

Ling, King Hwa; Hewitt, Chelsea A.; Beissbarth, Tim; Hyde, Lavinia; Banerjee, Kakoli; Cheah, Pike See; Cannon, Ping; Hahn, Christopher Norman; Thomas, Paul Quinton; Tan, Seong Seng; Thomas, Tim; Scott, Hamish Steele; Smyth, Gordon K.

Molecular networks involved in mouse cerebral corticogenesis and spatio-temporal regulation of Sox4 and Sox11 novel antisense transcripts revealed by transcriptome profiling, *Genome Biology*, 2009; 10(10):R104.

© 2009 Ling et al.; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The electronic version of this article is the complete one and can be found online at :

<http://genomebiology.com/2009/10/10/R104>

PERMISSIONS

<http://www.biomedcentral.com/about/license>

Anyone is free:

- to copy, distribute, and display the work;
- to make derivative works;
- to make commercial use of the work;

Under the following conditions: Attribution

- the original author must be given credit;
- for any reuse or distribution, it must be made clear to others what the license terms of this work are;
- any of these conditions can be waived if the authors gives permission.

13th August 2013

<http://hdl.handle.net/2440/58298>

Molecular networks involved in mouse cerebral corticogenesis and spatio-temporal regulation of *Sox4* and *Sox11* novel antisense transcripts revealed by transcriptome profiling

King-Hwa Ling^{*†‡#}, Chelsea A Hewitt^{*††}, Tim Beissbarth^{§‡‡}, Lavinia Hyde^{§§§}, Kakoli Banerjee[¶], Pike-See Cheah^{¶‡}, Ping Z Cannon^{*}, Christopher N Hahn[#], Paul Q Thomas[¶], Gordon K Smyth[§], Seong-Seng Tan^{**}, Tim Thomas^{*} and Hamish S Scott^{*†#}

Addresses: ^{*}Molecular Medicine Division, The Walter and Eliza Hall Institute of Medical Research, Royal Parade, Parkville, Victoria 3052, Australia. [†]The School of Medicine, The University of Adelaide, SA, 5005, Australia. [‡]Department of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor DE, Malaysia. [§]Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research, Royal Parade, Parkville, Victoria 3052, Australia. [¶]School of Molecular and Biomedical Science, Faculty of Sciences, University of Adelaide, Adelaide, SA 5005, Australia. ^{‡‡}Department of Human Anatomy, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor DE, Malaysia. ^{§§}Department of Molecular Pathology, The Institute of Medical and Veterinary Science and The Hanson Institute, Adelaide, SA 5000, Australia. ^{¶¶}Howard Florey Institute, The University of Melbourne, Parkville, Victoria 3010, Australia. ^{††}Current address: Pathology Department, The Peter MacCallum Cancer Centre, St Andrews Place, East Melbourne, Victoria 3002, Australia. ^{**}Current address: Department of Medical Statistics (Biostatistics), University of Göttingen, Humboldtalle 32, 37073 Göttingen, Germany. ^{§§}Current address: The Bioinformatics Unit, Murdoch Childrens Research Institute, Royal Children's Hospital, Melbourne, Victoria 3052, Australia.

Correspondence: Hamish S Scott. Email: hamish.scott@imvs.sa.gov.au

Published: 2 October 2009

Genome Biology 2009, **10**:R104 (doi:10.1186/gb-2009-10-10-r104)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2009/10/10/R104>

Received: 6 June 2009

Revised: 20 July 2009

Accepted: 2 October 2009

© 2009 Ling et al.; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Development of the cerebral cortex requires highly specific spatio-temporal regulation of gene expression. It is proposed that transcriptome profiling of the cerebral cortex at various developmental time points or regions will reveal candidate genes and associated molecular pathways involved in cerebral corticogenesis.

Results: Serial analysis of gene expression (SAGE) libraries were constructed from C57BL/6 mouse cerebral cortices of age embryonic day (E) 15.5, E17.5, postnatal day (P) 1.5 and 4 to 6 months. Hierarchical clustering analysis of 561 differentially expressed transcripts showed regionalized, stage-specific and co-regulated expression profiles. SAGE expression profiles of 70 differentially expressed transcripts were validated using quantitative RT-PCR assays. Ingenuity pathway analyses of validated differentially expressed transcripts demonstrated that these transcripts possess distinctive functional properties related to various stages of cerebral corticogenesis and human neurological disorders. Genomic clustering analysis of the differentially expressed transcripts identified two highly transcribed genomic loci, *Sox4* and *Sox11*, during embryonic cerebral corticogenesis. These loci feature unusual overlapping sense and antisense transcripts with alternative polyadenylation sites and differential expression. The *Sox4* and *Sox11* antisense transcripts were highly expressed in the brain compared to other mouse organs and are

differentially expressed in both the proliferating and differentiating neural stem/progenitor cells and P19 (embryonal carcinoma) cells.

Conclusions: We report validated gene expression profiles that have implications for understanding the associations between differentially expressed transcripts, novel targets and related disorders pertaining to cerebral corticogenesis. The study reports, for the first time, spatio-temporally regulated *Sox4* and *Sox11* antisense transcripts in the brain, neural stem/progenitor cells and P19 cells, suggesting they have an important role in cerebral corticogenesis and neuronal/glial cell differentiation.

Background

Complex behavioral tasks, from perception of sensory input and the control of motor output to cognitive functions such as learning and memory, are dependent on the precise development of innumerable interconnections of neuronal networks in the cerebral cortex. The development of the cerebral cortex (also known as cerebral corticogenesis) involves the specific influence of intrinsic and extrinsic mechanisms, which are triggered spatio-temporally [1-4]. Between embryonic day 11 (E11) and 18 (E18), the mouse cerebral cortex develops from a relatively homogenous band of mitotic multipotent progenitor cells into a complex laminated structure containing various classes of neuronal cells [2,4-6]. Cerebral corticogenesis involves: proliferation of multipotent progenitors (E11 to E16.5); migration of postmitotic cells (E11 to E17); cell morphogenesis (E13 to E18); gliogenesis and synaptogenesis (E16 until early postnatal period); and reorganization, elimination and stabilization of neuronal networks (up to adulthood).

The mouse cerebral cortex develops in the latero-medial and rostro-caudal axes [7,8]. At E11, the primordial plexiform layer begins to form in the most lateral part of the neural wall. Its growth continues in the latero-medial axis to the medial part of the telencephalon by E13. The primordial plexiform layer is also expanded in the rostro-caudal axis. The growth in this axis is always less than the growth in the latero-medial axis. The first wave of migratory neuronal cells form the cortical plate 2 days later after the development of the primordial plexiform layer. These events are followed by the development of the cortical plate into an organization of six distinct layers that forms the adult cerebral cortex. Generally, the rostral-most regions of the adult cerebral cortex consist of areas involved in executive functions and motor coordination, whilst the caudal-most regions consist of areas involved in sensory perception such as visual function. Although distinct functional arealization of the cerebral cortex do not fully apply to rodents, mounting evidence suggesting that regulated arealization exists has been shown in mice involving transcription factors such as empty spiracles homolog 2 (*Emx2*; *Drosophila*) [9,10], paired box gene 6 (*Pax6*) [9], COUP transcription factor 1 (*Coup-tf1*) [11], Sp8 transcription factor (*Sp8*) [12,13], distal-less homeobox 1/2 (*Dlx1/2*) and gastrulation brain homeobox 2 (*Gbx2*) [14]. The extensive cyto-architectural and anatomical changes occurring in a spa-

tio-temporal manner during the peak (E15) and at the end (E17) of embryonic cerebral neurogenesis as well as during early postnatal (P1) corticogenesis through to adulthood involves complex underlying molecular regulatory networks.

Complex molecular regulatory elements are important determinants in both spatial and temporal cerebral corticogenesis. These elements regulate gene expression at the chromatin, DNA or RNA, and protein levels through chromatin packaging or remodeling, histone acetylation and deacetylation, chromatin insulation, DNA methylation, post-transcriptional regulation and post-translational protein modification or degradation signaling pathways [15-17]. Other processes involved in such regulation include pre-mRNA processing and nuclear mRNA retention by nuclear-specific paraspeckle complexes [18,19], microRNAs (miRNAs) that interfere with mRNA translation [20-22], and a new class of under-characterized non-coding RNA transcripts known as natural antisense transcripts (NATs) [23,24]. These regulatory networks play a pivotal role in establishing when, where and how a multipotent progenitor cell should proliferate, migrate and settle at a designated position in the cortex. The information regarding regulatory networks during cerebral corticogenesis, however, remains incomplete and does not provide a comprehensive view of the underlying regulatory elements throughout this complex event.

In this study, we employed both short and long 3' serial analysis of gene expression (SAGE) technologies [25,26] to identify differentially expressed regulatory elements by comparing transcriptomes of cerebral cortices generated from four selected developmental stages: E15.5, E17.5, P1.5 and 4 to 6 months old. We also compared rostral to caudal regions of E15.5 and left to right regions of adult cerebral cortices. We report temporally co-regulated gene clusters, novel molecular networks and associated pathways, novel candidates in regionalized development and genomic clustering of SRY-box containing gene 4 (*Sox4*) and SRY-box containing gene 11 (*Sox11*) sense and antisense transcripts. The genomic clustering analysis led us to the discovery of spatio-temporal regulation of novel *Sox4* and *Sox11* antisense transcripts as well as differential regulation of these transcripts in proliferating and differentiating neural stem/progenitor cells (NSPCs) and P19 (embryonal carcinoma) cells.

Results

Generation and analysis of SAGE tags

We constructed 10 SAGE libraries from the cerebral cortex of E15.5, E17.5 and 4- to 6-month-old adult C57BL/6 mice (N = 10; Table 1). The data from two additional SAGE libraries generated from E15.5 and P1.5 cerebral cortices from Gunnensen *et al.* [27] were also incorporated into our analysis. These SAGE libraries represent key stages of cerebral corticogenesis and are accessible from the Gene Expression Omnibus (GEO) website [GEO: [GSE15031](#)] [28]. The libraries contain a total of 531,266 SAGE tags (approximately 44,000 tags per SAGE library), 223,471 of which are unique (non-redundant) after screening for artifacts and mapping of short SAGE tags to long SAGE tags (Table 1). The number of unique tags in each library ranges from approximately 7,200 to 32,000 due to the variation in library size (approximately 13,500 to 70,000). The distribution of tag abundance, however, is similar in all libraries (Figures S1 and S2 in Additional data file 1), in which the majority of tags were detected only once (58 to 76% or approximately 5,500 to approximately 24,000 tags), representing a trend comparable with previously reported SAGE analyses of mouse neocortices [27,29]. Of all unique tags, only 5,199 (approximately 2.4%) are common to all developmental stages. The low number of common unique tags is most probably due to the high abundance of tags that occur only once.

Analysis of differentially expressed transcripts/tags

To identify differentially expressed tags/transcripts (DETs), we considered only those 25,165 unique tags with a count >2 across all libraries. Under stringent analyses (Table S1 in Additional data file 1), we identified a total of 561 DETs in various comparisons between developmental stages (Table 2; Figure S3 in Additional data file 1). A full list of DETs with assigned IDs is provided in Additional data file 2. Greater numbers of DETs are observed when the interval of two comparative developmental stages becomes wider. We find the largest number of DETs (approximately 59% or 328 DETs) in the embryonic versus adult stages (E versus Ad) followed by P1.5 versus Ad (approximately 34% or 192 DETs), E15.5 versus P1.5 (approximately 6% or 36 DETs) and E15.5 versus E17.5 (approximately 7% or 38 DETs) comparisons. These indicate distinctive transcript signatures during cerebral cortex development. Comparisons between rostral and caudal E15.5 (R versus C), and left and right adult cerebral cortices (L versus Ri) are described in a different section below.

Approximately 69% of DETs have been assigned to known genes and 6% to expressed sequence tags (ESTs), while 10% of DETs have multiple matches (tag matching multiple gene identifiers) and 8% having ambiguous matches (tag matching the same gene identifier at multiple chromosome loci). Approximately 8% of DETs have no matches and it is most

Table 1

SAGE library information

SAGE library*	Sex	Age	Tissue	Generated sequences	Generated SAGE tag (library)	Unique tags	Unique tags (after scaling to 100,000 tags/library)	GEO accession number
E15_1†	U	E15.5	Rostral cerebral cortex	2,044	43,327	15,664	36,153	GSM375449
E15_2†	U	E15.5	Caudal cerebral cortex	2,925	39,314	19,929	50,692	GSM375450
E15_3‡	U	E15.5	Cerebral cortex	1,920	44,332	15,787	35,611	GSM375451
E17_1†	U	E17.5	Cerebral cortex	384	13,573	7,214	53,150	GSM375638
E17_2†	U	E17.5	Cerebral cortex	1,053	47,733	19,508	40,869	GSM375639
PI_1‡	U	P1.5	Cerebral cortex	4,194	42,869	20,465	47,738	GSM375640
Ad_1	M	4 months	Left cerebral cortex	2,016	50,760	19,032	37,494	GSM375458
Ad_2	M	4 months	Left cerebral cortex	1,536	52,476	19,157	36,506	GSM375459
Ad_3	F	5-6 months	Left cerebral cortex	2,688	30,914	15,998	51,750	GSM375460
Ad_4	F	5 months	Left cerebral cortex	1,728	43,592	17,262	39,599	GSM375461
Ad_5	F	5-6 months	Right cerebral cortex	2,684	53,292	21,693	40,706	GSM375462
Ad_6	F	6 months	Right cerebral cortex	3,264	69,084	31,762	45,976	GSM375463

*Each library was constructed by using an independent mouse cerebral cortex. †Short tags were generated from these libraries. ‡These libraries were obtained from [27]. F: female; M: male; U: undetermined sex.

Table 2**Summary of tag classification into various categories and comparisons**

Category	Number of DETs†	DETs in various comparisons*						Total
		E15.5 R versus C	Adult L versus Ri	E15.5 versus E17.5	E15.5 versus P1.5	P1.5 versus Ad	E versus Ad	
Gene	386 (68.8)	25 (56.8)	10 (58.8)	24 (63.2)	16 (44.4)	114 (59.4)	253 (77.1)	442
EST	33 (5.9)	2 (4.5)	2 (11.8)	1 (2.6)	1 (2.8)	9 (4.7)	24 (7.3)	39
Multiple matches	55 (9.8)	9 (20.5)	3 (17.6)	5 (13.2)	4 (11.1)	21 (10.9)	23 (7.0)	65
Ambiguous	44 (7.8)	4 (9.1)	2 (11.8)	6 (15.8)	4 (11.1)	14 (7.3)	20 (6.1)	50
No match	43 (7.7)	4 (9.1)	0 (0)	2 (5.3)	11 (30.6)	34 (17.7)	8 (2.4)	59
Total	561	44	17	38	36	192	328	655

Tag classification into various categories and comparisons was based on the mouse genome assembly released in July 2007. Values in parentheses represent the 'percentage' across a column. *Total number of tags for various comparisons is 655 (rather than 561) due to the same tag being statistically significant in more than one comparison. †Differentially expressed tags. Ad: adult stage; C: caudal region; E: embryonic day/stage; L: adult left hemisphere; P: postnatal days; R: rostral region; Ri: adult right hemisphere.

likely that these DETs belong to transcripts from less defined regions, such as centromeric and telomeric areas of a chromosome or assembly gaps of the mouse genome.

Hierarchical clustering of DETs

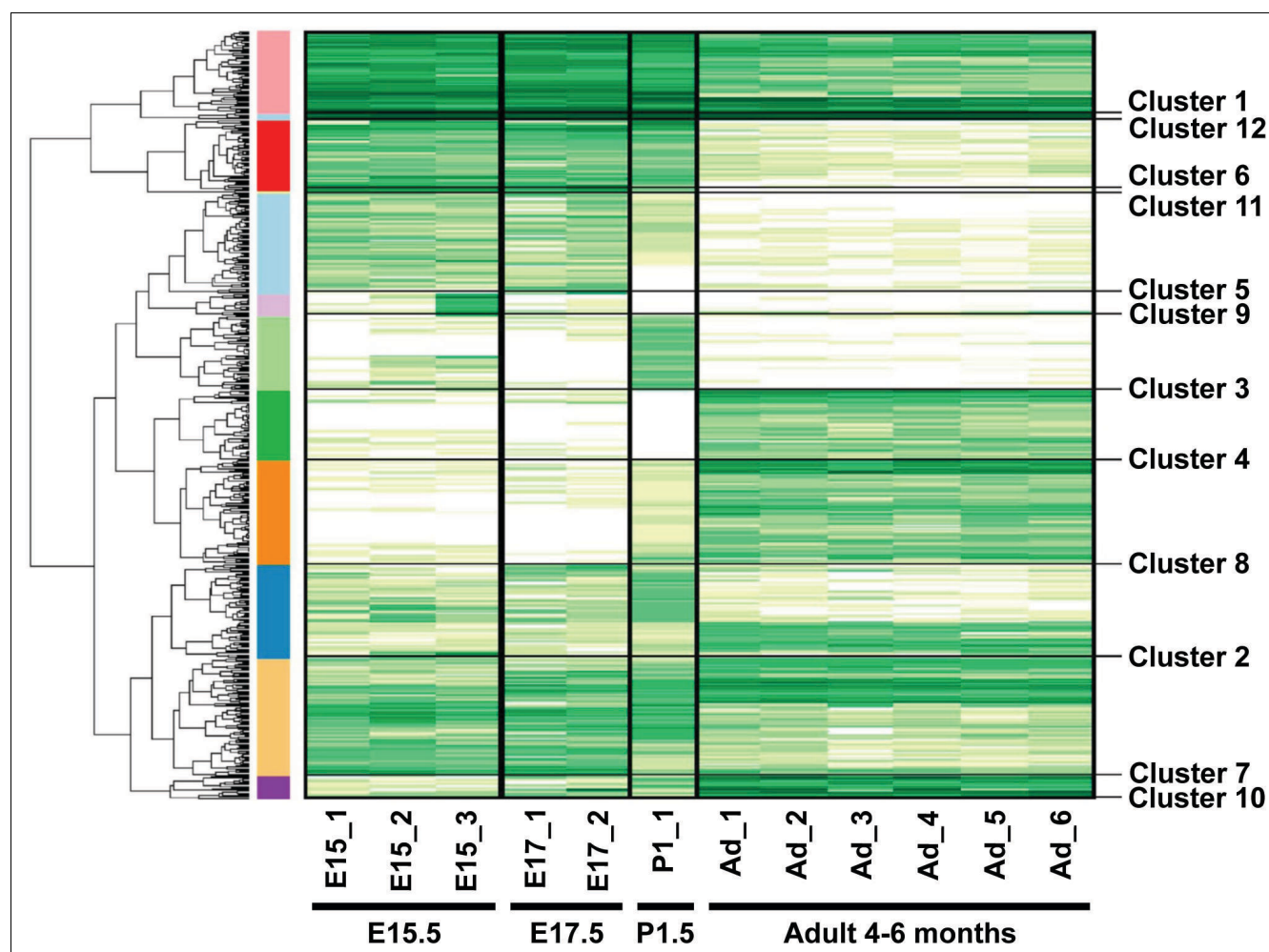
To identify co-regulated genes, all 561 DETs were hierarchically grouped into 12 clusters based on the calculation of the Euclidean distance of logged normalized counts (Figure 1). Clusters 1, 5 and 6 consist of embryonic-specific DETs that exhibit the highest expression during embryonic development of the cerebral cortex. DETs in cluster 1 are expressed throughout all stages of development but exhibit the lowest expression in the adult cortex. Expression of DETs in cluster 5 ceases prior to birth, whereas DETs in cluster 6 are expressed up to early postnatal stage. On the other hand, clusters 4, 8 and 10 consist of adult specific DETs, showing very similar temporal expression profiles, but with different magnitudes (for example, highest expression in cluster 10). Clusters 2 and 7 are termed 'gene-switching' clusters as they show interesting expression-switching profiles. Cluster 2 shows an expression switch between P1.5 and adult stages whilst cluster 7 shows an expression switch between embryonic and adult stages. Clusters 3 and 9 consist of DETs showing region-(caudal region of E15.5 cerebral cortex) and stage-specific (P1.5 only) expression. Clusters 11 and 12 were excluded from subsequent analyses because they contained very few annotated tags. DETs within the same cluster may be co-regulated and/or involved in similar biological functions during cerebral corticogenesis.

We performed a systematic gene ontology functional clustering using the standardized Gene Ontology term analysis tool DAVID (Database for Annotation, Visualization and Integrated Discovery) [30] (Additional data file 3). Functional analysis of these gene clusters reveals that they have distinctive roles during cerebral corticogenesis. Embryonic-specific

gene clusters (1, 5 and 6) are dominated by genes associated with cellular protein and macromolecule metabolic processes or biosynthesis, and nervous system and neuron development. These results match with the expected functional ontologies during embryonic cerebral cortex development in which neuronal migration, differentiation and axonogenesis events are at their peak. In contrast, adult-specific gene clusters (4, 8 and 10) consist of genes related to biological processes in the mature cerebral cortex, such as ion homeostasis, synaptic transmission and regulation of neurotransmitter level. In addition, these gene clusters are also enriched for ribonucleotide/nucleotide binding activity and components of cytoplasmic membrane-bound vesicles. These functional ontologies are in accordance with adult cerebral cortex function, which features synaptogenesis and nerve impulse transmission at synapses. Gene-switching clusters 2 and 7 are enriched with gene ontologies that are similar to both the embryonic- and adult-specific gene clusters. In addition, these gene clusters are also enriched for microtubule cytoskeleton organization and biogenesis, nucleotide biosynthesis and regulation of mRNA translation processes.

Quantitative RT-PCR validation of DETs and gene clusters

To ascertain the robustness of the SAGE datasets, we selected 136 candidate DETs and two additional genes of interests (ATPase, Cu⁺⁺ transporting, alpha polypeptide, *Atp7a*, and cullin-associated and neddylation-dissociated 2, *Cand2*) for validation after considering both stage-to-stage and hierarchical based analyses (Table S2 in Additional data file 1). The selected DETs are transcription regulators, chromatin modifiers or post-translational regulators, such as ubiquitination pathway related molecules. Seventeen DETs are ESTs, which have been identified in brain-related mouse cDNA libraries or transcriptomes. Independent quantitative RT-PCR (RT-qPCR) validation was carried out using three biological repli-

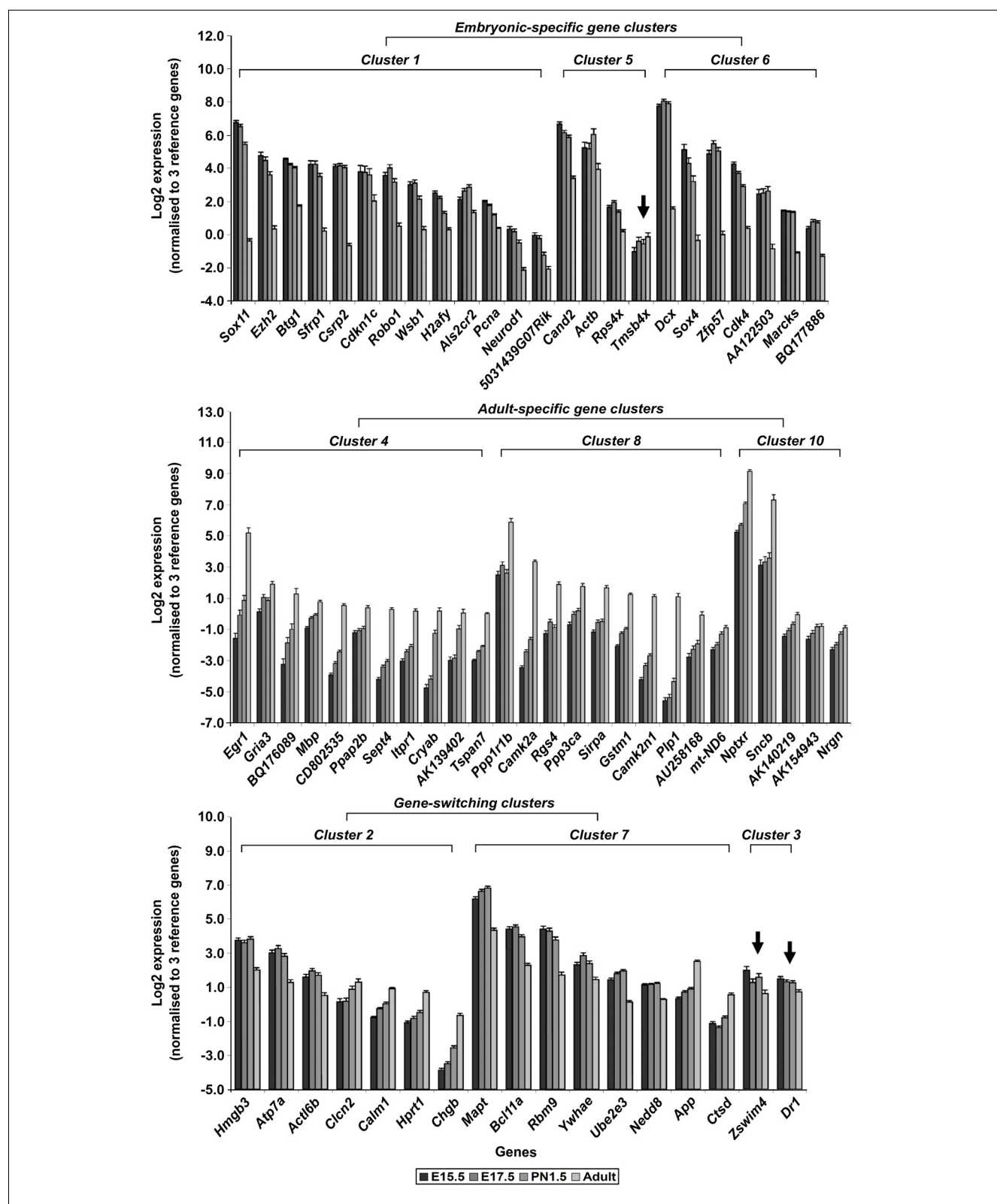
**Figure 1**

Hierarchical clustering of 561 differentially expressed transcripts/tags. Log₂ of normalized counts of DETs from cerebral cortices of various developmental stages were clustered. Dark green clusters denote high levels of expression whereas light green to white clusters denote low levels of expression. The x-axis represents the SAGE libraries whereas the y-axis represents the SAGE tags. The panel on the right shows the 12 different clusters.

cates of unpooled cerebral cortex total RNA for each developmental stage. We validated 70 DETs (including 10 ESTs) from SAGE profiles of comparisons between two developmental stages (Additional data file 3) after considering various stringent criteria and cutoffs. The RT-qPCR results for all the 70 validated DETs, *Atp7a* and *Cand2* are presented in Tables 3, 4, 5 and 6 (Additional data file 4). To validate the expression profiles of gene clusters from the hierarchical clustering analysis, we performed additional RT-qPCR analyses on 65 validated DETs (based on the analysis of two developmental stages), *Atp7a* and *Cand2* by including other developmental stages (Additional data file 5). The analysis validated 62 out of 65 DETs (Figure 2). Of these, 22 are from embryonic gene clusters, 26 from adult gene clusters, and 14 from gene-switching clusters. We assigned *Atp7a* and *Cand2* to clusters 2 and 5, respectively, based on their RT-qPCR expression profiles across all the developmental stages. See Tables 3, 4, 5 and 6 for complete list of DETs and their full gene names.

According to the comparison between two developmental stages, the most abundantly expressed and validated DETs in the E versus Ad analysis are *Camk2a*, *Egr1* and *Plp1* (Table 3). In the adult cerebral cortex, the expression of these DETs is more than 100-fold greater than that at E15.5. Other validated adult-specific DETs with expression levels of approximately tenfold or greater than those in the embryonic brain are *Camk2n1*, *Cryab*, *Nrgn*, *BO176089*, *CD802535*, *Sncb* and *Nptxr*. Conversely, *Sox11*, an embryonic specific DET, is expressed in the E15.5 and E17.5 cerebral cortices with an expression level of at least 100-fold greater than that in the adult. Other validated embryonic specific DETs are *Dcx*, *Zfp57*, *Ezh2*, *Sfrp1* and *Cdk4*, which have expression levels approximately tenfold or greater than those in the adult.

In the P1.5 versus Ad analysis, *Dcx* is expressed at a level approximately 80-fold greater in P1.5 compared to the adult cerebral cortex (Table 4). Other P1.5 enriched and validated

**Figure 2**

High-throughput RT-qPCR validation of gene clusters. All validations were based on DETs for canonical mRNA. Failed validation of DETs according to hierarchical clustering expression profiles is indicated by arrows. N = 3 and data are presented as mean \pm standard error of the mean.

Table 3**RT-qPCR validation of SAGE profile for E versus adult comparison**

SAGE tag	RefSeq accession	Gene ID	Fold change			
			Ad/E15.5 (SAGE)	Ad/E17.5 (SAGE)	Ad/E15.5 (RT-qPCR)	Ad/E17.5 (RT-qPCR)
gcttccccacccccctt	NM_177407	calcium/calmodulin-dependent protein kinase II alpha, <i>Camk2a</i>	As	As	111.59	54.74
ggatatgtgtgtgtac	NM_007913	early growth response I, <i>Egr1</i>	As	As	108.47	38.93
aaattattgggaatcc	NM_011123	proteolipid protein (myelin) I, <i>Plp1</i>	28.77	As	103.15	89.13
gtatttgcataaaaaaa	NM_025451	calcium/calmodulin-dependent protein kinase II inhibitor I, <i>Camk2n1</i>	69.59	59.76	40.08	21.46
gcttcctctccaggag	NM_009964	crystallin, alpha B, <i>Cryab</i>	As	11.59	30.47	20.58
ttaccatactgggttg	NM_022029	neurogranin, <i>Nrgn</i>	24.67	8.47	43.67	17.00
cctcatttccctgttc	BC176089	EST from adult C57BL/6 subfornical organ and postrema tissues	As	As	23.01	8.91
acccggcttagtgaa	NM_011129	septin 4, <i>Sept4</i>	18.97	16.29	22.38	12.96
ctcattataatcaagaa	CD802535	EST from 27-32 days C57BL/6 hippocampus tissue	As	9.05	22.03	13.25
aataaagccaatctgac	NM_033610	synuclein, beta, <i>Sncb</i>	20.89	7.66	18.20	15.83
gcttttgttaccatctc	NM_030689	neuronal pentraxin receptor, <i>Nptxr</i>	17.88	12.24	15.02	10.93
tccctcccttagtatcc	NM_144828	protein phosphatase I, regulatory (inhibitor) subunit 1B, <i>Ppp1r1b</i>	As	As	10.47	6.88
gcccccttctcattggc	NM_010358	glutathione S-transferase, mu I, <i>Gstm1</i>	As	As	9.87	5.58
tgactagcgtgacctgt	NM_007694	chromogranin B, <i>Chgb</i>	6.13	6.30	9.43	7.22
atttcttttctggatgg	NM_010585	inositol 1,4,5-triphosphate receptor I, <i>Itpr1</i>	14.34	6.16	9.29	6.12
actttgagattgtacct	NM_009062	regulator of G-protein signaling 4, <i>Rgs4</i>	12.23	21.01	8.84	5.32
aataattagccttaggt	AK139402	Mus musculus 10 days neonate cortex cDNA	As	As	8.28	7.55
ctagacagaggcattat	NM_019634	tetraspanin 7, <i>Tspan7</i>	13.08	5.61	7.99	5.38
tgtatacacacacgggt	NM_007547	signal-regulatory protein alpha, <i>Sirpa</i>	As	As	7.18	4.63
tgacaagacactgtggc	AU258168	EST from mouse brain	As	As	6.49	4.58
cttacctcaggtttcct	NM_008913	protein phosphatase 3, catalytic subunit, alpha isoform, <i>Ppp3ca</i>	As	5.43	5.48	3.45
atagctttctacacact	NM_007471	amyloid beta (A4) precursor protein, <i>App</i>	3.98	2.74	4.56	3.49
tttcagcagtggttgct	NM_013556	hypoxanthine guanine phosphoribosyl transferase I, <i>Hprt1</i>	6.90	8.87	3.45	2.85
aggtatgtacaaagttt	NM_016886	glutamate receptor, ionotropic, AMPA3 (alpha 3), <i>Gria3</i>	4.97	8.51	3.40	1.81*
tccaactgttaactata	NM_009790	calmodulin I, <i>Calm1</i>	4.23	3.17	3.24	2.27

Table 3 (Continued)**RT-qPCR validation of SAGE profile for E versus adult comparison**

cctcagcctgggtaga	NM_009983	cathepsin D, <i>Ctsd</i>	3.46	2.04	3.22	3.84
gcttcgtccacacagcg	NM_010777	myelin basic protein, <i>Mbp</i>	202.46	173.85	3.19	2.05
tattaaatgtgcttttt	NM_080555	phosphatidic acid phosphatase type 2B, <i>Ppap2b</i>	5.71	2.44	3.02	2.76
cttatcctcacctcagc	NC_005089	NADH dehydrogenase 6, mitochondrial, <i>mt-ND6</i>	As	As	2.65	2.13
caaacctccaaaacca	AK140219	Mus musculus adult male corpora quadrigemina cDNA	29.22	18.62	2.63	2.03
agtggctaattaggtgt	NM_009900	chloride channel 2, <i>Gln2</i>	14.55	4.16	2.23	2.18
accaatgaacaaaaaa	AK154943	Mus musculus NOD-derived CD11c +ve dendritic cells cDNA	109.40	51.13	1.76*	1.35 (NS)
ccagtacctgaaaaaaa	NM_008453	Kruppel-like factor 3 (basic), <i>Klf3</i>	As	-52.97	-1.471†	-1.27 (NS)
aagaaaacattttaata	NM_012010	eukaryotic translation initiation factor 2, subunit 3, structural gene X-linked, <i>Eif2s3x</i>	-7.70	-10.38	-1.72*	-1.74*
caccctgtgggagctca	NM_172656	amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 2 (human), <i>Als2cr2</i>	-11.03	-12.88	-1.75*	-2.48
cctccatcctttatact	NM_009536	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon polypeptide, <i>Ywhae</i>	-3.19	-2.52	-1.82*	-2.66
tgtgcttcctgtctta	NM_008683	neural precursor cell expressed, developmentally down-regulated gene 8, <i>Neddd8</i>	-4.73	-8.41	-1.83	-1.86
ctcctgaagcatagtt	NM_009454	ubiquitin-conjugating enzyme E2E 3, UBC4/5 homolog (yeast), <i>Ube2e3</i>	-4.14	-5.52	-2.51	-3.22
gtgaaactaaaaaaa	NM_009094	ribosomal protein S4, X-linked, <i>Rps4x</i>	-10.46	-19.19	-2.76	-3.35
aatgtttctgctttaca	NM_011045	proliferating cell nuclear antigen, <i>Pcna</i>	-14.21	-8.30	-3.09	-2.62
cgtgatccctctgtca	NM_009876	cyclin-dependent kinase inhibitor 1C (P57), <i>Cdkn1c</i>	-13.68	-7.16	-3.39*	-3.30*
cctttgtgacagtggcc	NM_025635	ZW10 interactor, <i>Zwint</i>	7.29	8.84	-3.47‡	-4.80‡
gaagccagtgggccatc	NM_001033273	RIKEN cDNA 5031439G07 gene, <i>5031439G07Rik</i>	-9.10	-10.00	-4.13	-3.65
gctgtgggtcgctgtgg	NM_010561	interleukin enhancer binding factor 3, <i>Ilf3</i>	As	As	-4.15‡	-2.97‡
accctgacccctgtgt	NM_016707	B-cell CLL/lymphoma 11A (zinc finger protein), <i>Bcl11a</i>	-7.88	-5.52	-4.42	-4.78

Table 3 (Continued)**RT-qPCR validation of SAGE profile for E versus adult comparison**

cggtgtccccacctcc	NM_012015	H2A histone family, member Y, <i>H2afy</i>	-21.48	-27.88	-4.62	-3.70
cagttgcaataaaaa	NM_010894	neurogenic differentiation 1, <i>Neurod1</i>	-7.54	-4.96	-5.62	-5.10
aagtttgcaagtctcca	NM_008538	myristoylated alanine rich protein kinase C substrate, <i>Marcks</i>	-24.87	-11.76	-5.72	-5.56
ttgctggcttttataa	NM_053104	RNA binding motif protein 9, <i>Rbm9</i>	-6.10	-7.82	-6.50	-5.97
ggttttgtttgttgac	NM_019653	WD repeat and SOCS box-containing 1, <i>Wsb1</i>	-7.88	-11.96	-6.54	-7.00
tatattgattgtgcaa	NM_007569	B-cell translocation gene 1, anti-proliferative, <i>Btg1</i>	-15.28	-16.48	-6.93	-5.58
taagaaacct	NM_019413	roundabout homolog 1 (<i>Drosophila</i>), <i>Robo1</i>	-28.64	-16.73	-8.23	-11.35
gctttgactgttctct	AA122503	EST from M2 cells of skin melanoma	-17.17	-18.68	-9.88	-10.15
tggagcgttggtgtat	NM_009870	cyclin-dependent kinase 4, <i>Cdk4</i>	-114.55	-80.88	-14.77	-10.00
ctttccctgccaatgta	NM_013834	secreted frizzled-related protein 1, <i>Sfrp1</i>	<i>Es</i>	<i>Es</i>	-16.48	-16.37
tgagctttctgttcaa	NM_007971	enhancer of zeste homolog 2 (<i>Drosophila</i>), <i>Ezh2</i>	-10.63	-5.52	-21.36	-17.59
cacgacacccccaccc	NM_009559	zinc finger protein 57, <i>Zfp57</i>	-30.72	-42.33	-29.13	-44.46
tgtgtgaggtgtgtga	NM_010025	doublecortin, <i>Dcx</i>	-44.41	-58.81	-73.76	-91.10
cagagtgtagtgtgtg	NM_009234	SRY-box containing gene 11, <i>Sox11</i>	<i>Es</i>	<i>Es</i>	-140.05	-120.35

All RT-qPCR data are statistically significant at $P < 0.001$ unless specified: * $P < 0.01$; † $P < 0.05$. ‡A disagreement between RT-qPCR and SAGE data. As, adult-specific expression; Es, embryonic-specific expression; NS, no statistically significant difference between two developmental stages. Fold change values of <1.0 are presented in a negative fold change format.

Table 4**RT-qPCR validation of SAGE profile for P1.5 versus adult stages**

SAGE tag	RefSeq accession	Gene ID	Fold change	
			Adult/P1.5 (SAGE)	Adult/P1.5 (RT-qPCR)
aaattattgggaaatcc	NM_011123	proteolipid protein (myelin) 1, <i>Plp1</i>	38.9	43.33
tttcagcagtggtgct	NM_013556	hypoxanthine guanine phosphoribosyl transferase 1, <i>Hprt1</i>	6.99	2.26
actcggagccaccagac	NM_009790	calmodulin 1, <i>Calm1</i>	4.57	1.83
gcttcgtccacacagcg	NM_010777	myelin basic protein, <i>Mbp</i>	As	1.79
tccccgtcat	NM_026106	down-regulator of transcription 1, <i>Dr1</i>	Ps	-1.46
gggaaactaaggagag	NM_172503	zinc finger, SWIM domain containing 4, <i>Zswim4</i>	-42.42	-1.96*
gaacgcaagttcagccc	NM_031404	actin-like 6B, <i>Actl6b</i>	-5.51	-2.26
gtgaaactaaaaaaaa	NM_009094	ribosomal protein S4, X-linked, <i>Rps4x</i>	-4.72	-2.27
Nil	NM_009726	ATPase, Cu ⁺⁺ transporting, alpha polypeptide, <i>Atp7a</i>	Nil	-2.91
agaagtgttgagttt	NM_008253	high mobility group box 3, <i>Hmgb3</i>	-20.99	-3.53
cctttgtgacagtggcc	NM_025635	ZW10 interactor, <i>Zwint</i>	4.65	-4.04†
acagtctatgttgagg	BQ177886/NM_010487	C57BL/6 whole brain E15.5 (or known as embryonic lethal, abnormal vision, Drosophila-like 3 (Hu antigen C), <i>Elavl3</i>)	-41.99	-4.06
gatacttgaatgacta	NM_007393	actin, beta, cytoplasmic, <i>Actb</i>	-18.49	-4.35
ctggcttctt	NM_008538	myristoylated alanine rich protein kinase C substrate, <i>Marcks</i>	Ps	-4.76
Nil	NM_025958	cullin-associated and neddylation-dissociated 2 (putative), <i>Cand2</i>	Nil	-5.62
gctttgactgttctt	NM_009870	cyclin-dependent kinase 4, <i>Cdk4</i>	-56.55	-5.77
ctcagtaatg	AA122503	EST from M2 cells of skin melanoma	-12.28	-11.12
tggagcgttggtgtat	NM_009238	SRY-box containing gene 4, <i>Sox4</i>	-31.51	-11.74
cacgacacccccaccc	NM_007792	cysteine and glycine-rich protein 2, <i>Csrp2</i>	Ps	-26.10
tgtgtgaggtgtgtga	NM_009559	zinc finger protein 57, <i>Zfp57</i>	-16.33	-32.74
gggacctcgtggaagcc	NM_010025	doublecortin, <i>Dcx</i>	-24.56	-82.11

All RT-qPCR data are statistically significant at $P < 0.001$ unless specified: * $P < 0.01$. †A disagreement between RT-qPCR and SAGE data. As: adult-specific expression; Nil: SAGE data not available; Ps: P1.5-specific expression. Fold change values of <1.0 are presented in a negative fold change format.

Table 5**RT-qPCR validation of SAGE profile for E15.5 versus P1.5 stages**

SAGE tag	RefSeq accession	Gene ID	Fold change	
			P1.5/E15.5 (SAGE)	P1.5/E15.5 (RT-qPCR)
Atttcttgggtgatttt	NM_010838	microtubule-associated protein tau, <i>Mapt</i>	5.51	1.53*
Gcactgttaacaagtgt	NM_009234	SRY-box containing gene 11, <i>Sox11</i>	-2.17	-3.32

All RT-qPCR data are statistically significant at $P < 0.001$ unless specified: * $P < 0.05$. Fold change values of <1.0 are presented in a negative fold change format.

Table 6**RT-qPCR validation of SAGE profile for rostral E15.5 and caudal E15.5 regions**

SAGE tag	RefSeq accession	Gene ID	Fold change	
			Caudal/rostral (SAGE)	Caudal/rostral (RT-qPCR)
gtgttcttccagtcgg	NM_016916	bladder cancer associated protein homolog (human), <i>Blcap</i>	2.68	1.43
gtcatagctgtctgtg	BC025816	EST sequence BC025816	Cs	1.31 (NS)
aagcttgacatttgaa	NM_026187	ankyrin repeat and zinc finger domain containing 1, <i>Ankzf1</i>	Rs	1.17 [†] (NS)
gatacttggaatgacta	NM_007393	actin, beta, cytoplasmic, <i>Actb</i>	-2.63	-1.32
ttggtgaaggaaaaaac	NM_021278	thymosin, beta 4, X chromosome, <i>Tmsb4x</i>	-2.33	-2.13 (NS)

All RT-qPCR data are statistically significant at $P < 0.05$ unless specified. [†]A disagreement between RT-qPCR and SAGE data. Cs: caudal region-specific expression; NS: no statistically significant difference between two developmental stages; Rs: rostral region-specific expression. Fold change values of <1.0 are presented in a negative fold change format.

DETs are (in descending order of enrichment) *Zfp57*, *Csrp2*, *AA122503*, *Cdk4*, *Sox4*, *Marcks*, *Actb*, [BQ177889](#), *Hmgb3*, *Rps4x*, *Actl6b*, *Zswim4* and *Dr1*, whose expression ranges from 33- to 1.4-fold greater than in the adult. On the other hand, the *Plp1* is expressed at a level 40 times greater in the adult cerebral cortex compared to P1.5. Other validated genes that are enriched in the adult cerebral cortex include (in descending order) *Hprt1*, *Calm1* and *Mbp*, with a 2.3- to 1.8-fold greater expression than the P1.5 cerebral cortex. Comparison between E15.5 and P1.5 shows that *Mapt* has a 1.5-fold greater expression level in the P1.5 cerebral cortex while *Sox11* expression is 3.3-fold lower (Table 5).

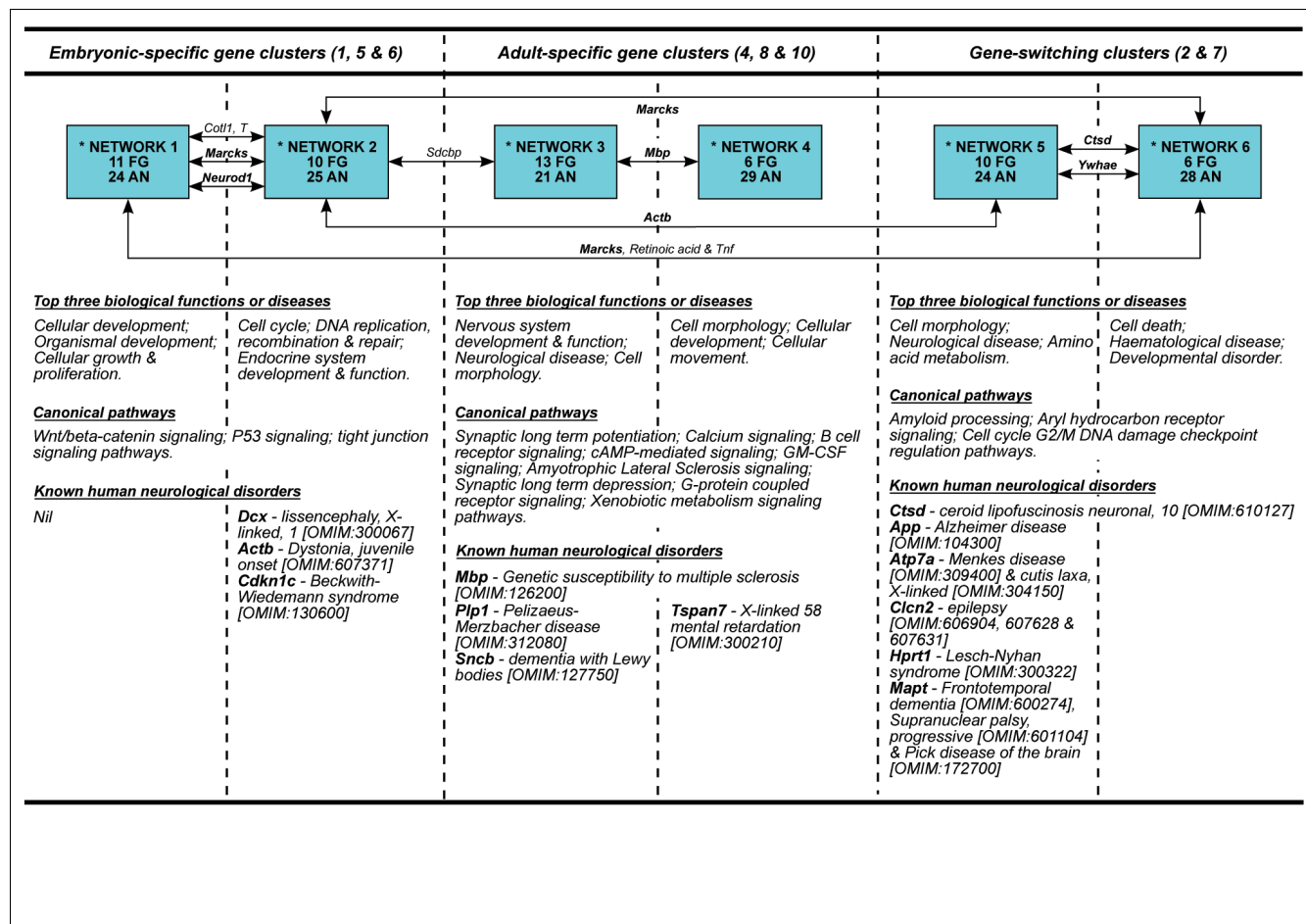
We were unable to validate all 17 DETs from L versus Ri regions, suggesting the left and right hemispheres of the adult mouse cerebral cortex are highly similar and indistinguishable at the molecular level. SAGE and RT-qPCR analyses for R versus C regions of E15.5 are discussed in a separate section below.

Functional analysis of validated gene clusters using Ingenuity Pathway Analysis

The validated DETs of embryonic, adult and gene-switching clusters were functionally characterized using proprietary software, Ingenuity Pathway Analysis (IPA) from Ingenuity Systems®, to identify enriched molecular networks and canonical pathways. Given a list of input genes (known as focus genes), IPA mapped these genes to a global molecular network developed from information contained in the Ingenuity knowledge base (a manually curated database of experimentally proven molecular interactions from published literature). Networks of these focus genes were then algorithmically generated based on their connectivity. IPA determined the most significantly enriched biological functions and/or related diseases by calculating the P -value using Fisher's exact test. Using similar methods, significantly represented canonical pathways in a set of focus genes were also determined using IPA (Section C in Additional data file 1).

From the embryonic-specific gene clusters, we identified two statistically significant molecular networks (made up of 19 focus genes and 47 associated nodes; networks 1 and 2 in Figure 3; Figures S4 and S5 in Additional data file 1). The networks are interconnected through two genes, *Marcks* and *Neurod1*. In general, these networks are associated with cell cycle and DNA replication, recombination and repair processes. Three statistically significant (using $P < 0.05$ as a cut-off) canonical pathways are enriched in these networks (Figure 3); Wnt/ β -catenin signaling (*Sox4*, *Sox11* and *Sfrp1*), P53 signaling (*Cdk4* and *Pcna*) and tight junction signaling (*Cdk4* and *Actb*) pathways. Validated DETs such as *Btg1*, *Cdk4*, *Cdkn1c*, *Csrp2*, *Ezh2*, *Neurod1*, *Pcna*, and *Rps4x* are associated with cell cycle control whereas *Actb*, *Ezh2*, *Als2cr2*, *Marcks*, *Robo1* and *Dcx* are associated with cellular assembly and organization. These processes are important in the formation of filopodia, membrane blebs and growth cones during neuronal growth, migration and axonogenesis. Known human neurological disorders associated with the networks, particularly network 2, include X-linked lissencephaly (Online Mendelian Inheritance in Man [OMIM:300067]; DCX), juvenile onset dystonia ([OMIM:607371]; ACTB) and Beckwith-Wiedemann syndrome ([OMIM:130600]; CDKN1C). All the DETs implicated in these networks are expressed in the cortical plate with the exception of *Pcna* (Table S5 in Additional data file 1) [31-42].

In adult-specific gene clusters, two molecular networks (18 focus genes and 50 associated nodes; networks 3 and 4 in Figure 3) were identified and interconnected via a single gene, *Mbp* (Figures S6 and S7 in Additional data file 1). These molecular networks enrich for nine statistically significant canonical pathways ($P < 0.05$) such as synaptic long-term potentiation and depression, calcium signaling, B cell receptor signaling, cAMP-mediated signaling, GM-CSF signaling, amyotrophic lateral sclerosis signaling, G-protein-coupled receptor signaling and xenobiotic metabolism signaling pathways. Validated DETs such as *Camk2a*, *Gria3*, *Itpr1*, *Egr1*,

**Figure 3**

Novel molecular networks involved in cerebral corticogenesis. The figure shows novel molecular networks, related biological functions/diseases, canonical pathways and known human neurological disorders based on Ingenuity Pathway Analysis and OMIM database. Detailed molecular interactions for all networks (indicated by asterisks) are illustrated in Figures S4, S5, S6, S7, S8 and S9 in Additional data file 1. Gene names next to arrow lines refer to common genes shared by two networks. Bold gene name refers to a focus gene. AN: associated nodes; FG: focused genes.

Rgs4, *Gstm1* and *Ppp3ca* are associated with cell-to-cell communication in the adult cerebral cortex. Biological functions such as cell proliferation, movement and death processes are also linked to these networks through the genes *Camk2a*, *Cryab*, *Egr1*, *Gstm1*, *Itp1*, *Mbp*, *Nrgn*, *Ppp1r1b*, *Ppp3ca*, *Rgs4*, *Sept4*, *Sirpa* and *Sncb*. Known human neurological disorders associated with Network 3 include Pelizaeus-Merzbacher disease ([OMIM:312080]; *PLP1*) and dementia with Lewy bodies ([OMIM:127750]; *SNCB*), whereas X-linked mental retardation ([OMIM:300210]; *TSPAN7* or *TM4SF2*) is associated with network 4. *MBP*, found in both network 3 and 4, is associated with genetic susceptibility to multiple sclerosis [OMIM:126200]. All the DETs implicated in these networks are expressed in layers I to VI of the cerebral cortex, with the exception of *Gstm1* (layers II to V), *Sirpa* and *Ppp3ca* (layers II to VI), and *Nptxr* (no expression data available) (Table S5 in Additional data file 1) [32-34,43-52].

Both networks 5 and 6 (Figure 3) are linked to gene-switching clusters. These two molecular networks (14 focus genes and 52 associated nodes) are associated with cellular morphogenesis, amino acid metabolism, cell death processes and developmental disorders (Figures S8 and S9 in Additional data file 1). These networks are connected by *Ctsd* and *Ywhae* and are implicated in amyloid processing (*App* and *Mapt*), aryl hydrocarbon receptor signaling (*Ctsd* and *Nedd8*) and cell cycle G2/M DNA damage checkpoint regulation pathways (*Ywhae*). *App*, *Atp7a*, *Hprt1*, *Mapt*, *Ctsd* and *Cln2* are involved in cell morphogenesis, assembly and organization. Among all the networks in the study, network 5 has the greatest number of DETs associated with known human neurological disorders with six out of ten focus genes being associated with known neurological disorders such as neuronal ceroid lipofuscinosis ([OMIM:610127]; *CTSD*), Alzheimer's disease ([OMIM:104300]; *APP*), Menkes disease ([OMIM:309400]; *ATP7A*), X-linked cutis laxa ([OMIM:304150]; *ATP7A*), epi-

lepsy ([OMIM:606904, 607628 and 607631]; *CLCN2*), Lesch-Nyhan syndrome ([OMIM:300322]; *HPRT1*), fronto-temporal dementia ([OMIM:600274]; *MAPT*), progressive supranuclear palsy ([OMIM:601104]; *MAPT*) and Pick disease of the brain ([OMIM:172700]; *MAPT*). All the DETs implicated in these networks are expressed in the cortical plate during embryonic development with the exception of *Ube2e3*. In the adult, all DETs are expressed in layers I to VI of the cerebral cortex except for *Bcl11a*, *Clcn2*, *Hprt1* (layers II to VI), *Atp7a* and *Ube2e3* (no expression was detected in the adult cerebral cortex) (Table S5 in Additional data file 1) [32-35,46,53-61].

To refine the functional analysis to a cellular level, we grouped all 40 RT-qPCR validated DETs from networks 3, 4, 5 and 6 (adult specific and gene-switching gene clusters) into three groups according to where they are expressed: only in cortical neurons (N group); only in cortical glia (G group); and in both cortical neurons and glia (B group). None of the DETs from networks 1 and 2 (embryonic-specific gene clusters) were analyzed because most of the cells within the cerebral cortex are committed to the neuronal lineage at E15.5. The methods used to tabulate and group validated DETs are detailed in Section D in Additional data file 1. DETs classified as part of the N group are *Bcl11a*, *Calm1*, *Camk2a*, *Camk2n1*, *Hprt1*, *Itpri1*, *Mapt*, *Nedd8*, *Nrgn*, *Ppp1r1b*, *Ppp3ca*, *Rbm9*, *Rgs4*, *Sncb* and *Ywhae*, whereas those in the G group include *Cryab*, *Mbp*, *Nptxr*, *Plp1*, *Ppap2b*, *Sept4* and *Tspan7*. The B group consists of *App*, *Atp7a*, *Chgb*, *Ctsd*, *Egr1*, *Gria3*, *Gstm1* and *Sirpa*. DETs without any cellular expression information, such as *Actl6b*, *Clcn2*, *Hmgb3*, *Ube2e3*, [AK138272](#), [AK139402](#), [AK140219](#), [AK154943](#), [AU258168](#) and [BQ176089](#), were placed into group B to facilitate downstream analysis. Of the 40 DETs, only 35 were mapped to the IPA knowledgebase and subjected to further analysis. Functional analysis of these 35 DETs show that the neuron-based DETs (N and B groups) are enriched for various human neurological disorders, such as schizophrenia (9 DETs; $P = 4.36E-8$), Huntington's disease (8 DETs; $P = 1.03E-5$), X-linked mental retardation (4 DETs; $P = 3.78E-7$), Parkinson's disease (3 DETs; $P = 6.99E-3$) and Alzheimer's disease (3 DETs; $P = 1.52E-2$) (Figure S10 in Additional data file 1). Thirteen DETs associated with these neurological disorders are also implicated in processes related to cell death. Of these, eight DETs are expressed in neurons only, two in glia only and three in both neurons and glia.

From the above IPA analysis, *Cand2* (embryonic gene clusters), *Camk2n1* (adult gene clusters), *Hmgb3* (gene-switching clusters) and 10 ESTs ([BQ176089](#), [CD802535](#), [AK139402](#), [AU258168](#), [AK138272](#), [AK140219](#), [AK154943](#), [AA122503](#), [NM_001033273](#) and [BQ177886](#)) are not currently connected to any networks.

Regionalized expression of DETs in the E15.5 cerebral cortex

Early regionalized development is an important event that could lay the foundations for adult arealization of cerebral functions. To identify genes with regionalized expression profiles, we compared SAGE libraries generated from rostral and caudal regions (equivalent to anterior and posterior regions of the human cerebral cortex) of the E15.5 cerebral cortex. We identified 44 DETs and selected 25 DETs (22 known genes, 1 EST and 2 ambiguous genes; Additional data file 2) and 2 genes of interest (bladder cancer associated protein, *Blcap*, and ankyrin repeat and zinc finger domain containing 1, *Ankzf1*) for RT-qPCR validation and further detailed region-based analysis. This was done using independent biological triplicates of clearly defined regions/quadrants, such as rostral-lateral (RL), rostral-medial (RM), caudo-lateral (CL) and caudo-medial (CM) of the E15.5 cerebral cortex (see Materials and methods). Two positive controls with known regionalised expression were included in the RT-qPCR: RAR-related orphan receptor beta (*Rorb*) and nuclear receptor subfamily 2, group F, member 1 (*Nr2f1*). *Rorb* is highly expressed in the rostral region whereas *Nr2f1* is highly expressed in the caudal region of the cerebral cortex [1,62].

An initial RT-qPCR analysis of combined RL and RM (rostral) as well as CL and CM (caudal) regions shows upregulation of *Rorb* and *Nr2f1* in rostral and caudal regions, respectively (based on fold change direction and magnitude of 1.3 times). The same analysis also confirmed the SAGE data for 3 out of 25 DETs (*Actb*, *Tmsb4x* and [BC025816](#)) and *Blcap* (Table 6; Additional data file 6). Both *Actb* and *Tmsb4x* have greater expression in the rostral region whereas *Blcap* and [BC025816](#) are greater in the caudal region. To identify expression profiles in a more refined area and prevent regional compensation due to combined quadrant analysis, we performed a quadrant versus quadrant multiple regions comparison. The largest number of DETs were found in the RL versus CM comparison, as they are the two developmentally most distinct regions within the cerebral cortex at E15.5 compared to others; RL versus CL > RM versus CL > RL versus RM = RM versus CM > CL versus CM. The region-specific expression profiles were plotted for each of the DETs (Figure 4) and grouped into two categories: RL-specific DETs, such as *Actb*, *Tmsb4x* and cytochrome b-245, beta polypeptide (*Cybb*); and CM-specific DETs, including *Blcap*, EST [BC025816](#), *Ankzf1* and cytochrome c oxidase I, mitochondrial (*Cox1*) (Additional data file 6).

To visualize the regionalized expression profiles, we performed *in situ* RNA hybridization (ISH) on all DETs validated by quadrant versus quadrant RT-qPCR analysis, *Blcap*, *Ankzf1* as well as the positive controls *Rorb* and *Nr2f1*. We performed ISH on sagittal and coronal sections (from rostral to caudal regions) of the E15.5 mouse brain (Figure 5). Under dark-field microscopic examination, we confirmed the regionalized expression of *Rorb* (at the rostral cortical plate;

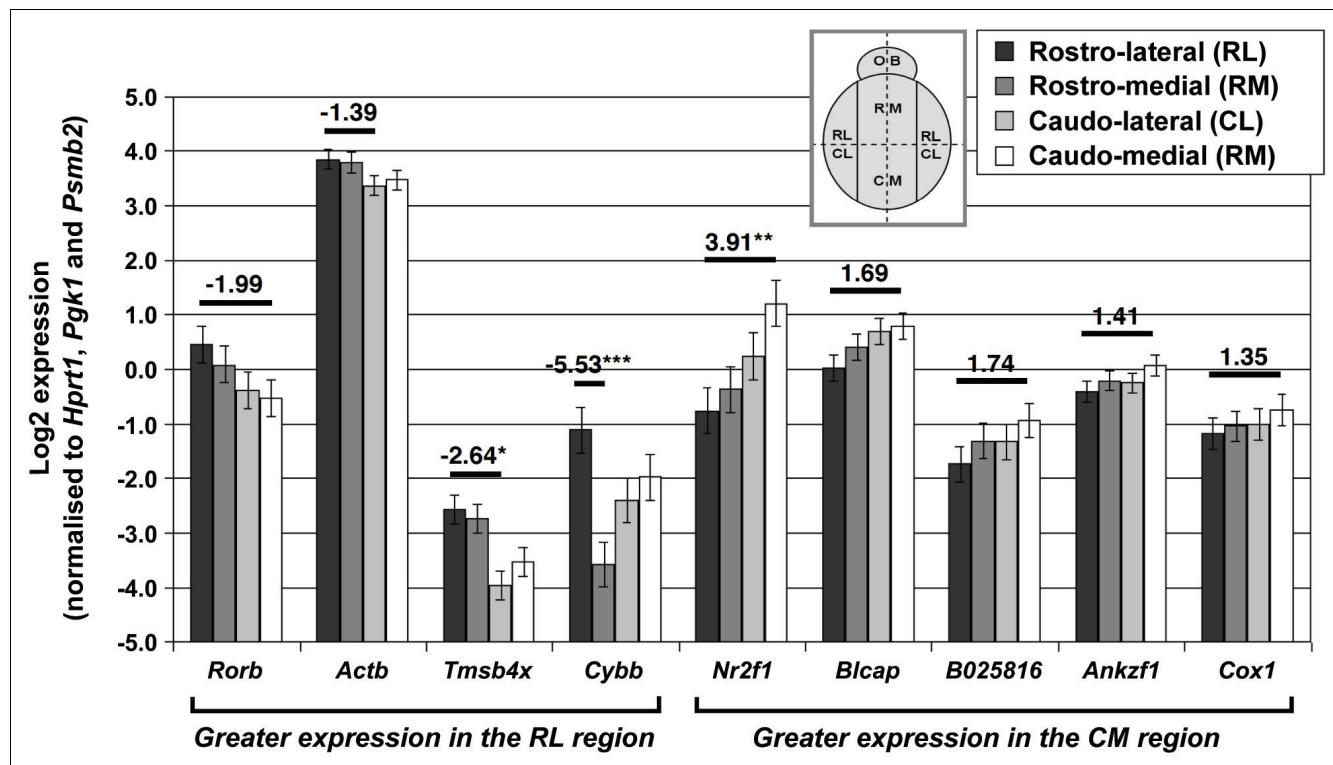


Figure 4

RT-qPCR analysis of all R versus C DETs based on quadrant versus quadrant analysis. RL, RM, CL and CM denote rostro-lateral, rostro-medial, caudo-lateral and caudo-medial regions of the cerebral cortex, respectively. OB denotes olfactory bulb. N = 3 per quadrant and data are presented as mean \pm standard error of the mean. Fold change values (normalized to RL) are presented above the comparative bar and any values <1 are presented in the negative fold change format. Only the most significant fold change value is presented for each target gene. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Figure 5a) and *Nr2f1* (at the caudal ventricular zone; Figure 5f). From the analysis, *Actb* is highly expressed at the cortical plate and the subplate (Figure 5b). *Tmsb4x* is highly expressed at the cortical plate and the intermediate zone (Figure 5c-e).

ISH analysis showed that both *Bcap* and *Ankzf1* are caudal specific. Serial coronal sections from rostral to caudal regions of the brain (Figure 5g-i) show *Bcap* is weakly expressed in the rostral cerebral cortex, particularly at the intermediate zone, subplate and the cortical plate, but is highly expressed in the hippocampus and thalamus (Figure 5i). *Ankzf1* is expressed specifically in the ventricular zone as well as the cortical plate towards the caudal region of the cerebral cortex (Figure 5j-m). The distinctive expression of *Ankzf1* in both the ventricular zone and cerebral cortex prompted us to extend our ISH analysis to various developmental stages starting from E11.5 to adulthood (Figure 5n-t). *Ankzf1* is expressed in the primordial plexiform and the ventricular zone layers of the telencephalon at E11.5 (Figure 5n). By E13.5, *Ankzf1* is weakly expressed in the ventricular zone, but is highly expressed in the cortical plate (Figure 5o). From E17.5 to P1.5, the expression of *Ankzf1* is maintained in the cerebral cortex

(Figure 5p, q). In the adult, *Ankzf1* expression is obvious in the piriform cortex, hippocampus and cerebellum (Figure 5r-t).

Genomic clustering of sense-antisense SAGE tags at the *Sox4* and *Sox11* loci

We performed genomic clustering analysis of the SAGE tags to determine any actively transcribed chromosomal loci throughout cerebral corticogenesis. Probabilities for chance occurrences of two, three, four, and five DETs being clustered within a window of ten adjacent tags present within each chromosomal location, irrespective of genetic distance, were calculated. This analysis was based on the DET lists described above (Additional data file 2). The analysis showed two over-represented chromosome loci at *Sox4* and *Sox11*, which derive from embryonic-specific gene clusters.

At both loci, we observed multiple SAGE tags with both sense and antisense orientations, which signify alternative polyadenylation sites, differential splicing and overlapping antisense transcription. As an initial validation of the antisense messages, we performed strand-specific RT-PCR (Figure 6) using cDNA synthesized from equally pooled total RNAs (three

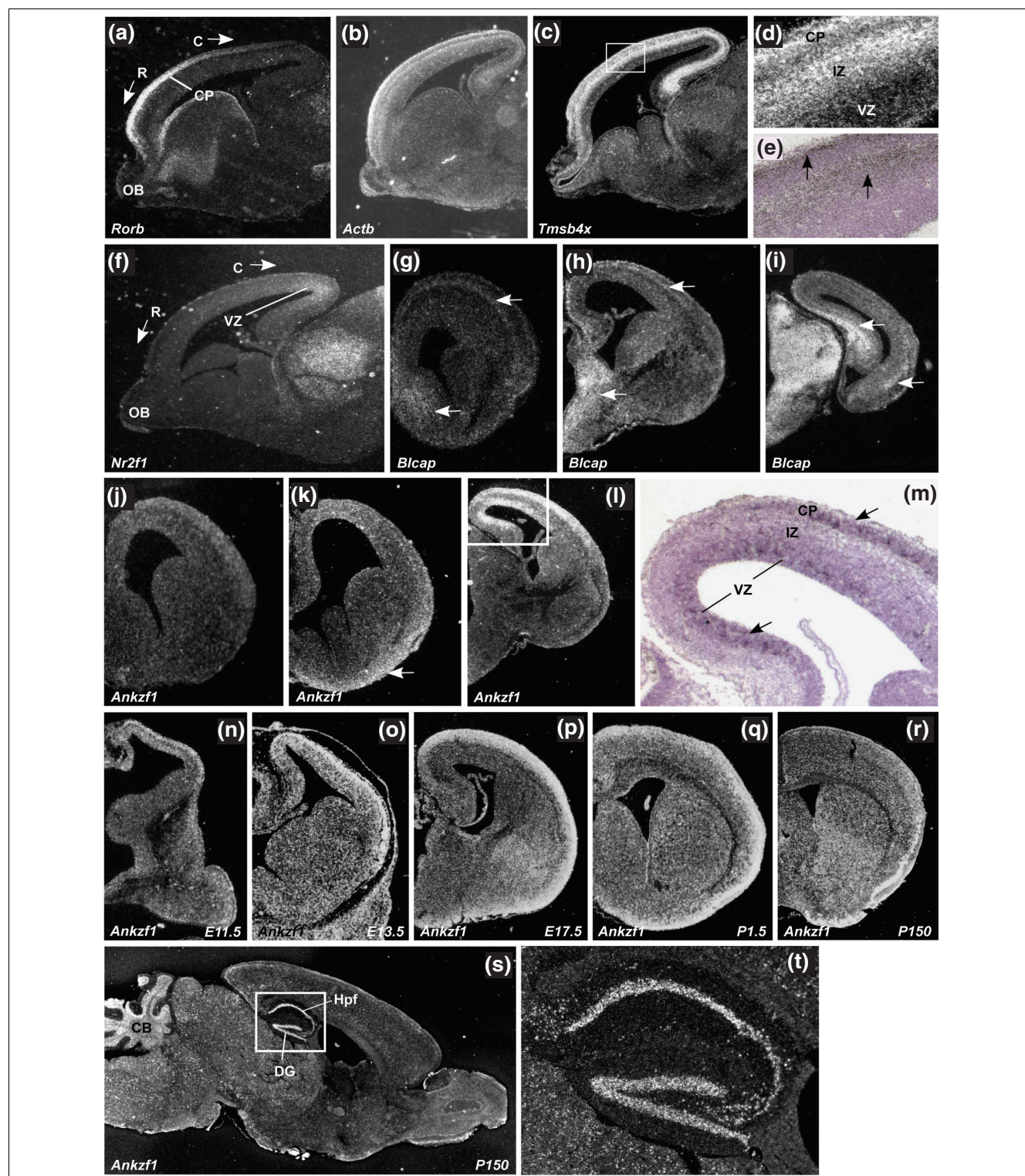
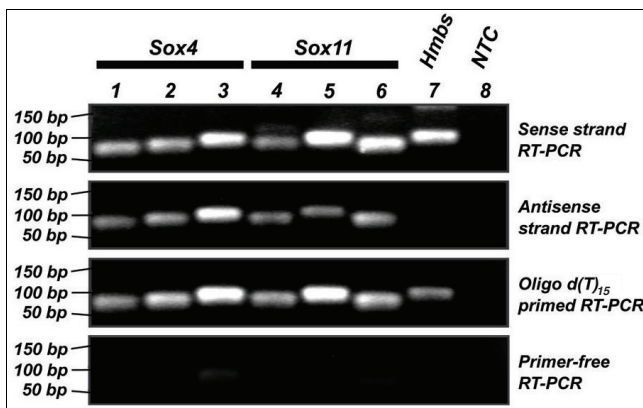


Figure 5

In situ RNA hybridization of selected R versus C regions of E15.5 DETs. ISH was performed on (a-m) E15.5 and (n-t) E11.5 to P150 brains. (g-m) Coronal sections that are generated from the rostral to caudal axis; (a-f, s-t) sagittal sections. Micrographs of higher magnification are presented directly after any micrographs with an inset box (d-e, m, t). All micrographs are in dark-field except for (e, m), which are bright-field micrographs. C: caudal; CB: cerebellum; CP: cortical plate; DG: dentate gyrus; HPf: hippocampal formation; IZ: intermediate zone; OB: olfactory bulb; R: rostral; VZ: ventricular zone. Arrows show the region with expression or silver grains.

**Figure 6**

Strand-specific RT-PCR. Lanes 1 to 6 represent the amplification of the *Sox4* and *Sox11* transcripts using more than one probe. Lane 7 shows the amplification of the *Hmbs* housekeeping gene whereas lane 8 represents the amplification of water as non-template control (NTC) using the primer set for *Hmbs*. The first panel consists of amplicons generated from the reverse-transcribed sense strand cDNA whereas the second panel consists of amplicons generated from the reverse-transcribed antisense strand cDNA. The third panel represents amplicons generated from oligo-d [T]₁₅ primed reverse-transcribed cDNA, which serves as a positive control. The last panel represents amplicons generated from primer-free reverse-transcription reactions. The numbers on the left indicate the size of the generated bands.

mice from each of E15.5, E17.5, P1.5 and adult stages). Three primer sets were used for each locus (designed at the middle, 5'- and 3'-ends of the canonical transcripts). For both *Sox4* and *Sox11* loci, the analysis showed positive signals from all primer sets used that were complementary to the antisense strand, therefore confirming the presence of one or more antisense transcript(s) that span the canonical transcript. Hydroxymethylbilane synthase gene (*Hmbs*) served as a negative control and there were no positive bands in the antisense strand RT-PCR, confirming the absence of antisense transcripts.

Genomic cluster at the *Sox4* gene locus

We mapped SAGE tags from the genomic cluster to the *Sox4* gene locus using the University of California Santa Cruz (UCSC) genome browser (Figure 7a; Additional data file 7) [63]. Only tags within and around the *Sox4* canonical transcript are shown. Evidence of mapped mouse mRNAs within this locus further justifies the existence of multiple SAGE tags in addition to the canonical *Sox4* transcript. Subsequent validation of the genomic clusters was solely based on the SAGE tags situated within the canonical transcript. Based on the SAGE tag information, 6 out of 12 tags are DETs. These include four DETs within the canonical transcript (Figure 7b), with *sox4_tag10* and *sox4_tag15* having greater expression in P1.5 compared to P150 (adult stage), whereas *sox4_tag12* and *sox4_tag16* are both abundantly expressed in the caudal

region of the E15.5 cerebral cortex. Based on the RT-qPCR analysis (Figure 7c), *sox4_tag10* has greater expression in E15.5 compared to E17.5 (1.76-fold change), P1.5 (3.72-fold change) and Ad (43.67-fold change). For *sox4_tag12*, *sox4_tag15* and *sox4_tag16*, differences are seen only in P1.5 (-2.00-fold, -1.57-fold and -1.84-fold changes, respectively) and Ad (-82.08-fold, -68.92-fold and -69.41-fold changes, respectively) when compared to E15.5. RT-qPCR analysis on the same tags did not find any differences between rostral and caudal cerebral cortices of E15.5. The differences in fold change between DETs suggest irregular overlapping of various transcript variants at different *Sox4* gene loci.

To further validate the expression profiles of the multiple *Sox4* DETs, we performed 3' rapid amplification of cDNA ends (RACE)-Southern analysis using pooled adaptor-oligo-d [T]₁₅ synthesized cDNAs from three mice at each developmental stage. Based on this method, we were able to semi-quantitatively and accurately measure the expression levels of individual SAGE tags at the locus. To show that the amplification was cDNA specific, we performed PCR by using the same primer sets on mouse genomic DNA under the same conditions. In all cases, no amplification was observed (data not shown). This analysis confirmed the presence of four out of seven alternative transcripts for *Sox4* (Figure 7d). Corresponding tags were determined by estimating the amplicon sizes between the strand-specific primers used, the next downstream AAUAAA/AUAAA polyadenylation signal (if any) and succeeding CATG sequence or SAGE tags. Figure 7d(1)-d(3) confirms the existence of *sox4_tag10*, *sox4_tag12* and *sox4_tag15*. Of these tags, SAGE expression profiles of *sox4_tag10* and *sox4_tag15* were validated (embryonic-specific and reduced expression after P1.5) but not *sox4_tag12* (E15.5 caudal region-specific). 3' RACE-Southern analysis using a sense probe detected bands in the rostral and caudal regions of E15.5, E15.5 and E17.5 cerebral cortices and, therefore, confirmed the existence of the *Sox4* antisense transcripts (Figure 7d(4)). Even though none of these tags were differentially expressed in between these regions based on the SAGE analysis, our findings show distinctive regionalization for *sox4_tag14* expression at the E15.5 rostral cerebral cortex. Proteasome (prosome, macropain) subunit, beta type 2 gene (*Psm2*) and *Hmbs* were used as controls and no antisense or alternative transcripts were identified at these gene loci (Figure 7d(5)-d(8)).

Since 3' RACE-Southern analysis was dependent on oligo-[dT]₁₅ priming, we could not rule out the possibility of amplicons that were generated by false priming on homopolymer-A stretches. Therefore, Northern analyses were performed on equally pooled total RNA extracted from the cerebral cortices of seven mice at E15.5, E17.5 and P150 (negative control). By using a double-stranded DNA probe at the 3' untranslated region (UTR) of *Sox4* (Additional data file 8), we identified six bands ranging from approximately 2 kb to approximately 4.7 kb (Figure 7e). *Sox4* sense transcripts are weakly expressed in

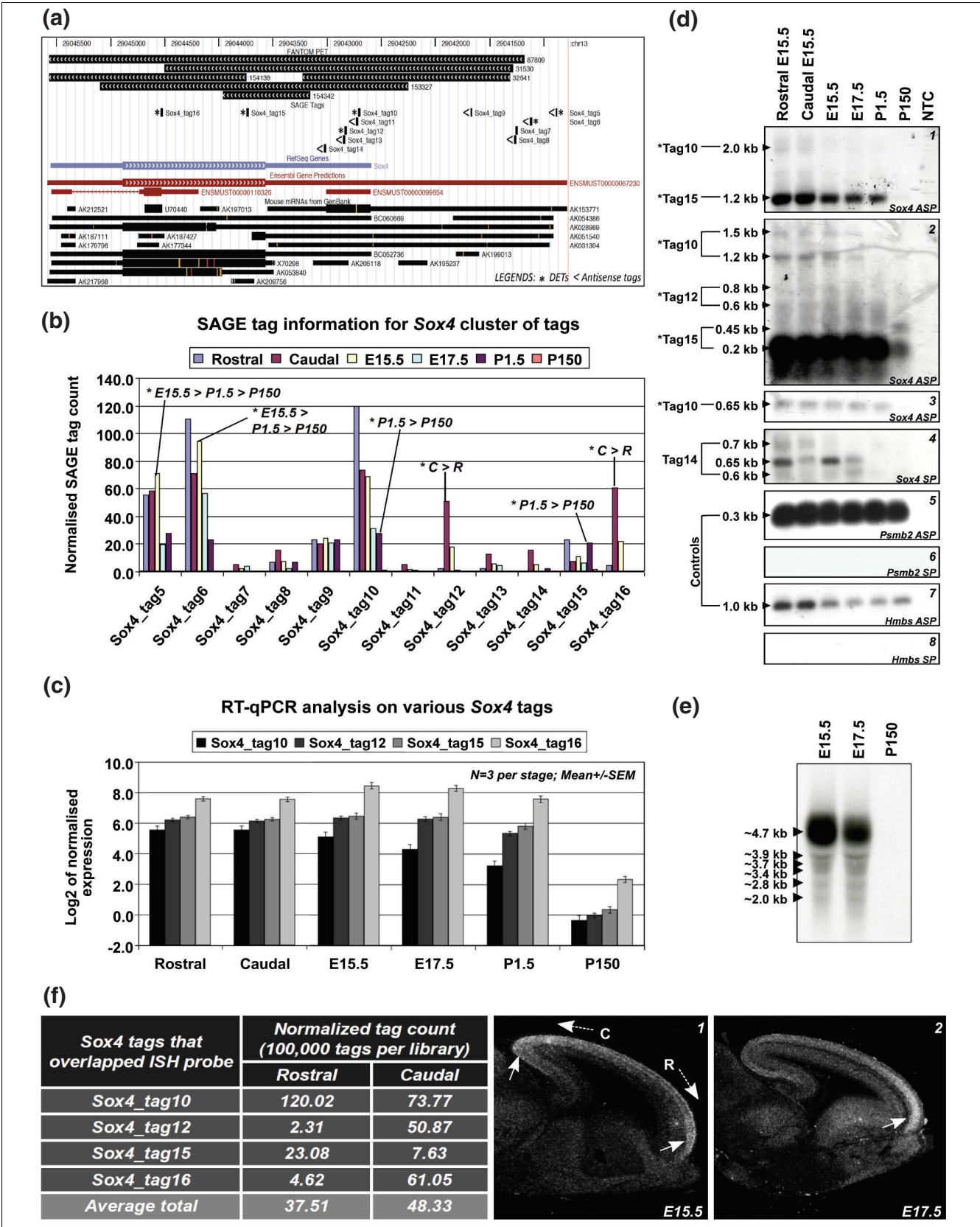


Figure 7 (see legend on next page)

Figure 7 (see previous page)

Genomic cluster at the *Sox4* locus. (a) The UCSC genome browser of the over-represented *Sox4* genomic locus. SAGE tags were found in both directions within the *Sox4* reference gene. (b) The SAGE expression profiles for each tag in the *Sox4* locus. (c) The RT-qPCR validations of selected DETs at various loci within the *Sox4* canonical gene. (d) The 3' RACE-Southern blotting analysis. The left panel of the figure shows the sizes of the bands and their corresponding tags. (d1-d3) Amplification of *Sox4* sense transcripts (ASP); (d4) amplification of *Sox4* antisense transcripts (SP). Tags with asterisks are DETs. Both (d5) and (d7) are positive controls exclusively generated from the sense strand of *Psmb2* and *Hmbs* housekeeping genes (ASP), respectively. The corresponding (d6) and (d8) are the antisense expression (negative control) of *Psmb2* and *Hmbs* (SP), respectively. (e) Northern analysis of total RNA isolated from pooled mouse cerebral cortices (N = 7). (f) The regionalized expression of *Sox4* sense transcripts determined by ISH. The table consists of SAGE information for related *Sox4* sense tags at rostral and caudal regions of the E15.5 cerebral cortex. (f1) and (f2) are sagittal sections obtained from E15.5 and E17.5 mouse brains, respectively. The figures show the expression of the *Sox4* sense transcript variants. Arrows show brain regions with greater *Sox4* sense expression. C: caudal region of the cerebral cortex; R: rostral region of the cerebral cortex.

the adult, but highly expressed in the embryonic cerebral cortex. The number of bands observed is similar and corresponds to overlapping mouse mRNAs as well as publicly available paired-end diTag (PET) sequences downloaded from Ensembl (Figure 7a; Table S6 in Additional data file 1) [64]. Taken together, the analysis confirmed the existence of multiple overlapping variants of *Sox4* sense and antisense transcripts at this gene locus.

To confirm the rostro-caudal expression of *Sox4* sense transcripts, we performed ISH on sagittal sections of mouse brains using a *Sox4* antisense riboprobe that spanned across the *sox4_tag10*, *sox4_tag12*, *sox4_tag15* and *sox4_tag16* SAGE tags. *Sox4* showed regionalized expression at E15.5 and E17.5 (Figure 7f). At E15.5, *Sox4* sense transcripts are expressed more in both the rostral- and caudal-end regions of the cortical plate compared to the intermediate region between them (Figure 7f(1)). By E17.5, expression of *Sox4* sense transcripts is obvious in the rostral cortical plate (Figure 7f(2)). At both stages of development, *Sox4* sense transcripts are uniformly expressed in the intermediate zone of the cerebral cortex. These findings correspond to the SAGE tag counts for E15.5 rostro-caudal regions of the cerebral cortex (Figure 7f). These observations explain the averaged total tag count per 100,000 tags for different *Sox4* sense transcripts, which are predominantly expressed in both rostral and caudal regions of the cerebral cortex. The regionalized expression of *Sox4* in the cortical plate is obvious only at E15.5 and E17.5, but not at other stages of development (Figure S11 in Additional data file 1).

In situ RNA hybridization of *Sox4* sense and antisense transcripts

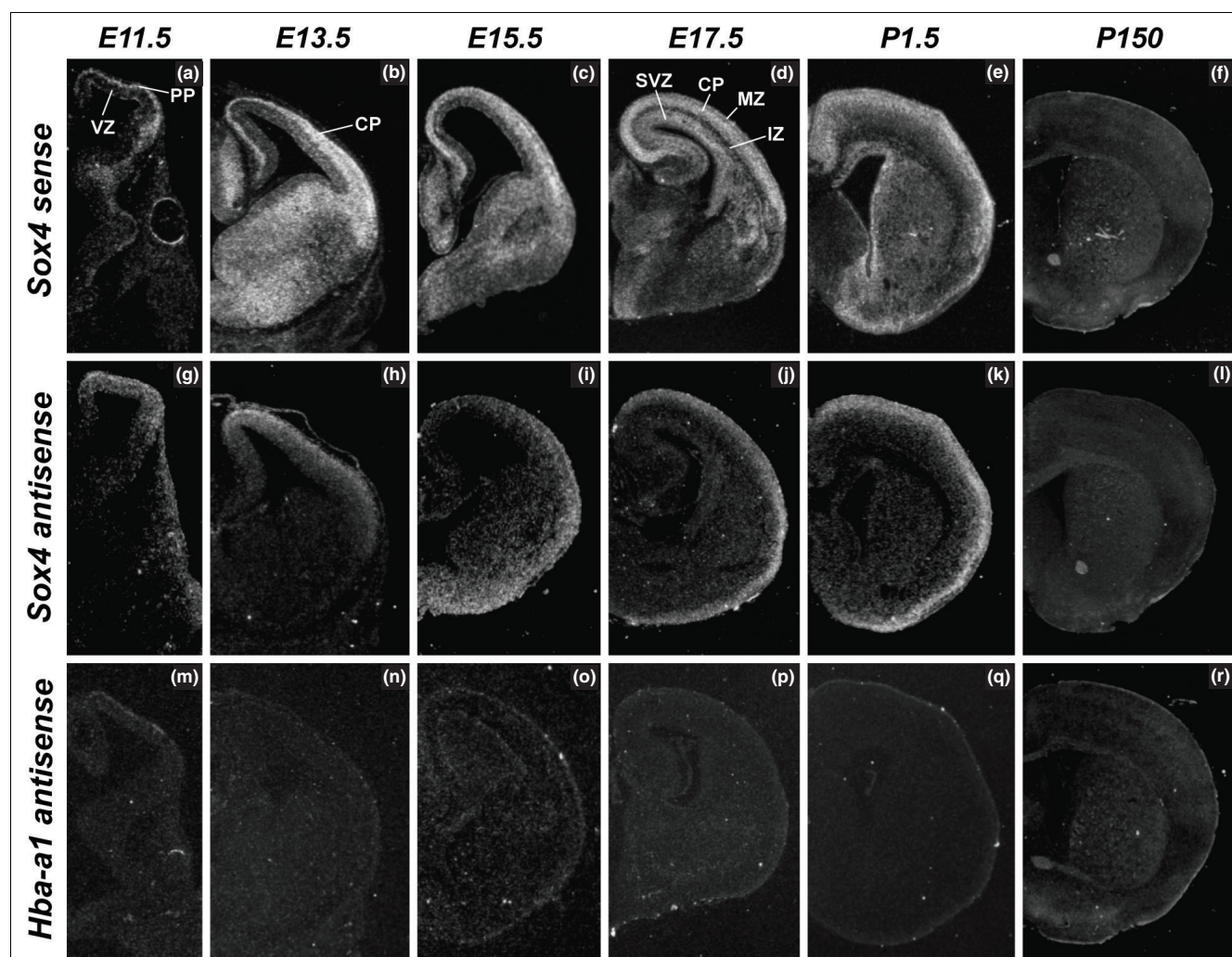
To further ascertain the antisense expression of *Sox4* in a spatio-temporal manner, we performed ISH on coronal sections obtained from E11.5, E13.5, E15.5, E17.5, P1.5 and P15.0 mouse brains. Sense and antisense RNA probes were generated from the same clone used in the Northern analysis. At E11.5, *Sox4* sense transcripts are confined to the primordial plexiform layer (Figure 8a). From E13.5 to P1.5, the sense transcripts are expressed throughout the cortical plate (Figure 8b-e). Expression of sense transcripts in the subventricular zones is observed at E17.5 and P1.5 only (Figure 8d, e). There is no observable sense expression in the adult stage

(Figure 8f). *Sox4* antisense transcripts are expressed throughout the telencephalon at E11.5 (Figure 8g). From E13.5 to P1.5, *Sox4* antisense expression is confined to the cortical plate only (Figure 8h-k). There is no obvious antisense expression in the cerebral cortex in the adult stage (Figure 8l). A microscopic examination at high magnification showed that *Sox4* antisense transcripts are predominantly localized in the nucleus whereas *Sox4* sense transcripts are found in both the nucleus and cytoplasm (Figures S12, S13 and S14 in Additional data file 1). We used hemoglobin alpha, adult chain 1 (*Hba-a1*) of the corresponding brain region and time-point as a control in the analysis (Figure 8m-r).

Furthermore, *Sox4* antisense expression occurs in the piriform cortex layer II (Figure 9a-c) and dentate gyrus (Figure 9g-i) in the adult brain; however, no sense expression is observed in these regions. At P1.5, we identified complementary expression between *Sox4* sense and antisense transcripts in the olfactory bulb (Figure 9d-f). *Sox4* sense expression was confined to the granular and glomerular layers of the olfactory bulb whereas antisense expression was found only in the outer plexiform layer. We used either *Sox11* or *Hba-a1* of the corresponding brain region and time-point as a control in the analysis (Figure 9c, f, i).

Analysis of the *Sox11* genomic cluster

SAGE tags, which represent multiple overlapping sense and antisense transcript variants at the *Sox11* genomic cluster, were validated using 3' RACE-Southern analysis as described above. See Section F in Additional data file 1 for a full description of the *Sox11* results. ISH analysis did not confirm the expression of antisense transcripts of *Sox11*, but the presence of PETs spanning three out of five antisense tags confirmed the existence of *Sox11* antisense transcripts (Table S8 and Figures S16, S17 and S18 in Additional data file 1). The discrepancy between ISH and RT-qPCR or 3' RACE-Southern analysis suggests that *Sox11* antisense transcripts might be expressed at low levels or at specific locations of the cerebral cortex, and hence can be detectable only by using serial sections or whole mount ISH.

**Figure 8**

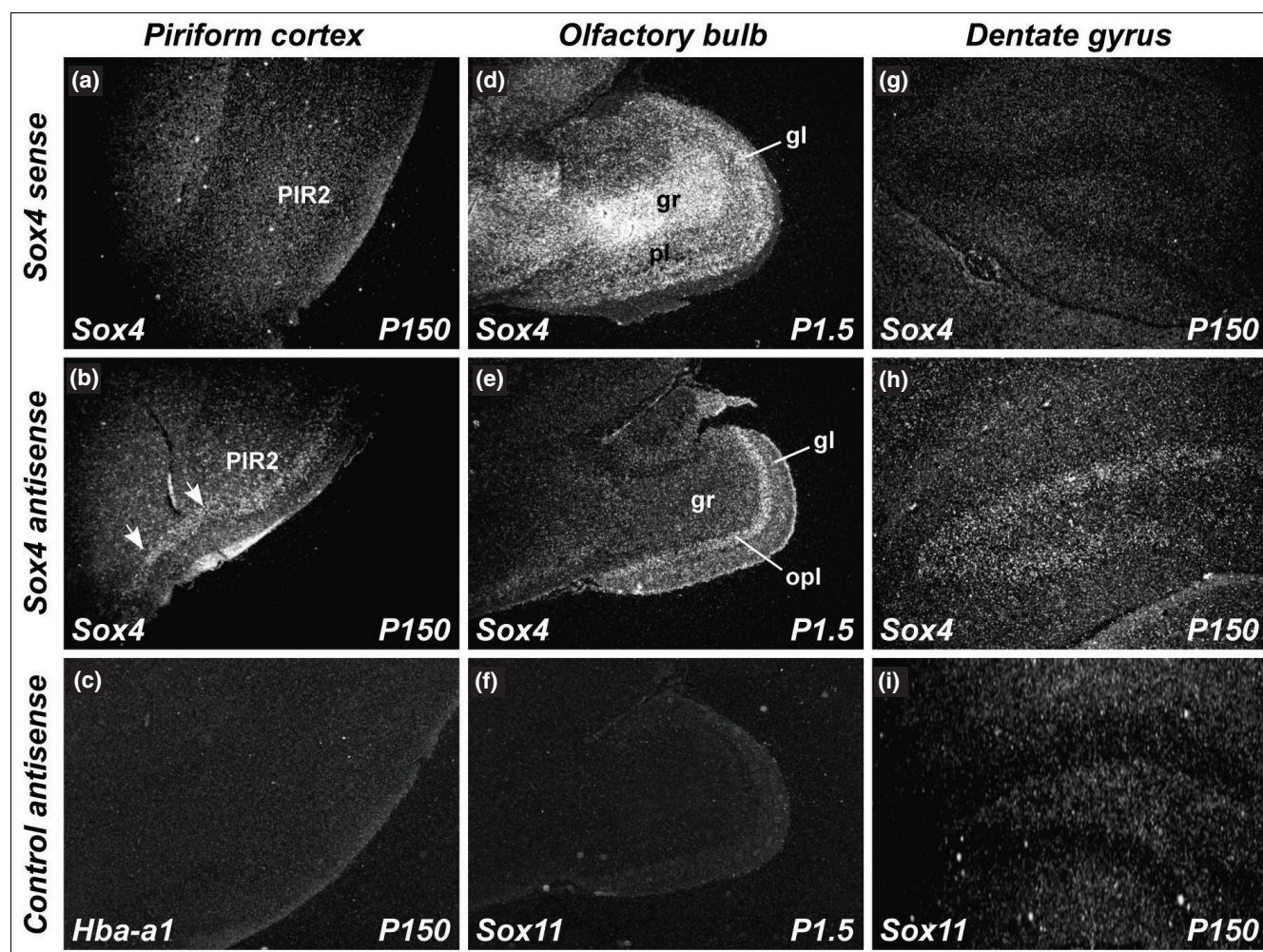
ISH analysis of *Sox4* transcripts in E11.5 to P150 mouse brains. (a-f) Expression of the sense transcript for *Sox4*. (g-l) The expression of the antisense transcript for *Sox4*. (m-r) *Hba-a1* antisense expression (negative control). All micrographs were taken from coronal sections. CP: cortical plate; IZ: intermediate zone; MZ: marginal zone; PP: primordial plexiform layer; SVZ: subventricular zone; VZ: ventricular zone.

Screening of *Sox4* and *Sox11* antisense transcripts in the adult mouse brain, organs, P19 cell line and neurospheres

We screened various adult brain regions (olfactory bulb, cerebellum, medulla, hippocampus, thalamus and cerebral cortex) and selected mouse organs (E15.5 whole brain, heart, kidney, liver, skeletal muscle, skin, spleen, stomach, testis and thymus) for the expression of *Sox4* and *Sox11* antisense transcripts by strand-specific RT-qPCR. Within the adult brain, *Sox4* sense and antisense transcripts are expressed in all regions, with the highest level found in the olfactory bulb, which is approximately four- to nine-fold greater than those in other brain regions (Figure 10a). Expression of *Sox4* antisense transcripts occurs in all mouse organs, with the highest level in the thymus followed by E15.5 whole brain, testis and skin (Figure 10b). *Sox4* sense and antisense expression profiles are similar throughout the entire series of samples

screened, with the sense transcripts being consistently expressed at a greater level than the antisense transcripts (approximately 1.7-fold in various brain regions and approximately 2- to 14-fold in various organ comparisons).

Sox11 sense transcripts are expressed at the highest level in the olfactory bulb, approximately two- to seven-fold greater than those in other brain regions (Figure 10a). *Sox11* antisense transcripts, on the other hand, are expressed in all brain regions screened and at a comparable level in the olfactory bulb, hippocampus, thalamus and cerebral cortex. In comparison to other adult mouse organs, *Sox11* sense and antisense transcripts are highly expressed in the E15.5 whole brain, with *Sox11* sense transcript levels at least 100-fold greater than those in other mouse organs (Figure 10b). On the other hand, *Sox11* antisense expression is observed only in the E15.5 whole brain, skin and stomach. Notably, *Sox11* sense

**Figure 9**

Sox4 antisense expression in other brain regions. (a-i) Sox4 antisense expression is also observed in other regions such as the piriform cortex (a-c; arrows), olfactory bulb (d-f) and dentate gyrus (g-i). (a, d, g) Sox4 sense expression; (b, e, h) Sox4 antisense expression. (c, f, i) *Hba-a1* or *Sox11* antisense expression (negative controls). All micrographs were taken from sagittal sections except (a-c), which were taken from coronal sections. gl: glomerular layer; gr: granule layer; Opl: outer plexiform layer; PIR2: piriform cortex layer II.

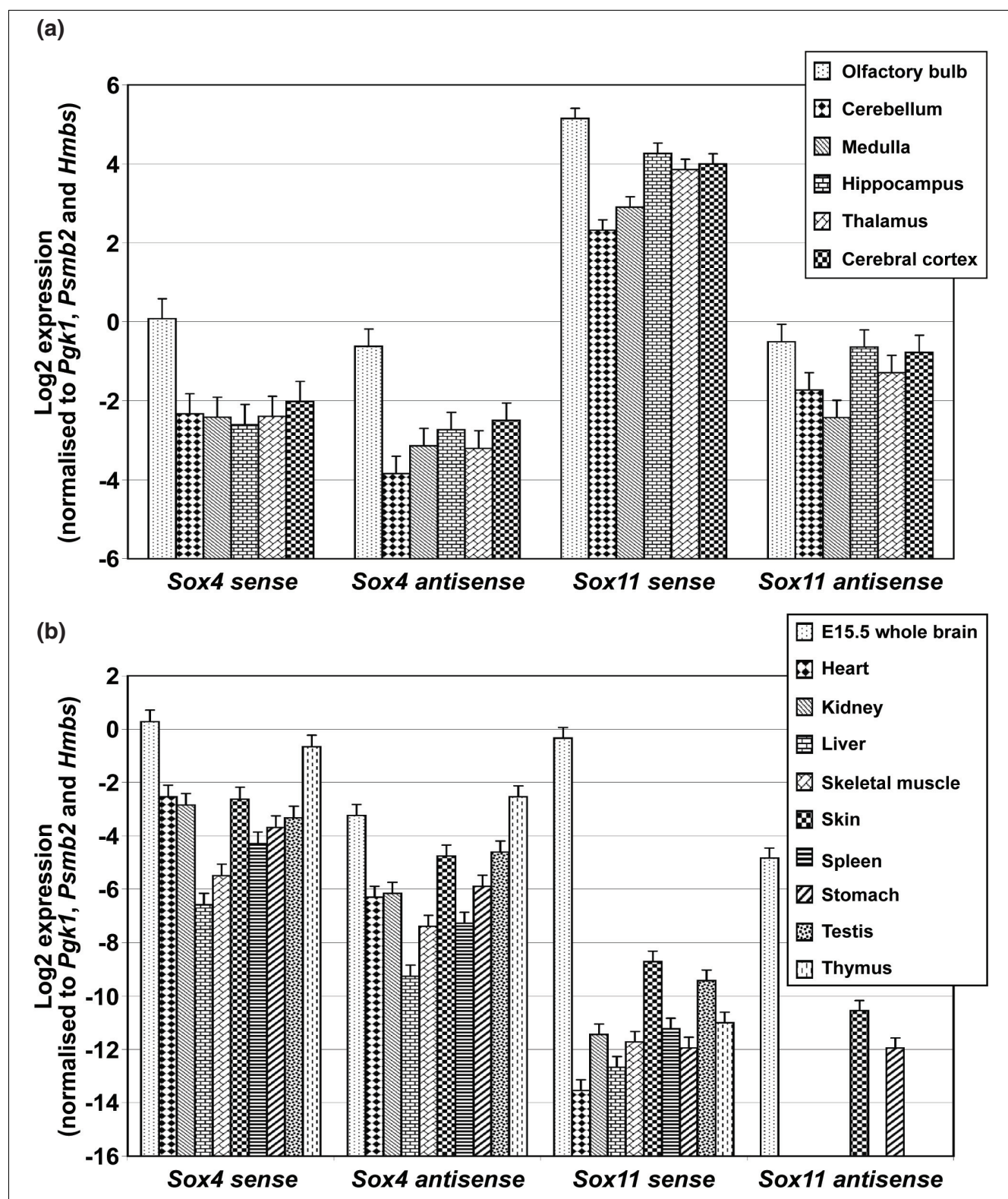
transcripts are expressed more highly than antisense transcripts in the E15.5 whole brain and skin (23- and 4-fold, respectively).

Since both *Sox4* and *Sox11* are implicated in neuronal differentiation and glial maturation processes [65,66], we examined both *Sox4* and *Sox11* sense and antisense transcript expression in proliferating and differentiating P19 (embryonal carcinoma cells) and in embryonic NSPCs grown as neurospheres. Both *Sox4* sense and antisense transcripts are upregulated during P19 cell differentiation (approximately 5.7- and 1.6-fold upregulation, respectively; Figure 11a) and neurosphere differentiation (approximately 1.9- and 1.8-fold upregulation, respectively; Figure 11b). For *Sox11*, both sense and antisense transcripts are upregulated in the differentiating compared to the proliferating P19 cells by approximately 2.3- and 4.2-fold, respectively (Figure 11c). Both the *Sox11*

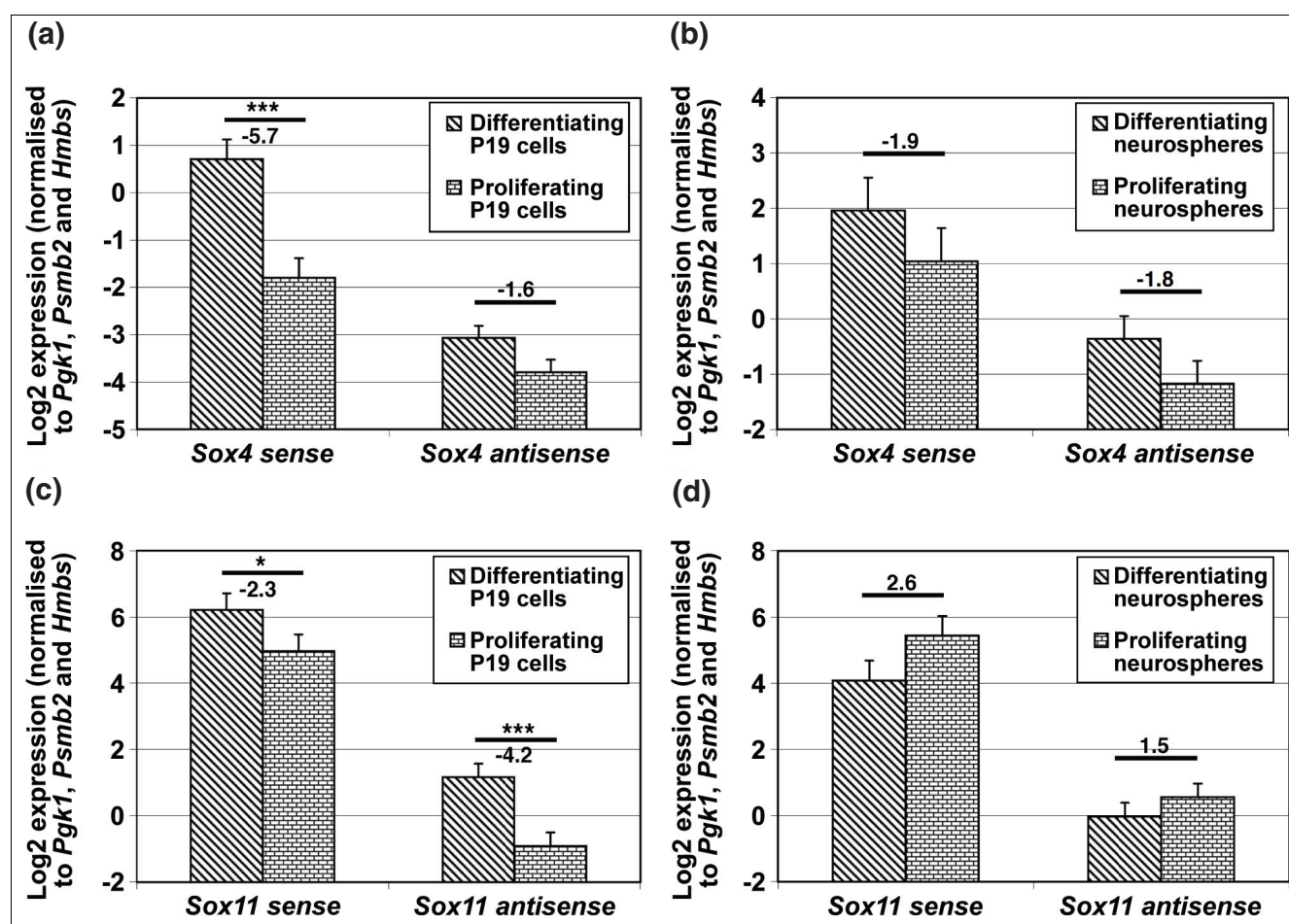
sense and antisense transcripts are, however, downregulated in the differentiating neurospheres (approximately 2.6- and 1.5-fold, respectively; Figure 11d).

Discussion

In this study, SAGE was used to analyze global gene expression in the normal mouse cerebral cortex at various developmental stages. We report validated spatio-temporal regulation of genes involved in mouse cerebral cortex development from embryo to adulthood. The study highlights four main findings: association of DETs from different gene clusters with known functional processes or signaling pathways and disease-causative genes that are involved in cerebral corticogenesis; *Ankzf1* and *Sox4* sense transcripts are regionally expressed in the E15.5 cerebral cortex; multiple overlapping *Sox4* and *Sox11* sense and antisense transcripts are spatio-

**Figure 10**

Expression of *Sox4* and *Sox11* transcripts in various mouse organs. (a) Strand-specific RT-qPCR screening of *Sox4* and *Sox11* sense and antisense transcript expression in various adult mouse brain regions. N = 2 and data are presented as mean \pm standard error of the mean (SEM). (b) Strand-specific RT-qPCR screening of *Sox4* and *Sox11* sense and antisense transcripts in various adult mouse organs. N = 3 and data are presented as mean \pm SEM.

**Figure 11**

Expression of *Sox4* and *Sox11* transcripts in neurospheres and P19 cells. (a-d) The figure shows strand-specific RT-qPCR screening of *Sox4* (a, b) and *Sox11* (c, d) sense and antisense transcripts expression in proliferating and differentiating P19 cells (a, c) and neurospheres (b, d). N = 3 for P19 cells and N = 2 for neurospheres. All data are presented as mean \pm standard error of the mean. Fold change values (normalization to proliferating cells) are presented above the comparative bar and any values <1 are presented in the negative fold change format. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

temporally regulated during cerebral corticogenesis; and *Sox4* and *Sox11* antisense transcripts are differentially regulated in both proliferating and differentiating embryonic-derived neurospheres and P19 cells.

We have shown that most tags generated in all libraries were singletons. The number of singletons could be reduced by increasing the number of tags sequenced. In mammalian cells, the number of additional unique transcripts identified approached zero when the number of SAGE tags sequenced reached approximately 600,000 [67]. Increasing the number of tags sequenced could improve the sensitivity of the technique to identify weakly expressed or novel transcripts, and the application of massively parallel signature sequencing [68] using a next-generation sequencer would be an ideal solution to accomplish this. In addition, one of the benefits of SAGE is that datasets generated from different groups or in public repositories such as SAGE Genie [69] and GEO [28] are readily comparable and, hence, can increase the tag count

and sensitivity of the technique in discovering DETs between SAGE libraries. However, any meta-analyses involving various SAGE datasets are affected by experimental and biological variation; thus, a careful selection of matching libraries is crucial to limit systematic error or biases.

Our SAGE analysis robustly detected DETs with a low false positive rate (for example, $<0.001\%$ for comparison between left and right hemispheres of the adult cerebral cortex). Of all the identified DETs, approximately 8% were not mapped to either a single locus in the mouse genome or any unique annotation. This problem could be overcome by generating additional information from the 5' end of the transcript through alternative techniques such as PET sequencing [70], cap analysis gene expression (CAGE) [71] and 5' LongSAGE [72].

We have identified functional ontologies, molecular interactions and enriched canonical pathways that are distinct to the

stage-specific gene clusters of validated DETs. The IPA network analysis generated connections between validated DETs across various developmental stages in relation to well-established proteins or molecules and neurological disorders. In the study, members of the Wnt/ β -catenin signaling pathway were enriched in networks 1 and 2 (embryonic-specific gene clusters). In neural development, Wnt/ β -catenin signaling plays an important role in regulating regional specification of the cortex along the rostro-caudal and dorso-ventral axes, and proliferation of cortical progenitors [73]. IPA highlighted three genes (*Sox4*, *Sox11* and *Sfrp1*) associated with this pathway. In humans, the SFRP1 protein (secreted frizzled-related protein 1) stabilizes β -catenin and increases transcription from β -catenin-responsive promoters [74]. In β -catenin-deficient mouse mutants, expression of both *Sox4* and *Sox11* is downregulated [75,76]. *Sox4* and *Sox11* proteins play an important role in establishing neuronal properties, pan-neuronal gene expression and proper myelination of the central nervous system [65,66]. This suggests that the role of the Wnt/ β -catenin signaling pathway in regulating neuronal development could be mediated, at least in part, by the *Sox4* and *Sox11* proteins.

The role of Wnt/ β -catenin signaling in regulating DETs (*Btg1*, *Cdk4*, *Cdkn1c*, *Csrp2*, *Ezh2*, *Neurod1*, *Pcna*, and *Rps4x*) involved in cell cycle and proliferation remains unclear. Gain- and loss-of-function studies have established that Wnt/ β -catenin signaling is essential to maintain the pool of precursors for proper development of the cerebral cortex [77,78]. To date, there is no direct evidence to show that *Ezh2*, *Pcna*, *Rps4x* and *Btg1* are involved in cell cycle regulation during early embryonic neurogenesis. But their expression in the ventricular/subventricular zone of the E15.5 developing mouse cerebral cortex [33,35,36] suggests that they may be involved in regulating cell proliferation during neurogenesis. The role of these DETs and their association with the Wnt/ β -catenin signaling pathway remains unclear and requires further experimentation.

Networks 3 and 4 from the adult gene cluster were associated with various canonical pathways, in particular, synaptic long term potentiation (LTP) and calcium signaling. LTP is a process of synapse enhancement, which is thought to underlie some forms of learning and memory [79]. This process depends on Ca^{2+} and calmodulin, which are major components of the calcium signaling pathway. We identified four DETs, *Gria3*, *Itp1r*, *Ppp3ca* and *Camk2a*, in both canonical pathways. In particular, the *Camk2a* enzyme is highly expressed in the brain and regulates mainly glutamatergic synapses during LTP [79]. DET products such as *Nrgn* and *Camk2n1* can directly or indirectly regulate *Camk2a* or Ca^{2+} /calmodulin and subsequently alter the outcome of the LTP pathway [80,81]. Therefore, they may serve as important candidate genes in the future analysis of the synaptic LTP pathway involving neurodegenerative diseases that cause the loss of cognitive function and memory.

Networks 5 and 6 are enriched for genes in the amyloid processing signaling pathway. *App* and *Mapt* are associated with this pathway. Under normal circumstances, the *App* protein is required for proper migration of neuronal precursors into the cortical plate in early embryonic corticogenesis [82]. The *Mapt* protein, on the other hand, plays an important role in maintaining the architecture of the neuronal cytoskeleton and intracellular trafficking. Overexpression of *App* protein and hyperphosphorylation of the *Mapt* protein have been implicated in the pathologies of Alzheimer's disease [83,84]. Interestingly, *Ctsd*, *Atp7a*, *Cln2* and *Hprt1*, the genes responsible for other human neurological disorders such as neuronal ceroid lipofuscinosis (*Ctsd*), Menkes disease (*Atp7a*), epilepsy (*Cln2*), and Lesch-Nyhan syndrome (*Hprt1*), are associated with *App* and *Mapt*. These candidate genes are also involved in cell morphogenesis, assembly and organization and could be linked to deterioration of neurons during the pathologic progression of these disorders.

Pathway analysis of DETs classified into the N, G and B groups showed DETs in the neuron (N and B groups) are associated with Huntington's disease and schizophrenia, which were not previously identified in networks 3 to 6. Our analysis showed that both disorders share three common DETs, namely *Rgs4*, *Ppp1r1b* and *Chgb*, whose expression is downregulated in humans with Huntington's disease or schizophrenia [85-88]. A proportion of patients with Huntington's disease also develop schizophrenia [89,90]. Taken together, downregulation of *Rgs4*, *Ppp1r1b* and *Chgb* expression in neurons may contribute to the common symptoms in these disorders. Our findings imply that many DETs (including *App*, *Hprt1* and *Snca*) associated with both Huntington's disease and schizophrenia are also involved in neuronal/cell death processes [91-93]. Other DETs in the N group, not previously implicated in neuronal cell death, may serve as novel potential candidates during pathologic development in these disorders.

Regionalized development of the cerebral cortex involves the differential regulation of cell cycle exit, early migration and attainment of positional identity in neuronal fated cells. To date, only few genes have been associated with regionalized development of the cerebral cortex [3,94]. In the regionalization analysis, we identified the highest number of DETs in the comparison of the RL and CM libraries, which signifies that these two regions of the cerebral cortex are the most different. This finding supports the notion that the cerebral cortex is developed in a latero-medial axis followed by a rostro-caudal axis [7,8]. At E15.5, both *Actb* and *Tmsb4x* were expressed greater in the rostral cerebral cortex than in the caudal region. Both *Actb* and *Tmsb4x* proteins are involved in the actin cytoskeleton-signaling pathway [95]. In particular, the *Tmsb4x* protein has been shown to promote cardiomyocyte migration [96] and axonal tract growth in zebrafish [97]. Therefore, co-expression of *Actb* and *Tmsb4x* in the E15.5 mouse cortical plate suggests that they may have a synergistic

role in early cortical cell development. Conversely, *Blcap* and *Ankzf1* were expressed more highly in the caudal than in the rostral region of the E15.5 cerebral cortex. To date, the function of both *Blcap* and *Ankzf1* in the cerebral cortex remains uncharacterized. This study provides the first comprehensive expression profile of *Ankzf1* and suggests it could be an important transcription factor in cerebral corticogenesis.

At E15.5, *Sox4* sense transcripts were expressed in a high-rostral and high-caudal manner with lesser expression within the intermediate region. By E17.5, *Sox4* expression becomes obvious at the rostral cortical plate, which is similar to *Rorb*. But, we did not find that the regionalized expression of *Sox4* sense transcripts resembles that of the restricted *Rorb* expression at E15.5 or in the postnatal brain [1,62]. This finding could be caused by the combined expression profiles of different *Sox4* sense transcripts that are present across the rostro-caudal axis of the cortical plate. The regionalized expression of *Sox4* sense transcripts occurs only between E15.5 and E17.5. Because the thalamic axon innervates the cortical plate after E17.5 [98], the regionalization of *Sox4* sense transcripts in early cortical development could be an outcome of an intrinsic instead of an extrinsic mechanism that regulates early patterning of the cerebral cortex.

Genomic clustering of DETs identified the differentially regulated *Sox4* and *Sox11* gene loci. These genomic clusters imply that there are multiple overlapping sense and antisense transcripts surrounding the same gene locus that are co-transcribed simultaneously during cerebral cortex development. Both *Sox4* and *Sox11* are single exon genes and these transcript variants are therefore likely to be generated due to alternative polyadenylation. The 3' UTRs of both *Sox4* and *Sox11* have tandem terminal polyadenylation signals on both sense and antisense strands (data not shown), which supports the occurrence of multiple transcript forms or SAGE tags. Multiple mRNA forms with different 3' UTRs can lead to cell-specific regulation, different nuclear or cytoplasmic mRNA stability and translation rates [99,100]. The 3' UTR of *Sox4* and *Sox11* may contain AU-rich elements that play an important role in determining mRNA stability through deadenylation, decapping or 3' → 5' decay [101]. Besides, different 3' UTR lengths may be targeted by different miRNAs, thus interfering with the translation process. Both *Sox4* and *Sox11* transcripts may be targeted by various miRNAs at different predicted positions across the 3' UTR (Tables S7 and S9 and Figures S15 and S18 in Additional data file 1). Therefore, 3' UTR lengths of *Sox4* and *Sox11* may be an important feature in the regulation of their protein expression during cerebral corticogenesis.

In the study, NATs were found at both the *Sox4* and *Sox11* gene loci overlapping the sense transcripts. Overlapping NATs may function as templates for the generation of pre-miRNA and mature miRNA with exceptional high sequence conservation that complement the overlapping sense protein-

coding transcripts [102]. To date, no mature or pre-miRNAs have been predicted on the *Sox4* and *Sox11* sense and antisense strands (data not shown) or have been reported in miR-Base [103]. In addition, NATs can self-complement to form double stranded RNA or pair with sense transcripts and function as templates for the generation of endogenous small interfering RNAs, which could subsequently interfere with translation or transcription of multiple protein-coding transcripts [104]. Because both *Sox4* and *Sox11* proteins are highly expressed in the cerebral cortex, the overlapping NATs do not seem to be involved in the regulation of *Sox4* and *Sox11* through miRNA- or small interfering RNA-mediated translation repression mechanisms, but rather through antisense-regulated sense transcription within the nucleus.

Our ISH analysis showed complementary cellular expression profiles of *Sox4* sense and antisense transcripts at the piriform cortex, olfactory bulb and dentate gyrus. This finding implies that *Sox4* antisense transcripts may be essential in intracellular and interlocus negative feedback loop regulation of the *Sox4* sense transcripts. Similar expression profiles of *Sox4* sense and antisense transcripts in multiple mouse organs and brain regions, however, suggest that these transcripts may be co-expressed. This observation is also supported by the temporal co-expression of *Sox4* sense and antisense transcripts in the cortical plate or layers I to III of the cerebral cortex. Taken together, the sense and antisense transcripts of *Sox4* are co-expressed in some cells and expressed complementarily in other cells, suggesting crucial cell-type-specific regulation.

Sox4 and *Sox11* have been shown to have redundant roles during mouse development [105], and *Sox11* may play a compensatory role in the absence of *Sox4* during brain development [106]. We demonstrated that *Sox4* and *Sox11* sense and antisense transcripts have a similar expression in the brain, but not in other organs, suggesting a compensatory role for *Sox11* only in the brain. *Sox11* antisense transcripts were expressed in the brain, skin and stomach only, suggesting organ-specific regulation.

Our data show upregulation of *Sox4* and *Sox11* sense transcripts in differentiating P19 cells, consistent with the findings of others [107,108], and demonstrate upregulation of antisense transcripts as well. We also find both *Sox4* and *Sox11* sense transcripts expressed in the NSPCs cultured as neurospheres, which is in agreement with Dy *et al.* [109]. Furthermore, we identify upregulation of both *Sox4* sense and antisense transcripts but downregulation of *Sox11* sense transcripts in differentiating neurospheres. Taken together, our findings show that there are potentially common and distinct roles for *Sox4* and *Sox11* sense and antisense transcripts during neuronal and non-neuronal cell proliferation and differentiation. The underlying regulatory mechanism of these transcripts, particularly the antisense ones, remains unknown and requires further investigation.

Conclusions

This study provides avenues for future research focus in understanding the fundamental processes and development of neurological disorders related to the cerebral cortex. We confirm the regionalized expression of new candidate genes in the E15.5 cerebral cortex as well as differential regulation of multiple overlapping sense and novel antisense transcripts within *Sox4* and *Sox11* gene loci during cerebral corticogenesis. We also report for the first time the spatio-temporal regulation of *Sox4* antisense transcripts in the brain as well as differential regulation of novel *Sox4* and *Sox11* antisense transcripts in various mouse organs and in proliferating and differentiating NSPCs and P19 cells. The finding provides an insight for future investigations into the role of antisense transcripts during cerebral corticogenesis and neuronal differentiation.

Materials and methods

Serial analysis of gene expression

Handling of animals and dissection of the cerebral cortex

All experiments that involved animal breeding and handling were performed according to protocols approved by the Melbourne Health Animal Ethics Committee (Project numbers 2001.045 and 2004.041). All animals involved in the study were C57BL/6 mice unless specified otherwise. All mice were kept under conditions of a 12-h light/12-h dark cycle with unlimited access to food and water. All mice were culled by cervical dislocation prior to dissection. Cortical tissue was procured in the following fashion. For adult samples, after removal of the meninges, coronal cuts were used to excise the olfactory bulb from the rostral region, and the superior colliculus from the caudal region. A sagittal cut to separate the two cortical hemispheres was performed. The cortical pallium was dissected from the subpallial striatum and the septum. The neocortex was then dissected away from the cingulate cortex and the entorhinal cortex. For embryonic samples, the cortical tissue was dissected free from the underlying ganglionic eminences at the pallial-subpallial border. An orthogonal cut was made to remove the presumptive striatum and the overlying piriform cortex. On the medial aspect, the medial limbic cortex was included for analysis, but the adjacent hippocampal primordium, including the cortical hem, was excluded. For the E15.5 cerebral cortex, the resulting hemispheres containing cortical tissue only were placed on the bottom of the Petri dish and, using a fine scalpel, divided into four equal quadrants per hemisphere, namely RL, RM, CL and CM. Rostral and caudal quadrants from both hemispheres were pooled for SAGE library construction but separately tracked for RT-qPCR analysis. Procurement of other adult brain tissues and related mouse organs for *Sox4* and *Sox11* antisense transcript screening was carried out according to the standard mouse necropsy protocol accessible at the National Institute of Allergy and Infectious Diseases (NIAID) website [110].

SAGE libraries and analysis of tags

Ten SAGE libraries were constructed from the cerebral cortex of E15.5, E17.5 and 4- to 6-month-old (Ad) mice according to either one of the two methods described previously [25,26], using I-SAGE™ or I-SAGE™ Long Kits (Invitrogen, Mulgrave, Victoria, Australia). Additional libraries from E15.5 and P1.5 of the cerebral cortex described previously [27] were also included in the analysis. These libraries contain a total of 26,436 traces. SAGE tags were preprocessed - that is, TAGs were extracted and corrected for sequencing errors, and artifacts like SAGE linkers, ribosomal RNA and duplicated ditags were removed using the 'sagenhaft' package, which is available from the Bioconductor website [111,112]. To compare libraries that contain long tags with those that contain short tags, all short tags were mapped to the existing long tags from the other libraries. A table for all libraries containing the unique long or short tags was generated and redundant tags were removed. Only tags with a total count >2 (across all libraries) were considered for subsequent comparisons. Each unique tag was mapped to the mouse genome using ESTgraph, which employs ESTs and their genomic position information. ESTgraph was created by Tim Beissbarth (unpublished) [113]. Identity was assigned to these tags and they were further grouped into the following categories: matching to a gene, a genomic sequence, or an EST, or ambiguous matches or no alignment at all. All annotations were based on the latest mouse assembly (mm9 released in July 2007) accessible from the UCSC Genome Bioinformatics website [63].

Identification of differentially expressed tags

Library comparisons were performed using two methods. Fisher's exact test was used to compare two individual SAGE libraries. In the analysis, multiple testing correction [114] was carried out to control for false-discovery rate and adjusted *P*-value cutoffs (*Q*-values) were used to select DETs. In cases where several libraries were combined to focus on a specific biological comparison (for example, different stages of development), a Bayesian model, as described previously [115], was used to integrate multiple libraries in pairwise comparisons involving biological replicates of libraries. The model accounts for within-class variability by means of mixture distributions. The resulting *E*-values were used to select DETs. A table of all relevant comparisons, the comparison method and *Q*- or *E*-value cutoffs is provided in Table S1 in Additional data file 1.

Hierarchical clustering of SAGE tags

To identify co-regulated genes, the clustering of DETs was performed based on the log₂ of normalized counts. Each library was normalized to 100,000 tags per library to account for size differences. A pseudocount of 0.5 was added before taking the log₂ of the normalized tag counts. The tag-wise mean was subtracted from the log₂ tag intensities before computing the Euclidean distance of the individual tag profiles. Hierarchical clustering was performed on the tags using

the 'hclust' function and complete linkage, which was implemented using the statistical computing environment of R [116].

Genomic clustering of SAGE tags

To assess whether there was any genomic clustering of tags, a method previously described [117] was adopted. In brief, first gene lists (based on all DETs in both pairwise and multiple library comparisons as well as gene lists from the hierarchical clustering analysis) were selected. The genomic clustering of either of these selections was compared to the total unique tag list (all 25,165 unique tags). The tags were mapped to the mouse genome. The number of selected tags in ten consecutive tag positions for each window of the chromosome was calculated. One thousand permutations were used to compute the null distribution of maximum tag counts per window. The method was implemented using the statistical computing environment of R [116].

Functional classification and characterization of DETs

Gene Ontology enrichment analysis

The DET lists generated from various comparisons were subjected to systematic functional annotation using the standardized Gene Ontology term analysis tools at the DAVID [30]. Functional clustering was performed using high stringency with a kappa similarity threshold of 0.85 and a minimum term overlap of 3. Classification was carried out using a multiple linkage threshold of 0.5 with both numbers of initial and final group members set to 3. A term was considered statistically significant when the computed *P*-value was < 0.05 . All queries were performed in September 2009.

Molecular interactions and pathway analysis

Identification of molecular network interactions and pathway analysis of validated DETs or co-regulated genes was completed using the IPA [118] tools from Ingenuity Systems® (Redwood City, California, USA). Accession numbers for all genes with their corresponding fold changes or normalized counts were imported into the IPA software. No focus genes were set at the beginning of the analysis. To start building networks, the application queries the list of input genes and all other gene objects stored in the Ingenuity knowledge base. Networks with a maximum of 30 genes or proteins were constructed, and scores were computed based on the likelihood of the genes being connected together due to random chance. A score of 2 indicates that there is a 1/100 chance that these genes are connected in a network due to random chance. Therefore, any networks with a score of 2 or above are considered statistically significant (with $>99\%$ confidence). The most significant novel networks and their interactions with existing canonical pathways were investigated further.

Quantitative RT-PCR

Total RNA isolation and first strand cDNA synthesis

RT-qPCR was carried out to validate all selected candidate or co-regulated DETs. Biological triplicates from E15.5, E17.5,

P1.5 and adult (5 to 6 months old) cerebral cortices were used. Genomic-free total RNA from independent mice was extracted using the RNeasy® Lipid Tissue Midi Kit (Qiagen, Doncaster, Victoria, Australia) according to the manufacturer's protocol (Section G in Additional data file 1). First strand cDNA was synthesized from high quality total RNA [119] using random hexamers (unless specified otherwise) and the SuperScript™ III Reverse Transcriptase Kit (Invitrogen, Mulgrave, Victoria, Australia), according to the manufacturer's protocol, under an RNase-free environment.

Primer design and RT-qPCR

All primers were designed using ProbeFinder 2.35 at the Assay Design Center for Universal ProbeLibrary Assay provided by Roche Applied Science [120]. All RT-qPCR reactions were prepared in 10 μ l volumes in a 384-well plate format consisting of LC480 Master Probe Mix (Roche Diagnostics, Castle Hill, New South Wales, Australia), Universal ProbeLibrary (Roche Diagnostics, Castle Hill, New South Wales, Australia) and forward and reverse primers (GeneWorks, Hindmarsh, South Australia, Australia or Bioneer Corporation, Daedeok-gu, Daejeon, Korea) according to the manufacturers' protocols. RT-qPCRs were performed using the LightCycler® 480 System (Roche Diagnostics, Castle Hill, New South Wales, Australia) according to the manufacturer's protocol (Section G in Additional data file 1). A full list of primers and probes used in this study is included in Additional data file 8.

Relative quantification using a standard curve method

The crossing point (Cp) from each signal was calculated based on the Second Derivative Maximum method [121]. A set of serially diluted cDNAs was used to construct a four-data point standard curve for every PCR system in each run. A total of three reference genes (from *Hprt1*, *Psmb2*, phosphoglycerate kinase 1 gene (*Pgk1*) or *Hmbs*) were used as endogenous controls. An estimated starting amount of each target gene was calculated and intra-samples multiple reference genes normalization was performed (Section G in Additional data file 1). A linear model was fitted to the time course of expression values for each gene. Genes differentially expressed between the various stages of development or regions were selected using empirical Bayesian moderated t-statistics, which borrow information between genes [122]. Standard errors for the mean expression at various developmental stages were obtained from the linear model. For each comparison, *P*-values were adjusted using the Benjamini and Hochberg [114] method to control the false discovery rate. See Section H in Additional data file 1 for the R code used to execute the analysis.

Validation of sense-antisense and multiple overlapping transcripts in genomic clusters

Strand specific RT-PCR

All RNA was prepared as described above. Total RNA from all developmental stages ($N = 3$ per developmental stage) was

equally pooled prior to cDNA synthesis. Four first strand cDNA synthesis reactions were prepared for each cluster as follows: with a primer complementary to the sense strand only; with a primer complementary to the antisense strand only; with oligo-d [T]₁₅ as a positive control; and without any primers as a negative control. In all four reactions, both primers were added in subsequent PCRs (Section G in Additional data file 1). PCR amplifications were carried out using FastStart PCR High Fidelity System (Roche Diagnostics, Castle Hill, New South Wales, Australia) according to the manufacturer's protocol. More than one primer set was used in the sense-antisense strand specific RT-PCR (Additional data file 8).

RACE

First strand cDNA synthesis was carried out using pooled total RNA extracted from three biological replicates of rostral and caudal E15.5 and whole E15.5, E17.5, P1.5 and adult (5 to 6 months old) cerebral cortices. Oligo-d [T]₁₅ with an adaptor sequence (5'-TACGACGTCTGCTAGGACTG-3') was used to prime the first strand cDNA synthesis. Second strand synthesis or PCR was then carried out using a strand-specific primer and the adaptor primers (Additional data file 8). All specific primers were designed to be complementary to the SAGE tags or their upstream sequences. PCR amplifications (Section G in Additional data file 1) were carried out using FastStart PCR High Fidelity System (Roche Diagnostics, Castle Hill, New South Wales, Australia) according to the manufacturer's protocol.

Southern blotting analysis

Amplified 3' RACE products were transferred to Hybond N⁺™ (GE Healthcare, Rydalmere, New South Wales, Australia) nylon membrane using the neutral transfer method. Prehybridization and hybridization steps were performed in Rapid-Hyb buffer (GE Healthcare, Rydalmere, New South Wales, Australia) according to the manufacturer's protocol. All oligonucleotides were designed to be complementary to sequence between the specific primer-priming site and the tag of interest. Synthetic oligonucleotides were 5' end-labeled using T4 polynucleotide kinase (Promega, Alexandria, New South Wales, Australia) and [γ -³²P]ATP (GE Healthcare, Rydalmere, New South Wales, Australia) with modifications to the manufacturer's protocol. After the hybridization step, the membrane was washed with 5× sodium chloride sodium citrate solution (with 0.1% v/v sodium dodecyl sulphate (SDS)) and 1× sodium chloride sodium citrate solution (with 0.1% w/v SDS) (Section G in Additional data file 1). See Additional data file 8 for primer sequences and oligonucleotides used for detection.

Northern blotting analysis

Independent preparations of total RNA from the cerebral cortex of seven mice at E15.5 and E17.5, and three adult mice were equally pooled to a final concentration of 20 µg per developmental stage. These pooled total RNAs were electro-

phoresed overnight and capillary transferred onto Hybond N⁺™ (GE Healthcare, Rydalmere, New South Wales, Australia) nylon membrane. Double-stranded DNA probes were radioactive-labeled using the Amersham Megaprime DNA Labeling System (GE Healthcare, Rydalmere, New South Wales, Australia) and [β -³²P]CTP, according to the manufacturer's protocol. Hybridization was carried out overnight (approximately 18 h) at 65°C in aqueous buffer (7% w/v SDS with 0.5 M phosphate). After hybridization, blots were washed using 1% w/v SDS at 65°C for 5 to 6 times until the background signal was low.

In situ RNA hybridization

ISH was carried out using paraffin sections (5 µm) of embryonic, postnatal and adult brains (E11.5, E13.5, E15.5, E17.5, P1.5 and P150) and a related [³⁵S]UTP-labeled complementary RNA probe (Additional data file 8) as described previously (Section G in Additional data file 1) [123].

Screening of Sox4 and Sox11 sense and antisense transcript expression in the adult mouse brain, organs and both the proliferating and differentiating P19 cells and neurospheres

Strand-specific RT-qPCR

Total RNA was extracted from harvested organs using the TRIzol® reagent (Invitrogen, Mulgrave, Victoria, Australia) according to the manufacturer's protocol. To avoid genomic DNA contamination, all isolated total RNA was treated with the recombinant DNase I enzyme provided by the DNA-free™ kit (Applied Biosystems, Scoresby, Victoria, Australia) according to the manufacturer's protocol. First strand cDNA synthesis was carried out using strand-specific primers followed by qPCR analysis as described above.

Embryonic neural stem cells grown as neurospheres

Mouse used for generation of neurospheres had a mixed genetic background including MF1, 129SvEv, C57BL/6 and CBA. Cerebral cortices from E14 embryos were dissected out into calcium-magnesium-free phosphate-buffered saline. The tissue was mechanically dissociated and centrifuged. The cells were plated in complete neuroculture medium (Section G in Additional data file 1) for 4 days followed by induction of neuronal differentiation. These cells were then plated on poly-D-lysine (catalogue number P6407, Sigma Aldrich, Castle Hill, New South Wales, Australia) and laminin (catalogue number 23017-015, Invitrogen, Mulgrave, Victoria, Australia) coated culture dishes in neuroculture medium with the presence of 2% (v/v) fetal bovine serum but not epidermal growth factor and basic fibroblast growth factor. The differentiation was allowed to proceed for 5 days. Total RNA was extracted from both proliferating and differentiating cells using TRIzol® reagent as described above.

P19 embryonal carcinoma cells

P19 mouse embryonal carcinoma cells were cultured and differentiated into neurons as described previously [124].

Briefly, P19 cell cultures were maintained in P19GM complete medium (Section G in Additional data file 1). For induction of neuronal differentiation, 1×10^6 P19 cells were cultured in suspension form using bacteriological Petri dishes. The P19GM medium with additional supplementation of 5×10^{-7} M all-trans retinoic acid (catalogue number R-2625; Sigma Aldrich, Castle Hill, New South Wales, Australia) was used for the induction. After 4 days, P19 cells formed embryoid body stages. Embryoid bodies were collected from suspension cultures and re-plated in adherent culture flasks in the P19GM medium with only 5% (v/v) fetal bovine serum and without retinoic acid supplementation. The cells were allowed to differentiate for 5 days. Total RNA was extracted from both proliferating and differentiating cells using TRIzol® reagent as described in above.

Abbreviations

CL: caudo-lateral; CM: caudo-medial; DAVID: Database for Annotation, Visualization and Integrated Discovery; DET: differentially expressed transcript/tag; E: embryonic day; EST: expressed sequence tag; GEO: Gene Expression Omnibus; IPA: Ingenuity Pathway Analysis; ISH: *in situ* hybridization; LTP: long term potentiation; miRNA: microRNA; NAT: natural antisense transcript; NSPC: neural stem/progenitor cell; OMIM: Online Mendelian Inheritance in Man; P: post-natal day; PET: paired-end diTag; RACE: rapid amplification of cDNA ends; RL: rostro-lateral; RM: rostro-medial; RT-qPCR: quantitative RT-PCR; SAGE: serial analysis of gene expression; SDS: sodium dodecyl sulphate; UCSC: University of California Santa Cruz; UTR: untranslated region.

Authors' contributions

KHL performed all the SAGE validation experiments. CAH, PZC and SST procured the mouse cerebral cortex and constructed the SAGE libraries. KHL, TB, LH and GKS designed, performed and supervised the SAGE, RT-qPCR and IPA analyses. KHL and TT performed all the ISH studies. KHL, KB, PSC, CNH and PQT carried out the expression studies on *Sox4* and *Sox11* transcripts. KHL, CAH and CNH drafted the manuscript. CAH, GKS, TT and HSS conceived of the study, and participated in its design and coordination. All authors read and approved the final manuscript.

Additional data files

The following additional data are available with the online version of this paper: analysis of SAGE, DETs, IPA, *Sox4* and *Sox11* genomic cluster analysis, and R script for implementing empirical Bayesian moderated *t*-test on multiple groups (Additional data file 1); SAGE tag information for 561 DETs (Additional data file 2); functional annotations clustering analysis using DAVID (Additional data file 3); RT-qPCR validation of DETs based on multiple comparisons between two developmental stages (E versus Ad, PN1.5 versus Ad and

E15.5 versus PN1.5) (Additional data file 4); RT-qPCR validation of gene clusters based on hierarchical clustering analysis (Additional data file 5); RT-qPCR validation of DETs based on the rostral versus caudal E15.5 cerebral cortex comparison (Additional data file 6); statistically significant over-represented genomic loci based on genomic clustering of tags (Additional data file 7); list of primers, probes, clones and assays designed for RT-qPCR, RACE, Southern, Northern and ISH analysis (Additional data file 8).

Acknowledgements

We thank Felix Ng, Sarah King and Caitlin Collins for technical assistance; Frank Weissenborn for assistance with tissue dissections; Ken Simpson and Keith Satterley for computational assistance. The P19 cell line was kindly provided by Dr K Jensen, School of Molecular and Biomedical Science, University of Adelaide, South Australia, Australia. This work was supported by a National Health and Medical Research Council fellowships (171601 and 461204 to HSS); National Health and Medical Research Council Grants 219176, 257501, 215201, and 257529 (to HSS, GKS, and S-ST), a fellowship from the Deutsche Forschungsgemeinschaft (to TB). PQT is a Pfizer Australia Research Fellow. K-HL was a recipient of the Melbourne International Fee Remission Scholarship (MIFRS) and Universiti Putra Malaysia Staff Training Scholarship (UPMSTS), and Adelaide International Fee Scholarship (AIFS) equivalent. The majority of DNA sequencing in production of SAGE libraries was performed by the Australian Genome Research Facility, which was established through the Commonwealth-funded Major National Research Facilities program.

References

- Miyashita-Lin EM, Hevner R, Wassarman KM, Martinez S, Rubenstein JL: **Early neocortical regionalization in the absence of thalamic innervation.** *Science* 1999, **285**:906-909.
- Molyneaux BJ, Arlotta P, Menezes JR, Macklis JD: **Neuronal subtype specification in the cerebral cortex.** *Nat Rev Neurosci* 2007, **8**:427-437.
- O'Leary DD, Chou SJ, Sahara S: **Area patterning of the mammalian cortex.** *Neuron* 2007, **56**:252-269.
- Gupta A, Tsai LH, Wynshaw-Boris A: **Life is a journey: a genetic look at neocortical development.** *Nat Rev Genet* 2002, **3**:342-355.
- Marin O, Rubenstein JL: **A long, remarkable journey: tangential migration in the telencephalon.** *Nat Rev Neurosci* 2001, **2**:780-790.
- Nadarajah B, Parnavelas JG: **Modes of neuronal migration in the developing cerebral cortex.** *Nat Rev Neurosci* 2002, **3**:423-432.
- Smart IH: **Histogenesis of the mesocortical area of the mouse telencephalon.** *J Anat* 1984, **138**:537-552.
- Uylings HBM, Van Eden CG, Parnavelas JG, Kalsbeek A: **The prenatal and postnatal development of rat cerebral cortex.** In *The Cerebral Cortex of the Rat* Edited by: Kolb B, Tees RC. Cambridge, MA: MIT Press; 1990:36-76.
- Bishop KM, Goudreau G, O'Leary DD: **Regulation of area identity in the mammalian neocortex by *Emx2* and *Pax6*.** *Science* 2000, **288**:344-349.
- Mallamaci A, Muzio L, Chan CH, Parnavelas J, Boncinelli E: **Area identity shifts in the early cerebral cortex of *Emx2*-/- mutant mice.** *Nat Neurosci* 2000, **3**:679-686.
- Armentano M, Chou SJ, Tomassy GS, Leingartner A, O'Leary DD, Studer M: **COUP-TFI regulates the balance of cortical patterning between frontal/motor and sensory areas.** *Nat Neurosci* 2007, **10**:1277-1286.
- Sahara S, Kawakami Y, Izpisua Belmonte JC, O'Leary DD: **Sp8 exhibits reciprocal induction with *Fgf8* but has an opposing effect on anterior-posterior cortical area patterning.** *Neural Dev* 2007, **2**:10.
- Zembrzycki A, Griesel G, Stoykova A, Mansouri A: **Genetic interplay between the transcription factors *Sp8* and *Emx2* in the patterning of the forebrain.** *Neural Dev* 2007, **2**:8.
- Bulfone A, Puelles L, Porteus MH, Frohman MA, Martin GR, Rubenstein JL: **Spatially restricted expression of *Dlx-1*, *Dlx-2* (Tes-**

- 1), Gbx-2, and Wnt-3 in the embryonic day 12.5 mouse forebrain defines potential transverse and longitudinal segmental boundaries. *J Neurosci* 1993, **13**:3155-3172.
15. Gaszner M, Felsenfeld G: **Insulators: exploiting transcriptional and epigenetic mechanisms.** *Nat Rev Genet* 2006, **7**:703-713.
16. Kadonaga JT: **Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines.** *Cell* 1998, **92**:307-313.
17. Ladd-Acosta C, Pevsner J, Sabuncian S, Yolken RH, Webster MJ, Dinkins T, Callinan PA, Fan JB, Potash JB, Feinberg AP: **DNA methylation signatures within the human brain.** *Am J Hum Genet* 2007, **81**:1304-1315.
18. Fox AH, Lam YW, Leung AK, Lyon CE, Andersen J, Mann M, Lamond AI: **Paraspeckles: a novel nuclear domain.** *Curr Biol* 2002, **12**:13-25.
19. Prasanth KV, Prasanth SG, Xuan Z, Hearn S, Freier SM, Bennett CF, Zhang MQ, Spector DL: **Regulating gene expression through RNA nuclear retention.** *Cell* 2005, **123**:249-263.
20. Schrott GM, Tuebing F, Nigh EA, Kane CG, Sabatini ME, Kiebler M, Greenberg ME: **A brain-specific microRNA regulates dendritic spine development.** *Nature* 2006, **439**:283-289.
21. Stark KL, Xu B, Bagchi A, Lai WS, Liu H, Hsu R, Wan X, Pavlidis P, Mills AA, Karayiorgou M, Gogos JA: **Altered brain microRNA biogenesis contributes to phenotypic deficits in a 22q11-deletion mouse model.** *Nat Genet* 2008, **40**:751-760.
22. Bak M, Silahatoglu A, Moller M, Christensen M, Rath MF, Skryabin B, Tommerup N, Kauppinen S: **MicroRNA expression in the adult mouse central nervous system.** *Rna* 2008, **14**:432-444.
23. Tochitani S, Hayashizaki Y: **Nlrx2.2 antisense RNA overexpression enhanced oligodendrocytic differentiation.** *Biochem Biophys Res Commun* 2008, **372**:691-696.
24. Smalheiser NR, Lugli G, Torvik VI, Mise N, Ikeda R, Abe K: **Natural antisense transcripts are co-expressed with sense mRNAs in synaptoneurosomes of adult mouse forebrain.** *Neurosci Res* 2008, **62**:236-239.
25. Saha S, Sparks AB, Rago C, Akmaev V, Wang CJ, Vogelstein B, Kinzler KW, Velculescu VE: **Using the transcriptome to annotate the genome.** *Nat Biotechnol* 2002, **20**:508-512.
26. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW: **Serial analysis of gene expression.** *Science* 1995, **270**:484-487.
27. Gunnersen JM, Augustine C, Spirkoska V, Kim M, Brown M, Tan SS: **Global analysis of gene expression patterns in developing mouse neocortex using serial analysis of gene expression.** *Mol Cell Neurosci* 2002, **19**:560-573.
28. **NCBI Gene Expression Omnibus.** [http://www.ncbi.nlm.nih.gov/geo/]
29. Boon WM, Beissbarth T, Hyde L, Smyth G, Gunnersen J, Denton DA, Scott H, Tan SS: **A comparative analysis of transcribed genes in the mouse hypothalamus and neocortex reveals chromosomal clustering.** *Proc Natl Acad Sci USA* 2004, **101**:14972-14977.
30. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA: **DAVID: Database for Annotation, Visualization, and Integrated Discovery.** *Genome Biol* 2003, **4**:P3.
31. Yuan X, Chittajallu R, Belachew S, Anderson S, McBain CJ, Gallo V: **Expression of the green fluorescent protein in the oligodendrocyte lineage: a transgenic mouse for developmental and physiological studies.** *J Neurosci Res* 2002, **70**:529-545.
32. **Allen Institute for Brain Science.** [http://mouse.brain-map.org/]
33. **Brain Gene Expression Map.** [http://www.stjudebgem.org/]
34. **Gene Expression Nervous System Atlas.** [http://www.genetax.org/]
35. **GenePaint.org.** [http://www.genepaint.org/]
36. Ino H, Chiba T: **Expression of proliferating cell nuclear antigen (PCNA) in the adult and developing mouse nervous system.** *Brain Res Mol Brain Res* 2000, **78**:163-174.
37. Lopez-Coviella I, Follett MT, Mellott TJ, Kovacheva VP, Slack BE, Diels V, Berse B, Thies RS, Blusztajn JK: **Bone morphogenetic protein 9 induces the transcriptome of basal forebrain cholinergic neurons.** *Proc Natl Acad Sci USA* 2005, **102**:6984-6989.
38. Bhat NR, Zhang P, Bhat AN: **The expression of myristoylated alanine-rich C-kinase substrate in oligodendrocytes is developmentally regulated.** *Dev Neurosci* 1995, **17**:256-263.
39. Calabrese B, Halpain S: **Essential role for the PKC target MARCKS in maintaining dendritic spine morphology.** *Neuron* 2005, **48**:77-90.
40. Lopez-Bendito G, Flames N, Ma L, Fouquet C, Di Meglio T, Chedotal A, Tessier-Lavigne M, Marin O: **Robo1 and Robo2 cooperate to control the guidance of major axonal tracts in the mammalian forebrain.** *J Neurosci* 2007, **27**:3395-3407.
41. Sher F, Rossler R, Brouwer N, Balasubramanian V, Boddeke E, Copray S: **Differentiation of neural stem cells into oligodendrocytes: involvement of the polycomb group protein Ezh2.** *Stem Cells* 2008, **26**:2875-2883.
42. Mermelstein F, Yeung K, Cao J, Inostroza JA, Erdjument-Bromage H, Egelson K, Landsman D, Levitt P, Tempst P, Reinberg D: **Requirement of a corepressor for Drl-mediated repression of transcription.** *Genes Dev* 1996, **10**:1033-1048.
43. Beiswanger CM, Diegmann MH, Novak RF, Philbert MA, Graessle TL, Reuhl KR, Lowndes HE: **Developmental changes in the cellular distribution of glutathione and glutathione S-transferases in the murine nervous system.** *Neurotoxicology* 1995, **16**:425-440.
44. Saha S, Datta K, Rangarajan P: **Characterization of mouse neuronal Ca2+/calmodulin kinase II inhibitor alpha.** *Brain Res* 2007, **1148**:38-42.
45. Wilson CA, Murphy DD, Giasson BI, Zhang B, Trojanowski JQ, Lee VM: **Degradative organelles containing mislocalized alpha- and beta-synuclein proliferate in presenilin-1 null neurons.** *J Cell Biol* 2004, **165**:335-346.
46. Bachoo RM, Kim RS, Ligon KL, Maher EA, Brennan C, Billings N, Chan S, Li C, Rowitch DH, Wong WH, DePinto RA: **Molecular diversity of astrocytes with implications for neurological disorders.** *Proc Natl Acad Sci USA* 2004, **101**:8384-8389.
47. Lencesova L, O'Neill A, Resneck WG, Bloch RJ, Blaustein MP: **Plasma membrane-cytoskeleton-endoplasmic reticulum complexes in neurons and astrocytes.** *J Biol Chem* 2004, **279**:2885-2893.
48. Zeng H, Chattarji S, Barbarosie M, Rondi-Reig L, Philpot BD, Miyakawa T, Bear MF, Tonegawa S: **Forebrain-specific calcineurin knockout selectively impairs bidirectional synaptic plasticity and working/episodic-like memory.** *Cell* 2001, **107**:617-629.
49. Kim PM, Aizawa H, Kim PS, Huang AS, Wickramasinghe SR, Kashani AH, Barrow RK, Haganir R, Ghosh A, Snyder SH: **Serine racemase: activation by glutamate neurotransmission via glutamate receptor interacting protein and mediation of neuronal migration.** *Proc Natl Acad Sci USA* 2005, **102**:2105-2110.
50. Mandir AS, Poitras MF, Berliner AR, Herring WJ, Guastella DB, Feldman A, Poirier GG, Wang ZQ, Dawson TM, Dawson VL: **NMDA but not non-NMDA excitotoxicity is mediated by Poly(ADP-ribose) polymerase.** *J Neurosci* 2000, **20**:8005-8011.
51. Yang X, Skoff RP: **Proteolipid protein regulates the survival and differentiation of oligodendrocytes.** *J Neurosci* 1997, **17**:2056-2070.
52. Southwood CM, Peppi M, Dryden S, Tainsky MA, Gow A: **Microtubule deacetylases, SirT2 and HDAC6, in the nervous system.** *Neurochem Res* 2007, **32**:187-195.
53. Niciu MJ, Ma XM, El Meskini R, Ronnett GV, Mains RE, Eipper BA: **Developmental changes in the expression of ATP7A during a critical period in postnatal neurodevelopment.** *Neuroscience* 2006, **139**:947-964.
54. Sanosaka T, Namihira M, Asano H, Kohyama J, Aisaki K, Igarashi K, Kanno J, Nakashima K: **Identification of genes that restrict astrocyte differentiation of midgestational neural precursor cells.** *Neuroscience* 2008, **155**:780-788.
55. Mbebi C, See V, Mercken L, Pradier L, Muller U, Loeffler JP: **Amyloid precursor protein family-induced neuronal death is mediated by impairment of the neuroprotective calcium/calmodulin protein kinase IV-dependent signaling pathway.** *J Biol Chem* 2002, **277**:20979-20990.
56. Jinnah HA, Hess EJ, Wilson MC, Gage FH, Friedmann T: **Localization of hypoxanthine-guanine phosphoribosyltransferase messenger RNA in the mouse brain by in situ hybridization.** *Mol Cell Neurosci* 1992, **3**:64-78.
57. Poet M, Kornak U, Schweizer R, Zdebek AA, Scheel O, Hoelter S, Wurst W, Schmitt A, Fuhrmann JC, Planells-Cases R, Mole SE, Hubner CA, Jentsch TJ: **Lysosomal storage disease upon disruption of the neuronal chloride transport protein CIC-6.** *Proc Natl Acad Sci USA* 2006, **103**:13854-13859.
58. Palfi A, Kortvely E, Fekete E, Kovacs B, Varszegi S, Gulya K: **Differential calmodulin gene expression in the rodent brain.** *Life Sci* 2002, **70**:2829-2855.
59. Toyo-oka K, Shionoya A, Gambello MJ, Cardoso C, Leventer R, Ward HL, Ayala R, Tsai LH, Dobyns W, Ledbetter D, Hirotsune S, Wynshaw-Boris A: **14-3-3epsilon is important for neuronal migration by binding to NUDEL: a molecular explanation for Miller-Dieker syndrome.** *Nat Genet* 2003, **34**:274-285.
60. Kuo TY, Hsueh YP: **Expression of zinc finger transcription fac-**

- tor BclIIA/Evi9/CTIP1 in rat brain. *J Neurosci Res* 2007, **85**:1628-1636.
61. Li Q, Lee JA, Black DL: **Neuronal regulation of alternative pre-mRNA splicing.** *Nat Rev Neurosci* 2007, **8**:819-831.
 62. Zhou C, Tsai SY, Tsai MJ: **COUP-TFI: an intrinsic factor for early regionalization of the neocortex.** *Genes Dev* 2001, **15**:2054-2059.
 63. **UCSC Genome Bioinformatics.** [http://genome.ucsc.edu/]
 64. **Ensembl Genome Browser.** [http://www.ensembl.org/]
 65. Bergsland M, Werme M, Malewicz M, Perlmann T, Muhr J: **The establishment of neuronal properties is controlled by Sox4 and Sox11.** *Genes Dev* 2006, **20**:3475-3486.
 66. Potzner MR, Griffel C, Lutjen-Drecoll E, Bosl MR, Wegner M, Sock E: **Prolonged Sox4 expression in oligodendrocytes interferes with normal myelination in the central nervous system.** *Mol Cell Biol* 2007, **27**:5316-5326.
 67. Velculescu VE, Madden SL, Zhang L, Lash AE, Yu J, Rago C, Lal A, Wang CJ, Beaudry GA, Ciriello KM, Cook BP, Dufault MR, Ferguson AT, Gao Y, He TC, Hermeking H, Hiraldo SK, Hwang PM, Lopez MA, Luderer HF, Mathews B, Petroziello JM, Polyak K, Zawel L, Kinzler KW, et al.: **Analysis of human transcriptomes.** *Nat Genet* 1999, **23**:387-388.
 68. Brenner S, Johnson M, Bridgman J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, George D, Eletr S, Albrecht G, Vermaas E, Williams SR, Moon K, Burcham T, Pallas M, DuBridge RB, Kirchner J, Fearon K, Mao J, Corcoran K: **Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays.** *Nat Biotechnol* 2000, **18**:630-634.
 69. **The Cancer Genome Anatomy Project: SAGE Genie.** [http://cgap.nci.nih.gov/SAGE]
 70. Ng P, Tan JJ, Ooi HS, Lee YL, Chiu KP, Fullwood MJ, Srinivasan KG, Perbost C, Du L, Sung WK, Wei CL, Ruan Y: **Multiplex sequencing of paired-end ditags (MS-PET): a strategy for the ultra-high-throughput analysis of transcriptomes and genomes.** *Nucleic Acids Res* 2006, **34**:e84.
 71. Shiraki T, Kondo S, Katayama S, Waki K, Kasukawa T, Kawaji H, Kodzius R, Watahiki A, Nakamura M, Arakawa T, Fukuda S, Sasaki D, Podhajski A, Harbers M, Kawai J, Carninci P, Hayashizaki Y: **Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage.** *Proc Natl Acad Sci USA* 2003, **100**:15776-15781.
 72. Wei CL, Ng P, Chiu KP, Wong CH, Ang CC, Lipovich L, Liu ET, Ruan Y: **5' Long serial analysis of gene expression (LongSAGE) and 3' LongSAGE for transcriptome characterization and genome annotation.** *Proc Natl Acad Sci USA* 2004, **101**:11701-11706.
 73. Chenn A: **Wnt/beta-catenin signaling in cerebral cortical development.** *Organogenesis* 2008, **4**:76-80.
 74. Bafico A, Gazit A, Pramila T, Finch PV, Yaniv A, Aaronson SA: **Interaction of frizzled related protein (FRP) with Wnt ligands and the frizzled receptor suggests alternative mechanisms for FRP inhibition of Wnt signaling.** *J Biol Chem* 1999, **274**:16180-16187.
 75. Morkel M, Huelsken J, Wakamiya M, Ding J, Wetering M van de, Clevers H, Taketo MM, Behringer RR, Shen MM, Birchmeier W: **Beta-catenin regulates Cripto- and Wnt3-dependent gene expression programs in mouse axis and mesoderm formation.** *Development* 2003, **130**:6283-6294.
 76. Fevr T, Robine S, Louvard D, Huelsken J: **Wnt/beta-catenin is essential for intestinal homeostasis and maintenance of intestinal stem cells.** *Mol Cell Biol* 2007, **27**:7551-7559.
 77. Machon O, Backman M, Machonova O, Kozmik Z, Vacik T, Andersen L, Krauss S: **A dynamic gradient of Wnt signaling controls initiation of neurogenesis in the mammalian cortex and cellular specification in the hippocampus.** *Dev Biol* 2007, **311**:223-237.
 78. Woodhead GJ, Mutch CA, Olson EC, Chenn A: **Cell-autonomous beta-catenin signaling regulates cortical precursor proliferation.** *J Neurosci* 2006, **26**:12620-12630.
 79. Lisman J, Schulman H, Cline H: **The molecular basis of CaMKII function in synaptic and behavioural memory.** *Nat Rev Neurosci* 2002, **3**:175-190.
 80. Pak JH, Huang FL, Li J, Balschun D, Reymann KG, Chiang C, Westphal H, Huang KP: **Involvement of neurogranin in the modulation of calcium/calmodulin-dependent protein kinase II, synaptic plasticity, and spatial learning: a study with knockout mice.** *Proc Natl Acad Sci USA* 2000, **97**:11232-11237.
 81. Lepicard EM, Mizuno K, Antunes-Martins A, von Hertzen LS, Giese KP: **An endogenous inhibitor of calcium/calmodulin-dependent kinase II is up-regulated during consolidation of fear memory.** *Eur J Neurosci* 2006, **23**:3063-3070.
 82. Young-Pearse TL, Bai J, Chang R, Zheng JB, LoTurco JJ, Selkoe DJ: **A critical function for beta-amyloid precursor protein in neuronal migration revealed by in utero RNA interference.** *J Neurosci* 2007, **27**:14459-14469.
 83. Alonso A, Zaidi T, Novak M, Grundke-Iqbal I, Iqbal K: **Hyperphosphorylation induces self-assembly of tau into tangles of paired helical filaments/straight filaments.** *Proc Natl Acad Sci USA* 2001, **98**:6923-6928.
 84. Rapoport M, Dawson HN, Binder LI, Vitek MP, Ferreira A: **Tau is essential to beta-amyloid-induced neurotoxicity.** *Proc Natl Acad Sci USA* 2002, **99**:6364-6369.
 85. Carter CJ: **eIF2B and oligodendrocyte survival: where nature and nurture meet in bipolar disorder and schizophrenia?** *Schizophr Bull* 2007, **33**:1343-1353.
 86. Craddock N, O'Donovan MC, Owen MJ: **The genetics of schizophrenia and bipolar disorder: dissecting psychosis.** *J Med Genet* 2005, **42**:193-204.
 87. Hodges A, Strand AD, Aragaki AK, Kuhn A, Sengstad T, Hughes G, Elliston LA, Hartog C, Goldstein DR, Thu D, Hollingsworth ZR, Collin F, Synek B, Holmans PA, Young AB, Wexler NS, Delorenzi M, Kooperberg C, Augood SJ, Faull RL, Olson JM, Jones L, Luthi-Carter R: **Regional and cellular gene expression changes in human Huntington's disease brain.** *Hum Mol Genet* 2006, **15**:965-977.
 88. Luthi-Carter R, Strand A, Peters NL, Solano SM, Hollingsworth ZR, Menon AS, Frey AS, Spektor BS, Penney EB, Schilling G, Ross CA, Borchelt DR, Tapscott SJ, Young AB, Cha JH, Olson JM: **Decreased expression of striatal signaling genes in a mouse model of Huntington's disease.** *Hum Mol Genet* 2000, **9**:1259-1271.
 89. Lovestone S, Hodgson S, Sham P, Differ AM, Levy R: **Familial psychiatric presentation of Huntington's disease.** *J Med Genet* 1996, **33**:128-131.
 90. Shiwach R: **Psychopathology in Huntington's disease patients.** *Acta Psychiatr Scand* 1994, **90**:241-246.
 91. Chandra S, Gallardo G, Fernandez-Chacon R, Schluter OM, Sudhof TC: **Alpha-synuclein cooperates with CSPA in preventing neurodegeneration.** *Cell* 2005, **123**:383-396.
 92. Evert BO, Wullner U, Klockgether T: **Cell death in polyglutamine diseases.** *Cell Tissue Res* 2000, **301**:189-204.
 93. Morishima Y, Gotoh Y, Zieg J, Barrett T, Takano H, Flavell R, Davis RJ, Shirasaki Y, Greenberg ME: **Beta-amyloid induces neuronal apoptosis via a mechanism that involves the c-Jun N-terminal kinase pathway and the induction of Fas ligand.** *J Neurosci* 2001, **21**:7551-7560.
 94. Mallamaci A, Stoykova A: **Gene networks controlling early cerebral cortex arealization.** *Eur J Neurosci* 2006, **23**:847-856.
 95. Philip D, Huff T, Gho YS, Hannappel E, Kleinman HK: **The actin binding site on thymosin beta4 promotes angiogenesis.** *Faseb J* 2003, **17**:2103-2105.
 96. Bock-Marquette I, Saxena A, White MD, Dimaio JM, Srivastava D: **Thymosin beta4 activates integrin-linked kinase and promotes cardiac cell migration, survival and cardiac repair.** *Nature* 2004, **432**:466-472.
 97. Roth LW, Bormann P, Bonnet A, Reinhard E: **beta-thymosin is required for axonal tract formation in developing zebrafish brain.** *Development* 1999, **126**:1365-1374.
 98. Lopez-Bendito G, Molnar Z: **Thalamocortical development: how are we going to get there?** *Nat Rev Neurosci* 2003, **4**:276-289.
 99. Winter J, Kunath M, Roepcke S, Krause S, Schneider R, Schweiger S: **Alternative polyadenylation signals and promoters act in concert to control tissue-specific expression of the Opitz Syndrome gene MIDI.** *BMC Mol Biol* 2007, **8**:105.
 100. Abdel Wahab N, Gibbs J, Mason RM: **Regulation of gene expression by alternative polyadenylation and mRNA instability in hyperglycaemic mesangial cells.** *Biochem J* 1998, **336**:405-411.
 101. Wilusz CJ, Wilusz J: **Bringing the role of mRNA decay in the control of gene expression into focus.** *Trends Genet* 2004, **20**:491-497.
 102. Lu C, Jeong D-H, Kulkarni K, Pillay M, Nobuta K, German R, Thatcher SR, Maher C, Zhang L, Ware D, Liu B, Cao X, Meyers BC, Green PJ: **Genome-wide analysis for discovery of rice microRNAs reveals natural antisense microRNAs (nat-miRNAs).** *Proc Natl Acad Sci USA* 2008, **105**:4951-4956.
 103. **miRBase.** [http://microrna.sanger.ac.uk/]

104. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, Surani MA, Sakaki Y, Sasaki H: **Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes.** *Nature* 2008, **453**:539-543.
105. Hoser M, Potzner MR, Koch JM, Bosl MR, Wegner M, Sock E: **Sox12 deletion in the mouse reveals nonreciprocal redundancy with the related Sox4 and Sox11 transcription factors.** *Mol Cell Biol* 2008, **28**:4675-4687.
106. Cheung M, Abu-Elmagd M, Clevers H, Scotting PJ: **Roles of Sox4 in central nervous system development.** *Brain Res Mol Brain Res* 2000, **79**:180-191.
107. Jin Z, Liu L, Bian W, Chen Y, Xu G, Cheng L, Jing N: **Different transcription factors regulate nestin gene expression during P19 cell neural differentiation and central nervous system development.** *J Biol Chem* 2009, **284**:8160-8173.
108. Lee MS, Jun DH, Hwang CI, Park SS, Kang JJ, Park HS, Kim J, Kim JH, Seo JS, Park WY: **Selection of neural differentiation-specific genes by comparing profiles of random differentiation.** *Stem Cells* 2006, **24**:1946-1955.
109. Dy P, Penzo-Mendez A, Wang H, Pedraza CE, Macklin WB, Lefebvre V: **The three SoxC proteins - Sox4, Sox11 and Sox12 - exhibit overlapping expression patterns and molecular properties.** *Nucleic Acids Res* 2008, **36**:3101-3117.
110. **Virtual Mouse Necropsy.** [<http://www3.niaid.nih.gov/labs/about/labs/cmb/InfectiousDiseasePathogenesisSection/mouseNecropsy/>]
111. Beissbarth T, Hyde L, Smyth GK, Job C, Boon WM, Tan SS, Scott HS, Speed TP: **Statistical modeling of sequencing errors in SAGE libraries.** *Bioinformatics* 2004, **20**(Suppl 1):131-139.
112. **Bioconductor.** [<http://www.bioconductor.org/>]
113. **Mapping SAGE Tags to Genes Using EST and Genomic Information.** [<http://www.mbgproject.org/tagmapping/>]
114. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.** *J R Statist Soc* 1995, **57**:289-300.
115. Vencio RZ, Brentani H, Patrao DF, Pereira CA: **Bayesian model accounting for within-class biological variability in serial analysis of gene expression (SAGE).** *BMC Bioinformatics* 2004, **5**:119.
116. **The R Project for Statistical Computing.** [<http://www.r-project.org/>]
117. Gotter J, Brors B, Hergenhausen M, Kyewski B: **Medullary epithelial cells of the human thymus express a highly diverse selection of tissue-specific genes colocalized in chromosomal clusters.** *J Exp Med* 2004, **199**:155-166.
118. **Ingenuity Pathway Analysis.** [<http://www.ingenuity.com>]
119. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, Ragg T: **The RIN: an RNA integrity number for assigning integrity values to RNA measurements.** *BMC Mol Biol* 2006, **7**:3.
120. **Assay Design Center for Universal ProbeLibrary Assay.** [<http://www.roche-applied-science.com/sis/rtpcr/upl/ezhome.html>]
121. Luu-The V, Paquet N, Calvo E, Cumps J: **Improved real-time RT-PCR method for high-throughput measurements using second derivative calculation and double correction.** *Biotechniques* 2005, **38**:287-293.
122. Smyth GK: **Linear models and empirical bayes methods for assessing differential expression in microarray experiments.** *Stat Appl Genet Mol Biol* 2004, **3**: Article3
123. Thomas T, Voss AK, Chowdhury K, Gruss P: **Querkopf, a MYST family histone acetyltransferase, is required for normal cerebral cortex development.** *Development* 2000, **127**:2537-2548.
124. Rudnicki MA, McBurney MW: **Cell culture methods and induction of differentiation of embryonal carcinoma cell lines.** In *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach* Edited by: Robertson EJ. Oxford: IRL Press; 1987:19-49.