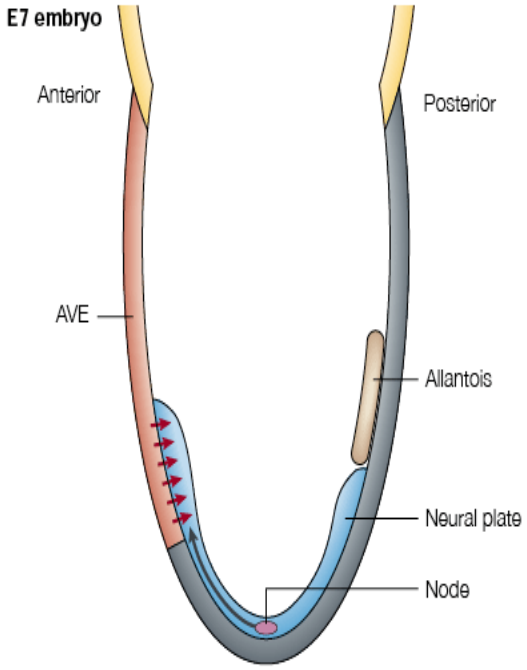


Fig. 1.2. Signals and tissues involved in inducing anterior neural character (Rallu et al. 2002).

Signals that come from the node establish gross anterior pattern (black arrow). The anterior visceral endoderm (AVE), together with the node, acts to induce and/or maintain anterior neural character. The AVE is located beneath the future neural plate and expresses molecules, such as Cerberus and dickkopf (red arrows), that inhibit factors that would otherwise act to posteriorize the anterior neural plate (Rallu et al. 2002).

As "posteriorizing"factors the molecules involved are Fgfs (fibroblast growth factors), Wnts, Retinoic acid (RA)(Agathon et al. 2003)(Muñoz-Sanjuán & H-Brivanlou 2001). Inhibitors of these diffusible signals are responsible for maintaining anterior neural identity.

The transforming growth factor-β (TGF- β)-related family proteins, such as bone morphogenetic proteins (BMPs) and Nodal, also seem to play a part in this process.

Mutants in which Nodal signaling is compromised have an enlarged telencephalon (Schier & M. M. Shen 2000).

Interestingly, Cerberus and Dickkopf (secreted by the AVE), which in addition to their ability to block Wnt signaling, also function as Nodal or Bmp antagonists (Bouwmeester et al. 1996)(Glinka et al. 1998), are only two of the many proteins that block these signaling pathways. These include chordin, noggin, follistatin and the Frizzled-related protein Frzb (Thomsen 1997)(Wessely & De Robertis 2002), all of which seem to act in the specification of the forebrain.

Each of these proteins has many functions in addition to their roles in antero-posterior (AP) patterning of the nervous system; FGF is a potent mesoderm inducer, BMPs have a role in dorso-ventral (DV) patterning of the early embryo, and both Wnt and Nodal signaling act in the establishment of the AP axis before neurulation (C D Stern 2001)(Wessely & De Robertis 2002).

Establishing whether anterior versus posterior neural structures can be neatly divided from their other functions is very difficult at present. It is probable that the combined activity of all these signaling pathways is necessary for the early R-C specification and then other signals act locally for the refining, within the neural plate, of the subdivisions of the central nervous system.

Morphogenetic controlling processes at specific locations of the developing neural primordium have led to the concept of secondary organizers, which regulate the identity and regional polarity of neighboring neuroepithelial regions (Ruiz i Altaba 1998)(Echevarría et al. 2003).

These organizers, secondary to those operating throughout the embryo during gastrulation, usually develop within the previously broadly regionalized neuroectoderm at given genetic boundaries (frequently where cells expressing different transcription factors are juxtaposed) and their subsequent activity refines local neural identities along the AP or DV axes, patterning the anterior neural plate and neural tube(Meinhardt

1983)(Figdor & C D Stern 1993)(J L Rubenstein et al. 1998) (Joyner et al. 2000).
Along the AP axis, secondary organizers are: the anterior neural ridge (ANR), at the anterior end of the neural plate/tube; the zona limitans intrathalamica (ZLI) and the isthmic organizer (IsO) at the mid-hindbrain boundary.

Soon after neural induction is initiated, the region of the rostral neural plate, the anterior neural ridge (ANR), the border between the anterior neural and non-neural ectoderm, play an important role in promoting telencephalic development within the forebrain territory (Shimamura & J L Rubenstein 1997)(Houart et al. 1998)(Tian et al. 2002)(Echevarría et al. 2003). It begins to express Fgf8 which in turn promotes the expression of Foxg1 in the early anterior plate cells that are destined to form the telencephalon and of En2 in the tectum of mesencephalon (Shimamura & J L Rubenstein 1997).

The specification of longitudinally aligned regions within the CNS involves patterning along the ML dimension of the neural plate. This ML patterning in the neural plate is topogically equivalent to DV patterning in the neural tube.
The DV patterning relies initially on external signals: at the earliest stages of neural tube formation, the neural epithelium is constrained between BMP-secreting tissue, defining the edge between neural and non-neural epithelium; at the same time, an external, mesodermic source, the notochord, secretes the morphogen Sonic Hedgehog (Shh), which is in turn also expressed in the adjacent midline neural tissue.
Notochordal SHH and ectodermal BMP sources, induce the formation of the floor and roof plate respectively, with opposite and mutually-repressing SHH and BMP signaling establishing a molecular patterning across the DV axis of the neural tube.

BMP function is further regulated by inhibitors, such as noggin and chordin, secreted by ventrally-placed mesoderm (Patten & Marysia Placzek 2002)(Lupo et al. 2006).

Shh signaling is crucial for ventral patterning at all levels of the nervous system (Briscoe & Ericson 1999).
Shh is first produced by the notochord (E7.5), and later its expression is induced in the overlying medial neural plate. Such early Shh expression commits the whole anterior neural plate to express Nkx2.1, a transcription factor specifying ventral identity (Ericson et al. 1995).
Loss- and gain-of-function analyses in several species have shown that the Shh protein is necessary and sufficient for the development and ventral neural structures and the expression of associated neural markers. Embryos that lack Shh fail to form normal ventral telencephalic structures, and they show markedly reduced expression of ventral markers (Chiang et al. 1996)(Rallu et al. 2002).
On the other hand, ectopic expression of Shh is sufficient to induce ventral telencephalic marker expression, both in vitro and in vivo (J D Kohtz et al. 1998)(Rallu et al. 2002).

BMP signaling is required for the formation of the dorsal midline, which gives rise to the choroid plexus and the cortical hem. As showed, in mice that lack both BMP receptor 1a and BMP receptor 1b (Bmpr1a□/□; Bmpr1b□/□ double mutants), the choroid plexus and cortical hem fail to form (Fernandes et al. 2007). Transgenic expression of an activated BMP receptor in the cortical neuroepithelium leads to an expansion of midline cell types at the expense of cortical ones (Panchision et al. 2001).
BMP function is further regulated by inhibitors, such as noggin and chordin, secreted by ventrally-placed mesoderm (Patten & Marysia Placzek 2002)(Lupo et al. 2006).

In addition to BMPs, Wnts are also involved in dorsal telencephalic development.
Wnts factors, secreted at the border between the neural folds and the ectodermal field, activate dorsal, i.e. cortical, genes (Gunhaga et al. 2003) and at the same time antagonize Shh effects, so inhibiting Nkx2.1 expression in the dorsal neural field.

After the closure of the neural tube, retinoic acid (RA) synthesized by the adjacent lateral ectoderm, has a crucial role in specifying telencephalic cells of intermediate character.
In particular it activates striatum-specific genes. Later, FGF signals derived from dorsal midline cells act together with WNT signals to induce definitive dorsal/precortical character in early dorsal cells (Gunhaga et al. 2003).

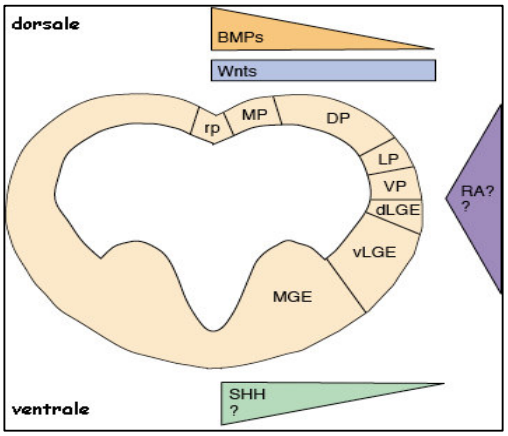


Fig. 1.3 specification of distinct dorso-ventral territories within the E10.0 mouse telencephalon (adapted from Campbell K. 2003).

Four main classes of diffusible ligands partition the telencephalic field along the dorso-ventral axis: BMPs and Wnts, more abundant in dorsal-most territories, promote pallial identity; retinoic acid, released from the ectoderm which covers the lateral part of the vesicle, stimulates striatal programs; Shh, more abundant in ventral-most territories, promotes pallidal morphogenesis. Abbreviations: dLGE, dorsal lateral ganglionic eminence; DP, dorsal pallium; LP, lateral pallium; MP, medial pallium; rp, roof plate; vLGE, ventral lateral ganglionic eminence; VP, ventral pallium; RA, retinoic acid.

Morphogens expressed in early patterning centers help establish the expression patterns of individual transcription factors (TFs) or combinations of TFs that correlate with morphologic boundaries within the telencephalon(E. Puelles et al. 2004). These TFs play a prominent role in regionalization of the telencephalon, including establishing and maintaining the identities of the ventral and dorsal telencephalon, and the general characteristics of specific cell types generated within them (Rallu et al. 2002).

The transcription factors Emx2 and Pax6 which are expressed in opposing and overlapping gradients in the dorsal telencephalon, are the key determinants of the proper development of cortical areas and are also required for establishing the identity of dorsal progenitors (Luca Muzio et al. 2002b). They have been proposed to work together in a cortical selector role. In mice deficient in either Pax6 or Emx2, cortical gene expression patterns shift along the A/P axis, and the shifts are complementary in two mutants (Bishop et al. 2000). At least one functional allele of Emx2 or Pax6 is necessary and sufficient to activate cortical fate and suppress ventral telencephalic fate. In the absence of both genes, the cortex does not form and ventral progenitor domains expand across the entire dorsal telencephalon (Luca Muzio et al. 2002a).

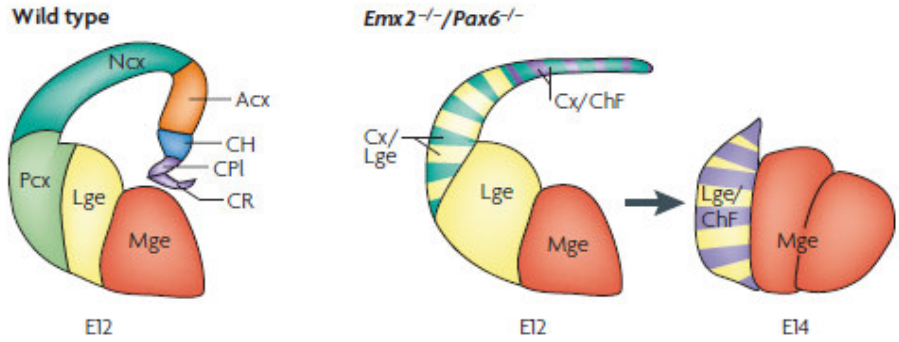


Fig. 1.4 ***Emx2* and *Pax6* have crucial roles in the specification of neocortical progenitors.**

Loss of both empty spiracles homologue 2 (Emx2) and paired box 6 (Pax6) results in ventralization of cortical progenitors and the loss of the neocortical domain (Ncx),

archicortex (Acx), cortical hem (CH) and choroid plexus (CPl) by embryonic day (E) 14. CR, choroidal roof; ChF, choroid field (choroid plexus and choroidal roof); Cx, cortex; Lge, lateral ganglionic eminence; Mge, medial ganglionic eminence; Pcx, paleocortex. Modified from Muzio et al., 2002.

1.1.2. Arealization

The developmental process that leads to the breaking up of the cortical sheet into anatomically, functionally and connectionally distinct areas is called arealization.

During the earlier stage, prior to the arrival of the thalamocortical projections, molecular regionalization of the cortical primordium would occur on the basis of intrinsic information to this primordium. During the latter stage, after the arrival of these projections (from E 13.5 onwards), cortical arealization would be refined, basing on information transported by thalamocortical fibers.

Special relevance to the whole process is given to a particular development window, from E 10.5 to E 12.5 when cortical neuroblasts are aerially committed or determined, i.e. their areal potencies become restricted in a progressively less reversible way.

At the moment, two main classes of molecules are supposed to be crucial for early regionalization of the cortical primordium:

-Secreted ligands, released around the borders of the cortical field.

-Transcription factors, gradually expressed within primary proliferative layers of this field.

Moreover, there are three candidate signaling centers lying at the borders of the cortical field, which are relevant for its arealization:

-The commissural plate (at the rostromedial pole of telencephalon).

-The cortical hem (which forms between the cortical and the choroidal fields, at the caudiomedial edge of the cortical neuroepithelial sheet).

-The cortical antihem (a recently discovered signaling structure, which forms on the lateral side of the cortical field, at the pallial-subpallial boundary).

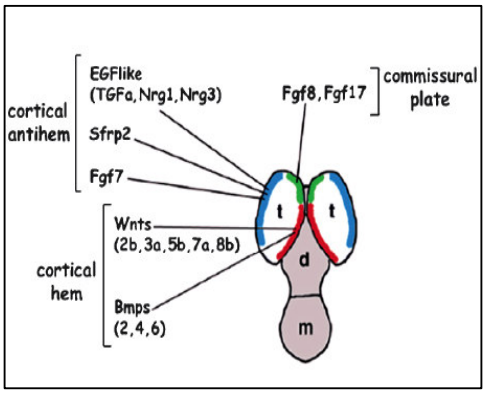


Fig. 1.5. Expression patterns of secreted ligands

(Antonello Mallamaci & Stoykova 2006)

Abbreviations: t, telencephalon; d, diencephalon; m, mesencephalon.

From earlier than E 10 to ~ E 12.5, the commissural plate (which derives from the ANR at the closure of the anterior neural folds), and its surrounding regions release Fgf 3, 8, 17 and 18 which it has been predicted, would promote rostral vs. caudal areal programs (Bachler & Neubüser 2001).

The second candidate signaling center is the cortical hem. From E10 it is a source of Wnts (Wnt2b, 3a, 5a, 7b, 8b) and bone morphogenetic proteins (Bmps; Bmp2, 4, 5, 6, 7), expressed in nested domains which also span the adjacent dorsomedial cortical field (Furuta et al. 1997)(S M Lee et al. 2000). Around E12.5 and afterwards, neural progenitors within the antihem specifically express five secreted signaling molecules: Fgf7, the Wnt-secreted inhibitor Sfrp2 and three Egf-related ligands, Tgf-a, Nrg1 and Nrg3 (Stavroula Assimacopoulos et al. 2003).

Secreted ligands, diffused through the cortical morphogenetic field, regulate the expression of cortical TF genes in dose-dependent manners. In this way they account for the further generation of concentration gradients of these factors. Graded and transient expression of these factors would finally encode for positional values, peculiar to distinctive regions of the cortical field.

Several transcription factor genes, including Emx2, Emx1, Lhx2, Pax6, Foxg1 and Coup-tf1, are expressed by neural progenitors within periventricular proliferative layers, in graded manners along the main tangential axes. These genes are crucial for imparting distinctive regional identities to neural progenitors.

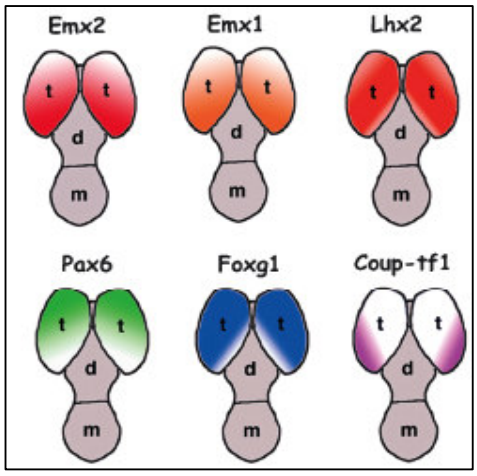


Fig. 1.6. Graded transcription factor genes in the early cortical primordium E12.5.

The homeobox gene Emx2, expressed highest in progenitors that generate posterior- medial areas of neocortex (such as V1) and lowest in progenitors that generate anterior-lateral areas (such as frontal and motor)(A Simeone et al. 1992)(Gulisano et al. 1996)(A Mallamaci et al. 1998) shapes the cortical areal profile as a promoter of caudomedial fates (D D O'Leary et al. 1994). In the absence of Emx2, the full repertoire

of areal identities is preserved but caudomedial areas are shrunken and rostrolateral ones expanded (Bishop et al. 2000)(A Mallamaci et al. 2000) However, the overall areal profile is finely tuned to the Emx2 dosage. Relative and absolute sizes of occipital areas of Emx2□/□ mutants are intermediate between null and wild-type mice and an expansion of caudal medial areas can be achieved by introducing one or, better, two alleles of a nestin-promoter-driven Emx2-expressing transgene into a wild-type genome. This suggests that Emx2 operates by a concentration-dependent mechanism in cortical progenitors to specify the sizes and positioning of the primary cortical areas and that higher level of Emx2 preferentially impart posterior-medial area identities (Hamasaki et al. 2004).

1.1.3. Cell cycle control, neural fate and cortical lamination

The mammalian neocortex is a complex, highly organized, six layered structure, divided into distinct areas according to their functions and cytological organization (Barbier et al. 2002).

The developmental progression of the cerebral cortex is unique to mammals and is fairly conserved throughout species.

During development, neural stem cells give rise to all neurons of the mammalian central nervous system (CNS). They are also the source of the two types of macroglial cells in the CNS – astrocytes and – oligodendrocytes (Doetsch, Caillé et al. 1999)(Alvarez-Buylla et al. 2001).

There are two broad classes of cortical neurons: interneurons, which make local connections; and projection neurons, which extend axons to distant intracortical, subcortical and subcerebral targets.

Projection neurons are glutamatergic neurons characterized by a tipical pyramidal morphology that transmit information between different regions of the neocortex and to other regions of the brain.

During development, they are generated from progenitors of the neocortical germinal zone located in the dorsolateral wall of the telencephalon (Anderson et al. 2002).

By contrast, GABA (γ-aminobutyric acid)-containing interneurons and Cajal-Retzius cells are generated primarily from progenitors in the ventral telencephalon and cortical hem, respectively, and migrate long distances to their final locations within the neocortex (Wonders & Anderson 2006).

During early development, there is a dramatic expansion of the neuroepithelium in the dorsolateral wall of the rostral neural tube that will give rise to neocortical projection neurons. The layer immediately adjacent to the ventricle is termed the ventricular zone (VZ). As neurogenesis proceeds, an additional proliferative layer known as the subventricular zone (SVZ) forms above the VZ (Bayer et al. 1991)(The Boulder Committee, 1970). Progenitors residing in the VZ and SVZ produce the projections neurons of the different neocortical layers in a tightly controlled temporal order from embryonic day (E) 11.5 to E 17.5 in the mouse (Angevine & Sidman 1961)(P Rakic 1974), and postmitotic neurons position themselves in the developing neocortex through defined modes of radial and tangential migration (P Rakic 1972)(Olga Britanova et al. 2006)(Pasko Rakic 2003).

The earliest born neurons appear around E 10.5 in the mouse and form a layered structure termed the preplate. The subsequent wave of neuronal migration (~E13) splits the preplate two layers: the more superficial marginal zone, which consists of the Cajal-Retzius cells born in the first wave of migration; and the deeply located subplate, which is constituted by the rest of the primordial cells. The cortical plate, which will give rise to the multilayered neocortex, begins to develop in between these two layers (Bayer et al. 1991), such that later born neurons arriving at the cortical plate migrate past earlier born neurons. (Angevine & Sidman 1961)(P Rakic 1974).

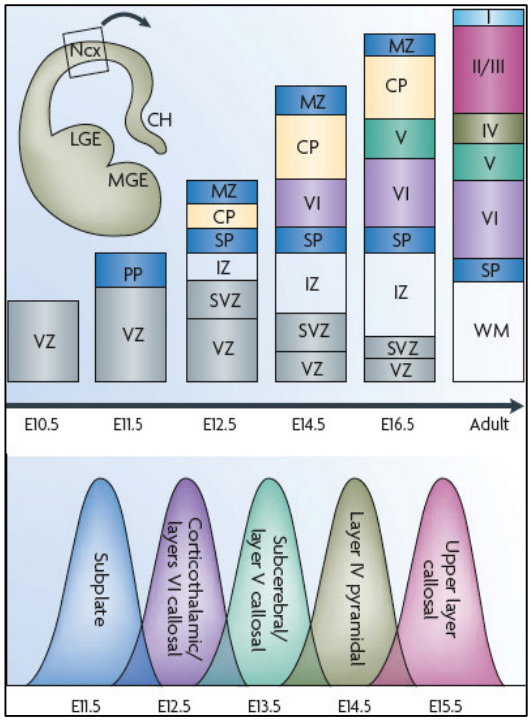


Fig. 1.7. Schematic depicting how progenitors residing in the VZ and SVZ in mice produce projection neuron in an "inside-out" fashion.

The earliest born neurons form the preplate (PP), which is later split into the more superficial marginal zone (MZ) and the deeply located subplate (SP). The cortical plate (CP), which will give rise to the multilayered neocortex, develops in between these two layers, such that later born neurons arriving at the cortical plate migrate past earlier born neurons. Different classes of projection neuron are born in overlapping temporal waves. All times listed are approximations given the neurogenic gradients that exist across the cortex, where caudomedial neurogenesis lags behind rostrolateral neurogenesis (Bayer et al. 1991). CH,cortical hem; E,embryonic day; Ncx,neocortex; IZ,intermediate zone; LGE,lateral ganglionic eminence; MGE,medial ganglionic eminence; SVZ,subventricular zone; VZ,ventricular zone; WM,white matter (Molyneaux et al. 2007).

There are four types of neurogenic progenitors within the developing neocortex:

Neuroepithelial cells (NEs)

Radial glia cells (RGCs)
Intermediate progenitors cells (IPCs)
Short neural precursors (SNPs)

Initially, there is a single sheet of pseudostratified neuroepithelial cells undergoing symmetric cell divisions to expand the pool of multipotent progenitors (proliferative divisions) as well as a smaller percentage of asymmetric divisions to generate a daughter stem cell plus a more differentiated cell such as a non-stem-cell progenitor or a neuron (Magdalena Götz & Huttner 2005)(Smart 1973).
As neurogenesis progresses, they transform into radial glia, which share some but not all antigenicity with the early neuroepithelial cells (E Hartfuss et al. 2001)(Paolo Malatesta et al. 2003). This radial glia cells, posses a long process that extend from the ventricular wall to the pial surface and have long been known to have crucial roles in guiding neurons to their final locations in the cortical plate by serving as migratory scaffolding (P Rakic 1972)(Pasko Rakic 2003).
In addition to the full-length radial glia, other neuron-producing precursors have been described in the VZ (Magdalena Götz & Huttner 2005)(Mo et al. 2007)(Gal et al. 2006). It is a subpopulation of progenitors that can be distinguished from radial glial cells by the absence of a full-length pial process, named "short neural precursors" (Gal et al. 2006).
During each cell cycle, this progenitor cells undergo a distinctive pattern of oscillation in the ventricular zone, termed interkinetic nuclear migration. Cells undergo S phase at the basal surface of the ventricular zone and mitosis at the apical surface.

RGCs undergo several types of symmetrical and asymmetrical divisions (Magdalena Götz & Huttner 2005)(Huttner & Kosodo 2005) including self-renewing ones or neurogenic divisions. Although it cannot be excluded

that SNPs are themselves derived RGCs, it is thought that SNPs are committed to symmetrical neurogenic divisions (Gal et al. 2006).
IPCs are neuronal progenitors derived from RGCs that divide away from the ventricular surface in the VZ and in the subventricular zone (SVZ). They have been reported to undergo mostly neurogenic divisions with a small fraction undergoing symmetrical proliferative divisions (Noctor et al. 2004).
The VZ generates lower layer neurons and the SVZ generates upper layer neurons.

The role of cell-cycle regulation in determining neuronal number in the adult cortex is consistent with findings elsewhere in the CNS (Oppenheim et al. 1989).
Specifically, two cell-cycle parameters determine neuron number: the rate of cell-cycle progression and the balance between cell-cycle re-entry or exit. Whereas proliferative divisions generate two progenitors that re-enter the cell cycle, differentiative divisions result in at least one daughter cell exiting the cell-cycle to undergo differentiation.
Although the molecular mechanisms that determine the tightly regulated occurrence of proliferative versus differentiative divisions are largely unknown, converging evidence suggest that the mode of division is correlated to cell-cycle components and, more specifically, to G1-phase regulation.

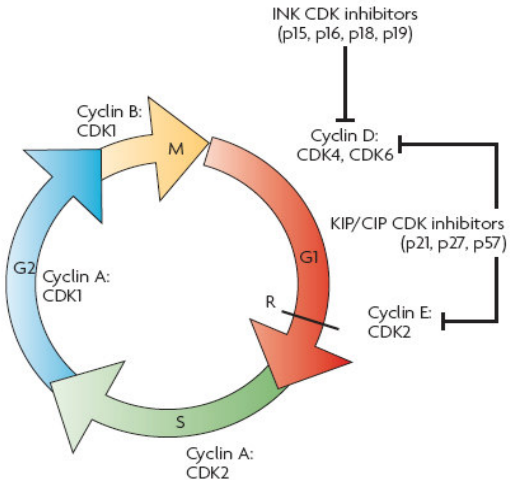


Fig. 1.8. Structure and regulation of the cell cycle (Dehay & Kennedy 2007a).

During mouse corticogenesis, there is a progressive increase in neuron production and in the frequency of differentiative divisions. Simultaneously, there is a slowing down of the rate of cell-cycle progression, which is mainly due to a lengthening of the G1 phase (Caviness et al. 1995). The G1 phase is a critical stage, allowing responses to extracellular cues that induce either commitment to a further round of cell divisions or withdrawal from the cell cycle to embark on a differentiation pathway (Zetterberg et al. 1995). A prolonged G1 phase could be a characteristic feature of differentiative divisions, facilitating the integration of extrinsic signals that influence cell fate and/or allowing an unequally inherited cell-fate determining factor(s) to act over a sufficient time period (Magdalena Götz & Huttner 2005)(Calegari & Huttner 2003). Birth-dating experiments coupled with manipulation of the cellular environment suggested that the cell fate is determined prior the migration (Caviness 1982)(Caviness & Sidman 1973)(K. F. Jensen & Killackey 1984).

The factors responsible for the timed generation of different neuronal phenotypes have been reinvestigated in lineage studies of isolated cortical precursors (Q. Shen et al. 2006). Together with earlier findings, these results show that there is a cell-intrinsic program (Q. Shen et al. 2006) that is influenced by extrinsic factors (S K McConnell & C E

Kaznowski 1991)(Frantz & S K McConnell 1996) (G Fishell 1995)(Qian et al. 1997) so that both extrinsic and intrinsic factors cooperate to determine cell fate. The temporal pattern of *in vitro* neuronal subtype generation matches that observed *in vivo*: Cajal-Retzius neurons are formed first, followed by cells expressing markers of initially lower (Foxp2, Tle4, ER81) and subsequently upper cortical layers (Cux1), confirming earlier findings that there is a progressive restriction of cell-fate potential. The finding that cell-cycle regulation and the developmental programs that generate sequential neuronal subtypes are maintained *in vitro* raises the possibility that cell-cycle control mechanisms are involved in fate determination(Q. Shen et al. 2006)(Federico Cremisi et al. 2003)(S Ohnuma et al. 2001). Cell-cycle mechanisms could be responsible for determining both the numbers and the phenotype of cortical neurons generated in each layer.

Emx2 is one of the earliest markers for the developing cerebral cortex. From E10 and during the formation of the cerebral cortex, the neuroepithelium becomes the only area of expression of Emx2; whereas its transcript is absent in most postmitotic neurons of the cortical plate suggesting a potential role for the gene in neuroblasts proliferation. At E12.5, when the developing cerebral cortex essentially coincides with the ventricular zone, the mRNA signal is present, and shows a gradient along the A/P axis, higher in posterior cortical regions than in anterior ones (A Simeone et al. 1992)(Gulisano et al. 1996). This antero-posterior gradient of expression of the transcript seems to follow the gradient in the maturation of neuroblasts during corticogenesis, the anterior regions containing mature cells earlier than the posterior ones. EMX2 has also been found in the nuclei of Cajal-Retzius (CR) cells, as well as in the most marginal cortical plate neurons, where it is restricted to their apical dendrites (A Mallamaci et al. 2000). These neurons are able to establish synaptic connections with CR cells, a transient cell population that consists of the first born neurons in neocortex and forms most superficial cell layer just underneath the pial membrane; they are thought to be

responsible for guiding radial migration of neurons from the VZ to their final destination across the CP (M. Ogawa et al. 1995). Taken together, these result indicate that not only it has a role in neuroblasts proliferation but also regulate their subsequent migration processes through the cortical plate during corticogenesis (S K McConnell 1995). Analysis of Emx2 null mice has revealed that these embryos lack CR cells from the MZ. As a consequence, the settling of radial glia is impaired and neurons display abnormal reeler-like migration patterns (A Mallamaci et al. 2000). Moreover, even in Emx2 null mice, between E11 and E13 the occipitohippocampal anlage expands less than normal, due to selective slowing down of DNA synthesis and exaggerated neuronogenesis in this region. This is associated with up-regulation of cyclin-dependent kinase 2 inhibitor genes *Kip*1 *p27* and *Kip*2 p57, exaggerated proneural : antineural gene expression ratio and depression of the Delta-Notch-Hes axis in the same region (L Muzio et al. 2005). However, for the role of Emx2 in controlling neural precursor's proliferation and differentiation kinetics, published data are not fully consistent.

It has been shown that Emx2 influences kinetics of cortical proliferating populations (Gangemi et al. 2001)(Heins et al. 2001)(Rossella Galli et al. 2002). However, these studies were run in vitro, on cell cultures, and, moreover their results were apparently controversial. High density in vitro cultures derived from Emx2 null mice cortexes show decreased proliferation, whereas Emx2 overexpressing cultures show enriched symmetrical divisions (Heins et al. 2001).

Thus, molecular and cellular mechanisms by which Emx2 acts remain unknown.

1.1.4. Wiring

Although once an intensely debated issue (P Rakic 1988)(D D O'Leary 1989), it is now widely held that the specification and differentiation of neocortical areas is controlled by an interplay between intrinsic

mechanisms, i.e., genetic mechanisms that operate within the cortex, and extrinsic mechanisms such as the sensory periphery and thalamocortical axon (TCA) input or information relayed by it.

The properties that distinguish cortical areas emerge gradually during development (D D O'Leary & Koester 1993). The nascent CP, before it acquires its mature functional abilities, lacks most of the anatomically based features that distinguish areas in the adult, even after all CP neurons have been generated and layers begin to differentiate within it. The cortex is initially a more or less uniform structure; many area-specific properties differentiate in parallel spatially and temporally to the development of TCA input (Chenn et al., 1997).

Each area is defined by both its structure and connectivity, which together determine its sensory, motor or cognitive function. There are three basic classes of cortical projections neurons:

-Associative projections neurons (neurons that extend axonal projections within a single cerebral hemisphere)

-Commissural projection neurons (neurons that extend axonal projections within the cortex to the opposite hemisphere via the corpus callosum ir the anterior commissure)

-Corticofugal projections neurons (neurons that extend axonal projections "away" from the cortex. These include subcerebral projections neurons and corticothalamic neurons)

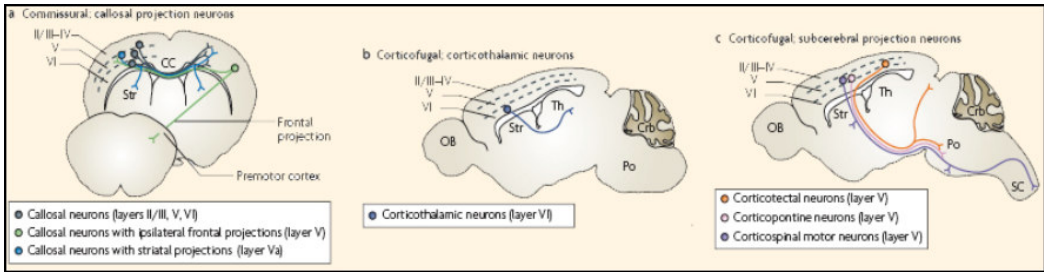


Fig. 1.9. Major subtypes of projections neurons within the neocortex (Molyneaux et al. 2007).

Area patterning and function exhibits considerable plasticity upon modification of sensory periphery or TCA input or performance o heterotopic transplantation (Chenn et al., 1997)(Dennis D M O'Leary & Y. Nakagawa 2002)(Sur & John L R Rubenstein 2005).

Heterotopic transplant experiments show that area-specific cytoarchitecture and axon/collateral elimination by layer 5 projection neurons is plastic during development. For example, transplants of embryonic occipital cortex, which will differentiate into visual areas, into S1 barrelfield in parietal cortex develop cytoarchitecture and the patterned expression of markers characteristic o the S1 barrelfield (Schlaggar & D D O'Leary 1991). Other studies show that developing layer 5 neurons transplanted from visual cortex to motor cortex permanently retain their normally transient spinal axon, whereas layer 5 neurons transplanted from motor cortex to visual cortex lose their normally permanent spinal axon and retain their transient axon collateral to the superior colliculus (Stanfield & D D O'Leary 1985)(D D O'Leary & Stanfield 1989). Thus, the projections retained by the transplanted layer 5 neurons are appropriate for the cortical area in which the transplanted neurons develop, not where they were born. These and other experimental manipulations reveal a plasticity in the development of the mature areal distributions of projections neurons from initially broad distributions, through mechanisms that are likely to be at least in part independent of the intrinsic specification of area identity.

With the possible exception of layer 6 corticothalamic neurons, cortical projection neurons initially exhibit "exuberant" areal distribution far more broad than those in the adult. In contrast to this lack of areal specificity in the early distribution of projections neurons, the area-specific projections of TCAs from the principal sensory thalamic nuclei is evident at the early stages in their development, prior to the emergence of the sharp cytoarchitectonic borders between areas that later become evident(D D O'Leary et al. 1994). Progress has been made in defining mechanisms of TCA pathfinding, particularly subcortically from dorsal thalamus to the

neocortex (Polleux 2005), but the molecular control of TCA targeting of specific areas remains relatively vague. Similar to the well-defined mechanisms that control development of topographically ordered retinal projections in the visual system (McLaughlin & Dennis D M O'Leary 2005) area specific TCA targeting is likely primarily controlled intracortically by graded axon guidance molecules (Dufour et al. 2003) and refined by neural activity (Catalano & Shatz 1998). Sub-plate neurons and their axons have also been implicated in the development of area-specific TCA targeting, but their role and its molecular basis are vague (Allendoerfer & Shatz 1994)(Molnár & Blakemore 1995).

Since TCAs are the sole source of modality-specific sensory information to the neocortex, the functional specializations of the primary sensory areas are defined in large part by, and dependent upon, TCA input.

The TCA projection exhibits area specificity throughout its development, and the gradual differentiation of areas within the CP parallels the elaboration of TCA projection within it (Chenn et al., 1997).

However, the role of TCAs in shaping cortical architecture is not limited to these later events in the differentiating CP. In vitro experiments using mouse tissue suggest that TCAs release a diffusible mitogenic activity that promotes the production of both glia and neurons by explants of the cortical VZ (Dehay & Kennedy 2007b).

As Emx2 is expressed in a high posteromedial to low anterolateral gradient, it is implicated in the spatial delineation of posterior functional areas (Hamasaki et al. 2004). Manipulation of Emx2 expression in mice by genetic knockout causes an expansion of frontal and lateral regions (i.e. S1 and M1) at the expense of the visual cortex (Hamasaki et al. 2004)(Leingärtner et al. 2007). Furthermore, viral overexpression of Emx2 causes ectopic TCA projection from the lateral geniculate nucleus to areas outside the area normally conforming to V1 (Leingärtner et al. 2003). This evidence support the establishment of a proper cortical map to instruct the correct spatial expression of attractive cues necessary for TCA path finding.

1.2. EMX2 in the adult central nervous system

Neurogenesis in the brain of adult mammals occurs throughout life, and has been clearly demonstrated at two locations under normal conditions: the subventricular zone (SVZ) of the lateral ventricles and the subgranular zone (SGZ) of the dentate gyrus in the hippocampus. Neurons born in the adult SVZ migrate over a great distance through the rostral migratory stream and become granule neurons and periglomerular neurons in the olfactory bulb. Neurons born in the adult SGZ migrate into the granule cell layer of the dentate gyrus and become dentate granule cells. Recent studies also showed that newborn neurons in the adult brain integrate into the existing circuitry and receive functional input.

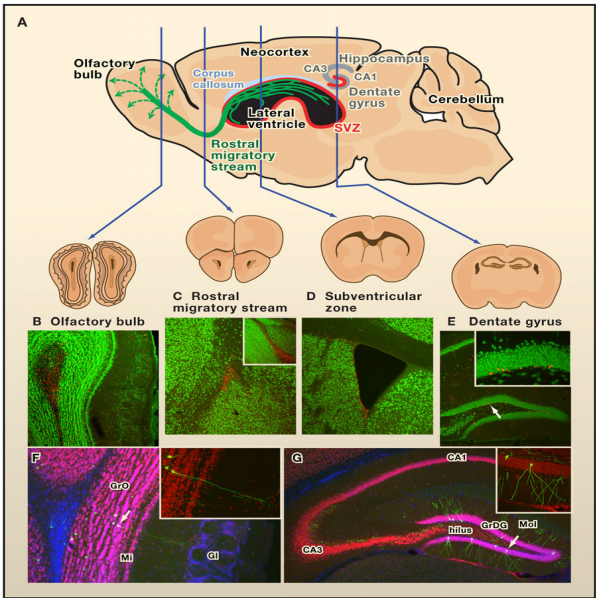


Fig. 1.10. Neurogenesis in the Adult Rodent Brain (C. Zhao et al. 2008).

(A) Depictions of sagittal and coronal views of mouse brain in areas where neurogenesis occurs. Red areas indicate the germinal zones in the adult mammalian brain: the subgranular zone (SGZ) of the dentate gyrus in the hippocampus and the subventricular zone (SVZ) of the lateral ventricles. (B–E) Neurogenesis revealed by BrdU incorporation in the olfactory bulb (B), rostral migratory stream (C), SVZ (D), and dentate gyrus (E). Colors indicate the following:red, BrdU; green, NeuN.(F and G). Newborn neurons in the olfactory bulb and dentate gyrus labeled by retrovirus-mediated expression of green fluorescent protein (GFP). red, NeuN; green, GFP; blue, DAPI

Adult neurogenesis is regulated by physiological and pathological activities at all levels, including the proliferation of adult neural stem cells (NSCs) or progenitors, differentiation and fate determination of progenitor cells, and the survival, maturation, and integration of newborn neurons. Furthermore, these cells may be required for certain forms of brain function involving the olfactory bulb and the hippocampus, which is important for some forms of learning and memory.

Adult NSCs are cells in the adult nervous system that can self-renew and differentiate into all types of neural cells, including neurons, astrocytes, and oligodendrocytes (F H Gage 2000).

Two types of neural progenitors can be identified in the SGZ according to their specific morphologies and expression of unique sets of molecular markers. Type 1 hippocampal progenitors have a radial process spanning the entire granule cell layer and ramify in the inner molecular layer. These cells express nestin, glial fibrillary acidic protein (GFAP), and the Sry-related HMG box transcription factor, Sox2 (Satoshi Fukuda et al. 2003)(Garcia et al. 2004)(Suh et al. 2007). Although expressing the astrocyte marker GFAP, these cells are morphologically and functionally different from mature astrocytes.

Type 2 hippocampal progenitors have only short processes and do not express GFAP. Type 2 cells may arise from type 1 cells, but direct evidence delineating this lineage relationship is still lacking.

A recent study showed that type 2 Sox2-positive cells can self-renew and that a single Sox2- positive cell can give rise to a neuron and an astrocyte, providing the first in vivo evidence of stem cell properties of hippocampal neural progenitors (Suh et al. 2007).

The transcription factor Sox2 is important for maintaining the ``stemness'' not only of certain types of adult stem cells including NSCs but also of embryonic stem (ES) cells.

The SVZ is located next to the ependyma, a thin cell layer that lines the lateral ventricles of the brain.

Ependymal cells have been suggested to be the adult NSCs responsible for neurogenesis in the SVZ (Johansson et al. 1999). Several studies have shown, however, that ependymal cells are quiescent and do not have the properties of NSCs in vitro (Capela & Sally Temple 2002)(Doetsch, García-Verdugo et al. 1999).

More importantly, cells within the SVZ (and less likely the ependyma itself) contribute to long-term neurogenesis in the olfactory bulb (Consiglio et al. 2004). Three types of precursor cells exist in the SVZ: type B GFAP-positive progenitors, type C transit amplifying cells and type A migrating neuroblasts.

Lineage tracing studies in adult mice have demonstrated that newborn neurons, astrocytes and sometimes oligodendrocytes can be derived from cells expressing a given molecular marker, such as Nestin, GFAP, GLAST and Sox2 (reviewed by Breunig et al., 2007). However, these markers are expressed in heterogeneous populations of cells and it is not clear whether cells expressing these markers are the primary progenitors.

Although NSCs can be isolated from many areas of the adult nervous system, adult neurogenesis has only been consistently found in the SVZ and SGZ in vivo. It is hypothesized that the microenvironments of the SGZ and SVZ, known as the neurogenic niche, may have specific factors that are permissive for the differentiation and integration of new neurons.

Any diffusible molecules produced by local cells can influence neural progenitors. Neighboring cells can also exert their influence through

direct cell-cell interactions. Furthermore, neural progenitors can be indirectly influenced by neurons outside of this microenvironment that are connected to neurons within the neurogenic niche through neural circuits. Therefore, adult neurogenesis is subject to complex extrinsic regulation.
Neurogenesis declines with aging in both the SVZ and SGZ (Reviewed by (Rossi et al. 2006)

Emx2 expression starts early in the mouse; the gene is expressed as early as at the three-somite stage in the latero-caudial primordial (Y Suda et al. 2001) and continue until long after birth - a time interval corresponding to major events in cortical neurogenesis, differentiation and migration, and to establishing of synaptic connections. The protein has also been detected in the proliferating cells of the SGZ of the dentate gyrus in postnatal mice (up to P15), suggesting that its role as inhibitor of neuronal proliferation and maturation could be maintained also in adult life (A Mallamaci et al. 1998). EMX2 has also been found in adult mice neural stem cells present in the subependymal layer of the lateral ventricles and in the dentate gyrus of the hippocampus, probably participating in the control of transition from symmetrically to asymmetrically dividing neural precursors (Gangemi et al. 2001). This hypothesis is supported by the presence of an increased number of precursors in the VZ of Emx2 mutant embryos, where this layer is thicker in contrast to the strong reduction in size of the cerebral hemispheres (Tole et al. 2000).

1.3. Emx2 transcriptional regulation in the cerebral cortex

Emx2 plays an essential role in each step and site of forebrain development. During development of the cerebral cortex, *Emx2* expression is confined to the ventricular proliferating compartment, forming - from E11.5 onward - an expression gradient a with the highest

expression levels in the caudal/medial domain. To elucidate the transcriptional regulation underlying this 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, unexpectedly directed all the *Emx2* expression in forebrain:

caudal forebrain primordium at E8.5, dorsal telencephalon at E9.5- E10.5 and the cortical ventricular zone after E12.5. However, it did not fire in cortical hem and its derivatives and 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 unique to mammalian telencephalon should exist.

1.4. Non-coding RNAs (ncRNAs):

For many years, RNAs have been considered to be just accessory molecules involved in mediating transcription and translation. This is an old-fashioned and simplistic view of how the molecular machinery works in eukaryotic cells. RNA molecules are very versatile, mainly due to their chemical properties, which allow them to form complex tertiary structures, capable of performing several roles that were thought to be under the exclusive domain of protein (Szymański et al. 2003).
RNAs can interact with different proteins forming ribocomplexes and can associate with specific DNA and/or RNA sequences, controlling several aspects of gene regulation and enhancing the plethora of molecular connections that might happen in eukaryotic cells (John S Mattick 2004).

In the past few years it has been revealed that the genomes of all studied eukaryotes are almost entirely transcribed; it has been estimated that ~98% of the transcriptional output of the human genome represents RNA that does not encode protein, thus generating an enormous number of non-protein-coding RNAs (ncRNAs)(Ewan Birney et al. 2007)(Core et al. 2008a)(Piero Carninci et al. 2006)

The increasing diversity of ncRNAs identified in the eukaryotic genome suggests a critical nexus between the regulatory potential of ncRNAs and the complexity of genome organization (Prasanth & Spector 2007). That is why this biological complexity is thought to be generally related to the proportion of the genome that is non-protein-coding (Taft et al. 2007).

These findings have directly challenged the traditional view of RNA as a simple intermediary between DNA and protein, and imply that the vast majority of the genome – long regarded as "junk" – encodes functional RNA species that orchestrate the development of complex organisms (John S Mattick 2007).

Indeed, it appears that RNA signaling is central to all complex genetic phenomena in the eukaryotes.

NcRNAs have been identified as –*cis* and *trans*-acting regulators of development and can operate by many different mechanisms at transcriptional or post-transcriptional level, regulating translation, mRNA stability and processing. It has also been demonstrated their important role in control of chromatin architecture and epigenetic processes, directing generic chromatin-modifying enzymes and complexes to their sites of action and regulating those pathways (reviewed in (Amaral & John S Mattick 2008).

Although there is increasing evidence for the functionality of large numbers of ncRNAs expressed from the mammalian transcriptome, there has also been controversy about whether these transcripts are mainly functional or simply represent "transcriptional noise". To date, it has not been confirmed yet that all of these ncRNAs transcripts are functional, particularly because most of them are transcribed at very low levels and pervasive transcription seems to be a common event in eukaryotic genomes (Ebisuya et al. 2008). However, low levels of transcription do not exclude the possibility that these non coding genomic regions are playing a role in epigenetic mechanisms, as it has been suggested.

Pervasive transcription is a widespread event in eukaryotic genomes and this is not a futile process, it might play a crucial role in controlling gene expression and genome plasticity (Berretta & Morillon 2009).

The evidence for general functionality of some ncRNAs is represented by their regulated expression.
First of all, most ncRNAs exhibit differential expression in different cells and tissues(Masaaki Furuno et al. 2006)(Kapranov et al. 2002), in patterns similar to those observed with mRNA, suggesting that they are similarly regulated and thus may participate in developmental or physiological processes in vivo (Ewan Birney et al. 2007). In particular, in situ hybridization analysis in human brain reveals that hundreds of ncRNAs are precisely expressed in restricted functional regions of the brain, and many show specific subcellular locations (Mercer et al. 2008) Indeed, specific subcellular localization appears to be a common feature of ncRNAs (Ginger et al. 2006)(Mercer et al. 2008), some of which mark novel domains (Royo et al. 2007)(Sone et al. 2007), adding weight to the proposition that these ncRNA have biological function.
Second, the expression of many intergenic and intronic ncRNAs responds to environmental signals (Cawley et al. 2004)(Louro et al. 2007); many ncRNAs transcription units have upstream binding sites for the transcription factors (Cawley et al. 2004); expression of individual ncRNAs has also responds to specific signaling pathways, including Sonic hedgehog (Jhumku D Kohtz & Gord Fishell 2004), Notch (Tsutsumi & Motoyuki Itoh 2007) and BMP (K. Takeda et al. 1998).

Similar observations have been made also in other species like in *C.elegans* (Deng et al. 2006), and in *Drosophila*, where thousands of noncoding transcripts exhibit independent and dynamically regulated expression patterns during development (Manak et al. 2006)(Stolc et al. 2004). It has also been demonstrated that a large numbers of this ncRNAs have precise temporal and spatial expression, some of which

have conserved expression patterns in different *Drosophila* species (Inagaki et al. 2005)(Tupy et al. 2005). Moreover, many of those that were expressed during embryogenesis have tissue-specific expression (Inagaki et al. 2005).

All these evidences show that there are ncRNAs with specific roles in eukaryotic cells and they are not all merely pervasive transcription representing transcriptional noise. However, just a small number of ncRNAs has been carefully examined in biological relevant systems.

Initially the term ncRNA was used primarily to describe eukaryotic RNAs that are transcribed by RNA polymerase II (RNA pol II) and have a 7-methylguanosine cap structure at their 5' end and a poly (A) tail at their 3' end, but lack a single long "Open Reading Frame" (ORF). More recently this classification has been extended to all RNA transcripts that do not have a protein-coding capacity, including cap-less and/or polyA-less ones (J S Mattick & Gagen 2001)(Shabalina & Spiridonov 2004).

To date, an exhaustive classification is quite impossible since the catalogue of ncRNAs seems to be ever-growing. Moreover, there is an ongoing lack of clarity regarding the true number of ncRNAs within the genome. This is at least partly due to the inherent difficulties in discriminating ncRNAs from mRNA and especially from artifacts. Computational methods developed for protein coding genes often fail when searching for ncRNAs. Novel ncRNA genes are difficult to be identified basing on sequence analysis due to their sequence divergence across phyla (Pang et al. 2006).

The nature of ncRNA genes, including their variation in length (20 nucleotides (nt) to >100kb), lack of ORFs, and the relative immunity to point mutations, makes them difficult targets for genetic screens.

It is reasonable to make a first division into "classical" and "non classical" ncRNAs. Although both of them are not completely known, the first ones are the best characterized ones, while "non classical" ones still require proper approach and classification that should lead to a complete acknowledgment.

1.4.1. Classical non coding RNAs:

They can be subdivided into two classes, basing on functional relevance:

1Housekeeping ncRNAs: they are generally constitutively expressed and are required for the normal function and viability of the cell.

2Regulatory ncRNAs (or riboregulators): include those ncRNAs that are expressed at certain stages of development, during cell differentiation or as a response to external stimuli, which can affect the expression of other genes at the level of transcription or translation.

Housekeeping ncRNAs mainly comprehend species that are associated to expression of polypeptide-encoding genes like:

-transfer RNAs (tRNAs): function as adapters in translation

-ribosomal RNAs (rRNAs): ribosome components, catalysis of peptide bond formation

-small nuclear RNAs (snRNAs): involved in pre-mRNA splicing, spliceosome components

Moreover, there are also other species like: RNAse P RNA, involved in maturation of 5'ends of pre-tRNA; telomerase RNA, involved in telomeric DNA synthesis as a component of telomerase.

Regulatory ncRNAs comprehend small ncRNAs, that are a substantial portion of the RNA output of cells and function in several pathways modulating gene expression. They can be classified into different groups,

based on their origin or the components to which they are functionally connected.

Classical small RNAs

Class	Length (nt)	Function
microRNA (miRNA)	19-25	Translational repression
small interfering RNA (siRNA)	19-21	Target mRNA cleavage
trans-acting RNA (tasiRNA)	21-22	mRNA cleavage
small-scan RNA (scnRNA)	~28	DNA elimination
repeat-associated siRNA (rasiRNA)	24-27	Transposon control transcriptional silencing
piwi-interacting RNA (piRNA)	26-31	Transposon control in germ cells
small nucleolar RNA (snoRNA)	~60-300	methylation and pseudouridylation of rRNAs, tRNAs and snRNAs.

The most intensely studied class of small RNAs in eukaryotes are **microRNA (miRNA)**. They are ~22 nt single-stranded molecules produced from imperfect hairpin structures present in long ncRNA precursors or introns of non-coding and protein-coding genes, generally processed in two consecutive cleavage steps by Drosha and Dicer. Mature miRNA molecules interact by base-pairing with target mRNAs to modulate translation or direct degradation by effector partners of the Argonaute/Piwi (Ago/Piwi) family(Brodersen et al. 2008)(Farazi et al. 2008)(Filipowicz et al. 2008). In the past ten years, hundreds of miRNA have been identified in animal and plants and shown to play central roles in the control of gene expression programs during development with at least a thousand predicted in humans (Jones-Rhoades et al.

2006)(John S Mattick & Makunin 2005)(Stefani & Slack 2008). miRNA have hundreds of targets, which may include ncRNAs (Y. Zhao et al. 2008), and regulate development in a range of ways, for example, by targeting genes in key signaling pathways (E. C. Lai et al. 2005)(Martello et al. 2007)(Flynt et al. 2007).

miRNA can also broadly control expression programs during development by targeting general regulators of splicing (Makeyev & Maniatis 2008), and transcription, such as homeotic proteins of the Hox family, which control body patterning in animals (Hornstein et al. 2005)(Mansfield et al. 2004)(Stark et al. 2007)(Stark et al. 2008); More recently, miRNA have also been shown to control DNA methylation in mouse embryonic stem cells by targeting regulators of DNA methyltransferases, with large impacts on gene expression and telomere-length homeostasis (Benetti et al. 2008)(Sinkkonen et al. 2008). Although miRNAs are thought as repressors, a potentially general role for miRNAs as direct gene activators has been defined (Vasudevan et al. 2007) opening a new point of view about miRNAs functions.

Short interfering RNAs (siRNAs) are ~21 nt long, similar in size to miRNAs, but derived from double-stranded RNA precursors (either bimolecular RNA:RNA duplexes or from long internal hairpins) directly processed by Dicer (Farazi et al. 2008). They have perfect matches with their target RNAs, which are directed for degradation. siRNAs can be produced from RNA transcribed in the nucleus (endogenous siRNAs), otherwise can be virally derived or experimentally introduced as chemically synthesized dsRNA (exogenous siRNAS)(Farazi et al. 2008). Endogenous siRNAs can be further subdivided in subclasses on the bases of their biogenesis. siRNA are involved in anti-viral defence, transposon silencing, chromatin remodeling and post-transcriptional gene regulation through Argonaute-mediated cleavage of target transcripts (reviewed in (Carthew & Sontheimer 2009)(Ghildiyal & Zamore 2009). Endogenous siRNAs are not restricted to nematodes, as originally believed (Farazi et

al. 2008), but have been also identified in mammals and insects (John S Mattick & Makunin 2005). The exposure of mammalian cells to long dsRNA induces an antiviral interferon response that leads to apoptosis. For such reason endogenous siRNAs are thought to play an important role in defending genomes against transgenes and trasposons, as well as against foreign nucleic acids, such as viruses (Dorsett & Tuschl 2004).

Piwi-interacting RNAs (piRNA) are ~26-31 nt in length, characterized by their association with Piwi protein family (O'Donnell & Boeke 2007). This class of ncRNAs is mainly found in clusters of repetitive sequences throughout the genome (Aravin et al. 2006)(Brennecke et al. 2008). piRNAs are not processed in a precise manner since tens of thousands of distinct piRNAs generated from the 50 to 100 defined primary transcripts (T. Watanabe et al. 2006). Several lines of evidence indicate that the function of these ncRNAs is to silence retrotransposable elements in the genome (Aravin et al. 2006)(Brennecke et al. 2008). piRNAs have also been shown to directly regulate DNA methylation in mice germ cells, indicating that this class of ncRNAs might have important function in epigenetic mechanisms (Brennecke et al. 2008)(Kuramochi-Miyagawa et al. 2008). Although the targets of piRNAs and their mechanism of action are largely unknown, some reports demonstrated their role in spermatogenesis and transposons regulation in mammal (O'Donnell & Boeke 2007).

Small nucleolar RNAs (snoRNAs) are ~60-300 nt trans-acting RNA molecules that function as guides of specific protein complexes that perform nucleotide modifications (methylation and pseudouridylation) on target RNAs by base-pairing near modification sites (Bachellerie et al. 2002). Over thousands of snoRNAs have been described from many organisms (J. Xie et al. 2007). Most snoRNAs are processed from introns of precursor transcripts, both protein-coding and non-protein-coding (Bachellerie et al. 2002). Originally, these RNAs were thought to have mainly housekeeping functions in the modification of infrastructural RNAs such

as rRNAs e tRNAs. However, many snoRNAs have no recognized targets and some are expressed in a tissue-, developmental- and even clade-specific fashion in different eukaryotes, especially in the brain (Aspegren et al. 2004)(Jian-Hua Yang et al. 2006). Some appear to play a role in the regulation of physiologic processes, including feeding and growth in mice (Ding et al. 2008), and in the regulation of alternative splicing at sequences that are also subject to RNA editing (Kishore & Stamm 2006).

1.4.2. Non classical ncRNAs:

At structural level this ncRNAs can be classified in two main categories that are:

- Small ncRNA, shorter than 200 bases
- Long ncRNAs, longer than 200 bases

Small ncRNA. Several classes of small transcripts, associated to termini of polypeptide-encoding genes have been detected, by a variety of methodological tools, including RNA hybridization to tiling arrays and large scale sequencing of RNA. Literature in this field is extremely large and precise relationships among these classes are often still to be clearly defined.

PASRs and TASRs. Multiple transcripts at the 5' boundaries of genes were originally reported by Carninci et al. (Piero Carninci et al. 2006), including unstable IRNAs, postulated to be involved in regulation of gene expression (C. A. Davis & 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 (Zhenyu Xu et al. 2009)(Berretta & Morillon 2009).

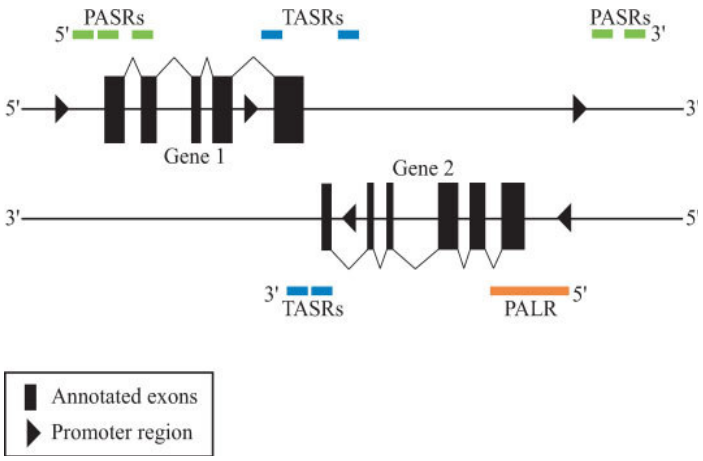


Fig. 1.11. A schematic representation of three novel RNA classes in a genomic context of protein-coding genes. promoter-associated short RNAs (PASRs; green), 3′ termini-associated short RNAs (TASRs; blue), and promoter-associated long RNAs (PALRs; orange)(Kapranov et al. 2007).

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,

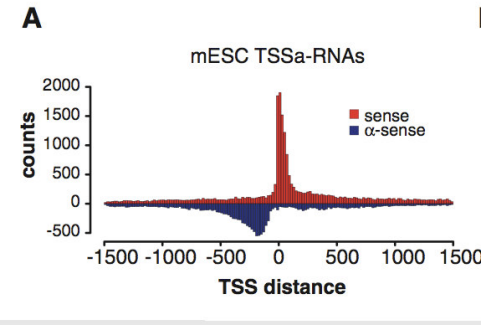
density of all PASRs and antisense TASRs positively correlates with expression levels of the associated gene; the vast majority of silent genes have no associated PASR/TASR. This points to a possible their involvement in fine transcriptional regulation of the associated polypeptide-encoding genes. It was known that siRNAs directed to promoter regions can have a regulatory impact, sometimes silencing (K. V. 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).

tiRNAs. Transcription initiation RNAs (tiRNAs) were discovered by deep sequencing of human, chicken and Drosophila transcriptomes (Taft et al. 2009). They have a modal lenght of 18 nt, are sense oriented and map within -60 to +120 nt of TSS, being preferentially associated with G+C-rich promoters. tiRNAs 5's are clustered at +10 to +30nt, suggesting that they are processed. tiRNAs show some similarities with the previously described PASRs, including low abundance, a distribution shifted to the 3' of the TSS, correlation with bidirectional transcription at particular promoters, and some association with highly expressed genes. However, there is a significant difference in size between them (18 nt for tiRNAs, 22 to >70 nt for PASRs), and it is unknown whether the two are directly related. tiRNAs are generally, although not exclusively, associated with highly expressed transcripts and sites of RNA polymerase II binding. It has been suggested that tiRNAs may be a signature of stalled or poised RNA Pol II, but the association is not strong. Alternatively, they might be a product of PolII `backtracking'/TFIIS cleavage. This hypothesis is based on previous reports that, at certain promoters, RNA Pol II, arrested at +20

to +32 from the TSS, in order to resume transcription, has to backtrack with the aid of the elongation and transcript cleavage factor TFIIS to approximately +12, which match well to the observed position of tiRNAs.

TSSa-RNAs . Transcription start site-associated RNAs (TSSa-RNAs) are 16 to 30 nt long, with a mean length of 20 nt, associated with more than half of all mouse genes (Seila et al. 2008a). They surround promoters in nonrandom and divergent orientations. Sense TSSa-RNAs map downstream of the associated promoter, overlapping genic transcripts and peaking in abundance between +0 and + 50 nt downstream of the TSS; while a quite similar percentage of TSSa-RNAs map upstream of the TSS and are oriented in the antisense direction relative to their associated genes, peaking between nucleotides -100 and – 300. Based on their direction and position relative to TSS, sense and antisense TSSa-RNAs arise from divergent transcription, defined as nonoverlapping transcription initiation events that proceed in opposite directions from the TSS (Seila et al. 2008a).

Fig. 1.12. Distribution of sense and antisense TSSa-RNAs around TSSs. Histogram of the distance from each TSSa-RNA to all associated gene TSSs (Seila et al. 2008a)



Northern analysis showed that these TSSa-RNAs are subsets of an RNA population 20 to 90 nucleotides in length, found at the majority of highly and moderately expressed genes, and 80% associate with promoters having high CpG dinucleotide frequency (CpG islands).

Examining their local chromatin environment it has been shown that promoter-associated RNAPII and H3K4- trimethylated histones, transcription initiation hallmarks, colocalize at sense and antisense TSSa-RNA positions. More exactly there are two distinct peaks for RNAPII detectable with a spacing of several hundred base pairs. These flanking peaks suggest that divergently paused RNAPII complexes may recruit H3K4 methyltransferase activity to mark active promoter boundaries. In contrast to the dual peaks of RNAPII and H3K4me3 surrounding TSS, H3K79me2, a chromatin mark found over RNAPII elongation regions, is solely enriched in the direction of productive transcription. These observations suggest that although divergent transcription initiation is widespread, productive elongation by RNAPII occurs primarily unidirectionally, downstream of TSSs (Seila et al. 2008a). Sense and antisense TSSa-RNAs with bound RNAPII are found at a large number of mammalian promoters suggesting that divergent initiation by RNAPII at TSSs is a general feature of transcriptional processes (Core et al. 2008b). Because TSSa-RNAs do not represent the 5'end of transcripts, they likely mark regions of RNAPII pausing rather than initiation. Pausing of RNAPII 20 to 50 nt downstream of the TSS has been observed at many genes and is thought to maintain a chromatin structure permissive to transcription initiation (Saunders et al. 2006)(Gilchrist et al. 2008). Instead, the position of paused antisense RNAPII, centers around 250 nt upstream of the TSS, as inferred by the presence of bound RNAPII and antisense short RNAs colocalizing at this location (Seila et al. 2008a). Considering that chromatin marks associated with elongating RNAPII are only found downstream of TSSs, it appears that antisense RNAPII frequently does not elongate after TSSa-RNA production (Guenther et al. 2007)(Mikkelsen et

al. 2007). This suggests the existence of an undefined mechanism that discriminates between the sense and antisense polymerase for productive elongation.

Several possible functions have been proposed for divergent transcription. First, the act of transcription itself could be crucial for granting access of transcription factors to control elements that reside upstream of core promoters, possibly by creating a barrier that prevents nucleosomes from obstructing transcription factor binding sites (Mavrich et al. 2008)(Gilchrist et al. 2008). Second, negative supercoiling produced in the wake of transcribing polymerases could facilitate initiation in these regions (Seila et al. 2008b). Third, these short nascent RNAs could themselves be functional, through either Argonaute-dependent (Han et al. 2007) or –independent (Xiangting Wang et al. 2008) pathways.

NRO-RNAs. Nuclear run-on RNAs (NRO-RNAs) were identified as a group of small RNAs generated by human promoters bound by RNAPII (Core et al. 2008a). These short ncRNAs might have a function in promoter activation and transcription orientation similar to TSSa-RNAs.

PROMPTs. Promoter upstream transcripts (PROMPTs) are unstable polyadenylated RNAs, whose identification by RNA hybridization to tiling arrays was allowed by inhibiting the exosome (Preker et al. 2008). They map to mammalian promoter regions (-2.5 to -0.5 from the TSS) and are both sense and antisense oriented. Their synthesis requires the downsream TSS and is positively correlated with expression levels of the associated gene.

Long ncRNAs

Long ncRNAs (lncRNAs), generally >200 and some >100 kb, are mRNA-like, non-protein-coding RNAs that are pervasively transcribed throughout eukaryotic genomes.

Tiling array studies of the human genome, for example, revealed that the majority of transcription, at least 80%, occurs as long ncRNAS, often overlapping with, or interspersed between multiple protein-coding and non-coding transcripts (Kapranov et al. 2007).

Recently, it has been developed a new approach for identifying large non-coding RNAs based on a distinctive chromatin signature that marks actively transcribed genes (Guttman et al. 2009). Genes actively transcribed by RNA polymerase II (Pol II) are marked by 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). This distinctive signature, indicative of transcriptional regulation and high levels of expression, could be referred as "K4-K36 domain". Searching for the "K4-K36 domain" in genome-wide chromatin-state maps across cell types and eliminating those corresponding to known protein-coding genes, it has been possible to identify many long intergenic non-coding RNAs (lincRNAs) with clear conservation of nucleotide sequence and chromatin structure. Moreover, transcription and processing of these lincRNAs appears to be similar to that for protein-coding genes – including Pol II transcription, 5'-capping and poly-adenylation (Guttman et al. 2009).

Long ncRNAs lack obvious features to allow a priori functional categorization or prediction; unlike protein-coding genes where sequence motifs are usually indicative of function, at least in the biochemical sense, the primary sequences of lncRNAs often contain insufficient information to predict their function. To date it is still difficult to discriminate between long ncRNAs and mRNAs. One of the most fundamental criteria used to distinguish long ncRNAs from mRNAs is open reading frames (ORFs) length. Since short putative ORFs can be

expected to occur by chance within long non-coding sequences, minimum ORF cutoffs are usually applied to reduce the likelihood of falsely categorizing ncRNAs as mRNAs. For instance, the FANTOM consortium originally used a cutoff of 300 nt (100 codons) to help identify putative mRNAs (Okazaki et al. 2002).

Although lncRNAs have frequently been disregarded as artifacts of chromatin remodeling or transcriptional "noise", it seems that some lncRNAs may be precursors for smaller RNAs, but many of which are detected as relatively stable polyadenylated and non-polyadenylated transcripts (J. Cheng et al. 2005). The biological significance of these long ncRNAs is controversial. Despite an increasing number of long ncRNAs have been shown to fulfill a diverse range of regulatory roles - chromatin remodeling (J. Zhao et al. 2008)(Nagano et al. 2008)(Rinn et al. 2007), transcriptional regulation (Feng et al. 2006)(Xiangting Wang et al. 2008), splicing, translation (H Wang et al. 2005), nuclear factor trafficking (A T Willingham et al. 2005), imprinting (Sleutels & Barlow 2002)(Thakur et al. 2004), genome rearrangement (Nowacki et al. 2008) and the integrity of subcellular compartments (Dinger et al. 2009)- the functions of the vast majority remain unknown and untested.

Comparative analyses of mouse long ncRNAs indicate that their promoters, primary sequence, and splice sites are under purifying selection (Ponjavic et al. 2007). Given the tissue- and cell-type specific (Kapranov et al. 2007) and dynamically regulated expression (Timothy Ravasi et al. 2006) of long ncRNAs, it seems, however, that the vast majority of mammalian long ncRNAs are intrinsically functional. Many long ncRNAs have been found to originate from complex loci, in which the ncRNAs are coordinately transcribed with their associated protein-coding transcripts (Engström et al. 2006), and several recent examples of characterized ncRNAs, support a functional relationship between the ncRNA and the associated or related protein-coding gene(s). Therefore, by examining the genomic context of ncRNAs relative to protein-coding

genes of known function, in conjunction with expression data, it may be possible to predict a related role for the associated nonprotein-coding transcript.

By the mechanisms of action, long ncRNAs can be subdivided in *cis*- and *trans*-active. Transcription of lncRNAs that regulate the expression of genes in close genomic proximity is called *cis*-acting regulation, while *trans*-acting regulation occurs when transcription of lncRNAs can target distant transcriptional activators or repressors via a variety of mechanisms.

***Cis*-active lncRNAs.** Unlike proteins and small RNAs, long ncRNAs may remain stably tethered to the site of transcription and can therefore uniquely represent a sequence-specific tag for the locus. In the case of *RepA* and *Xist* RNA, it is not a coincidence that the Repeat A motif is at the very 5'end of both molecules. This enables the RNAs to bind PCR2 co-transcriptionally and hold the RNA-protein complex in place for its exclusive *cis*-action. Once transcribed, it is rapidly degraded, via destabilization motifs at the 3' end that would be revealed once Pol II reached the terminus. In this way the RNA is prevented from diffusing away from the site of its synthesis. Moreover, long ncRNAs command a much larger sequence space and can therefore achieve very precise spatiotemporal control of development. This general RNA-based mechanism can rationally explain how a limited set of chromatin modifiers, which often lack subunits with sequence-specific DNA-binding activity but curiously possess putative RNA-binding domains (Denisenko et al. 1998)(Bernstein et al. 2006), can be directed to the mammalian genome in a spatially and temporally unique manner. These unique properties of long ncRNAs may explain why evolution, at the expense of protein-coding genes, has placed so many ncRNA genes throughout the genome. However, long ncRNAs bound to their birthplace can also serve as "ligands" for specific RNA-binding proteins, so causing allosteric

effects able to modulate their functional state. If ncRNAs are tethered to a specific chromatin region, such allosteric modifications will be mainly confined to the surroundings of this region, so adding a further layer of specificity to gene regulation. Obviously these two functions, of guide and allosteric modulator, are not mutually exclusive and may coexist in the same RNA molecule. In this way, cis-active lncRNAs may act as building blocks of a ncRNA/RNA-binding protein-based strategy of unexpected flexibility, suitable to integrate very complex and articulated transcriptional programs. To date, several categories of established or putatively cis-active lncRNAs may be distinguished.

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)(F. Koch et al. 2008). More recently, two teams showed that enhancer transcription actually is a quite general and genome-wide phenomenon. Kim et al (Tae-Kyung 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. 2010) 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.

X-chromosome inactivation linked cis-active lncRNAs. 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 & Jeannie T 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 & Isaacson 1967)(Rastan & Robertson 1985). To date, at least seven distinct non-coding genes have been found within the Xic and surrounding regions and several have been ascribed specific function during XCI. First of all, there must be chromosome counting, i.e. the determination of whether it possesses one (XY) or two Xs (XX) and whether it should therefore initiate XCI. An X: autosome (A) ratio of 1 trigger the XCI cascade (Kay et al. 1994)(Boumil & J T 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)(Jeannie T 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 (Jeannie T Lee 2002). 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* (Jeannie T 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 collinear 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 (J. 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 (J T Lee et al. 1999). 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 (J. 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 & 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 & 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 & 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 (Xiangting Wang et al. 2008).

***Trans*-active lncRNAs.** The evidence that knocking down some lincRNAs did not affect the expression level of nearby genes suggests that these lincRNAs are not likely to function via a *cis*-acting mechanism. Rather, it suggests that influence on gene regulation by these lincRNAs is likely exerted by a *trans* mechanism (Khalil et al. 2009). One of the persistent

challenges in the investigation of long trans-acting ncRNAs is that there is no unifying model that can explain their function or mechanism of action, although such models are expected to emerge over the next few years. The relatively few long ncRNAs that have been characterized to date appear to function by diverse mechanisms (Prasanth & Spector 2007). In one model, lncRNAs might recruit and "guide" its protein partners to proper chromosomal destinations. Specific sequences within the lncRNAs could recognize specific chromatin regions via sequence complementarity, therefore bringing the associated proteins to the targeted region. For example, *RepA* and *Xist,* recruit PRC2 to establish local chromatin modifications on the inactive X, *in cis* (J. Zhao et al. 2008) *;* while in the case of lncRNAs recruiting proteins at a distance (Nagano et al. 2008) or *in trans* (Rinn et al. 2007), the tertiary structure of the higher-order of chromatin might help bring distant chromatin regions together. AItrnatively, lncRNAs might induce allosteric structural modifications of their protein partners to either enhance (Feng et al. 2006) or suppress (Xiangting Wang et al. 2008) their normal activities. LncRNAs might also be able to both "guide" and "modify" its protein partner(s) during the same biological process (Xiangting Wang et al. 2008). It has been proposed and shown that trans-active lncRNAs may act as cofactors of classical transcription factors or may participate in multisubunit complexes in charge of modulating the epigenetic state of chromatin.

Trans-active lncRNAs as cofactors of transcription factors. One of the initial lines of evidence that ncRNAs could function as transcriptional coregulators to positively or negatively regulate gene transcription was provided by the identification of a ncRNA, termed SRA, in a screen for nuclear receptor coactivators (R B Lanz et al. 1999). This trans-active lncRNA works as docking site promoting assembling of transcription factors and coregulators. Steroid receptor RNA activator (SRA) is a nuclear receptor coregulator that is active as RNA; extensive studies

have demonstrated that SRA-mediated nuclear receptor (NR) coactivation does not require the expression of a SRA protein (R B Lanz et al. 1999). The SRA gene is well conserved across species. This RNA coregulator is thought to act as a scaffold bringing together NRs, coregulators and elements of the cell transcriptional machinery at NR target genes (Colley et al. 2008). In fact, secondary structure predictions suggest the existence of multiple stem-loops within SRA. Some of these stem-loops are critical for SRA's activity (Rainer B Lanz et al. 2002). It seems like multiple RNA substructures working together to effect SRA's overall coactivator function. The exact mechanism by which SRA functions to enhance ligand-dependent transcription of nuclear receptors however remains elusive. The identification of RNA-binding domains (RNA recognition motif, RRM) within multiple coregulators suggest that SRA it itself the target of both corepressor and coactivator molecule binding (Colley et al. 2008). Many proteins identified in screens for coactivators of nuclear receptors and other sequence specific transcription factors contain RNA-binding domain (Auboeuf et al. 2005). These domains have previously been considered to play roles in co-transcriptional mRNA processing (Puigserver & Spiegelman 2003) but the possibility must mow also be considered that these domains function as sensors of ncRNAs that work either *in cis* or *trans*.

Recent reports of transcription factors that bind DNA and RNA with distinct roles have been reported. Three demonstrated examples of this include the following: TFIIIA, a zinc finger-containing transcription factor that binds both 5S rDNA and 5S rRNA (Engelke et al. 1980)(Clemens et al. 1993), tra-1, another zinc finger transcription factor that regulates developmental genes and binds the tra-2 mRNA 3' untranslated region (UTR) (L. E. Graves et al. 1999), and bicoid, a homeodomain-containing transcription factor that regulates developmental genes and suppresses cad mRNA translation by binding to the cad mRNA 3'UTR (Dubnau & G. Struhl 1996). The notion that a DNA-binding transcription factor could recognize a specific RNA raises interesting structural and functional

questions. Little is known concerning the biological roles of these alternatives interactions, but the RNA binding by DNA-binding proteins is more common than currently appreciated. Although the same nucleotide sequence could be present in both targets (DNA and RNA), the folded structures of these nucleic acids are presumably very different.

Recent evidence suggest that the RNA bind in the DNA-binding cleft of the protein; perhaps the folded RNA assume a structure that mimics the three-dimensional array of charges and hydrogen bond presented to the protein-binding surface (Cassiday & L. J. Maher 2002).

Another case of a 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 & John L R 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 the ultraconserved enhancer is transcribed to a 3.8 kb ncRNA, *Evf*-2 and recent reports suggest that it is one of the several hundred ultraconserved sequences located close to key developmental regulators and DNA-binding proteins (Bejerano et al. 2004). *Evf*-2 specifically cooperates with homeodomain protein Dlx-2 to increase the activity of the Dlx-5/6 enhancer in a target and homeodomain-specific manner. The active form of *Evf*-2 is a single-stranded RNA. Thus, Dlx-2 binds and activates the Dlx-5/6 enhancer and cooperates with the *Evf*-2 ncRNA. Whether Dlx-2, binds both DNA and RNA during the cooperative interaction, or whether *Evf*-2 sequesters a transcriptional inhibitor independent of binding to Dlx-2 directly, remains to be determined. However, the presence of *Evf*-2/Dlx-2 complexes within the nucleus supports a direct role of the *Evf*-2 ncRNA on Dlx-2 transcriptional activity (Feng et al. 2004). *Evf*-2 is a developmentally regulated ncRNAs that affects transcriptional activity

by cooperation and complex formation with a developmentally regulated homeodomain protein rather than by affecting the general transcriptional machinery through interactions with RNA polymerase. This raises the possibility that also other conserved non-coding regions (UCRs) and enhancer sequences could be transcribed to generate lncRNAs capable of self-activation and transcription factor complex formation.

Trans-active lncRNAs as cofactors of chromatin modifier enzymes. The modulation of chromatin structure is one of the major hallmarks of eukaryotes and of gene regulation in multicellular development (Margueron et al. 2005). Chromatin architecture is dynamically altered by DNA methylation and by numerous compound patterns of covalent histone modifications (Kouzarides 2007). Mechanisms by which such modifications are differentially regulated and precisely targeted to tens of thousands of different genomic loci and positions in different cell lineages are mostly unknown. Recently, several lincRNAs have been found to associate with chromatin-modifying complexes and affect 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. Thus, lncRNA HOTAIR serves as a scaffold by providing binding surfaces to assemble select histone modification enzymes, and thereby specify the pattern of histone modifications on target genes (Miao-Chih Tsai et al. 2010). LncRNAs in the HOX loci have been shown to become deregulated during human cancer progression. In particular, altered HOTAIR expression is involved in breast cancer by promoting genomic relocalization of the Polycomb complex and H3K27 trimethylation. Interestingly, an increased HOTAIR expression in epithelial cancer cells reprograms the Polycomb binding profile inducing genome-wide re-targeting of PRC2 to an occupancy pattern more resembling embryonic fibroblasts, leading to altered histone H3 lysine 27 methylation, gene expression, and increased cancer invasiveness and metastasis in a manner dependent on PRC2. Conversely, loss of HOTAIR can inhibit cancer invasiveness, particularly in cells that possess excessive PRC2 activity (Gupta et al. 2010). These findings indicate that lncRNAs have active roles in modulating the cancer epigenome and thus may be important targets for cancer diagnosis and therapy.

To date, the majority of trans-acting lincRNAs have been found to associate with the H3K27 methyltransferase PRC2. 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.

2. MATERIAL AND METHODS

2.1. Embryonic tissue retrieval

Wild type mice (strains CD1 purchased from Harlan-Italy) used in this study were maintained at the SISSA-CBM mouse facility and 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 were harvested from pregnant dames killed by cervical dislocation and put in sterile ice-cold PBS supplemented with 0,6% glucose. Once collected they were dissected in order to recover cortexes from the rest of the brain, always preserved in ice-cold PBS supplemented with 0,6% glucose.

For the DT1 and DT2 expression investigation, cortexes were harvested from embryos at gestational day E11.5, 12.5, 14.5 and 18.5 while the retrieval of embryonic rombencephalon was made at E14.5; tissues have been processed immediately for RNA extraction.

As far as perturbative experiments are concerned, cortexes were harvested at E12.5 and used for cortex precursors cultures, as follows below.

2.2. Cell cultures

Primordial cortexes, dissected from E12.5 mouse embryos were mechanically dissociated to single cells, by gentle pipetting. The suspension was pipetted by a p200 Gilson pipette, till 8 times, avoiding making bubbles. Usually, big pieces of tissue persist, thus it is necessary to wait for two minutes until they go down; and then, the upper opaque

suspension can be harvested. Other medium must be added and the procedure is repeated 4-5 times; supernatants are joined together.
The number of cells is than quantified in a Bürker chamber and then plated at high density 1000 cell/µl in a multiwell plate (FALCON). The cells are cultured as floating neurospheres in a specific medium. Neurospheres are free floating spherical clusters of cells. Its composition is heterogenous: they contain neural stem cells, as well as proliferating progenitor cells of different lineages.

Medium composition:

	Final concentration
DMEM/F12/Glutamax (Invitrogen)	
Glucose 30%	0.6%
N2 100X (Gibco)	1X
BSA10%	0,1%
Heparin 2mg/ml (Sigma)	2µg/ml
Fungizone (Gibco)	0,0025X
Pen/Strept	0,01X
FGF2 100µg/ml (Invitrogen)	0,02µg/ml
EGF 1000 µg/ml (Invitrogen)	0,02µg/ml

Doxycyclin was added to the culture medium at 2µg/ml final concentration in order to obtain the expression of the transgene in the inducible Tet-on system.

Cellular lines cultures were performed for lentiviral production and titration.
HeLa Tet-Off cells were cultered in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% Tet System Approved FBS, Glutamax 200Mm, 100µg/ml of G418. They were used for fluorescent titration.

293T cells were cultured in Iscove's Modified Dulbecco's Medium (IMDM) (1X) supplemented with 10% serum FCS (Sigma). They were used for lentiviral production and quantitative titration.
Both are described below.

2.3. RNA extraction

RNA was extracted from samples using TRIzol Reagent (Invitrogen) according to manufacturer's instructions and resuspended in sterile deionized water.
Agarose gel electrophoresis and spectrophotometric measurements (NanoDrop ND-1000) were employed to estimate quantity, quality and purity of the resulting RNA. In case of RACE and DTs detection experiments, RNA was treated by DNase digestion (RQ1 RNase-free Dnase, Promega), followed by column purification (RNEasy Mini kit, Qiagen); sample quality was furthermore assayed with an Agilent 2100 Bioanalyzer, for a quantitative estimate of RNA degradation.

2.4. Reverse transcription

For the next studies, cDNA was produced by retrotranscription. 1.5 and 3µg RNA were retrotranscribed by SuperScriptIII™ (Invitrogen) according to manufacturer's instructions, in the presence of random hexamers or strand specific primers, for orientation analysis, respectively.

2.5. Quantitative RT-PCR

Quantitation of a chosen sequence within a DNA sample can be accomplished through Real-Time Polymerase Chain Reaction (RT-PCR). The technique follows the general principle of polymerase chain reaction; its key feature is that the amplified DNA is detected as the reaction progresses in *real time.* The SYBR Green technique makes use of a cyanine dye able to complex with double- stranded DNA and emit strong green fluorescence (λmax = 522 nm) after excitation from blue light (λmax = 488 nm). The fluorescence quantifies the amount of doublestranded DNA within the reaction which, if the primer pair has been chosen correctly, will represent target amplification. Fluorescence readings are then plotted against cycle number on a logarithmic scale; a threshold for detection of fluorescence above background is determined. The cycle at which the fluorescence from a sample crosses the threshold is called the cycle threshold, C_t. Its values can be transformed into absolute values comparing the results to a standard curve produced by real-time PCR of serial dilutions of a known amount of DNA. To accurately quantify gene expression, the measured amount of RNA from the gene of interest is divided by the amount of RNA from a housekeeping gene measured in the same sample to normalize possible variations in the amount and quality of RNA between different samples. This normalization permits accurate comparison of expression of the gene of interest between different samples, given that the expression of the reference (housekeeping) gene used in the normalization is very similar across all the samples.

For every PCR reaction 30 µg of corresponding RNA were analyzed. Reaction mix was prepared in a single batch and dispensed in the required wells for every plate. Each reaction was run in triplicate.

Reactions were assembled as suggested by the manufacturer:

iQ SYBR Green Supermix (Bio-Rad)	**(2X)**
Forward primer	**0.5 µM final concentration**
Reverse primer	**0.5 µM final concentration**
DNA template	**30 µg of corresponding RNA**
sterile deionized water	**to 10 µl final volume**

The reactions were run on a Bio-Rad Mini Opticon MJ Mini thermocycler with the following cycle parameters

***Tbp* protocol**

- Incubate at 95°C for 00:03:00
- Incubate at 95°C for 00:00:10
- Incubate at 60°C for 00:00:10
- Incubate at 72°C for 00:00:20
- Plate Read
- Incubate at 80°C for 00:00:01
- Plate Read
- Incubate at 81°C for 00:00:01
- Plate Read
- Go to line 2 for 39 more times
- Melting curve from 55°C to 95°C, read every 1.0°C, hold 00:00:01
- END

***Emx2* protocol**

- Incubate at 95°C for 00:03:00
- Incubate at 95°C for 00:00:10
- Incubate at 65°C for 00:00:20
- Incubate at 72°C for 00:00:20
- Plate Read
- Incubate at 80°C for 00:00:01
- Plate Read
- Incubate at 81°C for 00:00:01

• Plate Read
• Go to line 2 for 39 more times
• Melting curve from 60°C to 95°C, read every 0.5°C, hold 00:00:01
• END

2.6. RACE

Rapid amplification of cDNA ends (RACE) is a procedure for amplification of nucleic acid sequences from a messenger RNA template between a defined internal site and unknown sequences at either the 3' or the 5' -end of the mRNA.
This methodology of amplification with single-sided specificity has been described as "one-sided" PCR (Ohara et al. 1989) or "anchored" PCR (Loh et al. 1989).
Traditionally, cDNA sequence is obtained from clones isolated from plasmid or phage libraries. Frequently these clones lack sequences corresponding to the 5' ends of the mRNA transcripts. The missing sequence information is typically sought by repeatedly screening the cDNA library in an effort to obtain clones that extended further towards the 5' end of the message. The nature of the enzymatic reactions employed to produce cDNA libraries limits the probability of retrieving extreme 5' sequence even from libraries that are very high quality.
Products generated by the 3' and 5' RACE procedures may be combined to generate full-length cDNAs.

3' RACE takes advantage of the natural poly(A) tail in mRNA as a generic priming site for PCR amplification. In this procedure, mRNAs are converted into cDNA using reverse transcriptase (RT) and an oligo-dT adapter primer. Specific cDNA is then directly amplified by PCR using a

gene-specific primer (GSP) that anneals to a region of known exon sequences and the adapter primer that targets the poly(A) tail region. This permits the capture of unknown 3'-mRNA sequences that lie between the exon and the poly(A) tail.

5' RACE, or "anchored" PCR, is a technique that facilitates the isolation and characterization of 5' ends from low-copy messages.
Classic 5' RACE protocols vary slightly in design, but the general strategy remains consistent.
First strand cDNA is synthesized from either total or poly(A) RNA in a reverse transcription reaction
using a gene specific primer (GSP), Reverse Transcriptase and the deoxynucleotide mixture (dNTPs). This can be followed by an RNase Mix treatment (for example a mixture of RNase H, which is specific for RNA:DNA heteroduplex molecules, and RNase T1) and by a purification of unincorporated dNTPs, GSP, and proteins from cDNA.
Homopolymeric tails are then added to the 3' end of the first strand cDNA by tailing with TdT (Terminal deoxynucleotidyl transferase) or by ligation of an oligonucleotide adapter.
Finally, a gene specific primer is used in conjunction with a primer (i.e, Oligo dT-anchor primer) for the added 3' sequence to amplify the sequence between the adapter and the gene specific primer at the 5' end of the cDNA. The oligo dT-anchor primer is a mixture of oligonucleotides carrying a non-T nucleotide (*i.e.* A, C or G) at the 3' end following the dT-stretch. By this means the Oligo dT-anchor primer is forced to bind to the (5') start site of the poly(A)-tail).
Traditional 5' RACE is sometimes successful, but the major limitation of the procedure is that there is no selection for amplification of fragments corresponding to the actual 5' ends of mRNA: all cDNAs are acceptable templates in the reaction. Additionally, the PCR step selects the most efficient amplicons (e.g., the smallest), favoring amplification of less than full-length products.

One variation is constituted by the RNA Ligase Mediated Rapid Amplification of cDNA Ends (RLM-RACE) where total or poly(A) selected RNA is treated with Calf Intestine Alkaline Phosphatase (CIP) to remove free 5'-phosphates from molecules such as ribosomal RNA, fragmented mRNA, tRNA, and contaminating genomic DNA. The cap structure found on intact 5' ends of mRNA is not affected by CIP. The RNA is then treated with Tobacco Acid Pyrophosphatase (TAP) to remove the cap structure from full-length mRNA, leaving a 5'-monophosphate. An RNA Adapter oligonucleotide is then ligated to the RNA population using T4 RNA ligase. The adapter cannot ligate to dephosphorylated RNA because these molecules lack the 5'-phosphate necessary for ligation. During the ligation reaction, the majority of the full length, decapped mRNA acquires the adapter sequence as its 5' end; random-primed reverse transcription reaction and nested PCR then amplifies the 5' end of a specific transcript.

We decided to use the **SMARTer™ RACE cDNA Amplification kit** because it provides better sensitivity, less background and higher specificity. The cornerstone of the SMARTer RACE cDNA Amplification Kit is SMART technology, which eliminates the need for problematic adaptor ligation so it is possible use first-strand cDNA directly in RACE PCR (Chenchik *et al.*, 1998).

SMART technology provides a mechanism for generating full-length cDNAs in reverse transcription reactions (Y. Y. Zhu et al. 2001). This is made possible by the joint action of the SMARTer II A Oligonucleotide and SMARTScribe Reverse Transcriptase (a variant of MMLV RT). The SMARTScribe RT, upon reaching the end of an RNA template, exhibits terminal transferase activity, adding 3–5 residues to the 3′ end of the first-strand cDNA. The SMARTer oligo contains a terminal stretch of modified bases that anneal to the extended cDNA tail, allowing the oligo to serve as a template for the RT. SMARTScribe RT switches templates from the

mRNA molecule to the SMARTer oligo, generating a complete cDNA copy of the original RNA with the additional SMARTer sequence at the end. Since the template switching activity of the RT occurs only when the enzyme reaches the end of the RNA template, the SMARTer sequence is typically only incorporated into full-length, first-strand cDNAs.

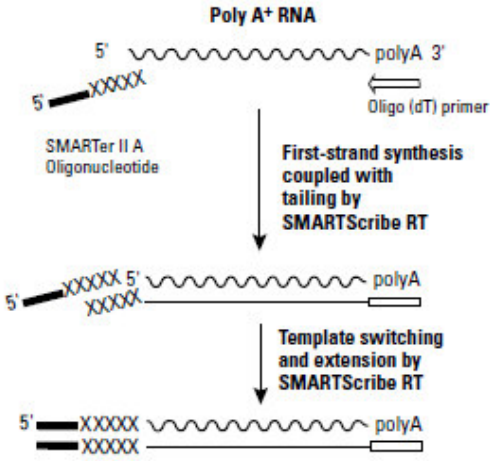


Fig. 2.1. Mechanism of SMARTer cDNA synthesis.

First-strand synthesis is primed using a modified oligo (dT) primer. After SMARTScribe Reverse Transcriptase (RT) reaches the end of the mRNA template, it adds several nontemplate residues. The SMARTer II A Oligonucleotide anneals to the tail of the cDNA and serves as an extended template for SMARTScribe RT.

Following reverse transcription, SMART technology allows first-strand cDNA to be used directly in 5'- and 3'-RACE PCR reactions. Incorporation of universal primer binding sites in a single-step during first-strand cDNA synthesis eliminates the need for tedious second-strand synthesis and adaptor ligation. This simple and highly efficient SMARTer cDNA synthesis method ensures higher specificity in amplifying target cDNA.

Moreover, suppression PCR & step-out PCR techniques are used in combination with SMARTer technology to decrease background amplification in RACE PCR.

In suppression PCR (Siebert et al. 1995), an inverted repeat is incorporated into the ends of DNA sequences to prevent amplification during PCR. The suppression effect occurs when these inverted repeats anneal intramolecularly to form panhandle structures which cannot be amplified by PCR. The SMARTer RACE Kit uses the technique of step-out PCR to add these inverted repeats and thus suppress the amplification of cDNA species that were synthesized by SMARTer II A oligo priming during reverse transcription. Step-out PCR uses a mixture of two primers to incorporate additional sequence at the end/s of template DNA (Matz et al. 1999). One of these primers is exceptionally long and contains the additional sequence as a non-annealing overhang. The overhang sequence is incorporated into template DNA ends in the early rounds of PCR. After overhang addition, the second primer, which is only complementary to the overhang sequence, takes over and serves as an efficient primer for PCR amplification. This short primer is essential because the bulky incorporation primer is inadequate for effective amplification. The short primer is included at a higher concentration than the long primer so that it out-competes the long primer in annealing to template DNA during PCR. In this same manner, the Universal Primer A Mix adds suppression PCR inverted repeat elements to ends of cDNAs in SMARTer RACE. One of the primers in the mix, the "Long" Universal Primer (UP), is identical to the SMARTer sequence at its 3' end and also has a 5' heel of 20 bp which contains the suppression sequence. During the early rounds of RACE PCR, this primer incorporates the suppression sequence on the 5' side of all SMARTer sequences present in the cDNA population. As a result, all cDNAs that were correctly primed by oligo(dT) and have only one SMARTer sequence at the 3' end of the first-strand cDNA will contain one suppression sequence at that end. Conversely, all cDNAs that were primed by the SMARTer II A oligo, and which were

consequently flanked by the SMARTer sequence, become flanked again by the inverted repeat and are subject to suppression PCR. Therefore, cDNAs that have the SMARTer sequence on only one end and the gene specific sequence will be amplified exclusively. As described above, the "Short" UP, which is present at five times the concentration of the Long UP, only contains the 5'-heel sequence of the Long UP, and simply serves as an efficient PCR primer after incorporation of the inverted repeat.

- RNA has been extracted from cortexes harvested from embryos at gestational day E.11.3 directly from the tissue using TRIzol Reagent (Invitrogen), quantified and qualitatively analyzed as described before.
- First-Strand cDNA synthesis has been performed starting from 1µg of RNA for each case.

Procedures have been carried out according to manufacturer's instructions.

The 5'-RACE cDNA has been synthesized using the protocol for first-strand synthesis with random priming as it was not known whether our RNA template lacked or not a polyadenylated tail.

The 3'-RACE cDNA has been synthesized using a traditional reverse transcription procedure, but
with a special oligo(dT) primer. This 3'-RACE primer includes the lock-docking nucleotide positions and also has a portion of the SMARTer sequence at its 5' end.
Lock-docking nucleotides consist in two degenerate nucleotide positions at the 3' end; its function is to position the primer at the start of the poly A+ tail and thus eliminate the 3' heterogeneity inherent with conventional oligo(dT) priming (Borson *et al.*, 1994).

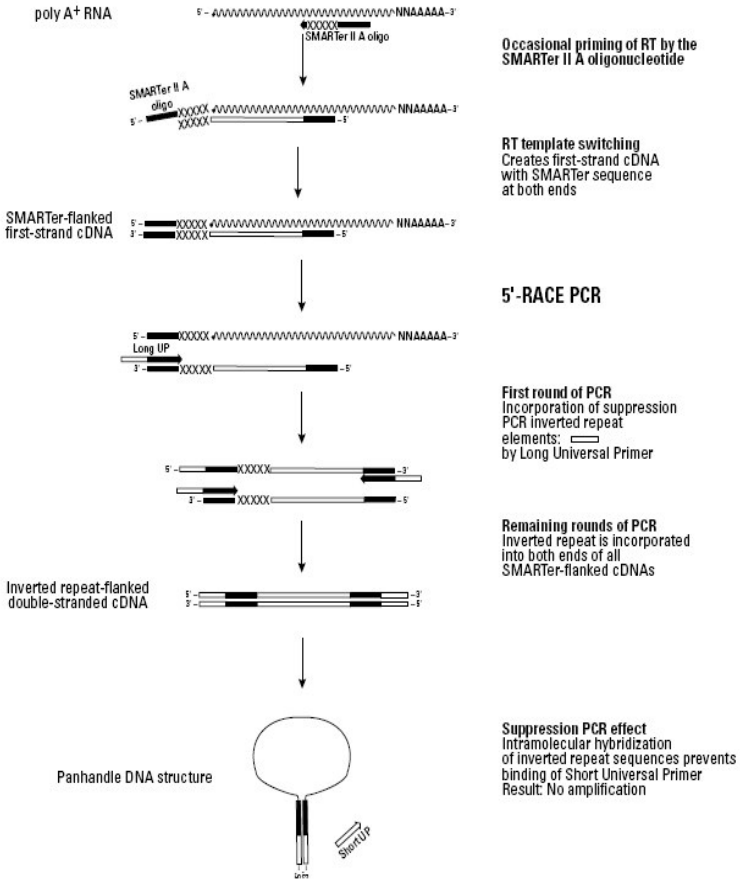


Fig. 2.2. Mechanisms of suppression PCR and step-out PCR.

On occasion, a reverse transcription reaction can be "nonspecifically" primed by the SMARTer II A oligonucleotide. This will result in the synthesis of a cDNA containing the SMARTer sequence at both ends. Through the technique of step-out PCR, suppression PCR inverted repeat elements are incorporated next to all SMARTer sequences. During PCR, these inverted repeats anneal to each other intramolecularly. This rapid first-order reaction out-competes the second-order binding of the Short Universal Primer to the cDNA. As a result, panhandle-like structures, which cannot be amplified, are formed.

By incorporating the SMARTer sequence into both the 5'- and 3'-RACE-Ready cDNA populations, it is possible to prime both RACE PCR reactions using the Universal Primer A Mix (UPM), which recognizes the SMARTer sequence, in conjunction with distinct gene-specific primers.

- 5’ and 3’ RACE PCR

Once RACE-Ready cDNA has been generated, 5’ and 3’ RACE PCR reactions have been performed using the Advantage 2 polymerase Mix. Advantage 2 is comprised of TITANIUM™ *Taq* DNA Polymerase – a nuclease-deficient N-terminal deletion of *Taq* DNA polymerase plus TaqStart® Antibody to provide automatic hot-start PCR (Kellogg et al. 1994) – and a minor amount of a proofreading polymerase. Advantage 2 technology enables to perform long distance PCR (LD PCR) reactions with confidence that products will have high fidelity to the original sequences (Barnes 1994)(S. Cheng et al. 1994). As a result, it is possible to amplify longer templates than were possible in traditional RACE procedures.

Gene specific primers for DT1 and DT2 have been designed.

Reactions were assembled as suggested by the manufacturer with minor modifications.

Component	1 5’-RACE Sample	4 UPM only (– Control)	5 GSP1 only (– Control)	Component	1 3’-RACE Sample
5’-RACE-Ready cDNA (experimental)	2.5 µl	2.5 µl	2.5 µl	3’-RACE-Ready cDNA (experimental)	2.5 µl
UPM (10X)	5 µl	5 µl	–	UPM (10X)	5 µl
GSP1 (10 µM)	1 µl	–	1 µl	GSP1 (10 µM)	–
GSP2 (10 µM)	–	–	–	GSP2 (10 µM)	1 µl
Control 5’-RACETFR Primer (10 µM)	–	–	–	Control 3’-RACETFR Primer (10 µM)	–
H_2O	–	1 µl	5 µl	H_2O	–
Master Mix	41.5 µl	41.5 µl	41.5 µl	Master Mix	41.5 µl
Final Volume	50 µl	50 µl	50 µl	Final Volume	50 µl

Master mix was composed of:

34.5 µl PCR-Grade Water

5.0 µl 10X Advantage 2 PCR Buffer

1.0 µl dNTP Mix (10 mM; in SMARTer RACE or Advantage 2 PCR Kit)

1.0 µl 50X Advantage 2 Polymerase Mix

Program 2 (used for DT1 as GSP Tm=60-70 ˚C; gradient PCR):

- 25 cycles (Total RNA):

  94°C 30 sec

  60-68°C 30 sec

  72°C 3 min*

DT1 primers sequences:

Oligo forward

DT1/RACE1 CCATTGATAGAGAAACCTCGTCTGTTCTGTTCGAGCTAC

DT1/RACE2 CAAAGCGTCCTGTGAGCTTTTGTGAAAG

Oligo reverse

DT1/RACE4 CTGATTTCACTTTCACAAAAGCTCACAGGACGC

DT1/RACE5 CACAGGACGCTTTGTAGCTCGAACAGAACAGACG

Program 1 (used for DT2 as GSP Tm>70 ˚C)

- 5 cycles:

  94°C 30 sec

  72°C 3 min

- 5 cycles:

  94°C 30 sec

  70°C 30 sec

  72°C 3 min

- 25 cycles (Total RNA):

  94°C 30 sec

  68°C 30 sec

72°C 3 min

DT2 primers sequences:

DT2/F TGGGAGTTTCAATCAAATCTTCCAAACAGGTCTGG

DT2/R TGACTACAAATTAGGACCACTAAACTCTTCACTCAG

Then, secondary or "nested" PCR have been performed using the NUP primer (Nested Universal Primer; supplied) and a NGSP (Nested Gene Specific Primer; designed).

- 5 µl of the diluted primary PCR product have been used in place of the RACE-Ready cDNAs.
- 1 µl of the NUP primer (instead of the UPM) and 1 µl of the nested GSP.
- 15- 20 cycles of Program 2 described before.

Primers sequences of NGSP used for DT1 and DT2 nested PCR:

Oligo forward

DT1/RACE2 CAAAGCGTCCTGTGAGCTTTTGTGAAAG

DT1/RACE3 GTGAAAGTGCAAATCAGTTTAAGCAATTATCATACCAGG

Oligo reverse

DT1/RACE5 CACAGGACGCTTTGTAGCTCGAACAGAACAGACG

DT1/RACE6 CAGAACAGACGAGGTTTCTCTATCAATGG

DT2/F-INT TGCCGCTCAAAGTCAATGGCTGTGATAC

DT2/R-INT CAGCATTAGCATCATTAGCATTAGTACGCTTG

RACE PCR fragments of interest have been gel-purified using the Nucleo Trap Gel Extraction Kit (Clontech) and QIAquick Gel Extraction Kit as

indicated in the protocols. Chosen fragments have been cloned directly into a T/A-type cloning vector as the pGEM®-T Easy.

Ligation assays have been performed according to manufacturer's instructions using the T4 DNA Ligase; ratio vector/insert used was 1:3. Reactions have been incubated overnight at 4°C for the maximum number of transformants.

Transformations have been made using One Shot TOP10 chemically competent *E.coli* cells, like previously described. 50-150µl of bacterial suspension was plated on LB Agar containing the antibiotic (75µg/ml) for resistance selection.

Colonies formed were incubated at 37 °C overnight in 4 ml LB medium, containing the same antibiotic used on the plates. On the next day, the extraction of the plasmidic DNA was carried out using the MiniPrep kit (SIGMA-ALDRICH).

Digestion assays have been performed in order to control the presence of inserts of interest and then sequenced.

2.7. Cloning techniques

DT1 (400bp) and DT2 (236bp) fragments, in sense and antisense orientation, have been amplified using standard Polymerase Chain Reaction (PCR) from genomic DNA.

Reactions have been carried out with these reagents:

Genomic DNA

Taq polymerase 0.5U/reaction (GoTaq Promega)

MgCl2 at final concentration 1.5mM (Promega)

Buffer GoFlexi final concentration 1X (Promega)

Deoxynucleotide Triphosphates (dNTPs) final concentration 0.2mM (Promega)

Primers at final concentration 0.2 mM

DT1 protocol and DT2 protocols

- 95°C 5min
- 35 cycles:

  95°C 1min

  Ta°C 30 sec

  72°C 1 min

- 72°C 10 min

DT1 Ta=60°C

DT2 Ta=58°C

Primers sequences:

DT1-S

Fw-DT1/AgeI

5' **CCC**ACCGGTAGCCTTATAAGGCAAGCATTCTGAGAGATCTTC 3'

Rev-DT1/XhoI

5'**CCG**CTCGAGCAGTGTAAAGGGTGGGGGAGATTGTCCTAAATTATGTCC 3'

DT1-OS

Fw-DT1/XhoI

5' **CCG**CTCGAGAGCCTTATAAGGCAAGCATTCTGAGAGATCTTC 3'

Rev-DT1/AgeI

5′ **CCC**ACCGGTCAGTGTAAAGGGTGGGGGAGATTGTCCTAAATTATGTCC 3′

DT2-S

FW-DT2 AgeI

5′ **CCC**ACCGGTCTGGTACAGAGAGGTGAGGGATCAAGTAAGTC 3′

Rev-DT2 BamHI

5′ **CGC**GGATCCGTGAGTCTTGTCTTGAATGGGTCTGCAATGCTG 3′

DT2-OS

FW-DT2 BamHI

5′ **CGC**GGATCCCTGGTACAGAGAGGTGAGGGATCAAGTAAGTC 3′

Rev-DT2 AgeI

5′ **CCC**ACCGGTGTGAGTCTTGTCTTGAATGGGTCTGCAATGCTG 3′

PCR fragments were purified by Qiagen column. Then, restriction enzyme digestion assays have been performed as indicated by the New England Biolabs and Promega producers.

Vector and PCR fragments have been digested at the temperature indicated as optimal for the enzyme for two hours, in order to provide the ligation assay.

After the digestion, vector and inserts have been purified and precipitated with phenol/chloroform/isoamilic alcohol (25:24:1) according to protocol procedures.

Ligation reactions have been performed using the Ligafast Rapid DNA Ligation System (Promega) designed for the efficient ligation of cohesive-ended DNA inserts into plasmid vectors in just 5 minutes. Rapid ligation is based on the combination of T4 DNA Ligase with a unique 2X rapid Ligation Buffer; 3U of enzyme for 100 ng vector and the ratio vector/insert used is 1:3.

Reaction has been incubated at room temperature for 5 minutes as indicated for cohesive-ended ligations.

2.8. Transforming Competent cells

All transformation has been made using One Shot TOP10 chemically competent *E.coli* cells.

These cells are preserved in glycerol 10% at a temperature of -80°C. From 25 to 50 µl of this suspension were incubated with no more than 100 ng of DNA (ligation mix) for 30 minutes on ice.

After that, a heat-shock has been made incubating the cells for 30 seconds at 42°C in order to produce temporary openings in the cell membrane through which DNA enters.

Quickly, 500µl of S.O.C. Medium (tryptone 20g/l, yeast extract 5g/l, NaCl 5 mM, KCl 2.5 mM, MgCl2 1 mM, D-glucose 2mM, MgSO4 5 mM) have been added.

Incubation proceeds for 1 hour at 37°C in order to allow the transforming bacteria to express the antibiotic resistant gene before the selective pressure is applied. Then, 50-150µl of bacterial suspension was plated on LB Agar containing the antibiotic (75µg/ml) for resistance selection.

2.9. Purification of plasmidic DNA and gel extraction

Some of the colonies formed on the plate were incubated in 4 ml LB medium, containing the same antibiotic used on the plates before, at 37 °C overnight. On the next day, the extraction of the plasmidic DNA is carried out using the MiniPrep kit (SIGMA-ALDRICH). When a greater amount of plasmidic DNA was required, the QIAGEN Plasmid Maxi Kit (QIAGEN) was used, starting from 500 ml of LB medium with antibiotic, in a shaking incubator at 37 °C per 16 hours.

Extracting DNA procedure contemplates the bacterial lyses in an alkaline environment, followed by an elution on a silicon matrix with an adequate super salf water volume (200-500 µl). The concentration and purity have been estimated through spectrophotometric reading while its integrity by electrophoretic course.

Agarose gels at variable concentrations have been used in order to estimate the size of DNA molecules. Agarose powder has been dissolved in TBE 1X (Tris-borate 0.09 M, EDTA 2mM).
To make DNA visible, ethidium bromide (EtBr) (50 ng/ml) must be added to the gel. It fluoresces under UV light when intercalated into DNA. DNA Samples were eluted in loading buffer 6X (10 mM Tris-HCl pH 7.6, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol, 60 mM EDTA). Electrophoretic courses have been made at constant voltage.
When necessary, DNA fragments of interest were excised after the electrophoretic course. DNA has been extracted and purified using the QIAquick Gel Extraction Kit as indicated in the protocol.

2.10. TET-ON Gene Expression System

To obtain the expression of the transgene in neuroblasts cultures the inducible Tet-On system has been used (Gossen et al. 1995).
In this system, expression is turned on by addition of Doxycycline (Dox).
In the Tet-On system the regulatory protein is based on a "reverse" Tet repressor (rTetR) created by four amino acid changes in TetR (Hillen & Berens 1994)(Gossen et al. 1995).
TetR, a 37-KDa protein, is fused with the C-terminal of Herpes Simplex virus VP16 activation domain (Triezenberg et al. 1988). Addition of VP16 activation domain converts the TetR from a transcriptional repressor to transcriptional activators, and the resulting hybrid protein is known as tetracycline-controlled transactivator (rTA).

The protein resulting from the aminoacid mutation is a "reverse" tTA (rtTA). Another critical component of the system is the vector which expresses the gene of interest under control of the tetracycline-response element or TRE. In this work, it has been used the TRE-tight promoter. It contain a modified TRE upstream of an altered minimal CMV promoter (PminCMVΔ), resulting in further reduced basal expression of the gene of interest. pTRE-Tight can fully minimize background expression in certain cell lines, and is especially useful in cases where background expression is unacceptable (April 2003 Clontechniques).

In the Tet-on system, rtTA binds the TRE and activate the transcription in presence of Dox in a precise and dose-dependent manner.

Tet-On system has several advantages:

- extremely tight on/off regulation
-no pleiotropic effects
-high inducibility and fast response
-high absolute expression level
-well-characterized inducers
-activation rather than repression of a promoter, to control expression.

2.11. Lentiviral transfer vector construction

Lentiviral vectors have great potential as gene therapy vectors because of their ability to transduce several types of target cells independently of their proliferation status both *ex vivo* and *in vivo.* Lentiviral vectors (LV) are replication-defective, hybrid viral particles made by the core proteins and enzymes of a lentivirus, and the envelope of a different virus, most often the vesicular stomatitis virus (VSV).

Lentiviruses have a complex genome. In addition to the structural genes (*gag*, *pol*, and *env*) common to all retroviruses, lentiviruses also contain two essential regulatory (*tat* and *rev*) and several accessory genes

involved in modulation of viral gene expression, assembly of viral particles, and structural and functional alterations in the infected cells.

The lentiviruses replication is mediated, in part, by *cis*-acting viral sequences, which do not encode proteins; most of these are essential for LV functioning and are usually included in the transfer construct (the part of LV which integrates in the host cell genome and encodes the gene of interest). The *trans*-acting sequences encode three groups of proteins: structural, regulatory, and accessory.

Lentiviral vectors are defective for replication, so only the early steps of the lentivirus life cycle (attachment, entry, reverse transcription, nuclear transport, and integration) are maintained.

Since the early steps do not depend on viral protein synthesis, all *trans*-acting genes could be excluded from the transfer vector that encodes only the gene of interest. So, the typical design of lentiviral vectors is based on the removal of all genes not necessary from the HIV-1 genome and on the separating of sequences acting in *cis* from those acting in *trans* (Delenda 2004).

Third-generation lentiviral vectors are produced by cotransfection of four types of plasmids into 293T cells (Dull et al. 1998a). The transfer vector contains all the *cis*-active sequences needed for packaging, reverse transcription, integration and transcription as well as the gene of interest. To improve the efficiency of gene delivery and expression in target cells some modifications have been made. One of these modifications involved inserting the posttranscriptional regulatory element from the genome of the woodchuck hepatitis virus (Wpre) into the 3' end of the transfer vector. The Wpre acts at the posttranscriptional level, by promoting nuclear export of transcripts and/or increasing the efficiency of polyadenilation of the nascent transcript, thus increasing the total amount of mRNA in cells (Zufferey et al. 1999).

The system safety is granted by the providing of the functions in *trans* which are required for the packaging by three distinct and additional plasmids (pMDL, pREV, pVSVG); besides, by the deletion in the enhancer region in the 3' LTR (Long Terminal Repeat) on the transfer vector. The first measure drastically reduces the accidental formation risk of viral genomes ready for replication, thanks to casual recombination. The second one, since the 3'LTR is used as a template to generate both copies of the LTR in the integrated proviral form of the vector, the deletion results in transcriptional inactivation of both LTRs and prevents its mobilization and recombination in transduced cells (SIN vector: self inactivating vector)(Bukovsky et al. 1999).

The expression constructs used for LV production are maintained in the form of bacterial plasmids and can be transfected into mammalian cells to produce replication-defectice virus stocks.

LV are traditionally produced by transient cotransfection of human embryonic kidney 293T cells (a continuous human embryonic kidney cell line transformed by shared Type 5 adenovirus DNA, by transfection with the tsA 1609 mutant gene of SV40 large T antigen and the Neo gene of *E.coli*), because these cells are good DNA recipients in transfection procedures and the backbones of the vector construct contain SV40 origin of replication. The four plasmids used are:

- *pMDLg/pRRE*, encode for proteins GAG-POL;
- *pRSV.REV*, encode for protein Rev;
- *pCCL-SIN18PPT.Prom.EGFPWpre*, transfer vector (selfinactivating)(Antonia Follenzi & Luigi Naldini 2002);
- *pMD2 VSV.G* encode for the envelope protein VSVG (L Naldini et al. 1996)(Dull et al. 1998b) (Zufferey et al. 1998)(A Follenzi et al. 2000);

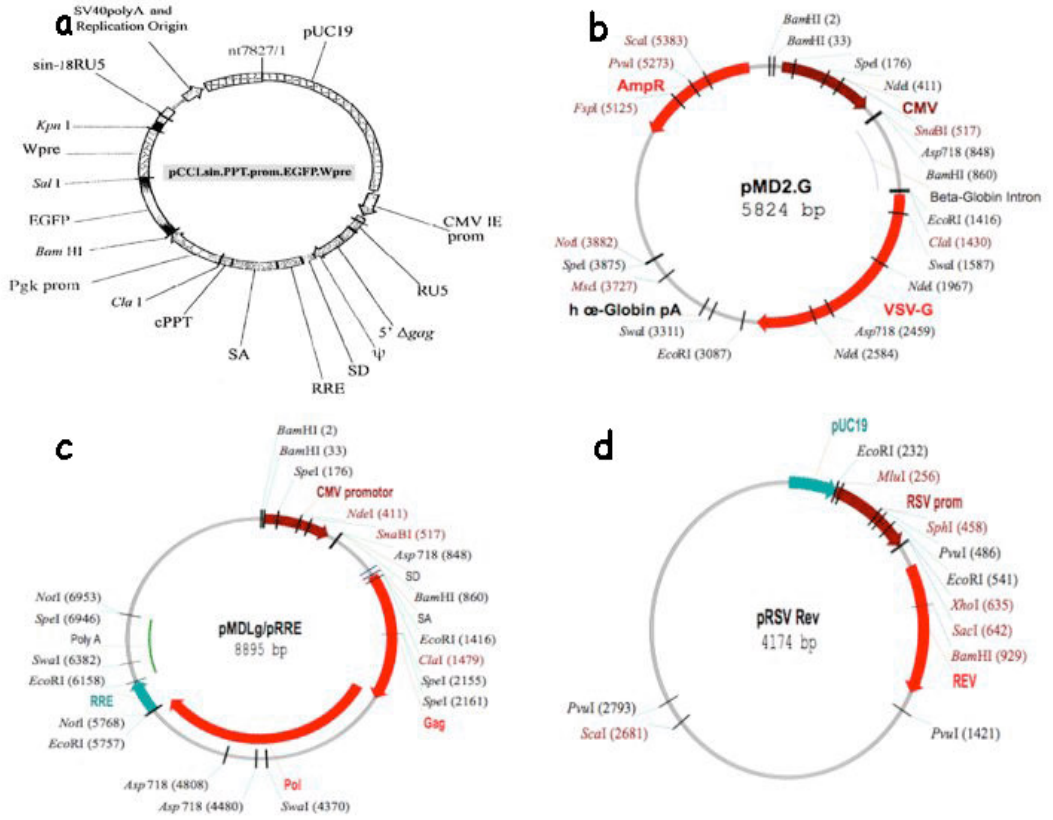


Fig. 2.3. Schematic drawing of the four constructs required for the production of self-inactivating vectors (SIN).

a)The selfinactivating (SIN) transfer construct containing HIV-1 cis-acting sequences and expression cassette for the transgene (enhanced green fluorescent protein or EGFP) driven by the internal promoter. b)The second construct encoding the heterologous protein of the envelope to pseudotype the vector, the protein G of the vesicular stomatitis virus (VSVG) under the control of the CMV promoter. c)The construct expressing the *gag* and *pol* genes driven by the CMV promoter. d)The construct for the expression of Rev protein under the RSV promoter.

2.12. Transient lipofection of Hek293T cells for Lentiviral Vectors Production

Approximately 16-24 hours before transfection 7.0*106 cells were plated in 10 cm (Nunc) plates supplied with 7 ml of Iscove's Modified Dulbecco's Medium (IMDM-glutaMAX GIBCO), 10% heat inactivated FCS (Sigma). Low passage cells are used (not more than P12-15).

The plasmid DNA mix is prepared by mixing:

3 µg / plate	ENV plasmid (VSV-G)
5 µg / plate 2.5µg / plate	Packaging plasmid (pMDLg/pRRE or CMV R8.74) pRSV-REV
16-18 µg / plate	Gene Transfer Plasmid/ or empty plasmid as negative control

The plasmid DNA mix is diluted in 1.5ml of medium (without serum and antibiotics)/plate. Another mix, containing 60 µl of Lipofectamine 2000 (Invitrogen) and 1.5ml of medium (without serum and antibiotics) for every plate is prepared. After 5 minutes of incubation at room temperature the two preparations are mixed and incubated for 20 minutes, after which it is possible to add this solution to the plates. 5 hours later the medium is removed and fresh medium is added. From this point Bio-safety level 2 plus (BSL-2+) procedures must be adopted because lentiviral vectors start being produced. After 14-16 hours, the medium is removed and additional 5 ml of medium per plate are added for the lentiviral collection to begin.

24 hours later, the cell surnatants is collected and filtered with a 0.45 µm filter (Corning) to remove the cellular debris, and stored at 4°C. Removed medium is replaced with 5ml of fresh medium/plate.

The same procedure is led after 48 hours. Surnatants after 24 hours and 48 hours are transferred in centrifuge polyallomer bottles (Beckman 357003), equilibrated and loaded in a centrifuge at 50000 RCF, 2 hours

and 30 minutes at 4°C (20500 rpm in a JA 25.50 rotor). After the centrifuge, supernatants were removed by inverting the bottles.
Pellet is resuspended in 200 µl of PBS 1x-MgCl2 –CaCl2 (Gibco). This first lentiviral suspension is used to resuspend also the second pellet. The pooled lentiviral suspension is than aliquotated and immediately frozen in dry ice. Aliquots were stored at -80°C.

2.13. Titering lentiviral vectors

The titer, TU/µl (Transducing Units/Volume Unit) is a measure of virus concentration per 1 µl. The transducing particles can vary within the different preparations.

Lentiviral vectors used in this work have been titered using two methods:

-fluorescence titration (it estimates the number of effective functional integrated virus particles)

- titration by Real Time quantitative PCR (it is an absolute quantification of integrated particles).

2.13.1. Fluorescence titration

EGFP-expressing lentiviral vectors were titrated on HeLa TET-off cells (Clontech) by fluorescence titration. HeLa Tet-Off cells encode the tetracycline—controlled transactivator (tTA) and are resistant to G418 (contain a neomycin-resistance) added to maintain the selection during their propagation. This cell line is designed for the "Tet-Off Inducuble Gene Expression System" thus it can be transfected with vectors containing the gene of interest under a Tet-respnsive promoter like the pTRE-Tight promoter. This leads to the expression of the gene of interest at maximal level without adding Doxycycline.

HeLa Tet-Off cells were cultered in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% Tet System Approved FBS, Glutamax 200Mm, 100µg/ml of G418.

Fluorescent titration consists in estimating the number of productive integrations of a lentiviral vector containing a fluorescent tracing detectable by microscope. Lentiviral vector expression is under control of an active promoter. Such a system results to be extremely dependent on both cell type and the type of promoter of the expression cassette.

The titer, TU/µl (Transducing Units/Volume Unit), is obtained estimating the number of transduced cells/total cells, in the appropriate serial dilution (positivity in the range of 1-15%).

Protocol steps are the follows:

- At day 0, $1*10_6$ cells/well are plated in a 6 well multi-well (Nunc). The number of wells, is proportional to the serial dilution considered, *plus* a supplementary well used for the cell count.
- At day 1, before the infection, cells belonging to one well are counted using a Bürker chamber, as an average of cells in each well. Then it is possible to proceed with the infection. The old medium is removed and additional 500 µl of new medium plus Polybrene (Hexadimethrine bromide; Sigma) are given at a final concentration of 9µg/ml, to maximize the transduction. Then, the lentiviral vectors, diluted in 500 µl of medium are added to each well.
- At day 3, cells are dissociated from each single well and transferred in 6 cm wells (Nunc).

On the next day, plate considered for the titration should contain about 10% of fluorescence cells. Cells are photographed in white and fluorescence, and counted with the Adobe Photoshop CS3 software.

The titer is calculated as follows:

TU/µl = (((N° positive/N°total)*N°at day 1/ µl added)*DF

N° positive = number of green fluorescence cells at day 3

N° total = number of total cells at day 3

N°at day 1 = number of cells at the time of infection

DF= dilution factor

µl added= µl of dilution added

2.13.2. Titration of lentiviral vectors by genomic Real Time PCR

One fluorescence-titrated LTV was included in each PCR titration session and PCR-titers were converted into fluorescence-equivalent titers.

Protocol steps are the follows:

At day 0, 1*106 293T cells are plated in 2 ml of medium (IMDM-10% serum on Nunc multiwell)

At day 1 wells are washed once with PBS1X (Gibco) and 500µl of medium with filtered polybrene (Sigma) at the final concentration of 9µg/ml are added. One of the wells is used for counting the cells in order to provide a valuation of the cell number at the time of infection. 1 and 2 µl of lentiviral suspension are both added to 500 µl of medium but in two different wells. As a mock, one well without lentiviral suspension is used.

At day 3, wells are washed twice with 2 ml of PBS1X and DNA is extracted using the FlexiGene DNA kit (Giagen). At the end the supernatants are kept by centrifugation (13000 rpm for 5 minutes) while the pellets are cells residual. The concentration is estimated through spectrophotometric reading; the final concentration used for real time PCR must be in the range of 100- 300ng/µl

Standards (std) used in the real time PCR reaction are:

Std "0": 40 ng/µl of plasmid pCCL-SIN18PPT.Prom.EGFPWpre linearized (5*109 molecules);

Std 1 to 7 : dilutes std "0" 1:10 in order to have 8 std that covers from 5*109 to 500 molecules;

Protocol used is the follows:

95 °C_ 4 minutes

(94°C_10 sec + 60°C_15 sec+ 72°C_20 sec+ plate read 72°C+ 78°C_1sec plate read) for 39 cycles

Melting curve 58 to 98°C reading every 0.5°C (1sec)

The plate is prepared with the following mix (10µl/well):

Mix #261 (each sample in triplicate)

5µl sybr green (BIO RAD)

0.5µl oligo forward #261 10µM

0.5µl oligo reverse #261 10µM

0.5 µl dna_sample (mock included)

3.5µl H2O

Mix #261 Std

5µl sybr green (BIO RAD)

0.5µl oligo forward #261 10µM

0.5µl oligo reverse #261 10µM

0.5 µl dna_mock

1µl std

2.5 µl H2O

Primer sequences:

#261/F2

5′ GGCAAGCAGGGAGCTAGAACGATTCGCAG 3′

#261/R2

5′ CTTCTGATCCTGTCTGAAGGGATGGTTGTAGCTGTCC 3′

These oligos permit the amplification in the ψ region (packaging sequence) in the lentiviral vector genome.

Calculating the titer:

The number of lentiviral amplicons found in the infected cells, that represents the number of integrated pro-virions, is estimated through the standard titrating curve.

The number of genomic DNA insert in each well is calculated with the formula:

(ng of genomic DNA)/((10^9)*1.8*10^12))= moli of DNA

10^9 is the conversion factor ng to g

1.8*10^12 is the PM of genomic DNA

Number genomic DNA=moli of DNA*6.022*10^23

Titer (TU/µl): (number of pro-virions/number of genomic DNA)*(number of cells)*dilution factor

Then, the ratio of titer between the examined LTV and a LTV at known concentration is calculated.

Dividing the titer of a LTV obtained by real time Pcr, for this ratio, the real titer is thus given.

2.14. Neurospheres infection

Cerebral cortex precursors were cultured as floating neurospheres. As previously described, cortexes, after the dissociation to single cell, were counted and plated at high density, 1000cell/ µl in multiwell plate (FALCON). They were infected by a mix of lentiviral vector (as in the scheme) to a total of 20 m.o.i (molteplicity of infection: number of viruses per cell).

DT1		**DT2**	
1. DT1-S	10 m.o.i	**1.** DT2-S	10 m.o.i
rt-TA	10 m.o.i	rt-TA	10 m.o.i
2. DT1-OS	10 m o.i	**2.** DT2-OS	10 m o.i
rt-Ta	10 m.o.i	rt-Ta	10 m.o.i
3. Neg.control	10 m.o.i	**3.** Neg.control	10 m.o.i
rt-Ta	10 m.o.i	rt-Ta	10 m.o.i

After 48 h the medium was refreshed for doxycycline (2X) and FGF_2/EGF (5X); after 72h the cell were collected for RNA extraction ad successive analysis.

2.15. Statistical analysis

A double normalization has been carried out: first, the data were normalized on Tbp mRNA; then, the resulting data were finally normalized against the experimental negative control. Anova and T-test were used for statistical analysis.

3. RESULTS

3.1. Time course analysis of DT2 and DT1 expression in vivo

Current investigations running in our lab showed that artificial miRNAs targeted against conserved non-coding elements of the mouse *Emx2* locus are able to alter *Emx2* expression levels, apparently by modulating its transcription rates. Among non-coding elements responding to these artificial miRNAs, there are DT2 and DT1, i.e. the two enhancers shown by Theil et al. (Theil et al. 2002) and Suda et al. (Yoko Suda et al. 2010) to master pro-encephalic transcription of *Emx2*. These phenomena obviously imply the capability by these miRNAs to accurately recognize their chromatin targets within the *Emx2* locus. Such recognition might rely on straight pairing of the ssRNA to genomic DNA (Schmitz et al. 2010); alternatively, it might involve nascent non coding transcripts originating from these targets, as miRNA-baits (J. C. Schwartz et al. 2008)(Yue et al. 2010). To cast light on this topic, we assayed endogenous transcription of DT2 and DT1 elements and investigated consequences of their artificial overexpression.

We profiled the developing neural tube for cumulative DT2 and DT1 transcript levels. To reduce results variability stemming from microdissection errors, we pooled biological samples from at least 4 embryos and further run our analysis on duplicates of these pools. Samples were profiled by random-primed RT, followed by quantitative PCR and normalization against *Tbp* mRNA.

Both DT transcripts were specifically detectable in the same structures which express *Emx2* mRNA, such as pallium and its derivatives, and were absent in neural regions devoided of it, such as rhombencephalon. Moreover, as it happens for *Emx2* mRNA, DT2- and DT1-ncRNA levels progressively declined from E11.5 to E18.5. Remarkably, this decline generally anticipated that of *Emx2* mRNA and was much more pronounced (Fig. 3.1).

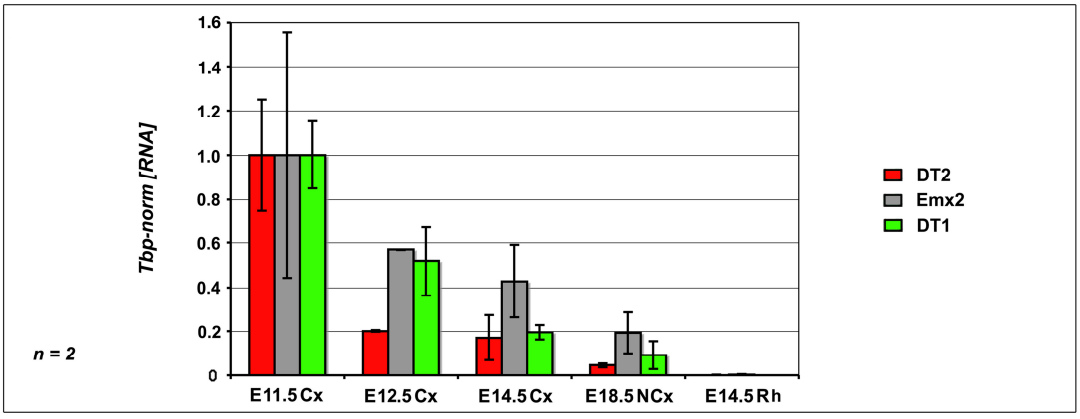


Figure 3.1. DTi-ncRNA expression in the developing CNS: cumulative time course profiles.

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.

3.2. Structural analysis of transcripts.

To reconstruct the transcriptional architecture at around DT elements, we addressed orientation and morphology of DT2 and DT1 transcripts at their peak expression time, i.e. E11.5.

To assess relative abundance of sense and antisense transcripts, per each DT element, we performed strand-specific retrotranscription (RT), PCR quantitation of an amplicon interposed between the two RT primers (AMPL(0)) and, finally, normalization of the PCR signals against *Tbp* mRNA. Noise due to aspecific RT was removed by subtracting the normalized PCR signal yielded by amplicons lying on the 5' end of the RT oligo in order, (AMPL(-1)) and (AMPL(+1)), from normalized (AMPL(0))

(Fig. 3.2). It resulted from this analysis that sense and antisense transcripts apparently coexisted at both DT2 and DT1 regions, albeit at different ratios (5:95 and 89:11, respectively) (Fig. 3.3).

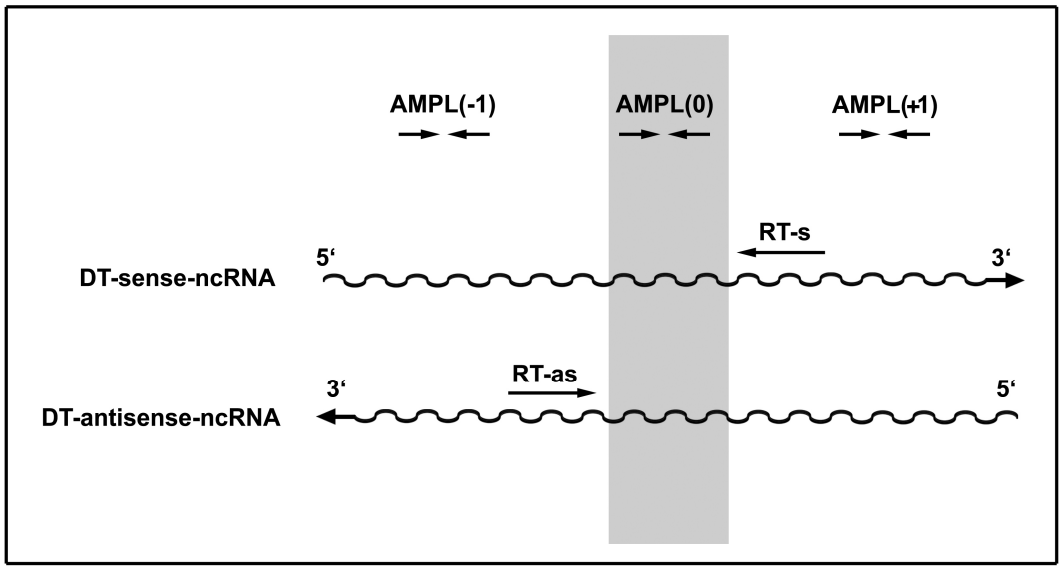


Figure 3.2. Strand-specific qRT-PCR. Per each DT element, retrotranscription (RT) was primed by either two oligos, RT-s and RT-as, the amplicon interposed between the two RT primers (AMPL(0)) was quantified by PCR and, finally, and the PCR signals was normalized against *Tbp* mRNA. Noise due to aspecific RT was removed by subtracting the normalized PCR signal yielded by amplicons lying on the 5' end of the RT oligo in order, (AMPL(-1)) and (AMPL(+1)), from normalized (AMPL(0)).

We performed 5' and 3' rapid amplification of cDNA ends (RACE), starting from the previously characterized DT expressed tags. Because of its simplicity and sensitivity, we used the Smartase-based technology. As for 3' RACE, this was not preceded by any polyA-tailing, so that we enriched the results for naturally polyA-tailed molecules. RACE analysis yielded multiple putative TSSs and 3' termini associated to DT tags. A subset of them was mapped (Fig. 3.3).

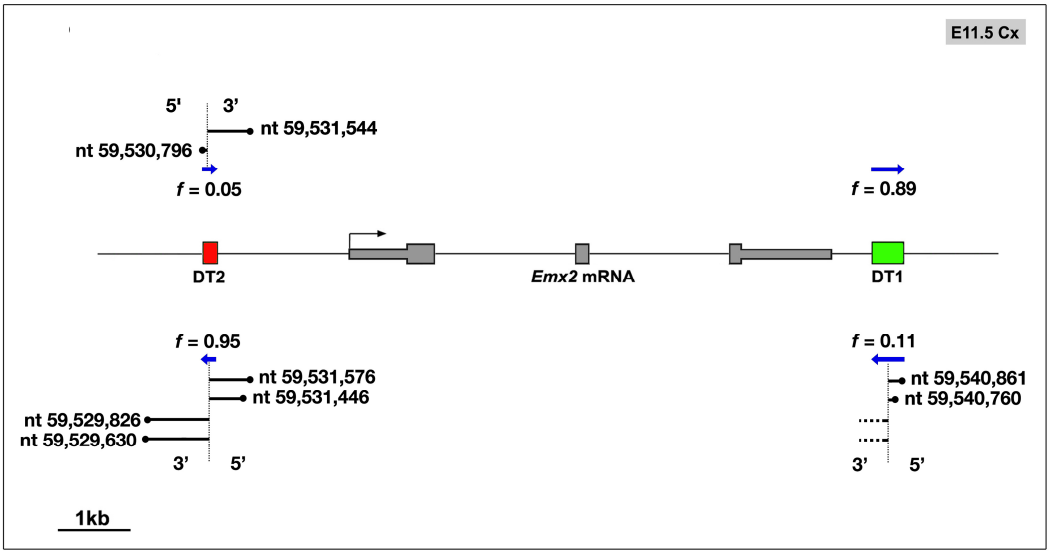


Figure 3.3. DTi-ncRNA expression in the developing CNS: quantitation of transcripts orientation, RACE analysis Relative abundance (f) of sense and antisense DT1 and DT2 transcript tags (oriented blue arrows), as assayed by specific strand-primed qRT-PCR of E11.5 cortico-cerebral total RNA. Associated to each transcript tag, shown are its 5' and 3' RACE extensions, with genomic coordinates of their ends (bullets).

The DT2-antisense expressed tag gave rise to two 5' extensions, ending >1kb upstream of *Emx2* TSS, at mmu_chr.19_nt59,531,576 and mmu_chr.19_nt59,531,446, TATA-less the former, preceded by a TATA sequence at -20 the latter. The same tag gave also rise to two 3' extensions, around 1kb long, ending at mmu_chr.19_nt59,529,826 and mmu_chr.19_nt59,529,630, both of them provided with canonical polyadenylation sites. The DT2-sense expressed tag yielded one short 5' extension, ending at mmu_chr.19_nt59,530,796 TATA-less, and one 3' extension, around 1kb long, ending at mmu_chr.19_nt59,531,544, not provided with a canonical polyadenylation site. The DT1-antisense expressed tag gave rise to two short 5' extensions, ending at mmu_chr.19_nt59,540,861 and mmu_chr.19_nt59,540,760, both of them TATA-less. The same tag gave also rise one 3' extension, around 1.5kb long, to be still mapped on the mouse genome. The DT1-sense expressed

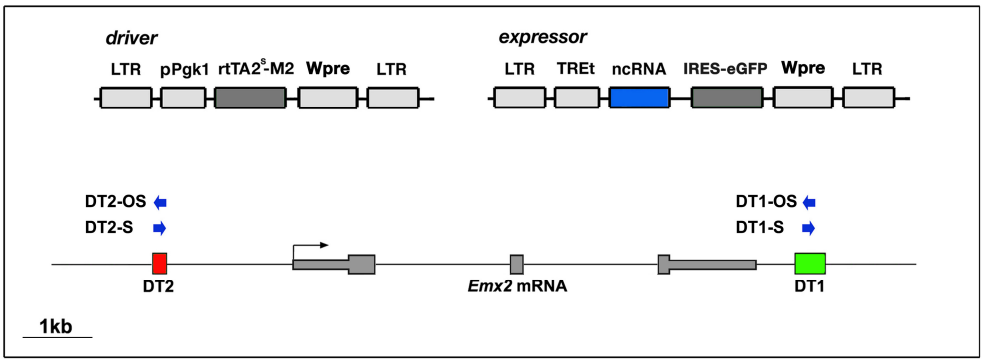
tag yielded a short 5' extension and one 3' 0.8kb long extension, both of them to be still mapped on the mouse genome.

3.3. Overexpression of artificial DT transcripts in embryonic cortico-cerebral precursors

As for biological meaning of endogenous transcription of DT2 and DT1 elements in the developing telencephalon, it is coincevable that transcription *per se* of these elements is crucial to proper tuning of *Emx2* mRNA synthesis and that DT-ncRNAs are simply by-products of such a transcription. Alternatively, the same ncRNAs might be involved in *Emx2* regulation, possibly as mediators, trans- or cis-active, of enhancers function.

To preliminary address this issue, we overexpressed DT2 and DT1 sense and antisense RNA tags in cortico-cerebral precursors, by TetON mediated lentiviral delivery (Fig. 3.4A). Remarkably, in two out of four cases (DT2 antisense and DT1 antisense), this treatment elicited a moderate, but statistically significant downregulation of *Emx2* mRNA (by about 25 and 50%, respectively), suggesting a functional implication of endogenous DT2 and DT1 transcripts in its natural proper tuning (Fig. 3.4B).

A



B

ncRNA	ncRNA coordinates (1)	[*Emx2* mRNA] (2)	*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 miR-NC treated samples

Figure 3.4. Emx2 mRNA modulation by overexpression of DT-ncRNA tags. (A) Lentiviral vectors driving constitutive overexpression of the artificial transactivator rtTA^{2S}-M2 ("driver") and rtTA^{2S}-M2/doxycyclin-dependent overexpression of the ncRNA in order ("expressor"); genomic localization and orientation of ncRNA overexpressed tags. Abbreviations: LTR, lentiviral long terminal repeat; pPgk1, human phosphoglycerkinese 1 promoter; rtTA^{2S}-M2, reverse tetracyclin-regulated trans-activator, type 2S-M2; Wpre, Woodchuck hepatitis virus posttranscriptional regulatory element; TREt, tetracyclin responsive element, tight; ncRNA, non coding RNA; IRES-eGFP, internal ribosome entry site-enhanced green fluorescent protein. (B) Table with: ncRNA tags subject of investigation; their genomic coordinates; relative Emx2 mRNA elicited by their overexpression; p, p-value, as assessed by t Student's test; n, number of independent test.

4. DISCUSSION

We wanted to investigate if and how two main enhancers promoting *Emx2* expression in the developing rostral CNS, DT2 and DT1, are transcribed. Moreover, we intended to preliminarly assay any possible involvement of ncRNAs originating from these enhancers in transcriptional regulation of the gene. We found that both DT2 and DT1 enhancers are transcribed in the deveolipng rostral CNS, mimicking the expression profile of *Emx2* mRNA. Moreover, both enhancers and their surroundings give rise to overlapping sense and antisense transcripts. Finally, lentiviral delivery of DT2 and DT1 antisense ncRNA tags may decrease *Emx2* mRNA levels, up to around 50% of normal levels.

At both DT2 and DT1, we detected small sets of transcripts, divergent and partially overlapping, provided or not with TATA boxes and polyadenylation sites. Such DT2 and DT1 transcripts are reminiscent of the oppositely oriented RNAs originating from activity-dependent neuronal enhancers (eRNAs), described by Kim et al. (Tae-Kyung Kim et al. 2010). They fall into the size range of these eRNAs, but, differently from them, are sometimes provided with polyA sites. DT2 and DT1 transcripts are conversely distinct from far enhancer-associated long non coding RNAs (lncRNAs) reported by Ørom et al. (Ørom et al. 2010), whose transcriptional organization is almost entirely unidirectional.

Remarkably, DT2 and DT1 ncRNAs are spatially restricted to regions of the developing CNS which also express the main *Emx2* mRNA. Moreover, time course progression of these ncRNAs mimes that of *Emx2* mRNA, whose levels progressively decline from the onset of cortical neuronogenesis, at around E11.5, up to perinatal stages. Actually, changes of ncRNA levels precede those of mRNA and, compared to them, are even more pronounced. All this suggests that a causative link may occur among the former and the latter ones. Actually, this scenario

is not novel. A genome-wide positive correlation among levels of ncRNAs and mRNAs transcribed from cis-associated regions has been already described. According to Kim et al. (Tae-Kyung Kim et al. 2010), *levels* of eRNAs, originating from enhancers associated to a large set of polypeptyde-encoding genes, positively correlate with mRNA *transcription levels* at nearby genes. However, in this case no formal proof is provided of a causative link proceding from the former to the latter ones. Conversely, the direction of this link seems to be opposite, as eRNA levels may drop to zero if promoters of associated polypeptide-encoding genes are ablated. A different scenario is reported by Ørom et al. (Ørom et al. 2010), who characterized another set of distal cis-active elements, genome-widely distributed and conserved among vertebrates. These elements, when associated to a minimal heterologous promoter, stimulate transcription from such promoter in an orientation independent way, like classical enhancers. Interestingly, they seem to act via their transcription products. In fact, in order to get cis-activation of their polypeptide-encoding partners: (a) they have to be transcribed (insertion of an artificial polyA site between their TSS and the body of their gene suppresses their transactivating properties); (b) integrity of the RNA product of their transcription is further required (its experimental depletion by siRNAs suppresses such properties as well).

We reasoned that overexpression of short RNA tags corresponding to DT elements, both sense and antisense, could be a fast and efficient approach to preliminalry assess any involvement of endogenous DT transcripts in *Emx2* regulation. To perform these assays, we paid particular care in the choice of the model system as well as in the technology selected to achieve such overexpression. It has been shown that the regulation of genes implicated in developmental processes is closely related to identity and spatio-temporal coordinates of the tissue under examination (Dennis D M O'Leary & Y. Nakagawa 2002). Moreover, it is often strictly dependent on the richness and complexity of cell-cell interactions occurring in living tissues (C. Zhao et al. 2008).

Finally, crucial to proper regulation of gene expression is not the mere presence or absence of specific gene regulators, but the quantitative levels at which they are expressed (Sansom et al. 2009). To properly deal with all these constraints, we performed our tests on primary cultures derived from dissociated embryonic cortices, dissected out at the time of neuronogenesis onset. Cortical precursors were grown at high density (1000 cells/microliter), as floating neurospheres, under a standard FGF2/EGF cocktail. In fact, Emx2 is specifically expressed by proliferating neural precursors, fading out in their neuronal progenies, and these culture conditions promote the retention by cortical precursors of their intermitotic state, largely recreating the richness of cell-cell interactions which occur in the natural *Emx2* expression domain. As for gene delivery, we ruled out somatic electroporation, which - even if fast and cheap - is paradoxically too efficient and does not allow for reproducible quantitative overexpression levels. We also excluded chemotransfection-based methods, poorly effective on embryonic CNS precursors. We delivered our "non coding genes" by recombinant lentiviruses, able to reliably transduce every kind of mammalian cell, with a reproducibly small number of transgene copies. Moreover, we put these genes under the control of a doxycyclin-modulatable system, TetON, allowing for further accurate quantitative control of transgene expression and perspectively more subtle temporal articulation of the tests.

By means of these tools, we found that overexpression of antisense transcript tags corresponding to DT2 and DT1 is able to decrease *Emx2* mRNA levels by up to 50%. These results, likely to originate from functional intereference of endogenous non coding DT transcripts, may be explained by distinct mechanisms. First, it is possible that they arose by "dominant negative" effects. In other words, the short artificial *antisense* tags we expressed might have mimicked key domains of endogenous *antisense* full length DT transcripts, so competing with them for the interaction with unknown factors promoting *Emx2* mRNA transcription. Stem-and-loop moieties often mediate specific RNA-protein interactions:

in this respect, the remarkable thermodynamical stability of stem-and-loop-rich tridimensional structures yielded in silico by both DT2 and DT1 antisense tags, might support this hypothesis (Fig. 4.1). Second, endogenous DT *sense* transcripts might be the true RNA molecules promoting *Emx2* mRNA synthesis. In such a case, the short artificial tags we overexpressed might have acted by chelating their endogenous *sense* counterparts and/or paving the way to their Dicer-mediated degradation (T. Watanabe et al. 2008). It should be possible distinguish between these two possibilities. If short antisense tags work as dominant negative, then siRNA-mediated depletion of endogenous full-length *antisense* transcripts should faithfully reproduce their effects. Should they work by antagonizing their endogenous *sense* partners, their effect should be mimicked by siRNAs against such partners. Remarkaby, these strand-specific siRNA assays should perform equally well, regardless of the cis- or trans-active properties of the endogenous molecule subject of investigation. Finally, should these siRNA tests not to reproduce downregulation of Emx2, this would suggest that the effect of our artificial antisense RNAs was due to different mechanisms, including a possible direct interaction of them with DT2 and DT1 chromatin.

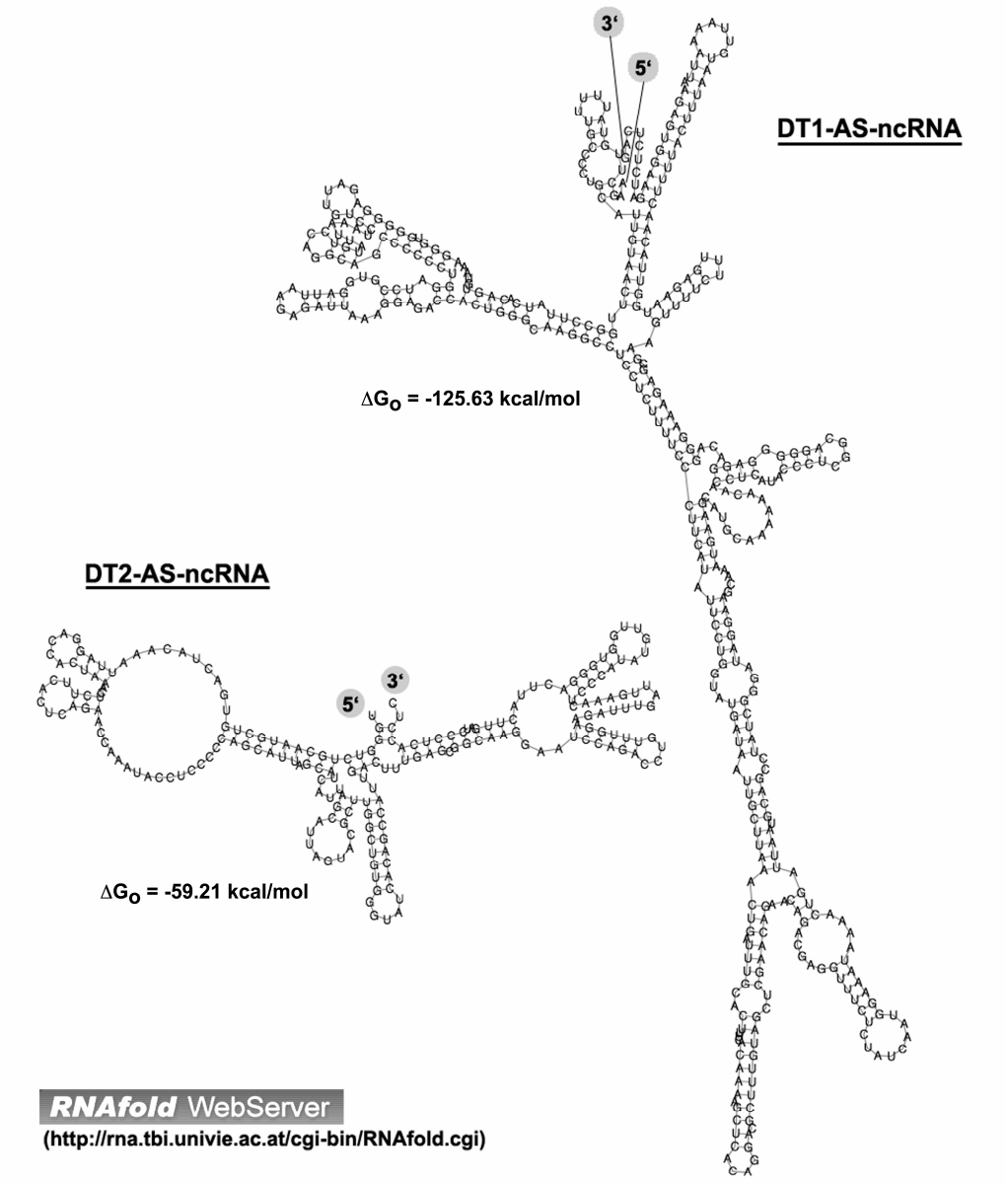


Figure 4.1. Presumptive tridimensional structure of DT2-AS-ncRNA and DT1-AS-ncRNA.
Reconstructed by RNA folder®, at http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi

5. FUTURE AIMS

They include:

(1) Completing fine characterization of DT1 RACE products

(2) Defining the actual impact of endogenous DT transcript(s) on *Emx2* transcription

(3) Studying necessity and sufficiency of such transcripts for *Emx2* modulation, paying special emphasis on cis- and/or trans-requirements.

6. REFERENCES

1970. Embryonic vertebrate central nervous system: revised terminology. The Boulder Committee. *The Anatomical Record*, 166(2), 257-261.

Agami, R. & Bernards, R., 2000. Distinct initiation and maintenance mechanisms cooperate to induce G1 cell cycle arrest in response to DNA damage. *Cell*, 102(1), 55-66.

Agathon, A., Thisse, C. & Thisse, B., 2003. The molecular nature of the zebrafish tail organizer. *Nature*, 424(6947), 448-452.

Allendoerfer, K.L. & Shatz, C.J., 1994. The subplate, a transient neocortical structure: its role in the development of connections between thalamus and cortex. *Annual Review of Neuroscience*, 17, 185-218.

Alvarez-Buylla, A., García-Verdugo, J.M. & Tramontin, A.D., 2001. A unified hypothesis on the lineage of neural stem cells. *Nature Reviews. Neuroscience*, 2(4), 287-293.

Amaral, P.P. & Mattick, J.S., 2008. Noncoding RNA in development. *Mammalian Genome: Official Journal of the International Mammalian Genome Society*, 19(7-8), 454-492.

Anderson, S.A. et al., 2002. Distinct origins of neocortical projection neurons and interneurons in vivo. *Cerebral Cortex (New York, N.Y.: 1991)*, 12(7), 702-709.

Angevine, J.B. & Sidman, R.L., 1961. Autoradiographic study of cell migration during histogenesis of cerebral cortex in the mouse. *Nature*, 192, 766-768.

Aravin, A. et al., 2006. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*, 442(7099), 203-207.

Aspegren, A. et al., 2004. Novel non-coding RNAs in Dictyostelium discoideum and their expression during development. *Nucleic Acids Research*, 32(15), 4646-4656.

Assimacopoulos, S., Grove, E.A. & Ragsdale, C.W., 2003. Identification of a Pax6-dependent epidermal growth factor family signaling source at the lateral edge of the embryonic cerebral cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 23(16), 6399-6403.

Auboeuf, D. et al., 2005. A subset of nuclear receptor coregulators act as coupling proteins during synthesis and maturation of RNA transcripts. *Molecular and Cellular Biology*, 25(13), 5307-5316.

Bachellerie, J.P., Cavaillé, J. & Hüttenhofer, A., 2002. The expanding snoRNA world. *Biochimie*, 84(8), 775-790.

Bachler, M. & Neubüser, A., 2001. Expression of members of the Fgf family and their receptors during midfacial development. *Mechanisms of Development*, 100(2), 313-316.

Barbier, E.L. et al., 2002. Imaging cortical anatomy by high-resolution MR at 3.0T: detection of the stripe of Gennari in visual area 17. *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 48(4), 735-738.

Barnes, W.M., 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates. *Proceedings of the National Academy of Sciences of the United States of America*, 91(6), 2216-2220.

Bayer, S.A. et al., 1991. Cell migration in the rat embryonic neocortex. *The Journal of Comparative Neurology*, 307(3), 499-516.

Beddington, R.S. & Robertson, E.J., 1999. Axis development and early asymmetry in mammals. *Cell*, 96(2), 195-209.

Bejerano, G. et al., 2004. Ultraconserved elements in the human genome. *Science (New York, N.Y.)*, 304(5675), 1321-1325.

Benetti, R. et al., 2008. A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. *Nature Structural & Molecular Biology*, 15(9), 998.

Bernstein, B.E. et al., 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, 125(2), 315-326.

Berretta, J. & Morillon, A., 2009. Pervasive transcription constitutes a new level of eukaryotic genome regulation. *EMBO Reports*, 10(9), 973-982.

Birney, E. et al., 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, 447(7146), 799-816.

Bishop, K.M., Goudreau, G. & O'Leary, D.D., 2000. Regulation of area identity in the mammalian neocortex by Emx2 and Pax6. *Science (New York, N.Y.)*, 288(5464), 344-349.

Boumil, R.M. & Lee, J.T., 2001. Forty years of decoding the silence in X-chromosome inactivation. *Human Molecular Genetics*, 10(20), 2225-2232.

Bouwmeester, T. et al., 1996. Cerberus is a head-inducing secreted factor

expressed in the anterior endoderm of Spemann's organizer. *Nature*, 382(6592), 595-601.

Brennecke, J. et al., 2008. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science (New York, N.Y.)*, 322(5906), 1387-1392.

Breunig, J.J., Macklis, J.D., and Pakic, P. (2007). Evolving methods for the labeling and mutation of postnatal neuronal precursor cells: a critical review. In Adult neurogenesis, F.H. Gage, G. Kempermann, and S. Hongjun, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 49–80.

Briscoe, J. & Ericson, J., 1999. The specification of neuronal identity by graded Sonic Hedgehog signalling. *Seminars in Cell & Developmental Biology*, 10(3), 353-362.

Britanova, O. et al., 2006. A novel mode of tangential migration of cortical projection neurons. *Developmental Biology*, 298(1), 299-311.

Brodersen, P. et al., 2008. Widespread translational inhibition by plant miRNAs and siRNAs. *Science (New York, N.Y.)*, 320(5880), 1185-1190.

Bukovsky, A.A., Song, J.P. & Naldini, L., 1999. Interaction of human immunodeficiency virus-derived vectors with wild-type virus in transduced cells. *Journal of Virology*, 73(8), 7087-7092.

Calegari, F. & Huttner, W.B., 2003. An inhibition of cyclin-dependent kinases that lengthens, but does not arrest, neuroepithelial cell cycle induces premature neurogenesis. *Journal of Cell Science*, 116(Pt 24), 4947-4955.

Capela, A. & Temple, S., 2002. LeX/ssea-1 is expressed by adult mouse CNS stem cells, identifying them as nonependymal. *Neuron*, 35(5), 865-875.

Carninci, P. et al., 2006. Genome-wide analysis of mammalian promoter architecture and evolution. *Nature Genetics*, 38(6), 626-635.

Carthew, R.W. & Sontheimer, E.J., 2009. Origins and Mechanisms of miRNAs and siRNAs. *Cell*, 136(4), 642-655.

Cassiday, L.A. & Maher, L.J., 2002. Having it both ways: transcription factors that bind DNA and RNA. *Nucleic Acids Research*, 30(19), 4118-4126.

Catalano, S.M. & Shatz, C.J., 1998. Activity-dependent cortical target selection by thalamic axons. *Science (New York, N.Y.)*, 281(5376), 559-562.

Cattanach, B.M. & Isaacson, J.H., 1967. Controlling elements in the mouse X chromosome. *Genetics*, 57(2), 331-346.

Caviness, V.S., 1982. Neocortical histogenesis in normal and reeler mice: a developmental study based upon [3H]thymidine autoradiography. *Brain Research*, 256(3), 293-302.

Caviness, V.S. & Sidman, R.L., 1973. Time of origin or corresponding cell classes in the cerebral cortex of normal and reeler mutant mice: an autoradiographic analysis. *The Journal of Comparative Neurology*, 148(2), 141-151.

Caviness, V.S., Takahashi, T. & Nowakowski, R.S., 1995. Numbers, time and neocortical neuronogenesis: a general developmental and evolutionary model. *Trends in Neurosciences*, 18(9), 379-383.

Cawley, S. et al., 2004. Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell*, 116(4), 499-509.

Cheng, J. et al., 2005. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science (New York, N.Y.)*, 308(5725), 1149-1154.

Cheng, S. et al., 1994. Effective amplification of long targets from cloned inserts and human genomic DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 91(12), 5695-5699.

Chenn, A., Braisted, J., McConnell, S., and O'Leary, D. (1997). Devlopment of the cerbral cortex: mechanisms controlling cell fate, laminar and areal patterning, and axonal connectivity. In Molecular and Cellular Approaches to Neural Devlopment, W.M. Cowan, L. Zipursky, and T. Jessel, eds. (Oxford: Oxford University Press), pp. 440-473.

Chiang, C. et al., 1996. Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature*, 383(6599), 407-413.

Clemens, K.R. et al., 1993. Molecular basis for specific recognition of both RNA and DNA by a zinc finger protein. *Science (New York, N.Y.)*, 260(5107), 530-533.

Colley, S.M., Iyer, K.R. & Leedman, P.J., 2008. The RNA coregulator SRA, its binding proteins and nuclear receptor signaling activity. *IUBMB Life*, 60(3), 159-164.

Consiglio, A. et al., 2004. Robust in vivo gene transfer into adult mammalian neural stem cells by lentiviral vectors. *Proceedings of the National Academy of Sciences of the United States of America*, 101(41), 14835-14840.

Core, L.J., Waterfall, J.J. & Lis, J.T., 2008a. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science (New York, N.Y.)*, 322(5909), 1845-1848.

Core, L.J., Waterfall, J.J. & Lis, J.T., 2008b. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science (New York, N.Y.)*, 322(5909), 1845-1848.

Cremisi, F., Philpott, A. & Ohnuma, S., 2003. Cell cycle and cell fate interactions

in neural development. *Current Opinion in Neurobiology*, 13(1), 26-33.

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

Davis, C.A. & Ares, M., 2006. Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, 103(9), 3262-3267.

Dehay, C. & Kennedy, H., 2007a. Cell-cycle control and cortical development. *Nature Reviews. Neuroscience*, 8(6), 438-450.

Dehay, C. & Kennedy, H., 2007b. Cell-cycle control and cortical development. *Nature Reviews. Neuroscience*, 8(6), 438-450.

Delenda, C., 2004. Lentiviral vectors: optimization of packaging, transduction and gene expression. *The Journal of Gene Medicine*, 6 Suppl 1, S125-138.

Deng, W. et al., 2006. Organization of the Caenorhabditis elegans small non-coding transcriptome: genomic features, biogenesis, and expression. *Genome Research*, 16(1), 20-29.

Denisenko, O. et al., 1998. Point mutations in the WD40 domain of Eed block its interaction with Ezh2. *Molecular and Cellular Biology*, 18(10), 5634-5642.

Di Stefano, L. et al., 2007. Mutation of Drosophila Lsd1 disrupts H3-K4 methylation, resulting in tissue-specific defects during development. *Current Biology: CB*, 17(9), 808-812.

Ding, F. et al., 2008. SnoRNA Snord116 (Pwcr1/MBII-85) deletion causes growth deficiency and hyperphagia in mice. *PloS One*, 3(3), e1709.

Dinger, M.E. et al., 2009. NRED: a database of long noncoding RNA expression. *Nucleic Acids Research*, 37(Database issue), D122-126.

Doetsch, F., Caillé, I. et al., 1999. Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell*, 97(6), 703-716.

Doetsch, F., García-Verdugo, J.M. & Alvarez-Buylla, A., 1999. Regeneration of a germinal layer in the adult mammalian brain. *Proceedings of the National Academy of Sciences of the United States of America*, 96(20), 11619-11624.

Dorsett, Y. & Tuschl, T., 2004. siRNAs: applications in functional genomics and potential as therapeutics. *Nature Reviews. Drug Discovery*, 3(4), 318-329.

Dubnau, J. & Struhl, G., 1996. RNA recognition and translational regulation by a homeodomain protein. *Nature*, 379(6567), 694-699.

Dufour, A. et al., 2003. Area specificity and topography of thalamocortical projections are controlled by ephrin/Eph genes. *Neuron*, 39(3), 453-465.

Dull, T. et al., 1998a. A third-generation lentivirus vector with a conditional packaging system. *Journal of Virology*, 72(11), 8463-8471.

Dull, T. et al., 1998b. A third-generation lentivirus vector with a conditional packaging system. *Journal of Virology*, 72(11), 8463-8471.

Dupin, E. et al., 2007. Neural crest progenitors and stem cells. *Comptes Rendus Biologies*, 330(6-7), 521-529.

Ebisuya, M. et al., 2008. Ripples from neighbouring transcription. *Nature Cell Biology*, 10(9), 1106-1113.

Echevarría, D. et al., 2003. Neuroepithelial secondary organizers and cell fate specification in the developing brain. *Brain Research. Brain Research Reviews*, 43(2), 179-191.

Edlund, T. & Jessell, T.M., 1999. Progression from extrinsic to intrinsic signaling in cell fate specification: a view from the nervous system. *Cell*, 96(2), 211-224.

Engelke, D.R. et al., 1980. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell*, 19(3), 717-728.

Engström, P.G. et al., 2006. Complex Loci in human and mouse genomes. *PLoS Genetics*, 2(4), e47.

Ericson, J. et al., 1995. Sonic hedgehog induces the differentiation of ventral forebrain neurons: a common signal for ventral patterning within the neural tube. *Cell*, 81(5), 747-756.

Farazi, T.A., Juranek, S.A. & Tuschl, T., 2008. The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development (Cambridge, England)*, 135(7), 1201-1214.

Feng, J. et al., 2006. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes & Development*, 20(11), 1470-1484.

Feng, J. et al., 2004. Synergistic and antagonistic roles of the Sonic hedgehog N- and C-terminal lipids. *Development (Cambridge, England)*, 131(17), 4357-4370.

Fernandes, M. et al., 2007. Mutations in the BMP pathway in mice support the existence of two molecular classes of holoprosencephaly. *Development (Cambridge, England)*, 134(21), 3789-3794.

Figdor, M.C. & Stern, C.D., 1993. Segmental organization of embryonic diencephalon. *Nature*, 363(6430), 630-634.

Filipowicz, W., Bhattacharyya, S.N. & Sonenberg, N., 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature Reviews. Genetics*, 9(2), 102-114.

Fishell, G., 1995. Striatal precursors adopt cortical identities in response to local cues. *Development (Cambridge, England)*, 121(3), 803-812.

Flynt, A.S. et al., 2007. Zebrafish miR-214 modulates Hedgehog signaling to specify muscle cell fate. *Nature Genetics*, 39(2), 259-263.

Follenzi, A. et al., 2000. Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nature Genetics*, 25(2), 217-222.

Follenzi, A. & Naldini, L., 2002. Generation of HIV-1 derived lentiviral vectors. *Methods in Enzymology*, 346, 454-465.

Frantz, G.D. & McConnell, S.K., 1996. Restriction of late cerebral cortical progenitors to an upper-layer fate. *Neuron*, 17(1), 55-61.

Fukuda, S. et al., 2003. Two distinct subpopulations of nestin-positive cells in adult mouse dentate gyrus. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 23(28), 9357-9366.

Furuno, M. et al., 2006. Clusters of internally primed transcripts reveal novel long noncoding RNAs. *PLoS Genetics*, 2(4), e37.

Furuta, Y., Piston, D.W. & Hogan, B.L., 1997. Bone morphogenetic proteins (BMPs) as regulators of dorsal forebrain development. *Development (Cambridge, England)*, 124(11), 2203-2212.

Gage, F.H., 2000. Mammalian neural stem cells. *Science (New York, N.Y.)*, 287(5457), 1433-1438.

Gal, J.S. et al., 2006. Molecular and morphological heterogeneity of neural precursors in the mouse neocortical proliferative zones. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 26(3), 1045-1056.

Galli, R. et al., 2002. Emx2 regulates the proliferation of stem cells of the adult mammalian central nervous system. *Development (Cambridge, England)*, 129(7), 1633-1644.

Gangemi, R.M. et al., 2001. Emx2 in adult neural precursor cells. *Mechanisms of Development*, 109(2), 323-329.

Garcia, A.D.R. et al., 2004. GFAP-expressing progenitors are the principal source of constitutive neurogenesis in adult mouse forebrain. *Nature Neuroscience*, 7(11), 1233-1241.

Ghildiyal, M. & Zamore, P.D., 2009. Small silencing RNAs: an expanding

universe. *Nature Reviews. Genetics*, 10(2), 94-108.

Gilchrist, D.A. et al., 2008. NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes & Development*, 22(14), 1921-1933.

Ginger, M.R. et al., 2006. A noncoding RNA is a potential marker of cell fate during mammary gland development. *Proceedings of the National Academy of Sciences of the United States of America*, 103(15), 5781-5786.

Glinka, A. et al., 1998. Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. *Nature*, 391(6665), 357-362.

Gossen, M. et al., 1995. Transcriptional activation by tetracyclines in mammalian cells. *Science (New York, N.Y.)*, 268(5218), 1766-1769.

Götz, M. & Huttner, W.B., 2005. The cell biology of neurogenesis. *Nature Reviews. Molecular Cell Biology*, 6(10), 777-788.

Graves, L.E., Segal, S. & Goodwin, E.B., 1999. TRA-1 regulates the cellular distribution of the tra-2 mRNA in C. elegans. *Nature*, 399(6738), 802-805.

Guenther, M.G. et al., 2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*, 130(1), 77-88.

Gulisano, M. et al., 1996. Emx1 and Emx2 show different patterns of expression during proliferation and differentiation of the developing cerebral cortex in the mouse. *The European Journal of Neuroscience*, 8(5), 1037-1050.

Gunhaga, L. et al., 2003. Specification of dorsal telencephalic character by sequential Wnt and FGF signaling. *Nature Neuroscience*, 6(7), 701-707.

Gupta, R.A. et al., 2010. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, 464(7291), 1071-1076.

Guttman, M. et al., 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*, 458(7235), 223-227.

Hamasaki, T. et al., 2004. EMX2 regulates sizes and positioning of the primary sensory and motor areas in neocortex by direct specification of cortical progenitors. *Neuron*, 43(3), 359-372.

Han, J., Kim, D. & Morris, K.V., 2007. Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. *Proceedings of the National Academy of Sciences of the United States of America*, 104(30), 12422-12427.

Hartfuss, E. et al., 2001. Characterization of CNS precursor subtypes and radial glia. *Developmental Biology*, 229(1), 15-30.

Heins, N. et al., 2001. Emx2 promotes symmetric cell divisions and a multipotential fate in precursors from the cerebral cortex. *Molecular and Cellular Neurosciences*, 18(5), 485-502.

Hillen, W. & Berens, C., 1994. Mechanisms underlying expression of Tn10 encoded tetracycline resistance. *Annual Review of Microbiology*, 48, 345-369.

Hooiveld, M.H. et al., 1999. Novel interactions between vertebrate Hox genes. *The International Journal of Developmental Biology*, 43(7), 665-674.

Hornstein, E. et al., 2005. The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development. *Nature*, 438(7068), 671-674.

Houart, C., Westerfield, M. & Wilson, S.W., 1998. A small population of anterior cells patterns the forebrain during zebrafish gastrulation. *Nature*, 391(6669), 788-792.

Huttner, W.B. & Kosodo, Y., 2005. Symmetric versus asymmetric cell division during neurogenesis in the developing vertebrate central nervous system. *Current Opinion in Cell Biology*, 17(6), 648-657.

Inagaki, S. et al., 2005. Identification and expression analysis of putative mRNA-like non-coding RNA in Drosophila. *Genes to Cells: Devoted to Molecular & Cellular Mechanisms*, 10(12), 1163-1173.

Janowski, B.A. et al., 2007. Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nature Chemical Biology*, 3(3), 166-173.

Jensen, K.F. & Killackey, H.P., 1984. Subcortical projections from ectopic neocortical neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 81(3), 964-968.

Johansson, C.B. et al., 1999. Identification of a neural stem cell in the adult mammalian central nervous system. *Cell*, 96(1), 25-34.

Jones-Rhoades, M.W., Bartel, D.P. & Bartel, B., 2006. MicroRNAS and their regulatory roles in plants. *Annual Review of Plant Biology*, 57, 19-53.

Joyner, A.L., Liu, A. & Millet, S., 2000. Otx2, Gbx2 and Fgf8 interact to position and maintain a mid-hindbrain organizer. *Current Opinion in Cell Biology*, 12(6), 736-741.

Kapranov, P. et al., 2002. Large-scale transcriptional activity in chromosomes 21 and 22. *Science (New York, N.Y.)*, 296(5569), 916-919.

Kapranov, P. et al., 2007. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science (New York, N.Y.)*, 316(5830), 1484-1488.

Kay, G.F. et al., 1994. Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development. *Cell*, 77(5), 639-

650.

Kellogg, D.E. et al., 1994. TaqStart Antibody: "hot start" PCR facilitated by a neutralizing monoclonal antibody directed against Taq DNA polymerase. *BioTechniques*, 16(6), 1134-1137.

Khalil, A.M. et al., 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 106(28), 11667-11672.

Kim, T. et al., 2010. Widespread transcription at neuronal activity-regulated enhancers. *Nature*, 465(7295), 182-187.

Kishore, S. & Stamm, S., 2006. The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science (New York, N.Y.)*, 311(5758), 230-232.

Koch, F. et al., 2008. Genome-wide RNA polymerase II: not genes only! *Trends in Biochemical Sciences*, 33(6), 265-273.

Kohtz, J.D. et al., 1998. Regionalization within the mammalian telencephalon is mediated by changes in responsiveness to Sonic Hedgehog. *Development (Cambridge, England)*, 125(24), 5079-5089.

Kohtz, J.D. & Fishell, G., 2004. Developmental regulation of EVF-1, a novel non-coding RNA transcribed upstream of the mouse Dlx6 gene. *Gene Expression Patterns: GEP*, 4(4), 407-412.

Kouzarides, T., 2007. Chromatin modifications and their function. *Cell*, 128(4), 693-705.

Kuramochi-Miyagawa, S. et al., 2008. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes & Development*, 22(7), 908-917.

Lai, E.C., Tam, B. & Rubin, G.M., 2005. Pervasive regulation of Drosophila Notch target genes by GY-box-, Brd-box-, and K-box-class microRNAs. *Genes & Development*, 19(9), 1067-1080.

Lanz, R.B. et al., 1999. A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell*, 97(1), 17-27.

Lanz, R.B. et al., 2002. Distinct RNA motifs are important for coactivation of steroid hormone receptors by steroid receptor RNA activator (SRA). *Proceedings of the National Academy of Sciences of the United States of America*, 99(25), 16081-16086.

Lee, J.T., Davidow, L.S. & Warshawsky, D., 1999. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nature Genetics*, 21(4), 400-404.

Lee, J.T., 2002. Homozygous Tsix mutant mice reveal a sex-ratio distortion and

revert to random X-inactivation. *Nature Genetics*, 32(1), 195-200.

Lee, J.T., 2005. Regulation of X-chromosome counting by Tsix and Xite sequences. *Science (New York, N.Y.)*, 309(5735), 768-771.

Lee, S.M. et al., 2000. A local Wnt-3a signal is required for development of the mammalian hippocampus. *Development (Cambridge, England)*, 127(3), 457-467.

Leingärtner, A. et al., 2003. Cloning and cortical expression of rat Emx2 and adenovirus-mediated overexpression to assess its regulation of area-specific targeting of thalamocortical axons. *Cerebral Cortex (New York, N.Y.: 1991)*, 13(6), 648-660.

Leingärtner, A. et al., 2007. Cortical area size dictates performance at modality-specific behaviors. *Proceedings of the National Academy of Sciences of the United States of America*, 104(10), 4153-4158.

Loh, E.Y. et al., 1989. Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. *Science (New York, N.Y.)*, 243(4888), 217-220.

Louro, R. et al., 2007. Androgen responsive intronic non-coding RNAs. *BMC Biology*, 5, 4.

Lunyak, V.V. et al., 2002. Corepressor-dependent silencing of chromosomal regions encoding neuronal genes. *Science (New York, N.Y.)*, 298(5599), 1747-1752.

Lupo, G., Harris, W.A. & Lewis, K.E., 2006. Mechanisms of ventral patterning in the vertebrate nervous system. *Nature Reviews. Neuroscience*, 7(2), 103-114.

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

Makeyev, E.V. & Maniatis, T., 2008. Multilevel regulation of gene expression by microRNAs. *Science (New York, N.Y.)*, 319(5871), 1789-1790.

Malatesta, P. et al., 2003. Neuronal or glial progeny: regional differences in radial glia fate. *Neuron*, 37(5), 751-764.

Mallamaci, A. et al., 1998. EMX2 protein in the developing mouse brain and olfactory area. *Mechanisms of Development*, 77(2), 165-172.

Mallamaci, A. et al., 2000. The lack of Emx2 causes impairment of Reelin signaling and defects of neuronal migration in the developing cerebral cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 20(3), 1109-1118.

Mallamaci, A. & Stoykova, A., 2006. Gene networks controlling early cerebral cortex arealization. *The European Journal of Neuroscience*, 23(4), 847-

856.

Manak, J.R. et al., 2006. Biological function of unannotated transcription during the early development of Drosophila melanogaster. *Nature Genetics*, 38(10), 1151-1158.

Mansfield, J.H. et al., 2004. MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. *Nature Genetics*, 36(10), 1079-1083.

Margueron, R., Trojer, P. & Reinberg, D., 2005. The key to development: interpreting the histone code? *Current Opinion in Genetics & Development*, 15(2), 163-176.

Martello, G. et al., 2007. MicroRNA control of Nodal signalling. *Nature*, 449(7159), 183-188.

Martianov, I. et al., 2007. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature*, 445(7128), 666-670.

Mattick, J.S. & Gagen, M.J., 2001. The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. *Molecular Biology and Evolution*, 18(9), 1611-1630.

Mattick, J.S., 2007. A new paradigm for developmental biology. *The Journal of Experimental Biology*, 210(Pt 9), 1526-1547.

Mattick, J.S., 2004. RNA regulation: a new genetics? *Nature Reviews. Genetics*, 5(4), 316-323.

Mattick, J.S. & Makunin, I.V., 2005. Small regulatory RNAs in mammals. *Human Molecular Genetics*, 14 Spec No 1, R121-132.

Matz, M. et al., 1999. Amplification of cDNA ends based on template-switching effect and step-out PCR. *Nucleic Acids Research*, 27(6), 1558-1560.

Mavrich, T.N. et al., 2008. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Research*, 18(7), 1073-1083.

McConnell, S.K., 1995. Constructing the cerebral cortex: neurogenesis and fate determination. *Neuron*, 15(4), 761-768.

McConnell, S.K. & Kaznowski, C.E., 1991. Cell cycle dependence of laminar determination in developing neocortex. *Science (New York, N.Y.)*, 254(5029), 282-285.

McLaughlin, T. & O'Leary, D.D.M., 2005. Molecular gradients and development of retinotopic maps. *Annual Review of Neuroscience*, 28, 327-355.

Meinhardt, H., 1983. Cell determination boundaries as organizing regions for

secondary embryonic fields. *Developmental Biology*, 96(2), 375-385.

Mercer, T.R. et al., 2008. Specific expression of long noncoding RNAs in the mouse brain. *Proceedings of the National Academy of Sciences of the United States of America*, 105(2), 716-721.

Mikkelsen, T.S. et al., 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, 448(7153), 553-560.

Mo, Z. et al., 2007. Human cortical neurons originate from radial glia and neuron-restricted progenitors. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 27(15), 4132-4145.

Molnár, Z. & Blakemore, C., 1995. How do thalamic axons find their way to the cortex? *Trends in Neurosciences*, 18(9), 389-397.

Molyneaux, B.J. et al., 2007. Neuronal subtype specification in the cerebral cortex. *Nature Reviews. Neuroscience*, 8(6), 427-437.

Morey, C. et al., 2004. The region 3' to Xist mediates X chromosome counting and H3 Lys-4 dimethylation within the Xist gene. *The EMBO Journal*, 23(3), 594-604.

Morris, K.V. et al., 2004. Small interfering RNA-induced transcriptional gene silencing in human cells. *Science (New York, N.Y.)*, 305(5688), 1289-1292.

Muñoz-Sanjuán, I. & H-Brivanlou, A., 2001. Early posterior/ventral fate specification in the vertebrate embryo. *Developmental Biology*, 237(1), 1-17.

Muzio, L. et al., 2005. A mutually stimulating loop involving emx2 and canonical wnt signalling specifically promotes expansion of occipital cortex and hippocampus. *Cerebral Cortex (New York, N.Y.: 1991)*, 15(12), 2021-2028.

Muzio, L. et al., 2002a. Conversion of cerebral cortex into basal ganglia in Emx2(-/-) Pax6(Sey/Sey) double-mutant mice. *Nature Neuroscience*, 5(8), 737-745.

Muzio, L. et al., 2002b. Emx2 and Pax6 control regionalization of the pre-neuronogenic cortical primordium. *Cerebral Cortex (New York, N.Y.: 1991)*, 12(2), 129-139.

Nagano, T. et al., 2008. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science (New York, N.Y.)*, 322(5908), 1717-1720.

Naldini, L. et al., 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science (New York, N.Y.)*, 272(5259), 263-267.

Noctor, S.C. et al., 2004. Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nature Neuroscience*, 7(2), 136-144.

Nowacki, M. et al., 2008. RNA-mediated epigenetic programming of a genome-rearrangement pathway. *Nature*, 451(7175), 153-158.

O'Donnell, K.A. & Boeke, J.D., 2007. Mighty Piwis defend the germline against genome intruders. *Cell*, 129(1), 37-44.

Ogawa, M. et al., 1995. The reeler gene-associated antigen on Cajal-Retzius neurons is a crucial molecule for laminar organization of cortical neurons. *Neuron*, 14(5), 899-912.

Ohara, O., Dorit, R.L. & Gilbert, W., 1989. One-sided polymerase chain reaction: the amplification of cDNA. *Proceedings of the National Academy of Sciences of the United States of America*, 86(15), 5673-5677.

Ohnuma, S., Philpott, A. & Harris, W.A., 2001. Cell cycle and cell fate in the nervous system. *Current Opinion in Neurobiology*, 11(1), 66-73.

Okazaki, Y. et al., 2002. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature*, 420(6915), 563-573.

O'Leary, D.D., 1989. Do cortical areas emerge from a protocortex? *Trends in Neurosciences*, 12(10), 400-406.

O'Leary, D.D. & Koester, S.E., 1993. Development of projection neuron types, axon pathways, and patterned connections of the mammalian cortex. *Neuron*, 10(6), 991-1006.

O'Leary, D.D., Schlaggar, B.L. & Tuttle, R., 1994. Specification of neocortical areas and thalamocortical connections. *Annual Review of Neuroscience*, 17, 419-439.

O'Leary, D.D. & Stanfield, B.B., 1989. Selective elimination of axons extended by developing cortical neurons is dependent on regional locale: experiments utilizing fetal cortical transplants. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 9(7), 2230-2246.

O'Leary, D.D.M. & Nakagawa, Y., 2002. Patterning centers, regulatory genes and extrinsic mechanisms controlling arealization of the neocortex. *Current Opinion in Neurobiology*, 12(1), 14-25.

Oppenheim, R.W., Cole, T. & Prevette, D., 1989. Early regional variations in motoneuron numbers arise by differential proliferation in the chick embryo spinal cord. *Developmental Biology*, 133(2), 468-474.

Ørom, U.A. et al., 2010. Long noncoding RNAs with enhancer-like function in

human cells. *Cell*, 143(1), 46-58.

Panchision, D.M. et al., 2001. Sequential actions of BMP receptors control neural precursor cell production and fate. *Genes & Development*, 15(16), 2094-2110.

Pang, K.C., Frith, M.C. & Mattick, J.S., 2006. Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends in Genetics: TIG*, 22(1), 1-5.

Panganiban, G. & Rubenstein, J.L.R., 2002. Developmental functions of the Distal-less/Dlx homeobox genes. *Development (Cambridge, England)*, 129(19), 4371-4386.

Patten, I. & Placzek, M., 2002. Opponent activities of Shh and BMP signaling during floor plate induction in vivo. *Current Biology: CB*, 12(1), 47-52.

Payer, B. & Lee, J.T., 2008. X chromosome dosage compensation: how mammals keep the balance. *Annual Review of Genetics*, 42, 733-772.

Penny, G.D. et al., 1996. Requirement for Xist in X chromosome inactivation. *Nature*, 379(6561), 131-137.

Polleux, F., 2005. Genetic mechanisms specifying cortical connectivity: let's make some projections together. *Neuron*, 46(3), 395-400.

Pombero, A. & Martinez, S., 2009. Telencephalic morphogenesis during the process of neurulation: an experimental study using quail-chick chimeras. *The Journal of Comparative Neurology*, 512(6), 784-797.

Ponjavic, J., Ponting, C.P. & Lunter, G., 2007. Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. *Genome Research*, 17(5), 556-565.

Prasanth, K.V. & Spector, D.L., 2007. Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes & Development*, 21(1), 11-42.

Preker, P. et al., 2008. RNA exosome depletion reveals transcription upstream of active human promoters. *Science (New York, N.Y.)*, 322(5909), 1851-1854.

Puelles, E. et al., 2004. Otx2 regulates the extent, identity and fate of neuronal progenitor domains in the ventral midbrain. *Development (Cambridge, England)*, 131(9), 2037-2048.

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

Qian, X. et al., 1997. FGF2 concentration regulates the generation of neurons and glia from multipotent cortical stem cells. *Neuron*, 18(1), 81-93.

Rakic, P., 1972. Mode of cell migration to the superficial layers of fetal monkey neocortex. *The Journal of Comparative Neurology*, 145(1), 61-83.

Rakic, P., 1974. Neurons in rhesus monkey visual cortex: systematic relation between time of origin and eventual disposition. *Science (New York, N.Y.)*, 183(123), 425-427.

Rakic, P., 1988. Specification of cerebral cortical areas. *Science (New York, N.Y.)*, 241(4862), 170-176.

Rakic, P., 2003. Developmental and evolutionary adaptations of cortical radial glia. *Cerebral Cortex (New York, N.Y.: 1991)*, 13(6), 541-549.

Rallu, M., Corbin, J.G. & Fishell, G., 2002. Parsing the prosencephalon. *Nature Reviews. Neuroscience*, 3(12), 943-951.

Rastan, S. & Robertson, E.J., 1985. X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. *Journal of Embryology and Experimental Morphology*, 90, 379-388.

Ravasi, T. et al., 2006. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Research*, 16(1), 11-19.

Rinn, J.L. et al., 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, 129(7), 1311-1323.

Rossi, C. et al., 2006. Brain-derived neurotrophic factor (BDNF) is required for the enhancement of hippocampal neurogenesis following environmental enrichment. *The European Journal of Neuroscience*, 24(7), 1850-1856.

Royo, H. et al., 2007. Bsr, a nuclear-retained RNA with monoallelic expression. *Molecular Biology of the Cell*, 18(8), 2817-2827.

Rubenstein, J.L. et al., 1998. Regionalization of the prosencephalic neural plate. *Annual Review of Neuroscience*, 21, 445-477.

Ruiz i Altaba, A., 1998. Combinatorial Gli gene function in floor plate and neuronal inductions by Sonic hedgehog. *Development (Cambridge, England)*, 125(12), 2203-2212.

Sansom, S.N. et al., 2009. The level of the transcription factor Pax6 is essential for controlling the balance between neural stem cell self-renewal and neurogenesis. *PLoS Genetics*, 5(6), e1000511.

Sasai, Y., 1998. Identifying the missing links: genes that connect neural induction and primary neurogenesis in vertebrate embryos. *Neuron*, 21(3), 455-458.

Saunders, A., Core, L.J. & Lis, J.T., 2006. Breaking barriers to transcription

elongation. *Nature Reviews. Molecular Cell Biology*, 7(8), 557-567.

Schier, A.F. & Shen, M.M., 2000. Nodal signalling in vertebrate development. *Nature*, 403(6768), 385-389.

Schlaggar, B.L. & O'Leary, D.D., 1991. Potential of visual cortex to develop an array of functional units unique to somatosensory cortex. *Science (New York, N.Y.)*, 252(5012), 1556-1560.

Schmitz, K. et al., 2010. Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes & Development*, 24(20), 2264-2269.

Schwartz, J.C. et al., 2008. Antisense transcripts are targets for activating small RNAs. *Nature Structural & Molecular Biology*, 15(8), 842-848.

Seidl, C.I.M., Stricker, S.H. & Barlow, D.P., 2006. The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *The EMBO Journal*, 25(15), 3565-3575.

Seila, A.C. et al., 2008a. Divergent transcription from active promoters. *Science (New York, N.Y.)*, 322(5909), 1849-1851.

Seila, A.C. et al., 2008b. Divergent transcription from active promoters. *Science (New York, N.Y.)*, 322(5909), 1849-1851.

Shabalina, S.A. & Spiridonov, N.A., 2004. The mammalian transcriptome and the function of non-coding DNA sequences. *Genome Biology*, 5(4), 105.

Shen, Q. et al., 2006. The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. *Nature Neuroscience*, 9(6), 743-751.

Shimamura, K. & Rubenstein, J.L., 1997. Inductive interactions direct early regionalization of the mouse forebrain. *Development (Cambridge, England)*, 124(14), 2709-2718.

Shin, J., Fitzpatrick, G.V. & Higgins, M.J., 2008. Two distinct mechanisms of silencing by the KvDMR1 imprinting control region. *The EMBO Journal*, 27(1), 168-178.

Siebert, P.D. et al., 1995. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Research*, 23(6), 1087-1088.

Simeone, A. et al., 1992. Two vertebrate homeobox genes related to the Drosophila empty spiracles gene are expressed in the embryonic cerebral cortex. *The EMBO Journal*, 11(7), 2541-2550.

Sinkkonen, L. et al., 2008. MicroRNAs control de novo DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. *Nature Structural & Molecular Biology*, 15(3), 259-267.

Sleutels, F. & Barlow, D.P., 2002. The origins of genomic imprinting in mammals. *Advances in Genetics*, 46, 119-163.

Smart, I.H., 1973. Proliferative characteristics of the ependymal layer during the early development of the mouse neocortex: a pilot study based on recording the number, location and plane of cleavage of mitotic figures. *Journal of Anatomy*, 116(Pt 1), 67-91.

Smilinich, N.J. 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. *Proceedings of the National Academy of Sciences of the United States of America*, 96(14), 8064-8069.

Smith, J.L. & Schoenwolf, G.C., 1997. Neurulation: coming to closure. *Trends in Neurosciences*, 20(11), 510-517.

Sone, M. et al., 2007. The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. *Journal of Cell Science*, 120(Pt 15), 2498-2506.

de Souza, F.S. & Niehrs, C., 2000. Anterior endoderm and head induction in early vertebrate embryos. *Cell and Tissue Research*, 300(2), 207-217.

Stanfield, B.B. & O'Leary, D.D., 1985. Fetal occipital cortical neurones transplanted to the rostral cortex can extend and maintain a pyramidal tract axon. *Nature*, 313(5998), 135-137.

Stark, A. et al., 2008. A single Hox locus in Drosophila produces functional microRNAs from opposite DNA strands. *Genes & Development*, 22(1), 8-13.

Stark, A. et al., 2007. Systematic discovery and characterization of fly microRNAs using 12 Drosophila genomes. *Genome Research*, 17(12), 1865-1879.

Stefani, G. & Slack, F.J., 2008. Small non-coding RNAs in animal development. *Nature Reviews. Molecular Cell Biology*, 9(3), 219-230.

Stern, C.D., 2001. Initial patterning of the central nervous system: how many organizers? *Nature Reviews. Neuroscience*, 2(2), 92-98.

Stern, C.D. et al., 2006. Head-tail patterning of the vertebrate embryo: one, two or many unresolved problems? *The International Journal of Developmental Biology*, 50(1), 3-15.

Stolc, V. et al., 2004. A gene expression map for the euchromatic genome of Drosophila melanogaster. *Science (New York, N.Y.)*, 306(5696), 655-660.

Suda, Y. et al., 2001. Emx2 directs the development of diencephalon in cooperation with Otx2. *Development (Cambridge, England)*, 128(13), 2433-2450.

Suda, Y. et al., 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 (Cambridge, England)*, 137(17), 2939-2949.

Suh, H. et al., 2007. In vivo fate analysis reveals the multipotent and self-renewal capacities of Sox2+ neural stem cells in the adult hippocampus. *Cell Stem Cell*, 1(5), 515-528.

Sur, M. & Rubenstein, J.L.R., 2005. Patterning and plasticity of the cerebral cortex. *Science (New York, N.Y.)*, 310(5749), 805-810.

Szutorisz, H., Dillon, N. & Tora, L., 2005. The role of enhancers as centres for general transcription factor recruitment. *Trends in Biochemical Sciences*, 30(11), 593-599.

Szymański, M. et al., 2003. Noncoding RNA transcripts. *Journal of Applied Genetics*, 44(1), 1-19.

Taft, R.J. et al., 2009. Tiny RNAs associated with transcription start sites in animals. *Nature Genetics*, 41(5), 572-578.

Taft, R.J., Pheasant, M. & Mattick, J.S., 2007. The relationship between non-protein-coding DNA and eukaryotic complexity. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, 29(3), 288-299.

Takeda, K. et al., 1998. Identification of a novel bone morphogenetic protein-responsive gene that may function as a noncoding RNA. *The Journal of Biological Chemistry*, 273(27), 17079-17085.

Thakur, N. et al., 2004. An antisense RNA regulates the bidirectional silencing property of the Kcnq1 imprinting control region. *Molecular and Cellular Biology*, 24(18), 7855-7862.

Theil, T. et al., 2002. Wnt and Bmp signalling cooperatively regulate graded Emx2 expression in the dorsal telencephalon. *Development (Cambridge, England)*, 129(13), 3045-3054.

Thomas, P. & Beddington, R., 1996. Anterior primitive endoderm may be responsible for patterning the anterior neural plate in the mouse embryo. *Current Biology: CB*, 6(11), 1487-1496.

Thomsen, G.H., 1997. Antagonism within and around the organizer: BMP inhibitors in vertebrate body patterning. *Trends in Genetics: TIG*, 13(6), 209-211.

Tian, E. et al., 2002. Otx2 is required to respond to signals from anterior neural ridge for forebrain specification. *Developmental Biology*, 242(2), 204-223.

Ting, A.H. et al., 2005. Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation.

Nature Genetics, 37(8), 906-910.

Tole, S. et al., 2000. Emx2 is required for growth of the hippocampus but not for hippocampal field specification. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 20(7), 2618-2625.

Triezenberg, S.J., Kingsbury, R.C. & McKnight, S.L., 1988. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes & Development*, 2(6), 718-729.

Tsai, M. et al., 2010. Long noncoding RNA as modular scaffold of histone modification complexes. *Science (New York, N.Y.)*, 329(5992), 689-693.

Tsutsumi, M. & Itoh, M., 2007. Novel transcript nort is a downstream target gene of the Notch signaling pathway in zebrafish. *Gene Expression Patterns: GEP*, 7(3), 227-232.

Tupy, J.L. et al., 2005. Identification of putative noncoding polyadenylated transcripts in Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, 102(15), 5495-5500.

Vasudevan, S., Tong, Y. & Steitz, J.A., 2007. Switching from repression to activation: microRNAs can up-regulate translation. *Science (New York, N.Y.)*, 318(5858), 1931-1934.

Vieira, C. et al., 2010. Molecular mechanisms controlling brain development: an overview of neuroepithelial secondary organizers. *The International Journal of Developmental Biology*, 54(1), 7-20.

Wang, H. et al., 2005. Roles of the polypyrimidine tract and 3' noncoding region of hepatitis C virus RNA in the internal ribosome entry site-mediated translation. *Archives of Virology*, 150(6), 1085-1099.

Wang, X. et al., 2008. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature*, 454(7200), 126-130.

Watanabe, T. et al., 2006. Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes & Development*, 20(13), 1732-1743.

Watanabe, T. et al., 2008. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature*, 453(7194), 539-543.

Wessely, O. & De Robertis, E.M., 2002. Neural plate patterning by secreted signals. *Neuron*, 33(4), 489-491.

Willingham, A.T. et al., 2005. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science (New York, N.Y.)*, 309(5740), 1570-1573.

Wittler, L. & Kessel, M., 2004. The acquisition of neural fate in the chick. *Mechanisms of Development*, 121(9), 1031-1042.

Woltering, J.M. & Durston, A.J., 2008. MiR-10 represses HoxB1a and HoxB3a in zebrafish. *PloS One*, 3(1), e1396.

Wonders, C.P. & Anderson, S.A., 2006. The origin and specification of cortical interneurons. *Nature Reviews. Neuroscience*, 7(9), 687-696.

Wutz, A., 2003. RNAs templating chromatin structure for dosage compensation in animals. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, 25(5), 434-442.

Xie, J. et al., 2007. Sno/scaRNAbase: a curated database for small nucleolar RNAs and cajal body-specific RNAs. *Nucleic Acids Research*, 35(Database issue), D183-187.

Xu, Z. et al., 2009. Bidirectional promoters generate pervasive transcription in yeast. *Nature*, 457(7232), 1033-1037.

Yang, J. et al., 2006. snoSeeker: an advanced computational package for screening of guide and orphan snoRNA genes in the human genome. *Nucleic Acids Research*, 34(18), 5112-5123.

Yue, X. et al., 2010. Transcriptional regulation by small RNAs at sequences downstream from 3' gene termini. *Nature Chemical Biology*, 6(8), 621-629.

Zetterberg, A., Larsson, O. & Wiman, K.G., 1995. What is the restriction point? *Current Opinion in Cell Biology*, 7(6), 835-842.

Zhao, C., Deng, W. & Gage, F.H., 2008. Mechanisms and functional implications of adult neurogenesis. *Cell*, 132(4), 645-660.

Zhao, J. et al., 2008. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science (New York, N.Y.)*, 322(5902), 750-756.

Zhao, Y. et al., 2008. MicroRNA regulation of messenger-like noncoding RNAs: a network of mutual microRNA control. *Trends in Genetics: TIG*, 24(7), 323-327.

Zhu, Y.Y. et al., 2001. Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *BioTechniques*, 30(4), 892-897.

Zufferey, R. et al., 1999. Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *Journal of Virology*, 73(4), 2886-2892.

Zufferey, R. et al., 1998. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *Journal of Virology*, 72(12), 9873-9880.