Telencephalic Embryonic Subtractive Sequences: A Unique Collection of Neurodevelopmental Genes

Alessandro Bulfone,² Pietro Carotenuto,^{1,3*} Andrea Faedo,^{2*} Veruska Aglio,^{1*} Livia Garzia,^{1,3} Anna Maria Bello,¹ Andrea Basile,² Alessandra Andr, ¹ Massimo Cocchia,³ Ombretta Guardiola,¹ Andrea Ballabio,³ John L. R. Rubenstein,⁴ and Massimo Zollo¹

¹Centro di Ingegneria Genetica e Biotecnologie Avanzate, 80145 Napoli, Italy, ²Stem Cell Research Institute–Hospital San Raffaele, Istituto Scientifico San Raffaele, 20132 Milan, Italy, ³Telethon Institute of Genetics and Medicine, 80131 Naples, Italy, and ⁴Nina Ireland Laboratory of Developmental Neurobiology, Center for Neurobiology and Psychiatry Genetics, University of California at San Francisco, San Francisco, California 94143-2611

The vertebrate telencephalon is composed of many architectonically and functionally distinct areas and structures, with billions of neurons that are precisely connected. This complexity is fine-tuned during development by numerous genes. To identify genes involved in the regulation of telencephalic development, a specific subset of differentially expressed genes was characterized. Here, we describe a set of cDNAs encoded by genes preferentially expressed during development of the mouse telencephalon that was identified through a functional genomics approach. Of 832 distinct transcripts found, 223 (27%) are known genes. Of the remaining, 228 (27%) correspond to expressed sequence tags of unknown function, 58 (7%) are homologs or orthologs of known genes, and 323 (39%) correspond to novel rare transcripts, including 48 (14%) new putative noncoding RNAs. As an example of this latter group of novel precursor transcripts of micro-RNAs, telencephalic embryonic subtractive sequence (TESS) 24.E3 was functionally characterized, and one of its targets was identified: the zinc finger transcription factor *ZFP9*. The TESS transcriptome has been annotated, mapped for chromosome loci, and arrayed for its gene expression profiles during neural development and differentiation (in Neuro2a and neural stem cells). Within this collection, 188 genes were also characterized on embryonic and postnatal tissue by *in situ* hybridization, demonstrating that most are specifically expressed in the embryonic CNS. The full information has been organized into a searchable database linked to other genomic resources, allowing easy access to those who are interested in the dissection of the molecular basis of telencephalic development.

Key words: telencephalon; EST; gene array; *in situ* hybridization; neurodevelopmental; diseases

Introduction

The availability of data on genome sequences, expressed regions of genomes, and gene expression in specific tissues is revolutionizing biological research. These resources are especially valuable in the neurosciences, particularly as ~35% of the genome is ex-

pressed in the developing and mature brain. Gene expression profiling has accelerated the ability to identify sets of genes that are differentially regulated in distinct biological contexts. However, additional layers of complexity are required to interpret gene expression-profiling data, such as alternative splicing (Mordrek and Lee, 2002) and noncoding RNAs (Lewis et al., 2003). Moreover, in the nervous system, these mechanisms are clearly associated with development and the transmission and regulation of signals (in particular, for the genes encoding transcription factors, receptors, and secreted molecules), as opposed to with metabolic or synthetic functions.

Recent studies have suggested that etiologies and abnormalities of neurodevelopmental disorders arise at the time of development, when neural cells go through definite stages of proliferation, migration, differentiation, and organization of their connectivity. The many expressed sequence tag (EST)-sequencing efforts of cDNA libraries have produced an impressive amount of data that can be retrieved from dedicated databases, and these large-scale EST projects have shown that the most abundant genes are indeed overrepresented in the cDNA tissue library collections. Thus, low-abundance transcripts are often underrepresented (see, for example, the Brain Molecular Anatomy Project, <http://trans.nih.gov/bmap/resources/resources.htm>), and for this reason, a subtractive strategy has been applied

Received Feb. 8, 2005; revised June 9, 2005; accepted June 10, 2005.

This work was supported by a Fondo Integrativo Ricerca di Base—Ministero Istruzione Universit e Ricerca Italiana RBAU01RW82 grant (A. Bulfone, M.Z.), Telethon Research grants (A. Bulfone, M.Z.), 2002 and 2004 Associazione Italiana per la Ricerca sul Cancro (AIRC)—Fondazione Italiana per la Ricerca sul Cancro (FIRC) grants (M.Z.), a Compagnia San Paolo Torino 2002 grant (M.Z.), a 2001 Telethon Institute of Genetics and Medicine (TIGEM)—Regione Campania grant (M.Z.), a European Union BRECOSM-LSH-CT-503234 FP6 grant (M.Z.), an AIRC—FIRC Fellowship (L.G.), the Molecular Oncology and Pharmacology PhD Program (University of Ferrara, Ferrara, Italy), and an Associazione Oncologia Pediatrica e Neuroblastoma Fellowship (V.A.). We thank Drs. Valeria Capra, Umberto Di Porzio, Elena Rugari, and Michele Studer for their fruitful discussions and critical reading of this manuscript. Additionally, we thank Virginia Avvantaggiato (Consiglio Nazionale delle Ricerche, Institute of Genetics and Biophysics, Naples, Italy) and Celia Pardini (Hospital San Raffaele, Dibat, Milan, Italy) for their technical assistance in the all-mount *in situ* analyses. The Sequencing Core at TIGEM deserves particular thanks, including Stefano Olivieri, Silvia Cocchia, and Anna Torino for their help with the sequencing reactions and the DNA analyses. We thank Mario Traditi (TIGEM), Roberto Martusciello (TIGEM), and Vittorio Lucignano (Centro di Ingegneria Genetica e Biotecnologie Avanzate) for database management and scripting.

*P.C., A.F., and V.A. contributed equally to this work.

Correspondence should be addressed to either of the following: Massimo Zollo, Centro di Ingegneria Genetica e Biotecnologie Avanzate, 80145 Napoli, Italy, E-mail: zollo@ceinge.unina.it; or Alessandro Bulfone, Stem Cell Research Institute—Istituto Scientifico San Raffaele, Via Olgettina 58, 20132 Milan, Italy, E-mail: bulfone.alessandro@hsr.it.

DOI:10.1523/JNEUROSCI.0522-05.2005

Copyright 2005 Society for Neuroscience 0270-6474/05/257586-15\$15.00/0

between the embryonic and adult murine telencephalons to enrich the rare developmentally regulated transcripts (Porteus et al., 1992).

Despite the success of cDNA cloning and bioinformatics in the identification of hundreds of microRNA (miRNA) genes, forward genetics remains one of the most fruitful approaches for detecting miRNAs, small RNAs that regulate the expression of complementary messenger RNAs. Important roles for miRNAs have been suggested in developmental timing (Lee and Ambros, 2001; Reinhart et al., 2002), cell death, cell proliferation (Bartel, 2004), and the patterning of the nervous system (Johnston and Hobert, 2003), and there is evidence that miRNAs are numerous and that their regulatory impact is more important than was previously thought.

The cDNAs of the subtractive hybridization library presented here are known as telencephalic embryonic subtraction sequences (TESS). The library has been optimized to detect rare or unique cDNAs, which correspond to genes that are expressed at levels from fivefold to 60-fold those in the embryonic telencephalon [embryonic day 14.5 (E14.5)] compared with the adult telencephalon (Porteus et al., 1992). This cDNA library was the source for the identification of various genes that have essential roles during embryogenesis (such as *Dlx2*, *Tbr1*, and *eomes-Tbr2*) (Porteus et al., 1991; Bulfone et al., 1995, 1999). We therefore decided to systematically characterize the TESS collection in a search for novel transcripts that, besides being preferentially expressed in the embryonic telencephalon, can be associated with the regulation of specific developmental processes (based on their regional and temporal patterns of expression; i.e., progenitor proliferation, cell-fate specification, differentiation, maturation, and survival) and with neuropsychiatric disorders of developmental origin. Our strategy, which is based on subtractive hybridization and expression profiling via arrays and *in situ* hybridization (ISH), has allowed the identification of mainly rare and novel transcripts, including several noncoding RNAs. Here, we provide evidence that this TESS collection of genes represents a useful resource for the identification of novel regulators of neural development and of candidates for neurodevelopmental diseases.

Materials and Methods

Cell culture and induction of neuronal differentiation

The Neuro2a murine neuroblastoma cell line (ATCC CCL-131; American Type Culture Collection, Manassas, VA) was cultured in DMEM (Invitrogen, San Diego, CA) supplemented with 10% FBS, 2.0 mM L-glutamine, 1.0 mM sodium pyruvate, 100 μ g/ml streptomycin, and 100 U/ml penicillin G in a 5% CO₂ and 95% air-humidified atmosphere at 37°C. The cells were seeded at 1×10^6 per 90-mm-diameter plastic dish and incubated for 24 h. The induction of neurogenesis was then performed as described previously (Tsuiji et al., 1988; Riboni et al., 1995). Briefly, differentiation was initiated by adding 20 μ M retinoic acid (Sigma, St. Louis, MO) in DMEM supplemented with 2% FBS. The culture medium was replaced every 2 d, and the incubation time was continued while taking different samples at 24 h, 48 h, 72 h, 96 h, and 7 d. The morphology of the differentiated cells (DCs) was analyzed by immunofluorescence.

RNA isolation

Total RNA was extracted from the Neuro2a cell line using TRIzol reagent (Invitrogen), according to the instructions of the manufacturer. Total RNA from E14.5 and adult dissected mouse telencephalon, as well as undifferentiated and differentiated neural stem cells (NSCs), was isolated using a method based on guanidinium lysis and phenol-chloroform extraction (ToTALLY RNA; Ambion, Austin, TX). The RNA concentrations were determined by absorbency at 260 nm, and their quality was verified by the integrity of the 18S and 28S rRNA after ethidium bromide

staining of the total RNA samples subjected to 1.2% agarose gel electrophoresis.

The TESS cDNA microarray

cDNA clones of validated library TESS sequences were amplified by PCR using primers that were complementary to the vector sequence. The PCR products were purified with 96-Sepharose columns and analyzed on 1% agarose gels. The PCR products were robotically arrayed (PIXISYS) onto lysine-coated slides (CMT-GAPS-coated slides; Corning, Corning, NY). The spotted DNA was cross-linked to the glass surface of the chips by UV irradiation (60 mJ). A total of 1026 cDNA clones were printed twice on each slide. Before each hybridization, the chips were blocked in blocking solution (50% formamide, 5 $\times$ SSC, 0.1% SDS, 10 mg/ml BSA) at 42°C for 45 min.

Probe preparation and hybridization array

Total RNA templates ranging from 15 to 20 μ g were labeled using the Amersham Biosciences (Piscataway, NJ) CyScribe First Strand cDNA Labeling kit, with the following modifications. The reverse transcription time was increased to 3 h at 42°C. The probe was then digested with RNase-H for 25 min at 37°C. To eliminate the residual RNA, the reaction was incubated with 0.25 M NaOH for 15 min and then neutralized with 10 μ l of HEPES-free acid. The total synthesized cDNA was purified using a Qiagen (Hilden, Germany) PCR purification kit, according to the instructions of the manufacturer. After purification, the amount of incorporated cyanine (Cy) dyes was calculated for both Cy3 and Cy5. The labeled Cy3 and Cy5 targets were coupled in equal amounts and hybridized onto the slides. Probes with <50 pmol of incorporated Cy dyes were not used for hybridizations. The hybridization reaction was performed in a final volume of 50 μ l, containing the Cy3- and Cy5-labeled cDNA probes mixed with 50% formamide, 25% microarray hybridization buffer (Amersham Biosciences CyScribe First Strand cDNA Labeling kit), 50 ng/ μ l polyd(A) (Sigma), and 25% RNase-free H₂O. After denaturation at 92°C for 2 min, the probe was added to the array, which was covered with a coverslip and placed in a sealed chamber to prevent evaporation. After hybridization for 48 h at 42°C, the arrays were washed in solutions of decreasing ionic strengths (one 10 min wash at 55°C in 1.0 $\times$ SSC and 0.2% SDS, two 10 min washes at 55°C in 0.1 $\times$ SSC and 0.2% SDS, two 1 min washes in 0.1 $\times$ SSC, and one 1 min wash in milliQ-deionized H₂O). Three independent sets of experiments were performed for every cellular induction data point.

Total RNA from E14.5 and adult dissected mouse telencephalon, as well as undifferentiated and differentiated neural stem cells, was isolated using a method based on guanidinium lysis and phenol-chloroform extraction (ToTALLY RNA; Ambion). Labeling of total RNA was performed using the dendrimer technology (3DNA Submicro Expression Array Detection kit; Genisphere, Hatfield, PA). The cDNA was then hybridized to the TESS array according to the methods described above.

Scanning and image analysis

The microarrays were scanned with a Medway Far 2000-I scanner, and the signal was converted to a 16 bits/pixel resolution, yielding a 65,536-count dynamic range. Image analysis and calculation of feature pixel intensities adjusted for the local channel-specific background were performed using Array Pro software from Media Cybernetics (Silver Spring, MD). Gridding, automated spot detection, manual and automated flagging, background subtraction, and normalization were all used. For each cDNA spot, the local background was subtracted from the total signal intensities of Cy5 and Cy3. The ratios of net fluorescence of the Cy5-specific channel to the Cy3-specific channel were calculated for each spot, which represented the relative expression in the cells with retinoic acid compared with that in the untreated cells. For every cellular induction data point, three independent sets of experiments were performed to reduce the variations related to the labeling and hybridization efficiencies across experiments.

For microarray experiments on neural stem cells and E14.5 and adult dissected mouse telencephalon, differential gene expression was assessed by scanning the hybridized arrays using a confocal laser scanner capable of interrogating both the Cy3- and Cy5-labeled probes and producing separate TIFF images for each (ScanArray Express; PerkinElmer Life Sci-

ences, Emeryville, CA). These images were then used to normalize and quantitate the data using the LOWESS method (locally weighted scatter plot smoothing; <http://stat-www.berkeley.edu/users/terry/zarray/TechReport/589.pdf>). After spot identification and local background determination, the background-subtracted hybridization intensities were calculated for each spot.

Data mining

To establish the statistical significance of any observed differences in expression between experimental conditions, an average of three replicates was performed for each experiment. To determine the significance and variability of our data, a *t* test was applied to the replicates. The Cyber-T tool (available at <http://visitor.ics.uci.edu/genex/cybert/index.shtml>) was used for the *t* test. A cutoff value of $p = 0.05$ was used. These stringent criteria were designed to minimize the number of false positives and to generate a list of informative genes. As a result of this analysis, we expected an ~10% false-positive rate in our data sets. Moreover, only the data sets that passed the *t* test were used for the clustering. Array expression data experiments and protocol were performed following Minimum Information About Microarray Experiment guidelines, and accession numbers for TESS were produced using Gene Expression Omnibus at National Center for Biotechnology Information (<http://www.ncbi.org/GEO/>).

Hierarchical cluster analyses were performed using the Cluster and TreeView software to extract the putative characteristic genes. For this hierarchical clustering, the signal values were used to calculate the uncentered Pearson correlation coefficients, which were used as the similarity metric. We used the average linkage method algorithm for clustering.

The significance of the enrichment during the differentiation process of genes from specific functional groups [nucleic acid metabolism and transcription and development gene ontology (GO) terms] was determined by comparisons between the percentages of the genes with altered expression in particular functional groups and the overall percentages of changes in the genes during differentiation. The statistics accompanying GO analyses and *p* values were undertaken by using *GFINDER* (<http://www.medinfopoli.polimi.it/GFINDER/>) and taken into consideration, as also discussed recently (Karsten et al., 2003).

Real-time quantitative PCR

Real-time quantitative PCR was performed using standard protocols with an Applied Biosystems (Foster City, CA) 7000 Sequence Detection system. Briefly, 4 μ l of a 1:5 dilution of cDNA in water was added to 12.5 μ l of the 2 $\times$ SYBR-green PCR master mix (Applied Biosystems), with a 400 nM concentration of each primer and water to 25 μ l. The reactions were amplified for 15 s at 95°C and 1 min at 60°C for 40 cycles. The thermal denaturation protocol was run at the end of the PCR to determine the number of products that were present in the reaction. All of the reactions were run in triplicate and included no-template and no-reverse-transcription controls for each gene. The cycle number at which the reaction crossed an arbitrarily placed threshold (C_T) was determined for each gene, and the amount of each TESS target relative to glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (used as reference) was calculated using the following equation: $2 \pm DC_T$, where $DC_T = (C_{T,TESS} \pm C_{T,GAPDH})$ (Schmittgen et al., 2004). To calculate the relative gene expression of miRNA precursors (pre-miRNAs), the relative amount of each pre-miRNA to U6 RNA was calculated using the equation $2 \pm DC_T$, where $DC_T = (C_{T,pre-miRNA} \pm C_{T,U6RNA})$ (Livak and Schmittgen, 2001). Real-time PCR primers for TESS genes were designed using Primer Express software version 2.0 (Applied Biosystems) with a T_m of 60°C and a primer length between 18 and 25 nt (details available on the TESS database website). Primers for pre-miR-9 (miRNA registry accession number MI0000720), pre-miR-24.E3(1), and pre-miR-24.E3(2) were designed using the criteria described by Schmittgen et al. (2004). The list of primers used is available on-line (<http://www.tess.tigem/tessdata.htm>).

In situ hybridization with digoxigenin-labeled probes

Tissue preparation. All of the studies were conducted in accordance with the principles and procedures outlined in the National Institutes of

Health Guidelines for Care and Use of Experimental Animals. Freshly frozen, postfixed sections were used for all of the experiments. At least one litter was used for each age group from which data are reported [E10.5, E12.5, E14.5, E16.5, postnatal day 0 (P0), and adult], with at least two embryos examined at each age. Pregnant CD1 mice were decapitated. For embryos at E10.5, E12.5, and E14.5, the entire uterus was freshly frozen and embedded in OCT compound for cryostat sectioning. Mouse embryos older than E14.5 were dissected quickly and staged according to their limb bud morphology, before being frozen and embedded in OCT. Embryonic trunk sagittal and transverse sections (14 μ m) were prepared to detail the expression patterns. Sections were mounted onto 3-aminopropyltriethoxy-silane-coated microscope slides and stored at -80°C until their use in the hybridization procedure.

Probes. Digoxigenin (DIG)-labeled RNA probes for the mouse TESS genes were synthesized and purified *in vitro* from the plasmid vectors harboring the appropriate cDNA sequences. PCR fragments were generated from T7 and T3 primers and were transcribed with T3 RNA polymerase to make the antisense probes. To produce sense probes, we transcribed PCR fragments with T7 RNA polymerase. The digoxigenin-labeled RNA probes were prepared with the DIG RNA Labeling kit from Roche (Roche Products, Welwyn Garden City, UK), according to the instructions of the manufacturer.

In situ hybridization procedure. Digoxigenin-labeled hybridizations were performed on tissue sections. Briefly, embryos were fixed in 4% paraformaldehyde in 0.1 M phosphate buffer (PB), pH 7.3, overnight at 4°C and then cryoprotected by immersion in 30% sucrose solution in PB. Cryostat sections (14 μ m) were mounted on 3-amino-propyltriethoxy-silane-coated slides and air dried. The probes were diluted immediately before use in the hybridization buffer (see below) and denatured at 75°C for 5 min, and an appropriate volume (see above; ISH with radiolabeled probes) of diluted probe was placed on each slide. After rinses in PBS, the sections were prehybridized for 1 h at 60°C in 50% formamide buffer, incubated with digoxigenin probes for 16 h at 60°C, and further washed at the same temperature. All of the incubations were performed in coplin jars. After overnight hybridization, the slides were incubated in wash buffer (1 $\times$ SSC, 50% formamide, 0.1% Tween 20) at 65°C for at least 15 min or until the coverslips fell off. Then, they were washed two additional times in wash buffer at 65°C for 30 min each. The slides were then incubated for 2 $\times$ 30 min in MABT [100 mM maleic acid, pH 7.5, 150 mM NaCl, 0.1% (v/v) Tween 20]. After incubation in 2% Roche blocking reagent solution for 1 h, the sections were incubated with an alkaline phosphatase-coupled anti-digoxigenin antibody (Roche Products) diluted 1:1000 in blocking solution, and the incubation continued overnight at 4°C. The slides were transferred to coplin jars and washed three times for 5 min in MABT and then two times for 10 min in prestaining buffer (100 mM Tris-HCl, pH 9.0, 100 mM NaCl, 5 mM MgCl₂). The prestaining buffer was replaced with staining buffer [100 mM Tris-HCl, pH 9.0, 100 mM NaCl, 5 mM MgCl₂, 5% (w/v) polyvinyl alcohol (mean molecular weight, 70–100 kDa; Sigma), 0.2 mM 5-bromo-4-chloro-3-indolyl-phosphate (Roche Diagnostics, Indianapolis, IN), and 0.2 mM nitroblue tetrazolium salt (Roche Diagnostics)] and incubated in the dark at room temperature until the signal reached a satisfactory intensity (usually a few hours to overnight, although, exceptionally, over a week-end). The slides were then washed in distilled water for 30 min, dehydrated through ascending ethanol concentrations (30, 60, 80, 95, and 100%; 1 min each), cleared in xylene, mounted under coverslips in XAM mountant (BDH Chemicals, Poole, UK), and analyzed under a Leica (Nussloch, Germany) DC500 compound microscope.

Identification of miR-24.E3 targets

The 3'-untranslated region (UTR) sequences for all mouse genes were retrieved using EnsMart version 15.1 (<http://www.ensembl.org/EnsMart>). Basic local alignment search tool (BLAST) analyses were performed to search the 3'-UTR sequences database with antisense matches to query of pri-miR-24.E3 by using the BLAST server, available on-line (<http://www.blast.tigem.it/>). Base pairing alignment between the miRNA and UTR were predicted by using RNAfold on a fold-back sequence consisting of an artificial stem loop (5'-GCGGGGACG-3') attached to the seed match, and the resulting free energy was used to

evaluate the binding (Kiriakidou et al., 2004), verifying the conservation of target site sequences in the 3'-UTRs of ortholog genes (i.e., between human and mouse and rat) using the following multiple sequence alignment programs: mVISTA (available at <http://genome.lbl.gov/vista/mvista/submit.shtml>) and CLUSTALW (available at <http://www.ebi.ac.uk/clustalw/>) analyses.

DNA constructs

The TESS 24.E3 pri-miRNA insert (536 bp) was digested from TESS 24.E3 pBlueScript K/S and cloned in the *HindIII*–*XhoI* sites of pcDNA3 modified vector (pREY2O-TESS 24.E3 vector). *ZFP9* (NM_011763) target-binding sites were cloned in tandem in the 3'-UTR of the pRL-cytomegalovirus (CMV) vector [outside the coding region of *Renilla* luciferase (RL); Promega, Madison, WI]. Briefly, for each target site, two cDNA oligos were synthesized (sense, 5'-CTAGAGGGGCTGGAGAGATGGCTCAGCAG-TTAAGAGAAAGGGGCTGGAGAGATGGCTCAGCAGTTAAGAG; antisense, 3'-TCCCCGACCTCTCTACCGAGTCGTCATTCTCTTCCCCGACCTCTCTACCGAGTCGTCATTCTCAGATC), annealed, and cloned in the *XbaI* sites of the pRL-CMV vector.

Transfections in P19 cells and luciferase assays

The ability of 24.E3 miRNAs to downregulate target mRNAs was tested by transfection of 24.E3-pREY2O in murine P19 cells (American Type Culture Collection CRL-1825). Quantitative PCR analyses of *HOMEZ* (AK129361), *ATF7* (NM_146065), *FBOX17* (NM_015796.1), and *ZFP9* (NM_011763) were measured by mRNA real-time PCR. Primer pairs were designed as described above; the nucleotide sequences are available on-line (<http://www.tess.tigem/tessdata.htm>).

To evaluate the activity of 24.E3 miRNAs on repression of the luciferase-*ZFP9* fused transcription product, pRL-CMV-*ZFP9* plasmids bearing target-binding sites in the 3'-UTR were constructed (Promega), sequence verified, and cotransfected along with 24.E3-pREY2O (0.5 μ g) and pRL-CMV-*ZFP9* (0.1 μ g) into P19 cells using Lipofectamine 2000 (Invitrogen). Luciferase activities were determined using a Dual Luciferase Reporter Assay system (Promega).

The TESS database

The developing Perl scripts, including the MySQL database, code (source and binaries), and documentation are available by download options on the open standard concurrent version system server (<http://cvsweb.tigem.it>).

Results

Sequence analysis of the TESS cDNAs

Approximately 3600 clones were picked at random, sequenced, and analyzed (BlastN and BlastX homology searches against the nonredundant, dbEST, and UniGene databases). This resulted in a set of 832 unique transcripts, which were annotated, mapped, and arrayed on coated glass slides (Fig. S1A, available at www.jneurosci.org as supplemental material).

The sequence analyses demonstrated that 27% (223) corresponded to known genes, 7% (58) were homologs or orthologs of known genes, 27% (228) corresponded to ESTs without known functions, and 39% (323) did not correspond to any ESTs that are publicly available and instead matched only genomic sequences in the public database (Mouse Genome Draft) (Fig. S1B, available at www.jneurosci.org as supplemental material). Therefore, these last presumably correspond to transcripts that have not been identified previously.

To categorize the TESS sequences into functional classes, we used the gene ontology annotation (www.geneontology.org) (Fig. S1, available at www.jneurosci.org as supplemental material). For instance, 23% of the TESS clones that corresponded to known genes belong to the transcription and nucleic acid metabolism group, 15% belong to the development group, and 10% belong to the signal transduction group. These categories include important transcription factors with functions that are critical for

neural development, such as genes containing structures like basic helix-loop-helix (*Neurogenin-2*, 32.C10; *NeuroD6*, 12.C4), homeobox (*Dlx-2*, 8.G10; *Dlx-5*, 6.B4), forkhead (*Fox-g1*, 24.F7; *Fox-p2*, 19.E3), high-mobility group (*Hmgb1*, 25.B6; *Hmgb2*, 29.E3; *Hmgb3*, 24.B10), leucine zipper (*Fxbl7*, 8.H5), and sex-determining region Y-box (*Sox4*, 21.C5; *Sox11*, 21.G10; *Sox21*, 14.F12) (Anderson et al., 1997; Guillemot, 1999; Miyata et al., 1999; Lai et al., 2001; Nieto et al., 2001; Sun et al., 2001; Bertrand et al., 2002; Nemeth et al., 2003; Pallier et al., 2003; Hanashima et al., 2004; Kan et al., 2004; Muller et al., 2004) (Table S1A, available at www.jneurosci.org as supplemental material).

Because the majority of the identified transcripts (66%) corresponded to uncharacterized ESTs (228) or novel transcripts (323) rather than to known genes, we interrogated the RIKEN full-length cDNA database (<http://fantom2.gsc.riken.go.jp/>), which contains >60,000 cDNAs (Bono et al., 2002; Okazaki et al., 2002). This analysis revealed that only 384, of the total 832, TESS cDNA clones were contained in the FANTOM2 set, confirming the uniqueness of the TESS collection of rare embryonic telencephalic transcripts.

Of the 323 different transcripts matching only genomic sequences, 143 (44%) map to introns with the same (sense) orientation as known or predicted genes, 66 (20%) map to introns but have an opposite (antisense) orientation, 86 (27%) are new homologs of ESTs found in other species, and 28 (9%) are transcripts corresponding to completely novel genes (Fig. S1B, available at www.jneurosci.org as supplemental material).

These 323 TESS cDNAs could correspond to novel genes, rare alternative splicing forms of already identified or putative genes, or to noncoding RNAs with potential regulatory functions (Houbaviy et al., 2003; Lewis et al., 2003; Bartel, 2004; Griffiths-Jones, 2004). To determine whether any of these TESS cDNAs corresponded to this last group, we interrogated the TESS library using the miRNA registry (Griffiths-Jones, 2004), which was developed at the Wellcome Trust Sanger Institute (www.sanger.ac.uk/Software/Rfam/mirna/search.shtml; release 3.1 April 2004). Several of these (48) that map to introns with sense and antisense orientations turned out to have significant homologies (selected by significance of *p* values) with the sequences of miRNA precursors (Table 1). The TESS precursors of pre-miRNAs that mapped to introns of known genes or ESTs with an antisense orientation included the following: 24.E3 [in the *Elavl-1* (*Etr-r3*) gene, miR-333; *p* = 0.00091], 12.A1 (in the *Raf-1* gene, miR166; *p* = 0.0011), 0.F3 (in the *Tankyrase* gene, miR-333; *p* = 0.00013), 9.E9 (in the *Nrxn-3* gene, miR-297-1; *p* = 0.073), 32.H5 (in the *UCH-L3* gene, miR-20; *p* = 0.00076), 6.A7 (in the *Atrophin-1* interacting protein 1, miR-157c; *p* = 0.00023), 21.D7 (in the EST CB849645, miR164b; *p* = 0.062), and 30.A7 (in the EST BY480859, miR169; *p* = 0.00049) (Table 1). To verify the reliability of this data, RNA ISH was performed on mouse embryos for the TESS transcripts and for the corresponding genes to which they map. As an example, the TESS clone 24.E3 has an intriguing expression pattern that overlaps with that reported for the *Elavl-1* gene at E14.5 (data not shown). TESS 24.E3 is specifically and highly expressed in the developing brain (E14.5) (Fig. 1A); at P0, its expression is mainly confined to the forebrain, whereas in the adult the transcript is still detectable in the hippocampus and cerebellum (Fig. 1A). Our analysis shows also that the transcript corresponding to TESS 24.E3 could represent a novel putative primary transcript (pri-miRNA) for two independent pre-miRNAs, which are both highly homologous to previously identified miRNAs in *Rattus norvegicus* and *Arabidopsis thaliana* (Fig. 1B). We evaluated also the ability of these pre-miRNAs to fold

into potential hairpin miRNA precursors. The two candidate miRNAs pre-miR-24.E3(1) and pre-miR-24.E3(2) were graphically and statistically represented by Mfold algorithm structure predictions, with the highly conserved sequence adopting a stem-loop characteristic of known miRNAs (Fig. 1C).

We additionally verified the relative expression levels of TESS 24.E3 pre-miRNAs by means of real-time PCR detection on several RNA tissues, both at E14.5, and in the adult brain and using as a positive control the pre-miR-9 (a known miRNA isolated in the brain; miRNA registry accession number MI0000720) (Schmittgen et al., 2004; Sempere et al., 2004) using the method described. As shown in Figure 1D, pre-miR-24.E3(1) is highly abundant in the embryonic neural, heart, and eye tissues, whereas its expression is lower in the corresponding adult tissues. For pre-miR-24.E3(2), the levels of expression are lower, but this transcript appears to be more specific for the embryonic brain if compared with pre-miRNA-24.E3(1), which is more represented in the tissues analyzed.

These data suggest that a number of “novel” noncoding RNAs are specifically expressed during telencephalic development and prove the significance of our approach, which paves the way to new challenges in the analysis of gene expression in neural development.

Systematic expression analysis by RNA ISH

The TESS transcripts were also characterized for their tissue expression distributions. ISH was performed on E14.5 mouse embryos using a selected set of TESS genes that received the highest priority mainly because of their novelty (overall, the expression pattern was assessed for 84 “known genes,” 26 “similar to known genes,” 94 ESTs, and 76 “matching genomic sequence”) (Table S2A,B, available at www.jneurosci.org as supplemental material). This analysis revealed a preferential distribution of TESS expression in the nervous system (almost all of the TESS genes analyzed are specifically expressed, or coexpressed, in the brain, even if at different levels). For the TESS genes that belong to the novel transcript category or that had specific neural expression patterns, the analysis was extended to the later stages of development and to the postnatal brain. Of the 188 TESS cDNAs analyzed, 28 (15%) are uniquely expressed in the proliferative and/or mantle zones of the forebrain, whereas the remaining 160 (85%) are also expressed in other parts of the CNS or embryonic tissues. Figure 2 shows some examples of TESS expression patterns at E14.5: TESS 32.G7, 13.F9, and 6.F7 are all expressed in postmitotic cells (in the cortical plate and intermediate zone of the cortex for TESS 32.G7 and 13.F9, whereas TESS 6.F7 expression is restricted to the cortical plate only); TESS 19.H5, 17.C12, and 20.C2 are instead expressed specifically in the cortical proliferative ventricular zone; TESS 12.A5, 21.D7, and 25.E5 exhibit instead more complex, but brain-specific, patterns of expression (telencephalic rostrocaudal gradient of expression for TESS 12.A5, restricted expression in specific forebrain domains for

Table 1. TESS cDNA sequences similar to miRNA precursors: sequences deposited in the miRNA database

TESS identity	Antisense intronic sequence		Representative similar miRNA precursor RNA	miRNA registry accession number and <i>p</i> value
	Gene/EST	mRNA accession number		
9.E9	Nrxn-3	BC060719	<i>Mus musculus</i> miR-297-1	MI0000395/0.073
32.H5	UCH-L3	AB033370	<i>Mus musculus</i> miR-20	MI0000568/0.00076
29.C12	EST	AK122421	<i>Mus musculus</i> miR-297-1	MI0000395/0.0061
13.A7	EST	CB844100	<i>Mus musculus</i> miR-297-1	MI0000395/0.00059
22.E12	EST	AK016309	<i>Mus musculus</i> miR-297-2	MI0000397/0.021
29.B10	EST	AK038447	<i>Mus musculus</i> miR-201	MI0000244/0.17
24.E3	Elavl-1	U65735	<i>Rattus norvegicus</i> miR-333	MI0000610/0.00091
0.F3	EST	AK047094	<i>Rattus norvegicus</i> miR-333	MI0000610/0.00013
19.E9	Tcf-20	AY007594	<i>C. elegans</i> miR-249	MI0000325/0.083
18.E7	No match		<i>C. elegans</i> miR-66	MI0000037/0.0090
12.A1	Raf-1	BC015273	<i>Arabidopsis</i> miR166e	MI0000205/0.0011
6.A7	Acvrnp-1	AK039336	<i>Arabidopsis</i> miR157c	MI0000186/0.00023
25.E2	Odz-2	AB025411	<i>Arabidopsis</i> miR166e	MI0000205/0.0044
18.D10	Rgs-7	BC051133	<i>Arabidopsis</i> miR165b	MI0000200/0.0014
14.F11	Adarb-2	BC052426	<i>Arabidopsis</i> miR169	MI0000212/0.53
23.C7	RAB-28	BC004580	<i>Arabidopsis</i> miR165b	MI0000200/0.00066
28.F12	NCoA-2	U39060	<i>Arabidopsis</i> miR169	MI0000212/0.0014
12.F10	Ube3a	NM_173010	<i>Arabidopsis</i> miR169	MI0000212/0.00069
18.B10	No match		<i>Arabidopsis</i> miR165a	MI0000199/0.13
8.E11	EST	AV318787	<i>Arabidopsis</i> miR169	MI0000212/0.0020
21.D7	EST	CB849645	<i>Arabidopsis</i> miR164b	MI0000198/0.062
30.A7	EST	BY480859	<i>Arabidopsis</i> miR169	MI0000212/0.00049
13.C4	EST	AA170468	<i>Arabidopsis</i> miR156b	MI0000179/0.019
30.A7	EST	BY480859	<i>Arabidopsis</i> miR169	MI0000212/0.00049
22.F6	EST	BU787350	<i>Arabidopsis</i> miR169	MI0000212/0.0043

TESS 21.D7, and regional expression along the neural tube for TESS 25.E5).

The ISH expression data have been organized and included in the TESS databases (<http://tess.tigem.it>; <http://tess.ceinge.unina.it>), which can be retrieved and queried on the basis of the regions of interest and/or the levels of expression. For instance, it is possible to interrogate and retrieve the expression patterns of the TESS clones that do not correspond to any known genes or ESTs and that are preferentially expressed in the postmitotic (mantle zone) or proliferative (ventricular zone) cells of the developing brain. The searchable expression archive within the TESS database (described further below) allows a rapid analysis of the putative involvement of any particular TESS clone in the regulation of particular aspects of neurogenesis, such as proliferation, migration, or differentiation of neural cells.

Gene expression profiling using the TESS array

To date, gene expression profiling studies in neurobiology have mainly provided lists of upregulated or downregulated genes, which are often without a complete validation and/or an assessment of their biological significance. We therefore decided to undertake a more comprehensive analysis in terms of the expression profile of our collection of TESS genes. The TESS collection of cDNAs corresponding to these 832 unique transcripts was arrayed on coated glass slides together with a number of control (house-keeping and developmental) genes (a total of 1026 cDNAs, in duplicate copies).

With the assumption that this restricted collection of TESS cDNAs corresponds to a tissue/time-specific transcriptome (E14.5 telencephalon) and comprises a number of unknown or uncharacterized genes, the corresponding array represents a new effective tool to assess the expression profile of the entire collection together in specific experimental settings.

Along these lines, we first investigated the level of preferential embryonic expression of the TESS genes versus their expression

Figure 1. Identification of two novel mouse miRNAs from TESS transcript 24.E3. **A**, RNA *in situ* hybridization expression analysis at E14.5 and P0 and in the adult brain, showing the downregulation of expression of TESS 24.E3. Although at E14.4 the primary transcript is detected in the entire postmitotic areas of the neural tube, at P0 it is mainly present in the diencephalic and telencephalic structures and in the adult brain in the internal granular layer of the cerebellum. Cb, Cerebellum; Cx, cortex; H, hindbrain; M, midbrain; St, striatum; T, thalamus. **B**, Multiple sequence alignment of the pre-miRNA segments corresponding to the hypothetical pre-miRNA-24.E3(1) and pre-miRNA-24.E3(2) and their homologs rno-mir-333, ath-mir166e, and ath-mir157d. The positions of mature miRNAs are shown in red. Conserved residues are highlighted with black boxes. **C**, Genomic organization of the hypothetical pre-miRNA-24.E3(1) and pre-miRNA-24.E3(2) clusters, together with the secondary structures of the proposed 24.E3 precursor miRNA(1) and miRNA(2). The TESS 24.E3 sequence corresponds to an intronic sequence of the Elavl-1 gene (RefSeq NM_010485). Rfam analysis shows that the sequence alignment of TESS 24.E3 matches the conserved noncoding region of an Elavl-1 intronic sequence and the target sequences of the precursor miRNA of mir-333, ath-MIR166e, and ath-MIR157d. Genomic coordinates are given as chromosome:start-end. A prediction of RNA secondary structure using the Mfold algorithm (<http://www.bioinfo.rpi.edu/applications/mfold/old/ma/form1.cgi>) shows that the TESS 24.E3 transcript contains a conserved sequence that adopts a stem-loop structure that is characteristic of miRNAs when selected for best statistical *p* values and predicted folding free energies. Left, The TESS 24.E3(2) sequence matches a precursor of miRNA from *A. thaliana* ath-MIR166e (*dG* = −30.3). Right, The TESS 24.E3(1) sequence matches a precursor miRNA from *Rattus mir-333* (*dG* = −23.5). **D**, Expression analysis of TESS 24.E3 precursor miRNA(1) and miRNA(2) by real-time PCR. Expression levels of pre-miR-24.E3(1), pre-miR-24.E3(2), and pre-miR-9 in adult and embryonic tissues are shown as histograms. Gene expression is presented relative to mouse U6 RNA expression. Mean $\pm$ SD is from triplicate real-time PCRs from a single cDNA sample preparation.

in the adult telencephalon using microarray hybridization. The total RNA was isolated from six dissected E14.5 and adult mice telencephalons and the corresponding labeled cDNA hybridized to the TESS microarray. The analysis allowed the detection and comparison of the levels of expression between these two tissues and confirmed that the TESS cDNAs are mostly preferentially expressed in the embryonic forebrain, thus confirming the subtractive approach (Fig. 3A). Considering a >1.5-fold differential expression between the two tissues, 60% were upregulated in the embryo, and only 10% were preferentially expressed in the adult brain. Similarly, although 15% of the TESS cDNAs were more than fourfold more expressed in the embryo, only 1% were expressed to such a level in the adult brain (Table 2).

After this screening, tissue RNA ISH was used to validate the array results using tissue at several stages of development along with adult brain tissue. Here, several of the TESS genes were chosen from among those with the lowest to the highest (fold) upregulation in the E14.5 telencephalon. Figure 3 highlights two examples of TESS cDNAs (21.D7 and 21.C5) where the expression levels are significantly higher (more than fourfold) in the E14.5 telencephalon compared with the adult telencephalon. ISH analysis performed using both the TESS clones as probes on embryonic and adult telencephalic sections confirmed the array data, showing the preferential embryonic expression of these genes (Fig. 3B); it also allowed a more detailed assessment of their patterns of expression. TESS 21.D7 (corresponding to a novel putative pre-miRNA) has a specific expression in the developing cortex and in only one of the two primordia of the basal ganglia (lateral ganglionic eminence), whereas TESS 21.C5 (corresponding to *Sox4*) is expressed throughout the mantle layer of the developing telencephalon. This analysis confirmed that known developmental genes are differentially expressed between the two tissues (e.g., including, among others, *Casp3*, *Foxg1-Bfl*, *Sox4*, and *Sox11*), and it also revealed that there are still a number of uncharacterized genes, or completely novel genes, preferentially expressed in the embryonic forebrain (Table 2).

These results indicate that we have identified a number of uncharacterized and novel genes that have a temporally and spatially restricted expression in the developing forebrain. Additional investigations will be needed to characterize their func-

tions and to further assess their roles in telencephalic development and neural differentiation.

TESS gene expression profiling experiments on *in vitro* differentiation cell models

To identify new regulatory genes involved in neurogenesis, we used the entire TESS transcriptome collection as a source for the selection of genes associated with proliferation/maintenance of neural precursors or with early neural differentiation. We performed gene expression profiling using the TESS microarray during *in vitro* differentiation of embryonic NSCs and the neuroblastoma cell line Neuro2a.

Neurospheres were isolated as described in the supplemental Material and Methods (available at www.jneurosci.org as supplemental material) and induced to terminal differentiation by using conventional procedures. The neural differentiation process was followed by using immunofluorescence analyses, showing the presence of all types of neural cells (positivity to glutamate, GABA, GFAP, and GalC neural markers) (Fig. S2, available at www.jneurosci.org as supplemental material).

The TESS array hybridization experiments were performed on bona fide undifferentiated embryonic NSCs (ENSCs) (cell “doublets”), in comparison with terminally DCs (Fig. S2, available at www.jneurosci.org as supplemental material), and they allowed the identification of a core set of 24 TESS transcripts that were selectively upregulated or downregulated in the progenitor cells, compared with their differentiated progeny (Table 3). These comprised genes that were already known to be upregulated in these cells (e.g., the TESS clones corresponding to *Foxg1*, *Emx2*, *Sox11*, *Rad23b*, and *Hmgb2*), and other genes corresponding instead to uncharacterized and novel ESTs.

Figure S2, *G* and *H* (available at www.jneurosci.org as supplemental material), shows two examples of TESS genes differentially expressed during NSC differentiation, analyzed for their tissue expression distribution at E14.5 by ISH: TESS 29.G5 (3.3-fold higher expression in ENSCs vs DCs) is mainly expressed in the proliferative zone of the telencephalon, whereas TESS 26.G3 (threefold higher expression in DCs vs ENSCs) is mainly detected in the postmitotic mantle zone (Fig. S2, available at www.jneurosci.org as supplemental material) (Table 3).

Similarly, using the Neuro2a mouse neuroblastoma cell line as a model of neu-

Figure 2. Examples of expression patterns in E14.5 embryos of TESS clones corresponding to uncharacterized ESTs or matching only genomic sequences in the public databases. The E14.5 sagittal sections show the expression distribution of the TESS clones, which varies from being restricted to postmitotic cells (*A–C*, and the cortical region at a higher magnification in *A'–C'*) and the cortical proliferative ventricular zone (*D–F*, and at a higher magnification in panels *D'–F'*) to exhibiting instead more complex but brain-specific patterns of expression (telencephalic rostrocaudal gradient of expression for TESS 12.A5 in *G* and *G'*; restricted expression in specific forebrain domains for TESS 21.D7 in *H* and *H'*; regional expression along the neural tube for TESS 25.E5 in *I* and *I'*). FM, Forebrain mantle; FP, forebrain proliferative; OE, olfactory epithelium; M, midbrain; H, hindbrain; SC, spinal cord; CG, cranial ganglia; G, gut; VF, vibrissa follicles; Cx, cortex; GE, ganglionic eminence; TB, tooth buds; MZ, marginal zone; CP, cortical plate; IZ, intermediate zone; VZ, ventricular zone.

Figure 3. **A**, Scatter-plot analysis of embryonic (E14.5) versus adult telencephalon gene expression profiling of the TESS clones collection. The mean expression levels (arbitrary units) were calculated from two hybridizations (dye-swap analyses). Genes that showed a difference of at least twofold are below (E14.5 genes) and above (adult genes) the dotted lines. The genes that are between the two dotted lines are considered equally expressed. Two examples of TESS cDNAs are also highlighted, corresponding to a novel putative pre-miRNA (21.D7) and Sox4 (21.C5), the expression levels of which are significantly higher (more than fourfold) in the E14.5 telencephalon compared with the adult telencephalon. **B**, **B'**, **C**, **C'**, The RNA *in situ* hybridization shows a specific laminar expression in the developing cortex and a broader expression in the ganglionic eminence for TESS 21.D7 (**B**), whereas TESS 21.C5 (Sox4) is expressed throughout the mantle layer of the developing telencephalon (**C**). For both genes, expression in the adult cerebral cortex was not detectable (**B'**, **C'**).

ronal differentiation that can be induced to terminal differentiation by exposure to retinoic acid, a large number of TESS genes were found to be differentially expressed (upregulated >1.5-fold or downregulated <0.7-fold; as shown in Table S3, available at www.jneurosci.org as supplemental material). Hierarchical cluster analysis (see Material and Methods) was performed on the gene expression data at six different time points after the exposure of these cells to retinoic acid (days 0, 1, 2, 3, 4, and 7), and several clusters grouping the TESS with similar expression profiles after differentiation were identified. The three main categories are represented (Fig. 4A–C).

To cross-verify whether the TESS genes found differentially

expressed in Neuro2a cell differentiation show an embryonic expression pattern consistent with the cellular model, we performed mRNA *in situ* hybridization analyses at the different stages of development (E14.5, P0, and adult cortex). The first cluster category (Fig. 4Aa, cluster 5) represents TESS genes that are downregulated after differentiation of Neuro2a cells, such as TESS 29.E8 (corresponding to *Foxg1*), which is highly expressed in the cortex at E14.5 (Fig. 4Ab), detectable in the ventricular zone at P0 (Fig. 4Ad) and negative in the adult cortex (Fig. 4Af). For TESS 21.B5 (an uncharacterized EST), which belongs to the same cluster in terms of its expression profile in differentiating Neuro2a cells, the pattern of expression is slightly different and it shows an upregulation in the P0 cortex (Fig. 4Ae), whereas in the adult it is undetectable (Fig. 4Ag). The second cluster category (Fig. 4Ba, cluster 9) includes TESS genes that are upregulated after “*in vitro*” terminal differentiation of the Neuro2a cells and thus putatively involved in the acquisition of a mature neural phenotype (Fig. 4Ba). Two examples of TESS genes belonging to cluster 9 are TESS 19.E2, which corresponds to the ring finger factor *Rfn2*, which is involved in ubiquitin ligase activity and transcription, and TESS 6.F7, which corresponds to a novel gene of unknown function. Both TESS 19.E2 and TESS 6.F7 are expressed in differentiated postmitotic neurons throughout cortical development and postnatally (Fig. 4Bd–Bf). In particular, TESS 6.F7 expression is confined to the cortical plate at E14.5 (Fig. 4Bc) and in the adult cortex to cortical layers 2–3 (Fig.

Table 3. Examples of the TESS genes that are differentially expressed between embryonic neural stem cells and the same cells induced to complete differentiation

TESS identity	Description	Log ₂ ENSC/DC	Accession number
18.A11	Foxg1-Bf1	4.7 ± 1.51	BC046958
21.E12	EST match	4.38 ± 0.73	AK011225
25.B1	Mest/Peg1	4.21 ± 0.78	BC006639
Emx2	Emx2	3.97 ± 0.87	
23.D3	EST match	3.96 ± 0.84	BC046562
29.E3	Hmgb2	3.74 ± 0.46	AK012568
29.G5	EST match	3.38 ± 0.8	BC021353
T-bet	T-bet	3.24 ± 0.49	
Pcna	Pcna	3.01 ± 0.23	
13.H6	Sox11	2.66 ± 0.99	XM_196071
18.A9	Rad23b	2.38 ± 0.34	BC027747
29.E11	Cct6a	2.05 ± 0.41	NM_009838
29.D8	Ptma	2 ± 0.25	X56135
23.G5	Cct2	1.87 ± 0.2	BC007470
13.D2	Hat1 homolog	1.53 ± 0.66	BX284624
18.D3	Cugbp2	−1.55 ± 0.16	AF090697
28.B1	Cnn3	−1.9 ± 0.14	BC005788
30.E12	EST match	−1.91 ± 0.25	AK053243
23.C11	EST match	−1.99 ± 0.27	BC030351
26.G3	Rcn2	−2.92 ± 0.91	NM_011992
29.C11	EST match	−3.12 ± 0.49	AK044903
16.H3	Basp1	−3.38 ± 1.54	XM_127954
Latexin	Latexin	−3.8 ± 0.96	
32.D4	Connexin 43	−6.3 ± 0.69	BC011324

4Be–Bg). A third category of TESS clustered on the basis of their expression profiles (Fig. 4Ca, cluster 12) includes TESS genes that are transiently upregulated during *in vitro* differentiation. Examples are TESS 30.H11 (*Sox5*) and TESS 13.F9 (novel gene of unknown function), which are both expressed at E14.5 (Fig. 4Cb, Cc), even if at different levels, not only in the cortical plate but also in the differentiating and migrating cells of the intermediate zone, whereas postnatally their expression levels are significantly downregulated (Fig. 4Cf, Cg).

These gene-expression results have been validated for a representative number of TESS genes (46 TESS) by real-time PCR in undifferentiated and differentiated cells of both NSCs and Neuro2a cells (the list of genes and primers used for the real-time PCR are available on the TESS database website, <http://tess.tigem.it>; <http://tess.ceinge.unina.it>).

To understand the biological significance of the overall changes in gene expression, the identified TESS corresponding to known genes were categorized on the basis of their involvement in specific biological processes. This analysis was performed by the functional grouping of the TESS genes using the gene ontology database (<http://www.geneontology.org/godatabase>), and the GFinder statistical analysis showed that the major overall mRNA/gene expression variation occurred in the more proliferative and less differentiated cells. Across the functional categories in the undifferentiated NSCs, three categories were considerably enriched and significantly differentially expressed when compared with the others: nucleic acid metabolism/transcription ($p < 0.04$), cell activation ($p < 0.024$), and response to external stimulus ($p < 0.026$) (Table 4). In the undifferentiated Neuro2a cells, only the response to external stimulus category was found to be statistically relevant ($p < 0.048$). Additionally, to verify the statistical relevance of the data obtained through the gene ontology and GFinder clustering, we assessed the levels of expression by real-time detection PCR in the Neuro2a differentiation mod-

els (0, 4, and 7 d) (Fig. 4Da–Dc) and by mRNA *in situ* hybridization on E14.5 mouse embryos of two TESS genes belonging to the nucleic acid metabolism/transcription category, TESS 32.F4 and TESS 12.C4 (Fig. 4Db–Dd).

The TESS 32.F4 (corresponding to *SpF*) transcript was downregulated during differentiation of the Neuro2a cells, and it had a similar expression profile to *BMP4*, a marker of neural progenitors (Linker and Stern, 2004). In the E14.5 cortex, it was mainly expressed in the proliferating ventricular and subventricular zones. In contrast, the expression levels of TESS 12.C4 (corresponding to *NeuroD6*) increased after Neuro2a differentiation, like with *MAP2*, a marker of terminal neuronal differentiation (Porteus et al., 1994), and its transcript was consistently and clearly detected in the E14.5 cortical plate and intermediate zone by *in situ* hybridization (Fig. 4Db–Dd).

A similar analysis was performed to determine whether, in our three independent gene expression data sets (telencephalon, NSC, and Neuro2a), there was any gene upregulated or downregulated in undifferentiated neurons or progenitor cells. Figure S3A (available at www.jneurosci.org as supplemental material) shows that only three genes were found to be upregulated in progenitors. No common TESS genes were found to be downregulated in all three models used here (Fig. S3B, available at www.jneurosci.org as supplemental material). With similar approaches, we grouped genes found upregulated in the telencephalon and upregulated or downregulated in undifferentiated Neuro2a cells (Fig. S3C, available at www.jneurosci.org as supplemental material) and downregulated in NSC cells (Fig. S3D, available at www.jneurosci.org as supplemental material). These analyses will help in the identification of new genes that have functions that are potentially involved in the maintenance of progenitor fate and in an undifferentiated mode. Additional studies will need to be performed to address these latest issues.

In our TESS collection, we also identified a specific subset of transcripts represented by the 48 putative miRNA precursors (Table 1). To identify the TESS miRNA precursors that might be involved in neural differentiation, we hierarchically clustered them on the basis of the similarities of their expression profiles in the Neuro2a cell differentiation model.

A subset of TESS corresponding to putative miRNA precursors showed a sharp increase in expression only after 4 d of retinoic acid treatment of the Neuro2a cells (Fig. 5A), implying that they are involved in regulation of terminal neural differentiation. A second group of TESS miRNA precursors exhibited instead a decreasing expression after 4 d of differentiation (Fig. 5B). In this latter group, TESS 24.E3 is present, which is a precursor transcript of two different pre-miRNAs (Fig. 1A–C). Because of the presence of two putative miRNAs in the TESS 24.E3 primary transcript, we evaluated independently their expression profiles in the Neuro2a model (by real-time PCR and selecting primer pairs specific for each independent pre-miRNA), and we showed that 24.E3 pre-miRNA(2) is expressed at higher levels than 24.E3 pre-miRNA(1) after differentiation from days 1 to 7 (Fig. 5C). As positive controls, we used pre-miR9, a known miRNA that is highly expressed in brain development (Sempere et al., 2004), and the *MAP2* gene, the expression of which correlates with neural differentiation (Porteus et al., 1994). When this was repeated with NSCs and DCs using the stem cell factor (*SFC*) gene (Jin et al., 2002) as a reference marker, the expression of 24.E3 pre-miRNA(1), as with the *SFC* gene, was upregulated in the NSCs and downregulated in the DCs, whereas 24.E3 pre-miRNA(2) was constantly less represented (Fig. 5D). These results underline the relevance of assessing the expression profiles of this particular

class of new transcripts to gain an understanding of the biology of the neural progenitor cells and to determine their full characterization and their target genes.

TESS 24.E3 miRNA gene-target identification and *in vitro* function validation analyses

To verify the role of 24.E3 miRNA *in vitro*, we performed an “*in silico*” genome-wide search analysis on the 3′-UTR sequences to find significant homology matches by the BLAST algorithm. Four candidate genes with their 3′-UTR significantly matching the consensus sequence of the 24.E3 miRNA were identified (Fig. 6*Aa,Ab*), and we therefore verified their levels of expression in Neuro2a following the retinoic acid differentiation treatment (Fig. 6*Ac*). As observed here, 24.E3 pre-miR(1) was abundant at day 0 (Neuro2a noninduced cells), whereas it decreased to almost one-tenth of this level at day 2; the levels of *ZFP9*, corresponding to the zing finger *Kruppel*-like *ZFP9* gene (also known as the *Krox-4* gene), instead showed an increase at day 2 of differentiation, when compared with the undifferentiated non-induced cells. Additionally, the level of expression of the *ATF7* gene showed an increased level of expression at day 7, whereas the level of 24.E3 pre-miR(1) expression decreased. Of note, *Homez* mRNA level of expression at these stages of differentiation was not detected. All together, these data in the Neuro2a differentiation model shows that *ZFP9* and *ATF7*

Figure 4. Analysis of three clusters of differentially expressed TESS after differentiation of the mouse Neuro2a neuroblastoma cell line. ***Aa***, The different expression profiles of cluster 5. ***Ba***, The different expression profiles of cluster 9. ***Ca***, The different expression profiles of cluster 12. All of the clusters that were present during *in vitro* neuronal differentiation (***Aa***, ***Ba***, ***Ca***) and the expression patterns of two examples for each cluster (cluster 5, ***Ab–Ag***, ***Bb–Bg***, ***Cb–Cg***) by RNA *in situ* hybridization on E14.5, P0, and adult cerebral cortex are reported. ***Da, Dc***, The levels of expression of two TESS genes belonging to the nucleic acid metabolism/transcription category, TESS 32.F4 and TESS 12.C4, by real-time PCR detection in the Neuro2a differentiation models (0, 4, and 7 d). ***Db, Dd***, mRNA *in situ* hybridization on the E14.5 cortex. The TESS 32.F4 (corresponding to *Spf*) transcript is downregulated during differentiation of the Neuro2a cells, and in the E14.5 cortex, it is mainly expressed in the proliferating ventricular zone (VZ) and subventricular zone. Alternatively, the expression levels of TESS 12.C4 (corresponding to *NeuroD6*) increase after Neuro2a differentiation, and its transcript is consistently, clearly detected in the E14.5 cortical plate and intermediate zone by *in situ* hybridization. *BMP4* and *MAP2* were used as control markers. MZ, Marginal zone; CP, cortical plate; SP, subplate; IZ, intermediate zone; 2–4, layers 2–4; 4, layer 4; 5, layer 5; 6, layer 6; 2–3, layers 2–3.

Table 4. Gene ontology categorization of TESS genes differentially expressed in cellular models of neural differentiation (Neuro2a and NSC)

Functional group	TESS	NSCs undifferentiated	<i>p</i> value	NSCs differentiated	<i>p</i> value	TESS	Neuro2a undifferentiated	<i>p</i> value	Neuro2a differentiated	<i>p</i> value
Signal transduction	21	5	0.986	1	0.179	13	6	0.100	3	0.726
Response to external stimulus	9	4	0.029			3	1	0.048	2	0.416
Regulation of cell differentiation	6	2	0.731							
Protein metabolism	31	8	0.956	2	0.328	15			5	0.418
Nucleic acid metabolism	64	11	0.045	3	0.124	40	11	0.257	8	0.086
Cell growth and/or maintenance	40	8	0.933	2	0.407	23	6	0.106		
Cell activation	5	2	0.024						2	0.416
Response to stress	10	3	0.059			10	4	0.645		
Organogenesis	30	4	0.240							

The TESS genes are grouped functionally using the gene ontology database and the GFinder statistical analyses.

mRNA might be regulated by 24.E3 pre-miR(1), the first in a more sensitive manner over early days of differentiation. Thus, we decided to verify this in the P19 cell line (both wild-type cells and in cells transiently overexpressing the 24.E3 pre-miRNAs). The levels of one of the putative target transcripts (*ZFP9*) were significantly reduced after TESS24.E3 overexpression (*ZFP9* mRNA expression was reduced 3.2-fold) (Fig. 6*Ad*), thus indicating a potential downregulation of the expression levels of *ZFP9* mRNA by 24.E3 miRNAs. The downregulation of *ZFP9* coincided with an increase in *BMP4* gene expression, thus indicating a potential role for these miRNAs in the differentiation process of the neural cell progenitors (data not shown). Although this analysis shows that *ZFP9* might be a direct target for 24.E3 miRNA, the analyses for *ATF7* did not show any comparative alteration on its level of expression in the P19 overexpression cell model. For this reason, we continued our analyses on the *ZFP9* target validation. Thus, we verified by real-time PCR analyses the levels of expression in several embryonic and adult tissues of both pre-miR-24.E3(1) and *ZFP9* (Fig. 6*Ba*). As shown in Figure 6*Ba*, the levels of expression between the tissues analyzed here (embryonic and adult) were similarly represented for both genes. In addition, we found that in tissues derived from embryonic brain structures (telencephalon, mesencephalon, and eye), the levels of expression were quantitatively similar, although in adult tissues the levels were found to be significantly reduced (~5- to 10-fold). These results were further confirmed for embryonic coexpression by comparing parallel sections and the expression of *ZFP9* and TESS 24.E3 in mouse brain development (Fig. 6*Bb*) and by confirming the *in situ* hybridization data deposited at Mouse Genomics Informatics brain atlas (www.informatics.jacks.org), as described previously (Gray et al., 2004), and the data presented in Figure 1*A*. Then, following our strategy, a luciferase reporter assay was performed (Fig. 6*Bc*) by cloning two tandem 3'-UTR binding sites of *ZFP9* at the 3' site of the *Renilla* luciferase cDNA construct to validate *in vitro* the silencing of *ZFP9*

Figure 5. Expression analysis of pre-miRNAs identified in the TESS cDNA library. *A*, Clustering of TESS transcripts similar to pre-miRNAs. The graph shows the expression levels during the retinoic acid-induced differentiation, in which these TESS transcripts are upregulated at day 7 of Neuro2a differentiation. *B*, Clustering of TESS transcripts similar to pre-miRNAs. The graph shows the expression levels during the retinoic acid-induced differentiation, in which these TESS transcripts are downregulated at day 7 of Neuro2a differentiation. *C*, miRNA-24.E3(1), miRNA-24.E3(2), and miR-9 precursor levels of expression during retinoic acid-induced Neuro2a differentiation by real-time PCR analyses. Gene expression is expressed in relation to U6 RNA gene expression. The results are given as means $\pm$ SD of triplicate real-time PCR experiments from two independent cDNA sample sources. As a control for gene expression, the neuronal marker MAP2 has been used. *D*, miRNA-24.E3(1), miRNA-24.E3(2), and miR-9 precursor levels of expression during retinoic acid-induced Neuro2a differentiation by real-time PCR analyses. Gene expression is expressed in relation to U6 RNA gene expression. The results are given as means $\pm$ SD of triplicate real-time PCR experiments from two independent cDNA sample sources. As a control for gene expression of NSC differentiation, the stem cell factor gene (*SCF*) was used.

expression by pri-miR-24.E3 overexpression. As shown in Figure 6*Bc*, miR-24.E3(1) recognizes the *ZFP9* binding site, thus decreasing the luciferase reporter protein levels. This demonstrated the functional relevance of the 48 novel miRNAs identified and reinforces our findings for the assessment of their potential roles in neurogenesis and forebrain development. The search for the targets of the miRNAs identified here is an ongoing effort.

The cumulative gene expression analysis using the TESS microarray on embryonic and adult telencephalon and on the un-

brain development and *in vitro* differentiation of the neural cell lines (E14.5 vs adult telencephalon, undifferentiated vs differentiated neural cell lines), are as follows: TESS 20.D2 (corresponding to the *Cdon* gene, encoding for a cell adhesion molecule), which is upregulated sevenfold in the E14.5 telencephalon, 2.11-fold in NSCs, and 2.43-fold in undifferentiated Neuro2a cells; and TESS 24.D10 (corresponding to an uncharacterized EST, BC016236), which shows twofold higher expression in the embryonic telencephalon, as well as in the undifferentiated NSCs and in Neuro2a cells.

The analysis of the TESS collection required the generation of a dedicated database to store, annotate, assign predictions of gene function, and disseminate the data to the scientific community

(<http://tess.tigem.it/>; <http://tess.ceinge.unina.it/>). These data are publicly available by access through the websites. To achieve this, we produced a web interface using open source software (PHP and MySQL) that links in-house data to existing public databases. Our resource allows BLAST analysis against the TESS sequences, as well as extrapolation of sequence analysis information from other linked databases (GenBank, Unigene, EST at NCBI, Fantom2, GeneOntology, GeneTrap, BayGenomics, and miRNA registry).

The dedicated database also contains information on the genomic location of the TESS genes in both the mouse and the human genome and information on their expression patterns and profiles in the developing mouse embryo (E14.5) and the neural cell lines (TESS array data on embryonic NSCs and the Neuro2a cell line). The database also includes the TESS genes that have already been trapped (Table S5, available at www.jneurosci.org as supplemental material). To date, a total of 63 TESS are already present in the ES gene-trap database, including 14 TESS, which correspond to previously uncharacterized transcripts (e.g., TESS 6.B5), confirming again that the TESS correspond to genes that are actively expressed.

Discussion

The complexity of brain cell types and circuits is reflected in the complexity of gene expression patterns in the brain. Here, we describe a novel approach to efficiently scan gene expression profiles of the developing telencephalon that has taken advantage of a subtractive hybridization strategy coupled with ISH and microarray analysis. We used an EST sequencing approach to catalog and array the repertoire of genes represented in a subtractive library (subtraction between adult and E14.5 mouse telencephalon), optimized to select rare or unique cDNAs (named TESS) preferentially, or exclusively, expressed in the E14.5 mouse telencephalon.

The rationale is that genes preferentially expressed during embryogenesis are likely to be specifically involved in the generation of the telencephalic tissue and structure, and in our case in the developmental processes that are taking place at E14.5. These include the following: (1) proliferation of neural progenitors in the ventricular zone, (2) differentiation and migration of the deep cortical layers, and (3) differentiation and migration of the striatal patch cells.

A set of 832 unique transcripts was identified, annotated, mapped, and arrayed on coated glass slides ("TESS cDNA microarray"). Through these analyses, the TESS genes have been classified into 13 categories according to their annotation and putative functional roles. The GO functional group with the highest number of TESS genes was the DNA metabolism and transcription group (which includes, for example, several members of the *sry-box*-, *homeobox*-, and *forkhead-box*-containing gene families). This is not surprising, because many transcriptional activators have been implicated in the regulation of the complex spatial and temporal expression of genes involved in neural development.

The TESS collection contains, for example, members of the *Dlx* transcription factors, which are expressed preferentially in the embryonic forebrain, such as *Dlx2* and *Dlx5*, and not the forms that are also expressed in the adult brain (*Dlx1* and *Dlx6*) (Eisenstat et al., 1999). As has been described previously (Faedo et al., 2004), during the systematic analysis of the TESS transcripts, we found a putative noncoding RNA TESS 31.E5, an ortholog of the rat *Evf-1* (AY518691) (Kohtz and Fishell, 2004). Therefore,

this confirms the relevance of our approach and the importance of the isolation and characterization of the TESS cDNA library.

The TESS cDNAs corresponding to already known genes present in the array are also significant, for at least two reasons. First, they can be considered as markers that are useful for validating experimental results. For instance, *Sox4* (TESS 21.C5) is a transcription factor expressed at various stages of embryonic telencephalic development, and *CD24a* (TESS 19.G4) is a glycosylphosphatidylinositol-anchored molecule that is expressed in migrating neurons of the developing mouse nervous system and in a small area of the adult subventricular zone. Second, there are known genes that were not known to be involved in telencephalic development but that now have to be considered as being implicated in this process. For instance, TESS 29.G5 and TESS 12.B6 corresponded to *Hmgb2* and *Mest/Peg1*, respectively. Interestingly, for both of them, because they were expressed in the embryonic brain and in undifferentiated adult NSCs, their role in development needs to be addressed. *Hmgb2* is known to be involved in chromatin modification, and chromatin remodeling may have a crucial role in stem cell biology and differentiation by regulating the intrinsic state of responsiveness of a cell (Doetsch, 2003). *Mest/Peg1* may instead have a role in the transformation of toxic compounds, thus controlling cell homeostasis. Finally, the presence in the TESS library array of *Cdon* (TESS 20.D2) and *Crb1* (TESS 11.3E) further shows that the microarray is a development-specific gene collection, because both of these genes have been characterized extensively as having roles in development. *Cdon* expression was found to be spatially and temporally restricted during embryogenesis (Mulieri et al., 2000). *Crb1*, a mouse homolog of the *Drosophila crumbs*, is exclusively expressed in the embryonic eye and CNS, whereas in the adult its expression is confined to the regions of active neurogenesis.

For the TESS group that did not correspond to any publicly available ESTs and, at present, only match genomic sequences in the public database (323), additional analyses were performed. Of this brain-specific transcript collection, 48 map to introns with sense or antisense orientations and have significant sequence homologies (selected according to significant matches and *p* values) with sequences corresponding to RNA precursors of miRNA primary transcripts identified in mouse and other species. The data presented here suggest that a number of novel noncoding RNAs are specifically expressed during telencephalic development. These findings have also been confirmed by several other studies supporting the tissue-specific or developmental stage-specific expression and their evolutionary conservation (Pasquinelli et al., 2000; Aravin et al., 2001; Lagos-Quintana et al., 2001, 2002, 2003; Lau et al., 2001; Lee and Ambros, 2001; Carington and Ambros, 2003; Lewis et al., 2003). Moreover, it is of extreme interest that our TESS collection has several novel genes mapping into introns of known genes (for example, 18.D10 in an intron of *Regulator of G protein signaling 7* in an antisense orientation, 12.A1 in an intron of the *RAF-1* oncogene, 25.D1 in an intron of the *Odz-2* gene in an antisense orientation, 23.C7 in an intron of the *Rab 28* similar gene), and that they correspond to putative clusters of multiple precursors of miRNAs, as has been shown recently be the case with several other newly discovered miRNAs.

To characterize new mammalian miRNAs that might have roles in neural development, we analyzed here the expression of the pri-miRNA corresponding to TESS 24.E3 during murine CNS development (Fig. 1A). The expression analysis by ISH shows the high and specific expression of TESS 24.E3 in the embryonic mouse brain, suggesting an important role in the devel-

opment of the telencephalon, with specific regard to the differentiation and migration of the neural precursors that, once they have left the proliferative ventricular zone, express TESS 24.E3. The expression is maintained also at P0, mainly in the telencephalon, whereas in the adult brain, the expression is restricted only to the dentate gyrus of the hippocampus. We also demonstrated that TESS 24.E3 aligns with two different families of miRNAs, the *Arabidopsis sativa* miR-166 and miR-156 families, and the rat miR-333 family (Fig. 1B). The analysis of the genomic locus reveals that the TESS 24.E3 transcript corresponds to a polycistronic precursor transcript (pri-miRNA) containing two different pre-miRNAs organized in tandem [pre-miRNA-24.E3(1) and pre-miRNA-24.E3(2)], and that the pri-miR-24.E3 is transcribed from an intronic sequence of the ELAVL1 gene. These findings reflect two prominent characteristics of vertebrate miRNAs: (1) the miRNA genes are organized in tandem, and the clustered miRNAs are often processed from a single messenger RNA molecule (Ambros, 2004); and (2) miRNA genes often reside within introns of protein-coding genes (Rodriguez et al., 2004).

Furthermore, the two candidate TESS 24.E3 miRNA sequences fold into potential hairpin miRNA precursors, which, together with the phylogenetic conservation of the hairpin fold, represent good evidence for the existence of a putative miRNA (Ambros et al., 2003). As shown in Figure 1, B and C, whereas the putative pre-miR-24.E3 form (1) has a high homology with miR-333, the pre-miR-24.E3 form (2) is highly homologous to miR-166 and miR-157. Both of these last miRNAs are located on chromosome 2 of *A. thaliana* (Reinhart et al., 2002), but although miR-166 appears to target mRNAs coding for HD-Zip transcription factors (including Phabulosa-PHB and Phavoluta-PHV), regulating axillary meristem initiation, and leaf development (Rhoades et al., 2002), miR-157 is thought to target 10 mRNAs coding for proteins containing the Squamosa-promoter binding protein box. Moreover, the two pre-miR-24.E3 forms adopt a different stem-loop structure (which is a structural characteristic of the known miRNAs), and they have different tissue expression profiles, as assessed by real-time PCR analysis (Fig. 1D), which was performed using the computationally identified pre-miRNA sequences.

Interestingly, as shown here (Fig. 6Bc), we determined that TESS 24.E3 silenced the *in vitro* expression of *ZNF9* (Chavrier et al., 1988), a mouse multigene appertaining to the zinc finger family that was fully characterized by ISH in the developing brain by Gray et al. (2004). Our data indicate that the TESS 24.E3 function might have an *in vivo* control of the expression of an important transcription factor influencing the brain development processes. Unfortunately, little is known about the role of this uncharacterized zinc finger protein, *ZFP9*, in brain development processes, and therefore it will be of great interest to investigate its *in vivo* function and regulation by TESS 24.E3 miRNA.

In a complementary manner, we also used the selected transcriptome contained in the unique microarray of subtracted TESS cDNAs for gene expression profiling experiments on the following: (1) embryonic and adult telencephalon, and (2) two sets of neural cell lines, NSCs and Neuro2a cells, which were analyzed as undifferentiated and actively proliferating or as their terminally differentiated progeny. Furthermore, by clustering the results of the gene expression patterns (ISH) and profiles (TESS array), we were able to highlight specific subsets of the TESS genes that are upregulated or downregulated after differentiation in the neural cell lines or in specific regions of the developing forebrain.

This allowed us to further select for novel and rare transcripts involved in forebrain development and the regulation of neural progenitor cell properties.

In conclusion, because many of the TESS transcripts do not correspond to any known genes or ESTs, this transcriptome represents a complementary collection of genes for expression profiling. Furthermore, the large number of novel noncoding RNAs identified in the TESS transcriptome highlights the potential roles for these molecules in regulating gene expression during brain development. Among these, the identified subclass of miRNAs are highly expressed during development, as recently seen by other studies (miRNA and disease) (Calin et al., 2004; John et al., 2004), and they will need to be further analyzed for their function in brain development.

The TESS EST sequences have been deposited in the GenBank EST database using the acronym TESS, and they are available for BLAST searches at <http://tess.tigem.it>; <http://tess.ceinge.unina.it>. In the on-line supplemental material (available at www.jneurosci.org), we discuss the hypothesis that TESS genes are potentially involved in some neurodevelopmental diseases on the basis of our comprehensive expression analyses.

References

- Ambros V (2004) The functions of animal microRNAs. *Nature* 431:350–355.
- Ambros V, Lee RC, Lavanway A, Williams PT, Jewell D (2003) MicroRNAs and other tiny endogenous RNAs in *C. elegans*. *Curr Biol* 13:807–818.
- Anderson SA, Eisenstat DD, Shi L, Rubenstein JL (1997) Interneuron migration from basal forebrain to neocortex: dependence on *Dlx* genes. *Science* 278:474–476.
- Aravin AA, Naumova NM, Tulin AV, Vagin VV, Rozovsky YM, Gvozdev VA (2001) Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr Biol* 11:1017–1027.
- Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116:281–297.
- Bertrand N, Castro DS, Guillemot F (2002) Proneural genes and the specification of neural cell types. *Nat Rev Neurosci* 3:517–530.
- Bono H, Kasukawa T, Furuno M, Hayashizaki Y, Okazaki Y (2002) FANTOM DB: database of functional annotation of RIKEN mouse cDNA clones. *Nucleic Acids Res* 30:116–118.
- Bulfone A, Smiga SM, Shimamura K, Peterson A, Puelles L, Rubenstein JL (1995) T-brain-1: a homolog of *Brachyury* whose expression defines molecularly distinct domains within the cerebral cortex. *Neuron* 15:63–78.
- Bulfone A, Martinez S, Marigo V, Campanella M, Basile A, Quaderi N, Gattuso C, Rubenstein JL, Ballabio A (1999) Expression pattern of the *Tbr2* (Eomesodermin) gene during mouse and chick brain development. *Mech Dev* 84:133–138.
- Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M, Croce CM (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci USA* 101:2999–3004.
- Carrington JC, Ambros V (2003) Role of microRNAs in plant and animal development. *Science* 301:336–338.
- Chavrier P, Lemaire P, Revelant O, Bravo R, Charnay P (1988) Characterization of a mouse multigene family that encodes zinc finger structures. *Mol Cell Biol* 8:1319–1326.
- Doetsch F (2003) A niche for adult neural stem cells. *Curr Opin Genet Dev* 13:543–550.
- Eisenstat DD, Liu JK, Mione M, Zhong W, Yu G, Anderson SA, Ghattas I, Puelles L, Rubenstein JL (1999) *DLX-1*, *DLX-2*, and *DLX-5* expression define distinct stages of basal forebrain differentiation. *J Comp Neurol* 414:217–237.
- Faedo A, Quinn JC, Stoney P, Long JE, Dye C, Zollo M, Rubenstein JL, Price DJ, Bulfone A (2004) Identification and characterization of a novel transcript down-regulated in *Dlx1/Dlx2* and up-regulated in *Pax6* mutant telencephalon. *Dev Dyn* 231:614–620.
- Gray PA, Fu H, Luo P, Zhao Q, Yu J, Ferrari A, Tenzen T, Yuk DI, Tsung EF, Cai Z, Alberta JA, Cheng LP, Liu Y, Stenman JM, Valerius MT, Billings N,

- Kim HA, Greenberg ME, McMahon AP, Rowitch DH, et al. (2004) Mouse brain organization revealed through direct genome-scale TF expression analysis. *Science* 306:2255–2257.
- Griffiths-Jones S (2004) The microRNA registry. *Nucleic Acids Res* 32:D109–D111.
- Guillemot F (1999) Vertebrate bHLH genes and the determination of neuronal fates. *Exp Cell Res* 253:357–364.
- Hanashima C, Li SC, Shen L, Lai E, Fishell G (2004) Foxg1 suppresses early cortical cell fate. *Science* 303:56–59.
- Houbaviy HB, Murray MF, Sharp PA (2003) Embryonic stem cell-specific microRNAs. *Dev Cell* 5:351–358.
- Jin K, Mao XO, Sun Y, Xie L, Greenberg DA (2002) Stem cell factor stimulates neurogenesis in vitro and in vivo. *J Clin Invest* 110:311–319.
- John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS (2004) Human microRNA targets. *PLoS Biol* 2:e363.
- Johnston RJ, Hobert O (2003) A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. *Nature* 426:845–849.
- Kan L, Israsena N, Zhang Z, Hu M, Zhao LR, Jalali A, Sahni V, Kessler JA (2004) Sox1 acts through multiple independent pathways to promote neurogenesis. *Dev Biol* 269:580–594.
- Karsten SL, Kudo LC, Jackson R, Sabatti C, Kornblum HI, Geschwind DH (2003) Global analysis of gene expression in neural progenitors reveals specific cell-cycle, signaling, and metabolic networks. *Dev Biol* 261:165–182.
- Kiriakidou M, Nelson PT, Kouranov A, Fitziev P, Bouyioukos C, Mourelatos Z, Hatzigeorgiou A (2004) A combined computational-experimental approach predicts human microRNA targets. *Genes Dev* 18:1165–1178.
- Kohtz JD, Fishell G (2004) Developmental regulation of EVF-1, a novel non-coding RNA transcribed upstream of the mouse *Dlx6* gene. *Gene Expr Patterns* 4:407–412.
- Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. *Science* 294:853–858.
- Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T (2002) Identification of tissue-specific microRNAs from mouse. *Curr Biol* 12:735–739.
- Lagos-Quintana M, Rauhut R, Meyer J, Borkhardt A, Tuschl T (2003) New microRNAs from mouse and human. *RNA* 9:175–179.
- Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413:519–523.
- Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294:858–862.
- Lee RC, Ambros V (2001) An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294:862–864.
- Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. *Cell* 115:787–798.
- Linker C, Stern CD (2004) Neural induction requires BMP inhibition only as a late step, and involves signals other than FGF and Wnt antagonists. *Development* 131:5671–5681.
- Livak KJ, Schmittgen TD (2001) Analysis of representative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25:402–408.
- Miyata T, Maeda T, Lee JE (1999) NeuroD is required for differentiation of the granule cells in the cerebellum and hippocampus. *Genes Dev* 13:1647–1652.
- Modrek B, Lee C (2002) A genomic view of alternative splicing. *Nat Genet* 30:13–19.
- Mulieri PJ, Okada A, Sassoon DA, McConnell SK, Krauss RS (2000) Developmental expression pattern of the *cdo* gene. *Dev Dyn* 219:40–49.
- Muller S, Ronfani L, Bianchi ME (2004) Regulated expression and subcellular localization of HMGB1, a chromatin protein with a cytokine function. *J Intern Med* 255:332–343.
- Nemeth MJ, Curtis DJ, Kirby MR, Garrett-Beal LJ, Seidel NE, Cline AP, Bodine DM (2003) Hmgb3: an HMG-box family member expressed in primitive hematopoietic cells that inhibits myeloid and B-cell differentiation. *Blood* 102:1298–1306.
- Nieto M, Schuurmans C, Britz O, Guillemot F (2001) Neural bHLH genes control the neuronal versus glial fate decision in cortical progenitors. *Neuron* 29:401–413.
- Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, Nikaido I, Osato N, Saito R, Suzuki H, Yamanaka I, Kiyosawa H, Yagi K, Tomaru Y, Hasegawa Y, Nogami A, Schonbach C, Gojobori T, Baldarelli R, Hill DP, et al. (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* 420:563–573.
- Pallier C, Scaffidi P, Chopineau-Proust S, Agresti A, Nordmann P, Bianchi ME, Marechal V (2003) Association of chromatin proteins high mobility group box (HMGB) 1 and HMGB2 with mitotic chromosomes. *Mol Biol Cell* 14:3414–3426.
- Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, Hayward DC, Ball EE, Degnan B, Muller P, Spring J, Srinivasan A, Fishman M, Finnerty J, Corbo J, Levine M, Leahy P, Davidson E, Ruvkun G (2000) Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* 408:86–89.
- Porteus MH, Bulfone A, Ciaranello RD, Rubenstein JL (1991) Isolation and characterization of a novel cDNA clone encoding a homeodomain that is developmentally regulated in the ventral forebrain. *Neuron* 7:221–229.
- Porteus MH, Brice AE, Bulfone A, Usdin TB, Ciaranello RD, Rubenstein JL (1992) Isolation and characterization of a library of cDNA clones that are preferentially expressed in the embryonic telencephalon. *Brain Res Mol Brain Res* 12:7–22.
- Porteus MH, Bulfone A, Liu JK, Puellas L, Lo LC, Rubenstein JL (1994) DLX-2, MASH-1, and MAP-2 expression and bromodeoxyuridine incorporation define molecularly distinct cell populations in the embryonic mouse forebrain. *J Neurosci* 14:6370–6383.
- Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP (2002) MicroRNAs in plants. *Genes Dev* 16:1616–1626.
- Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of plant microRNA targets. *Cell* 110:513–520.
- Riboni L, Prinetti A, Bassi R, Caminiti A, Tettamanti G (1995) A mediator role of ceramide in the regulation of neuroblastoma Neuro2a cell differentiation. *J Biol Chem* 270:26868–26875.
- Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A (2004) Identification of mammalian microRNA host genes and transcription units. *Genome Res* 14:1902–1910.
- Schmittgen TD, Jiang J, Liu Q, Yang L (2004) A high-throughput method to monitor the expression of microRNA precursors. *Nucleic Acids Res* 32:e43.
- Sempere LF, Freemantle S, Pitha-Rowe I, Moss E, Dmitrovsky E, Ambros V (2004) Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. *Genome Biol* 5:R13.
- Sun Y, Nadal-Vicens M, Misono S, Lin MZ, Zubiaga A, Hua X, Fan G, Greenberg ME (2001) Neurogenin promotes neurogenesis and inhibits glial differentiation by independent mechanisms. *Cell* 104:365–376.
- Tsuji S, Yamashita T, Tanaka M, Nagai Y (1988) Synthetic sialyl compounds as well as natural gangliosides induce neuritogenesis in a mouse neuroblastoma cell line (Neuro2a). *J Neurochem* 50:414–423.